University of Novi Sad
Center for Applied Statistics

Vesna Jevremović

Rečnik statističkih termina

Drugo izdanje

Glossary of Statistical Terms

MMXIX

TEMPUS

SERBIA

Vesna
Jevremović

REČNIK
STATISTIČKIH
TERMINA

DRUGO IZDANjE

TEMPUS „Master in Applied Statistics“

Beograd
MMXIX

REČNIK STATISTIČKIH TERMINA
Autor: Vesna Jevremović

RECENZENTI
Prof. dr Jovan Mališić, Beograd, i prof. dr Andrija Tepavčević, Novi Sad

LEKTURA I OBRADA TEKSTA
Vladimir Živanović

IZDAJE I ŠTAMPA
Republički zavod za statistiku, Beograd, Milana Rakića 5

TIRAŽ
100 primeraka

Prvo izdanje publikacije objavljeno je 2013. godine, u okviru Tempus projekta „Master program iz primenjene statistike“ 511140-Tempus-1-2010-1-RS-Tempus-JPCR.

Drugo izdanje publikacije objavljuje Republički zavod za statistiku 2019. godine.

„Ovaj projekat je finansiran od strane Evropske komisije. Ova publikacija odražava samo stavove autora i Komisija ne može biti odgovorna za bilo kakvu upotrebu informacija koje se u publikaciji sadrže“.

“This project has been funded whit support from the European Commission. This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.”

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

311(038)=00
821.163.41'276.6:331

ЈЕВРЕМОВИЋ, Весна, 1955-
Rečnik statističkih termina / Vesna Jevremović. - 2. izd. - Beograd : Republički zavod za statistiku, 2019 (Beograd : Republički zavod za statistiku). - 318 str. : ilustr. ; 21 cm

Na nasl. str.: TEMPUS "Master in Applied Statistics". - Tiraž 100 . - Napomene uz tekst. - Oznake korišćene u knjizi: str. 266-273.

ISBN 978-86-6161-188-9

а) Статистика -- Вишејезички речници
б) Српски језик -- Статистичка терминологија

COBISS.SR-ID 280513804

Rečnik statističkih termina

Речник статистичких термина

Glossary of statistical terms

Glossaire des termes statistiques

Словарь статистических терминов

统计学术语辞典

معجم المصـــطلحات الاحصـــائية

Λεξικό στατιστικών όρων

Wörterbuch der statistischen Begriffe

Glosario de términos estadísticos

Glossario dei termini statistici

Slovník štatistických pojmov

Glosar statističnih izrazov

Statisztikai kifejezések szótára

Glossarium terminorum statisticorum

Predgovor

„**Rečnik statističkih termina**“ je odraz autorove ljubavi prema stranim jezicima, u koje možemo ubrojati i matematiku i njenu važnu oblast – statistiku. Zaista, svaka struka ima svoje specifične termine koji se ponekad uklapaju u standardne termine običnog jezika, ali su koji put i različitih, pa čak i suprotnih značenja.

Na srpskom govornom području je malo publikacija objavljenih na temu matematičkih termina, a još manje na temu statističke terminologije.

Zamisao autora je da rečnik ima više tomova, pri čemu bi se ovaj, prvi tom odnosio na osnovne pojmove, pretežno teorijske statistike, ali obuhvatao i neke termine deskriptivne statistike. U sledećem tomu bi svoje mesto našli većinom termini vezani za primenu računara u statistici, kao i termini raznih disciplina koje koriste statistiku.

Izabrani statistički termini su prevedeni na dvanaest jezika – pet velikih svetskih jezika (engleski, francuski, ruski, kineski i arapski), grčki kao jedan od jezika koji su postavili temelje evropske civilizacije i šest jezika zemalja partnera Srbije u okviru Tempus projekta MAS (nemački, španski, italijanski, slovački, slovenački i mađarski).

Korist od ovog rečnika, koji je zapravo pojmovnik, tj. nešto između rečnika, priručnika i udžbenika će, nadam se, imati ne samo oni koji se prvi put sreću sa statistikom, nego i oni koji u izvesnoj meri poznaju materiju o kojoj je reč.

Iako je ovo rečnik, bilo je neophodno da sadrži i matematičke formule, makar u najmanjoj mogućoj meri. Osim izdvojenih odrednica postoje i pojmovi koji se nalaze u okviru objašnjavanja drugih termina, a takođe se neki pojmovi ilustruju jednostavnim primerima. Detalji biografija pomenutih naučnika su takođe uvršteni u rečnik, a data su i etimološka objašnjenja stranih reči koje se koriste u statističkoj terminologiji na srpskom jeziku.

Svaka odrednica je data na dve ili eventualno četiri strane, tako što se na levoj, parnoj strani, nalaze termini latiničnim i ćiriličnim pismom, prevodi, povezani termini, njihova etimologija i druge napomene.

Alternativna hipoteza
Алтернативна хипотеза

EN	Alternative hypothesis
FR	Hypothèse alternative
RU	Альтернативная гипотеза
ZH	备择假设, bèizé jiǎshè
AR	الفرضية البديلة
EL	Εναλλακτική υπόθεση
DE	Gegenhypothese
ES	Hipótesis alternativa
IT	Ipotesi alternativa
SK	Alternatívna hypotéza
SL	Alternativna domneva
HU	Ellenhipotézis

Povezani termini:

Funkcija moći testa (86), greška druge vrste (96), greška prve vrste (98), kritična oblast (136), moć testa (175), nulta hipoteza (174), prag značajnosti (186), testiranje hipoteze (216)

Etimologija:

Alternativni < alternāre^{LA} (raditi na smenu) < alter^{LA} (drugi)

Hipoteza < ὑπόθεση^{EL} (staviti ispod, pretpostaviti) < ὑπό^{EL} (ispod) + τίθημι^{EL} (postaviti)

8

U testiranju statističkih hipoteza **alternativna hipoteza** je pretpostavka koju, na određeni način, suprotstavljamo hipotezi koju testiramo, tzv. nultoj hipotezi. Alternativna hipoteza može biti i prosta (*simple hypothesis*)[1] i složena[2], a njome se utvrđuje oblik kritične oblasti posmatranog testa. Alternativna hipoteza se često označava sa H_1. U slučaju da se hipoteza odnosi na parametar, onda je moguće postaviti više različitih alternativnih hipoteza, a za koju ćemo se opredeliti zavisi od same pojave koju proučavamo. Na primer, ako je nulta hipoteza da je matematičko očekivanje m posmatranog obeležja jednako m_0, onda alternativna može biti: $m > m_0$ ili $m \neq m_0$ ili $m = m_1$, gde je m_1 neki broj različit od m_0, itd.

Ako se nulta hipoteza odnosi na raspodelu posmatranog obeležja, onda je alternativna hipoteza često negacija nulte. Na primer, ako je nulta hipoteza da su posmatrana obeležja nezavisna, alternativna će biti da su obeležja zavisna, ili, ako je nulta hipoteza da je raspodela obeležja normalna raspodela $N(m, \sigma^2)$, onda je alternativna hipoteza da nije takva raspodela – što može da znači da je u pitanju nešto od navedenog: a) raspodela jeste normalna, ali matematičko očekivanje nije jednako navedenoj vrednosti, dok disperzija jeste, b) raspodela jeste normalna, očekivanje jeste jednako navedenoj vrednosti, ali disperzija nije, c) raspodela jeste normalna, ali ni očekivanje ni disperzija nisu jednaki navedenim vrednostima i d) raspodela nije normalna.

[1] Hipoteza je prosta ako jednoznačno određuje raspodelu obeležja iz familije dopustivih (mogućih) raspodela.
[2] Hipoteza je složena ako ne određuje jednoznačno raspodelu obeležja iz familije dopustivih (mogućih) raspodela.

9

Na desnoj strani su objašnjenja glavnih termina, kao i onih s njima povezanih, koji nemaju svoje posebno mesto u Rečniku. Takvi termini su takođe prevedeni, ali samo na engleski jezik.

Sve osnovne odrednice koje se pominju u tekstovima na neparnoj strani podvučene su punom linijom. Termini podvučeni isprekidanom linijom takođe predstavljaju deo statističke terminologije, ali nisu izdvojeni kao zasebne odrednice već su objašnjeni na toj stranici gde su spomenuti. Napomene i primeri za ilustraciju pojmova su dati u fusnotama.

Na kraju knjige nalazi se i spisak oznaka korišćenih u Rečniku i njihovi prevodi na engleski.

U radu na Rečniku veliku podršku i pomoć pružile su mi kolege: dr Fivos Papadimitriu iz Atine, prof. dr Tibor Poganj sa Pomorskog fakulteta u Rijeci, prof. dr Vladimir Janiš sa Univerziteta Mateja Bela u Banskoj Bistrici, prof. dr Anuška Ferligoj sa Univerziteta u Ljubljani, mr Ehfajed Šeneina i Halima Elfagihe iz Libije i druge kolege i prijatelji. Posebnu zahvalnost dugujem svom profesoru i mentoru prof. dr Jovanu Mališiću i svom sinu Vladimiru Živanoviću.

15. februar 2013.
U Tunisu, na obali Sredozemnog mora

Autor

Umesto predgovora drugom izdanju

Zahvaljujem Republičkom zavodu za statistiku i njegovom direktoru dr Miladinu Kovačeviću što su uvideli značaj ove publikacije i omogućili štampanje drugog izdanja.

20. oktobar 2019.

Autor

Jezici

Језици

EN	Engleski	English
FR	Francuski	Français
RU	Ruski	Русский
ZH	Kineski	中文
AR	Arapski	العربية
EL	Grčki	Ελληνικά
DE	Nemački	Deutsch
ES	Španski	Español
IT	Italijanski	Italiano
SK	Slovački	Slovenčina
SL	Slovenački	Slovenščina
HU	Mađarski	Magyar nyelv
LA	Latinski	Lingua latina

Alternativna hipoteza

Алтернативна хипотеза

EN	Alternative hypothesis
FR	Hypothèse alternative
RU	Альтернативная гипотеза
ZH	备择假设, bèizé jiǎshè
AR	الفرضية البديلة
EL	Εναλλακτική υπόθεση
DE	Gegenhypothese
ES	Hipótesis alternativa
IT	Ipotesi alternativa
SK	Alternatívna hypotéza
SL	Alternativna domneva
HU	Ellenhipotézis

Povezani termini:

Funkcija moći testa (86), greška druge vrste (96), greška prve vrste (98), kritična oblast (136), moć testa (175), nulta hipoteza (174), prag značajnosti (186), testiranje hipoteze (216)

Etimologija:

Alternativni < alternāre^{LA} (raditi na smenu) < alter^{LA} (drugi)

Hipoteza < ὑπόθεση^{EL} (staviti ispod, pretpostaviti) < ὑπό^{EL} (ispod) + τίθημι^{EL} (postaviti)

U testiranju statističkih hipoteza **alternativna hipoteza** je pretpostavka koju, na određeni način, suprotstavljamo hipotezi koju testiramo, tzv. nultoj hipotezi. Alternativna hipoteza može biti i prosta (*simple hypothesis*)[1] i složena[2], a njome se utvrđuje oblik kritične oblasti posmatranog testa. Alternativna hipoteza se često označava sa H_1. U slučaju da se hipoteza odnosi na parametar, onda je moguće postaviti više različitih alternativnih hipoteza, a za koju ćemo se opredeliti zavisi od same pojave koju proučavamo. Na primer, ako je nulta hipoteza da je matematičko očekivanje m posmatranog obeležja jednako m_0, onda alternativna može biti: $m > m_0$ ili $m \neq m_0$ ili $m = m_1$, gde je m_1 neki broj različit od m_0, itd.

Ako se nulta hipoteza odnosi na raspodelu posmatranog obeležja, onda je alternativna hipoteza često negacija nulte. Na primer, ako je nulta hipoteza da su posmatrana obeležja nezavisna, alternativna će biti da su obeležja zavisna, ili, ako je nulta hipoteza da je raspodela obeležja normalna raspodela $N(m, \sigma^2)$, onda je alternativna hipoteza da nije takva raspodela – što može da znači da je u pitanju nešto od navedenog: a) raspodela jeste normalna, ali matematičko očekivanje nije jednako navedenoj vrednosti, dok disperzija jeste, b) raspodela jeste normalna, očekivanje jeste jednako navedenoj vrednosti, ali disperzija nije, c) raspodela jeste normalna, ali ni očekivanje ni disperzija nisu jednaki navedenim vrednostima i d) raspodela nije normalna.

[1] Hipoteza je prosta ako jednoznačno određuje raspodelu obeležja iz familije dopustivih (mogućih) raspodela.
[2] Hipoteza je složena ako ne određuje jednoznačno raspodelu obeležja iz familije dopustivih (mogućih) raspodela.

Aposteriorna verovatnoća

Апостериорна вероватноћа

EN	Posterior probability
FR	Probabilité à posteriori
RU	Апостериорная вероятность
ZH	后验概率, hòuyàn gàilǜ
AR	الاحتمال اللاحق
EL	A posteriori πιθανότητα
DE	A-posteriori-Wahrscheinlichkeit
ES	Probabilidad a posteriori
IT	Probabilità a posteriori
SK	Aposteriórna pravdepodobnosť
SL	Posteriorna verjetnost
HU	A posteriori (utólagos) valószinűség

Povezani termini:

Apriorna verovatnoća (12), Bajesova formula (28), bajesovsko zaključivanje (30), uslovna verovatnoća (222)

Etimologija:

A posteriōrī^{LA} (ono što dolazi posle) < posterus^{LA} (sledeći)

Neka su $H_1, H_2,..., H_n$ slučajni događaji koji čine potpun sistem događaja[3] i neka je A neki događaj iz istog prostora elementarnih ishoda. Može se posmatrati realizacija događaja A u funkciji realizacije događaja (hipoteza) $H_1, H_2,..., H_n$. Takođe može da se posmatra i realizacija hipoteza, ako se zna da se realizovao događaj A. Tada možemo odrediti uslovne verovatnoće $P(H_k \mid A)$, za svako $k = 1, 2,..., n$. Te verovatnoće[4] se nazivaju **aposteriorne verovatnoće** hipoteza[5] i računaju se po Bajesovoj formuli[6]:

$$P(H_k \mid A) = \frac{P(H_k)P(A \mid H_k)}{\sum_{j=1}^{n} P(H_j)P(A \mid H_j)}.$$

Iz ove definicije jasno proizilazi da je zbir aposteriornih verovatnoća svih hipoteza u odnosu na isti događaj A, jednak 1.

U bajesovskom pristupu pri donošenju zaključaka o populaciji (tj. bajesovsko zaključivanje) bitnu ulogu imaju uslovne verovatnoće i s njima povezane uslovne raspodele.

[3] Događaji $H_1, H_2,..., H_n$ su disjunktni u parovima a njihova unija je siguran događaj. Npr. pri bacanju kockice možemo posmatrati potpun sistem događaja koji čine: H_1 kada je dobijen broj koji pri deljenju sa tri daje ostatak 1, H_2 kada je dobijen broj koji pri deljenju sa tri daje ostatak 2 i H_3 kada je dobijen broj deljiv sa 3.

[4] Svaka od njih može biti manja, veća ili jednaka verovatnoći hipoteze H_k, $k = 1, 2,..., n$.

[5] Zbir svih aposteriornih verovatnoća je jednak 1, što znači da u zadacima možemo izračunati n–1 jednostavniju verovatnoću, a preostalu dobiti kao komplement do 1.

[6] Izraz u imeniocu je tzv. formula potpune verovatnoće i vrednost tog izraza jednaka je verovatnoći događaja A.

Apriorna verovatnoća
Априорна вероватноћа

EN Prior probability
FR Probabilité à priori
RU Априорная вероятность
ZH 先验概率, xiānyàn gàilǜ
AR الاحتمال الأولي
EL A priori πιθανότητα
DE A-priori-Wahrscheinlichkeit
ES Probabilidad a priori
IT Probabilità a priori
SK Apriórna pravdepodobnosť
SL Apriorna verjetnost
HU A priori (előzetes) valószínűség

Povezani termini:

Aposteriorna verovatnoća (10), Bajesova formula (28), bajesovsko zaključivanje (30), uslovna verovatnoća (222)

Etimologija:

A priōrī^{LA} (ono što dolazi pre) < prior^{LA} (prethodni). Latinska reč *prior* i srpska reč *pre* imaju zajednički koren u proto-indoevropskom *prŏ (pre, preko).

Neka su H_1, H_2, ..., H_n slučajni događaji koji čine potpun sistem događaja[7] i neka je A neki događaj iz istog prostora elementarnih ishoda. Može se posmatrati realizacija događaja A u funkciji realizacije događaja H_1, H_2,..., H_n, koji se nazivaju hipoteze. Kada se smatra da su verovatnoće $P(H_j)$, $j = 1,..., n$ poznate, one se nazivaju **apriorne verovatnoće** hipoteza.

U matematičkoj statistici pri posmatranju grupnog ili stratifikovanog uzorka se populacija razlaže na disjunktne delove (analogon potpunog sistema događaja), a obimi tih grupa ili stratuma u odnosu na obim populacije odgovaraju apriornim verovatnoćama.

Kad se nepoznati parametar u raspodeli obeležja ocenjuje po Bajesovoj metodi, tada se pretpostavlja da on predstavlja slučajnu veličinu sa izvesnom, unapred datom raspodelom, koja se naziva apriorna raspodela.

---

[7] Događaji H_1, H_2,..., H_n čine potpun sistem događaja ako su disjunktni u parovima, tj. ako je $H_j \cap H_k = \Phi$, $j \neq k$, a njihova unija je siguran događaj. Npr. pri kontroli kvaliteta, ako su svi proizvodi u jednakim kutijama i imamo n kutija, a na slučajan način biramo jednu radi kontrole, onda hipotezama H_1, H_2,..., H_n odgovaraju redni brojevi tih kutija, a $P(H_j) = 1/n$, $j = 1,...,$ n.

Apsolutno srednje odstupanje

Апсолутно средње одступање

EN Mean absolute deviation
FR Erreur absolue moyenne
RU Среднее абсолютное отклонение
ZH 平均绝对误差, píngjūn juéduì wùchā
AR متوسط الانحرافات المطلق
EL Μέση απόλυτη απόκλιση
DE Durchschnittliche Abweichung
ES Desviación promedia
IT Scarto medio assoluto
SK Stredná absolútna odchýlka
SL Povprečni absolutni odklon
HU Átlagos abszolút hiba

Povezani termini:

Medijana (154), srednja vrednost, uzorak (240)

Etimologija:

Apsolutan < absolūtus^{LA} (slobodan od, nevezan) < ab^{LA} (od) + solvō^{LA} (odvezati, rešiti)

Druga napomena:

Formulacija „apsolutno srednje odstupanje“ je uobičajena, a njeno značenje je potpuno jasno tek kada se napiše i odgovarajuća formula.

Među mnogobrojnim numeričkim karakteristikama <u>populacije</u> i <u>uzorka</u> je i **apsolutno srednje odstupanje**. Ono pokazuje stepen raširenosti podataka u odnosu na <u>sredinu</u> ili <u>medijanu</u> (bilo uzorka, bilo populacije). Apsolutno srednje odstupanje predstavlja bitan pokazatelj zajedno sa sredinom ili medijanom, zato što različiti uzorci (iz iste ili iz različitih populacija) mogu imati istu vrednost sredine[8] ili medijane[9], ali biti u manjoj ili većoj meri rašireni u odnosu na sredinu ili medijanu.

Apsolutno srednje odstupanje podataka iz uzorka (x_1, x_2,..., x_n) u odnosu na sredinu $\bar{x}_n = \frac{1}{n}(x_1 + ... + x_n)$ je $\frac{1}{n}\sum_{j=1}^{n}\left|x_j - \bar{x}_n\right|$, a u odnosu na medijanu m_e je $\frac{1}{n}\sum_{j=1}^{n}\left|x_j - m_e\right|$.

Ako su podaci iz uzorka (x_1, x_2,..., x_n) dati tabelarno tako da su izdvojene samo različite vrednosti y_1,..., y_k među njima, (nebitno da li su poređani u rastući poredak ili ne) i ako su date frekvencije pojavljivanja

vrednosti	$y_1\ y_2\ ...\ y_k$
frekvencije	$f_1\ f_2\ ...\ f_k$

onda je apsolutno srednje odstupanje u odnosu na srednju vrednost $\frac{1}{n}\sum_{j=1}^{k} f_j\left|y_j - \bar{x}_n\right|$, a u odnosu na medijanu $\frac{1}{n}\sum_{j=1}^{k} f_j\left|y_j - m_e\right|$. Tu je $n = \sum_{j=1}^{k} f_j$.

[8] *Primer.* Skupovi podataka 1,–1,–1,1,1,–1 i 50,–100,50 imaju iste sredine, jednake 0. Raširenost oko sredine je u drugom slučaju jednaka 200/3 i veća nego u prvom slučaju kad je jednaka 1 (odn. 6/6).

[9] *Primer.* Skupovi podataka 4,3,1,2,5 i –11,3,46,5,2 imaju iste medijane, jednake 3, ali je u drugom slučaju raširenost oko medijane veća i jednaka 60/5 a u prvom 6/5.

Aritmetička sredina

Аритметичка средина

EN	Arithmetic mean
FR	Moyenne arithmétique
RU	Арифметическое среднее
ZH	算术中项, suànshù zhōngxiàng
AR	المتوسط الحسابي
EL	Αριθμητικός μέσος
DE	Arithmetisches Mittel
ES	Media aritmética
IT	Media aritmetica
SK	Aritmetický priemer
SL	Aritmetična sredina
HU	Számtani közép

Povezani termini:

Uzorački momenti (238), uzorak (240)

Etimologija:

Aritmetika < ἀριθμός^EL (broj) – oblast matematike koja se bavi elementarnim operacijama, sabiranje, množenje, deljenje.

Neka su $x_1, x_2, \ldots, x_n$ proizvoljni realni brojevi. Njihova **aritmetička sredina** je broj

$$\overline{x} = \frac{x_1 + \ldots + x_n}{n}.$$

Aritmetička sredina se nalazi uvek između najmanje i najveće vrednosti iz posmatranog skupa podataka[10], a može se lakše izračunati korišćenjem translacije[11] podataka. U matematičkoj statistici se aritmetička sredina često koristi za izračunavanje uzoračkih vrednosti nekih statistika[12].

Osim (aritmetičke) sredine u statistici se koriste i ponderisane (težinske) sredine (*weighted mean*)

$$\overline{x}_p = \sum_{j=1}^{n} c_{pj} x_j ,$$

gde su c_{pj} tzv. ponderi (težinski koeficijenti). U slučaju kad je $c_{pj} = 1/n$, dobija se aritmetička sredina. To znači da svi podaci imaju ravnopravan tretman, dok kod ponderisane sredine neki podaci mogu imati veći uticaj na rezultat od drugih.

U Analizi vremenskih serija, na primer, koriste se aritmetičke sredine kad se od polaznog niza podataka formira nov niz – niz pokretnih proseka. Primer: umesto elementa x_2 posmatramo aritmetičku sredinu elemenata x_1, x_2 i x_3, a umesto x_3 aritmetičku sredinu elemenata x_2, x_3 i x_4, itd. Novi niz podataka ima manju disperziju od polaznog. U primenama se ovaj postupak naziva izravnanje.

---

[10] Ako je $a \le x_j \le b, j = 1,n$, onda je $a \le \overline{x} \le b, j = 1,n$.

[11] Ako se umesto podataka $x_1, x_2, \ldots, x_n$ posmatraju

$$y_j = \alpha x_j + \beta, j = 1,n,$$

onda je $\overline{y} = \alpha\overline{x} + \beta$.

[12] Uzoračka sredina, uzorački momenti, centralni uzorački momenti, koeficijent asimetrije, koeficijent spljoštenosti, koeficijent varijacije, koeficijent korelacije, i dr.

Asimetrija

Асиметрија

EN Skewness
FR Asymétrie
RU Асимметрия
ZH 偏斜度, piānxiédù
AR الالتواء
EL Ασυμμετρία
DE Schiefe
ES Asimetría
IT Asimmetria
SK Šikmosť
SL Asimetrija
HU Ferdeség

Povezani termini:

Gustina raspodele (84), histogram (108), matematičko očekivanje (152), medijana (154), raspodela, zakon raspodele

Etimologija:

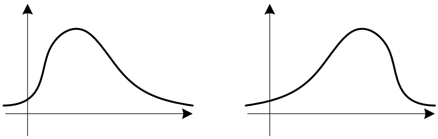
Starogrčki prefiks ἀ označava negaciju. U tom smislu ovaj termin možemo prevesti i opisno kao „odsustvo simetrije“.

Simetrija < συμμετρία^EL < σύν^EL (sa) + μέτρον^EL (mera)

Engleski termin je izveden od reči *skew* (kosina).

Ukoliko raspodela posmatrane slučajne veličine (obeležja) nije simetrična oko matematičkog očekivanja ili oko medijane, govori se o asimetričnoj raspodeli.

Raspodela je pozitivno asimetrična ako je, u nekom smislu, veći skup vrednosti desno od srednje vrednosti ili medijane (v. sl. levo, tu je mod raspodele manji od medijane, a ona manja od očekivanja). U suprotnom je raspodela negativno asimetrična (v. sl. desno, tu je matematičko očekivanje manje od medijane, a ona manja od moda raspodele).



Asimetrija je karakteristika zakona raspodele ili gustine raspodele i može se meriti. Da bi se tim merenjem omogućilo poređenje različitih raspodela, mera asimetrije mora biti broj koji ne zavisi od mernih jedinica. Mogući su razni načini merenja asimetrije, a svaki od njih treba da zadovoljava uslov: ako je raspodela simetrična mera asimetrije treba da bude jednaka 0.

Osim koeficijenta asimetrije, ponekad se kao merilo asimetrije posmatra količnik:

$$(\textit{očekivanje} - \textit{mod})/\textit{standardna devijacija},$$

ili količnik (sa obaveznim faktorom 3):

$$3(\textit{očekivanje} - \textit{medijana})/\textit{standardna devijacija}.$$

Od češće korišćenih raspodela asimetrične su binomna (kad je $p \neq ½$), hi-kvadrat i eksponencijalna raspodela, a simetrične su: normalna, ravnomerna, Studentova raspodela.

→ *Nastavak na str. 256*

Asimptotski efikasna ocena

Асимптотски ефикасна оцена

EN Asymptotically efficient estimator
FR Estimateur asymptotiquement efficace
RU Асимптотически эффективная оценка
ZH 渐近有效估计量, xījìn yǒuxiào gūjìliàng
AR المقدر الكافئ المتقارب
EL Ασυμπτωτικά αποτελεσματική εκτιμήτρια
DE Asymptotisch effizienter Schätzer
ES Estimador asintóticamente eficiente
IT Stimatore asintoticamente efficiente
SK Asymptoticky efektívny odhad
SL Asimptotično učinkovita cenilka
HU Aszimptotikusan hatékony becslés

Povezani termini:

Disperzija ocene, efikasna ocena (62), informanta uzorka, kvalitet ocene, nejednakost Rao-Kramera (21), ocenjivanje parametara (178)

Etimologija:

Asimptotski < ἀσύμπτωτος^{EL} (onaj koji ne pada zajedno) < ἀ^{EL} (ne) + σύν^{EL} (zajedno) + πίπτω^{EL} (pasti). Asimptota je kriva linija koja se približava grafiku funkcije; postoje pravolinijske i krivolinijske asimptote; *asimptotski* ovde označava svojstvo koje se postiže kad se obim uzorka neograničeno uvećava.

Efikasan < efficiēns^{LA} (onaj koji izvršava) < ē^{LA} (iz) + faciō^{LA} (raditi). U engleskom se jeziku ova reč u značenju „delotvornog“ prvi put javlja 1787. godine.

Efikasnost ocene je količnik disperzije posmatrane ocene i donje granice disperzija nepristrasnih ocena prema nejednakosti Rao-Kramera[13]. Ocena čija efikasnost teži jedinici kad se obim uzorka neograničeno povećava naziva se **asimptotski efikasna ocena**.

Kaljampadi Radakrišna Rao (1920), svetski poznat statističar indijskog porekla. Studirao je u Engleskoj i bio Fišerov student. Posle penzionisanja u Indiji 1980. godine nastavio je karijeru u SAD. Mnoge teoreme nose njegovo ime.

Karl Harald Kramer (1893–1985), švedski matematičar i statističar. Izuzetno je poznata i citirana njegova knjiga „Matematičke metode u statistici".

Popravljena uzoračka disperzija je primer statistike koja predstavlja asimptotski efikasnu ocenu.[14]

---

[13] Ukoliko je θ nepoznati parametar u raspodeli obeležja X, T nepristrasna ocene tog parametra, a L funkcija verodostojnosti, onda je donja granica disperzije za T jednaka $1/I(\theta)$, gde je

$$I(\theta) = E(\partial \ln L/\partial\theta)^2.$$

$I(\theta)$ se naziva informaciona funkcija Fišera.

[14] Neka je $(X_1,..., X_n)$ prost slučajni uzorak obima n za obeležje X sa normalnom raspodelom disperzijom σ^2. Statistika

$$\frac{1}{n-1}\sum_{j=1}^{n}(X_j - \overline{X})^2$$

je nepristrasna ocena za σ^2, a njena disperzija je $2\sigma^4/(n-1)$, dok je donja granica disperzija nepristrasnih ocena za σ^2 jednaka $2\sigma^4/n$, pa je efikasnost posmatrane ocene $(n-1)/n$, što teži 1, kad $n \to \infty$.

Asimptotska raspodela

Асимптотска расподела

EN Asymptotic distribution
FR Distribution asymptotique
RU Асимптотическое распределение
ZH 渐近分布, xījìn fēnbù
AR التوزيع التقاربي
EL Ασυμπτωτική κατανομή
DE Grenzverteilung
ES Distribución asintótica
IT Distribuzione asintotica
SK Asymptotické rozdelenie
SL Asimptotičnim distribucija
HU Aszimptotikus eloszlás

Povezani termini:

Centralna granična teorema (46), raspodele ekstremnih vrednosti, statistika (167)

Etimologija:

Asimptotski < ἀσύμπτωτος[EL] (onaj koji ne pada zajedno) < ἀ[EL] (ne) + σύν[EL] (zajedno) + πίπτω[EL] (pasti). Asimptota je kriva linija koja se približava grafiku funkcije; postoje pravolinijske i krivolinijske asimptote; *asimptotski* ovde označava svojstvo koje se postiže kad se obim uzorka neograničeno uvećava.

Granična raspodela neke statistike kad obim uzorka neograničeno raste se naziva **asimptotska raspodela**. Zapravo posmatramo niz slučajnih veličina koje zavise (između ostalog) od obima uzorka, npr. uzoračke sredine $\overline{X}_n$, prema centralnoj graničnoj teoremi imaju normalnu raspodelu kao graničnu raspodelu. Tada za konkretni obim uzorka, pod uslovom da je dovoljno veliki, na primer, veći od 50, smatramo da $\overline{X}_n$ ima normalnu raspodelu $N(m,\sigma^2/n)$, što koristimo pri testiranju hipoteza ili određivanju intervala poverenja. Tu su m i σ^2 očekivanje i disperzija obeležja X.

Dalje, asimptotska raspodela vezana za uzoračku disperziju je, takođe po centralnoj graničnoj teoremi, u vezi sa χ^2-raspodelom. Tačnije, raspodela za

$$\sum_{j=1}^{n}(X_j - \overline{X}_n)^2 / \sigma^2$$

je χ^2-raspodela sa $(n-1)$ stepeni slobode.

Važne granične raspodele su i raspodele ekstremnih vrednosti vezanih za statistike poretka, tj. raspodele za $X_{\max} = \max\limits_{1\le j\le n} X_j$ ili za $X_{\min} = \min\limits_{1\le j\le n} X_j$ kao i raspodele linearnih funkcija ovih statistika.

Prema centralnoj teoremi matematičke statistike funkcija raspodele obeležja je asimptotska raspodela za empirijsku funkciju raspodele.

Asimptotska raspodela test-statistike u χ^2-testu je χ^2-raspodela sa odgovarajućim brojem stepeni slobode. Prilikom određivanja asimptotskih raspodela nekih test-statistika dobijene su nove, do tada nepoznate raspodele, npr. raspodela Kolmogorova.

Asimptotska relativna efikasnost

Асимптотска релативна ефикасност

EN Asymptotic relative efficiency
FR Efficacité relative asymptotique
RU Асимптотическая относительная эффективность
ZH 渐近相对效率, xījìn xiāngduì xiàolǜ
AR الكفاءة النسبية التقربية
EL Ασυμπτωτική σχετική αποτελεσματικότητα
DE Asymptotische relative Effizienz
ES Eficiencia relativa asintótica
IT Efficienza relativa asintotica
SK Asymptotická relatívna efektívnosť
SL Asimptotičnim relativna učinkovitost
HU Aszimptotikus relatív hatékonyság

Povezani termini:

Disperzija (54), efikasna ocena (62)

Etimologija:

Asimptotski < ἀσύμπτωτος^EL (onaj koji ne pada zajedno) < ἀ^EL (ne) + σύν^EL (zajedno) + πίπτω^EL (pasti). Asimptota je kriva linija koja se približava grafiku funkcije; postoje pravolinijske i krivolinijske asimptote; *asimptotski* ovde označava svojstvo koje se postiže kad se obim uzorka neograničeno uvećava.

Relativno < relativum^LA (povezano, ono što se odnosi na nešto) < referō^LA (nositi nazad) < re^LA (ponovo) + ferre^LA (nositi)

Efikasan < efficiēns^LA (onaj koji izvršava) < ē^LA (iz) + faciō^LA (raditi). U engleskom se jeziku ova reč u značenju „delotvornog“ prvi put javlja 1787. godine.

Asimptotska relativna efikasnost je granična vrednost količnika disperzija[15] dveju ocena nekog parametra kad se obim uzorka neograničeno uvećava.

Ako asimptotska relativna efikasnost teži jedinici kada se obim uzorka neograničeno uvećava, onda se opredeljujemo za onu od posmatranih ocena koja ima jednostavniji izraz ili neko drugo povoljno svojstvo[16].

Takođe se može posmatrati relativna efikasnost dve ocene pri istom obimu uzorka.

Efikasnost ocene predstavlja jedan od kriterijuma za upoređivanje različitih nepristrasnih ocena istog parametra. U nekim klasama ocena određene su najefikasnije ocene.

Takođe se, u klasi ocena regularnih u smislu Rao-Kramera, može da odredi donja granica disperzije nepristrasnih ocena posmatranog parametra, pa se pod efikasnošću ocene smatra količnik disperzije te ocene i dobijene donje granice iz nejednakosti Rao-Kramera.

---

[15] *Primer.* Neka je dat uzorak obima $2n$ za obeležje X. Posmatramo statistike $U = (X_1 + X_3 + ... + X_{2n-1})/n$ i $V = (X_1 + X_2 + ... + X_{2n})/2n$. Ako je disperzija obeležja X označena sa σ^2, onda su disperzije statistika U i V redom jednake σ^2/n i $\sigma^2/2n$, pa je statistika V dvaput efikasnija od statistike U.

[16] Pod povoljnim svojstvima smatramo: nepristrasnost, postojanost, dovoljnost, ali ima i drugih kriterijuma (kompletnost).

Autokorelacija

Аутокорелација

EN	Autocorrelation
FR	Autocorrélation
RU	Автокорреляция
ZH	自相关, zìxiāngguān
AR	لترابط التلقائي
EL	Αυτοσυσχέτιση
DE	Autokorrelation
ES	Autocorrelación
IT	Autocorrelazione
SK	Autokorelácia
SL	Avtokorelacija
HU	Autokorreláció

Povezani termini:

Korelacija, kovarijansa (132), linearna zavisnost

Etimologija:

Prefiks auto- < αὐτός^{EL} (sebe, se) označava samostalnost; ovde prefiksom menjamo smisao drugog dela termina, tako da autokorelacija znači posmatranje neke veze među elementima jedne pojave a ne u odnosu na elemente neke druge pojave.

Korelacija < correlatio^{LA} (međusobna povezanost) < co-^{LA} (sa) + relatio^{LA} (odnos) < referō^{LA} (nositi nazad) < re^{LA} (ponovo) + ferre^{LA} (nositi)

Autokorelacija je mera zavisnosti članova u nekom nizu slučajnih (međusobno zavisnih) veličina. Posebno se taj pojam koristi u analizi vremenskih serija[17], kada se utvrđuje stepen linearne zavisnosti među vrednostima vremenske serije koje su na određenom rastojanju. Ako je data vremenska serija $\{X_t, t \in T\}$ onda je vrednost autokorelacije sa korakom k jednaka

$$Cov(X_t, X_{t+k}) / \sqrt{D(X_t)D(X_{t+k})}\,,$$

gde je $Cov(X_t, X_{t+k})$ kovarijacija koja je jednaka

$$E(X_t - E(X_t))(X_{t+k} - E(X_{t+k}))$$

a $D(X_t)$ je disperzija.

Kao i korelacija, i autokorelacija ima vrednosti u intervalu $[-1,1]$, a u slučaju linearne zavisnosti ima vrednost 1 ili -1.

Koeficijent autokorelacije koristimo u nekim testovima slučajnosti[18], kao i za identifikaciju modela vremenske serije.

Pri modeliranju realnih pojava često se koriste modeli kod kojih autokorealcija opada sa povećavanjem koraka k.

[17] Vremenska serija je niz podataka beleženih u toku vremena: temperatura vazduha u toku dana, cene određenog proizvoda na berzi, itd. Zbog međusobne zavisnosti podataka zabeleženih u bliskim vremenskim trenucima, metode matematičke statistike ne mogu biti primenjene, jer su teorijski zasnovane na konceptu prostog slučajnog uzorka – niza nezavisnih slučajnih veličina sa istom raspodelom.

[18] Test slučajnosti je test kojim se proverava da li je dati niz podataka slučajan (i odgovara teorijskom konceptu prostog slučajnog uzorka). Tada se podaci mogu koristiti bez obzira na redosled kojim su dobijeni, što u slučaju odbacivanja hipoteze o slučajnosti nije moguće, jer bi dovelo do pogrešnih rezultata.

Bajesova formula

Бајесова формула

EN	Bayes’ formula
FR	Formule de Bayes
RU	Формула Бейеса
ZH	贝叶斯公式, Bèiyèsī gōngshì
AR	صيغة بايز
EL	Κανόνας του Bayes
DE	Bayes-Formel
ES	Formula de Bayes
IT	Formula di Bayes
SK	Bayesova formula
SL	Bayesova formula
HU	Bayes-tétel

Povezani termini:

Aposteriorna verovatnoća (10), uslovna verovatnoća (222)

Etimologija:

Formula^{LA} (kalup, izraz, princip) < fōrma^{LA} (oblik).

Druga napomena:

Mada je kod nas uobičajeno da se kaže Bajes, pravilan izgovor je Bejz /ˈbeɪz/, odnosno Bejzova formula.

Neka su $H_1, H_2,..., H_n$ slučajni događaji koji čine potpun sistem događaja i neka je A neki događaj iz istog prostora elementarnih ishoda. Može se posmatrati realizacija događaja A u funkciji realizacije događaja $H_1, H_2,..., H_n$, koji se nazivaju hipoteze. Takođe može da se posmatra i realizacija hipoteza, ako se zna da se realizovao događaj A. Tada možemo odrediti verovatnoće $P(H_j \mid A)$, za svako $j = 1, 2,..., n$[19]. Te verovatnoće se računaju po **Bajesovoj formuli**[20]:

$$P\left(H_k \mid A\right) = \frac{P\left(H_k\right)P\left(A \mid H_k\right)}{\sum_{j=1}^{n} P\left(H_j\right)P\left(A \mid H_j\right)}.$$

Često srećemo i varijantu sa samo dva elementa u potpunom sistemu događaja (B i B^c) kada je

$$P(B / A) = \frac{P(B)P(A / B)}{P(B)P(A / B) + P(B^c)P(A / B^c)},$$

na osnovu koje ocenjujemo šansu dešavanja događaja B ukoliko znamo da se događaj A desio.

Tomas Bajes/Bejz (1701–1761), engleski prezbiterijanski sveštenik iz Tunbridž Velsa bio je imenovan za člana Kraljevskog društva 1742. zbog izuzetnih rezultata postignutih na polju matematike. Formula koja nosi njegovo ime bila je osnov za drukčiji pristup u zaključivanju nego što je klasični statistički pristup. Pristalice takvog pristupa se nazivaju Bajesovci (Bajezijanci).

[19] Zbir svih aposteriornih verovatnoća je jednak 1: $\sum_{k=1}^{n} P(H_k / A) = 1$, što znači da u zadacima možemo izračunati n–1 jednostavniju verovatnoću, a preostalu iz gornje jednakosti.

[20] Imenilac ove formule je formula potpune verovatnoće.

Bajesovsko zaključivanje

Бајесовско закључивање

EN Bayesian inference
FR Inférence de Bayes
RU Бейесовский вывод
ZH 贝叶斯推论, Bèiyèsī tuīlùn
AR معادلة بيز
EL Συμπερασματολογία κατά Bayes
DE Bayessche Schlußweise
ES Inferencia de Bayes
IT Inferenza Bayesiana
SK Bayesovská inferencia
SL Bayesovo sklepanje
HU Bayes-féle következtetés

Povezani termini:

Funkcija gubitaka (82), funkcija odluke (31), ocenjivanje parametara (178), raspodela, uzorak (240)

Druga napomena:

Mada je kod nas uobičajeno da se kaže Bajes, pravilan izgovor je Bejz /ˈbeɪz/, odnosno Bejzovsko zaključivanje.

U bajesovskom pristupu smatra se da je nepoznati parametar u raspodeli posmatranog obeležja takođe slučajna veličina[21], sa unapred datom raspodelom, koja se naziva apriorna raspodela. Obično se podrazumeva da se skup mogućih ocena (vrednosti koje se dobijaju na osnovu uzorka) poklapa sa parametarskim prostorom (pretpostavljenim ili dopustivim skupom vrednosti parametra). Kad se dobiju podaci iz uzorka, pošto oni sadrže informacije o parametru, oni utiču na raspodelu parametra, tako da se dobija uslovna raspodela parametra u odnosu na realizovani uzorak i ona se naziva aposteriorna raspodela.

Za ocenu parametra se uzima veličina koja minimizira tzv. srednji rizik, tj. očekivanu vrednost funkcije gubitaka, u odnosu na aposteriornu raspodelu parametra. **Bajesovsko zaključivanje** je posebno primenljivo na uzorke malog obima, ali i u slučaju uzoraka velikog obima, pod relativno opštim pretpostavkama, Bajesove ocene predstavljaju dovoljne statistike, asimptotski su nepristrasne, asimptotski najefikasnije, imaju asimptotski normalnu raspodelu, a od ocena dobijenih metodom maksimalne verodostojnosti se razlikuju za red veličine $1/\sqrt{n}$, gde je n obim uzorka.

Uz funkciju gubitka i raspodelu parametra u bajesovskom odlučivanju je važna i funkcija odluke[22].

---

[21] Za razliku od bajesovskog pristupa, klasične metode ocenjivanja parametara u statistici zapravo podrazumevaju da su sve vrednosti parametara iz parametarskog prostora međusobno ravnopravne i jednakoverovatne. Na taj način je parametar slučajna promenljiva sa uniformnom raspodelom, mada se ta činjenica ne koristi u daljem radu, jer se na osnovu datog uzorka, realizovana vrednost odgovarajuće statistike uzima za ocenu parametra.

[22] Funkcija odluke (*decision function*) je u slučaju ocenjivanja parametra dobijena ocena, u slučaju testiranja hipoteza odluka o prihvatanju ili odbacivanju hipoteze...

Basuova teorema

Басуова теорема

EN	Basu's theorem
FR	Théorème de Basu
RU	Теорема Басу
ZH	巴苏定理, Bāsū dìnglǐ
AR	نظرية باسو
EL	Θεώρημα Basu
DE	Basuscher Lehrsatz
ES	Teorema de Basu
IT	Teorema di Basu
SK	Basuova veta
SL	Basu izrek
HU	Basu-tétel

Povezani termini:

Dovoljna statistika (58), kompletna dovoljna statistika

Etimologija:

Teorema < θεώρημα^EL (opažanje, tvrđenje, iskaz) < θεωρέω^EL (posmatrati) < θέα^EL (pogled) + ὁράω^EL (gledati)

Teoreme Basua govore o dovoljnim statistikama. Navodimo jednu od njih:

Ako je statistika U kompletna dovoljna statistika za parametar θ i V druga statistika koja nije funkcija samo od U i ako raspodela za V ne zavisi[23] od θ, onda su statistike U i V nezavisne[24] slučajne veličine.

Debabrata Basu (1924–2001), rođen u Bangladešu, živeo u Indiji i Americi. Više tvrđenja o dovoljnim statistikama, uslovnoj nezavisnosti (*conditional independence*) i pomoćnim statistikama nose njegovo ime.

Teorema je korisna u primenama za utvrđivanje nezavisnosti nekih statistika[25].

[23] Treba obratiti pažnju na činjenicu da θ u tom slučaju nije ni u izrazu za V, ali nije ni u oblasti definisanosti za V.

[24] Sreće se i formulacija prema kojoj je zaključak teoreme da su U i V „stohastički nezavisne".

[25] *Primer.* Neka je raspodela obeležja X normalna raspodela $N(m, \sigma^2)$, gde je disperzija nepoznata, očekivanje parametar koji ocenjujemo na osnovu uzorka obima n. Tada je uzoračka sredina $\overline{X}_n$ kompletna dovoljna statistika za m, a raspodela statistike $n\overline{S}^2_n/\sigma^2$ je χ^2-raspodela sa $(n–1)$ stepeni slobode i ne zavisi od m. Stoga su uzoračka sredina i uzoračka disperzija nezavisne veličine. Zaključak je važan stoga što se u izrazu za uzoračku disperziju javlja uzoračka sredina, ali je veza takva da su posmatrane statistike nezavisne.

Bernulijev zakon velikih brojeva

Бернулијев закон великих бројева

EN Bernoulli law of large numbers
FR Loi des grandes nombres de Bernoulli
RU Закон больших чисел Бернулли
ZH 伯努利大数定律, Bónǔlì dàshù dìnglǜ
AR برنولي قانون الأعداد الكبيرة
EL Νόμος των μεγάλων αριθμών του Bernoulli
DE Bernoullisches Gesetz der großen Zahlen
ES Ley de los grandes números de Bernoulli
IT Legge dei grandi numeri di Bernoulli
SK Bernoulliho zákon veľkých čísel
SL Bernoullijev zakon velikih števil
HU A nagy számok Bernoulli-féle törvénye

Povezani termini:

Relativna frekvencija (211), zakon velikih brojeva (252)

Za svako $\varepsilon > 0$, za slučajnu veličinu X sa $B(n, p)$ raspodelom važi

$$P\left\{\left|\frac{X}{n} - p\right| \geq \varepsilon\right\} \to 0, n \to \infty .$$

Navedeno tvrđenje se naziva **Bernulijev zakon velikih brojeva** i ukazuje na to da se relativne frekvencije događaja *A*, čija je verovatnoća p, grupišu upravo oko p.

Jakob Bernuli (1654–1705), jedan od predstavnika švajcarske porodice Bernuli, koja je dala veliki broj izuzetno uspešnih i poznatih matematičara. Filozofiju i teologiju je studirao po želji roditelja, a matematiku i astronomiju po svojoj želji. Navedeni zakon velikih brojeva je iz 1689. godine, a njegov doprinos teoriji verovatnoće i statistici je i binomna (Bernulijeva) raspodela.

Ovo se neposredno koristi za ocenjivanje verovatnoće nekog događaja njegovom relativnom frekvencijom, tako što se za uzorak dovoljno velikog obima (npr. $n > 50$) uzima da je ocena nepoznate verovatnoće p jednaka relativnoj frekvenciji X/n.

Zakon velikih brojeva Bernulija dokazuje se primenom nejednakosti Čebiševa[26].

Pafnutij Ljvovič Čebišev (1821–1894), ruski matematičar koji je najveći deo karijere proveo u Sankt Petersburgu. Osim navedene nejednakosti, značajno je što je na svom univerzitetu razvio matematičku školu međunarodne reputacije. Pri određivanju koeficijenata u polinomijalnoj regresiji mogu se koristiti i tzv. Čebiševljevi polinomi.

[26] $P(|X - E(X)| \geq \varepsilon) \leq D(X)/\varepsilon^2$, za svaki pozitivan broj ε. Često se sreće i u obliku $P(|X - E(X)| \geq k\sigma) \leq 1/k^2$, gde je $D(X) = \sigma^2$.

Beta raspodela

Бета расподела

EN	Beta distribution
FR	Distribution bêta
RU	Бета-распределение
ZH	贝它分布, bèitā fēnbù
AR	توزيع بيتا
EL	Κατανομή Βήτα
DE	Betaverteilung
ES	Distribución beta
IT	Distribuzione beta
SK	Beta rozdelenie
SL	Beta porazdelitev
HU	Béta-eloszlás

Povezani termini:

Familije raspodela, uniformna raspodela (194)

Etimologija:

Beta (βῆτα) je drugo slovo grčkog alfabeta. Pošto je grčki alfabet nastao iz feničanskog pisma, nazivi slova u grčkom jeziku su takođe preuzeti iz ovog semitskog jezika. Naziv *beta* potiče od proto-semitskog korena *bajt- (kuća, dom).

Slučajna veličina X ima **beta raspodelu** sa parametrima a i b ako je njena gustina raspodele

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1},\ 0 < x < 1 .$$

Parametri a i b su pozitivni, a $B(a, b)$ označava beta funkciju, koja ovde služi kao normirajući faktor, tj. obezbeđuje da važi

$$\int_0^1 f(x)dx = 1 .$$

Očekivana vrednost slučajne veličine sa beta raspodelom navedenog oblika je $a/(a+b)$, dok je disperzija $ab/(a+b)^2(a+b+1)$. Za različite vrednosti parametara se dobijaju razne familije gustina, a kao specijalni slučaj za $a = 1$, $b = 1$ se dobija ravnomerna raspodela.

Polazeći od uniformne odn. ravnomerne raspodele se može dobiti beta raspodela. Naime, ako obeležje X ima uniformnu raspodelu, tada statistike poretka imaju beta raspodele sa odgovarajućim parametrima.

Naziv beta raspodele je uveo Korado Đini. Naziv dolazi od toga što je gustina raspodele $f(x)$ proporcionalna sa podintegralnom funkcijom u definiciji beta funkcije $B(a, b)$.

Korado Đini (1884–1965), italijanski statističar koji je u okviru studija na pravnom fakultetu upoznao statističke metode. On je 1936. godine osonovao prvi odsek za statistiku u Italiji. U vezi sa Lorencovom krivom definisao je jedan koeficijent koji se danas naziva Đinijev indeks.

Binomna raspodela

Биномна расподела

EN	Binomial distribution
FR	Distribution binômiale
RU	Биномиальное распределение
ZH	二项式分布, èrxiàngshì fēnbù
AR	توزيع ذي الحدين
EL	Διωνυμική κατανομή
DE	Binomialverteilung
ES	Distribución binomial
IT	Distribuzione binomiale
SK	Binomické rozdelenie
SL	Binomska porazdelitev
HU	Binomiális eloszlás

Povezani termini:

Normalna raspodela (172), Puasonova raspodela

Etimologija:

Binomni < binōminis^{LA} (koji ima dva imena) < bis^{LA} (dvaput) + nōmen^{LA} (ime). Zanimljivo je da su u latinskom jeziku inicijalni *b* i *du* često zamenjivani jedno s drugim, tako da je *bis* nastalo transformacijom od priloga *duis* (dvaput), koje sa našim brojem *dva* ima zajednički koren u proto-indoevropskom *dwṓu (dva).

Slučajna veličina X ima **binomnu raspodelu** sa parametrima n i p, gde je n prirodan broj, a p verovatnoća, ako je

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\ k\in\{0,1,...,n\}.$$

Navedene verovatnoće zadovoljavaju relaciju

$$P\{X=k+1\}=\frac{p}{1-p}\cdot\frac{n-k}{k+1}\cdot P\{X=k\},$$

pa se na osnovu te jednakosti mogu računati potrebne verovatnoće jedna za drugom.

Neka je verovatnoća događaja A (često se kaže „*verovatnoća uspeha*“) u svakom eksperimentu jednaka p i neka je izvedeno n eksperimenata pod istim uslovima, nezavisno jedan od drugog. Slučajna veličina X jednaka je broju pojavljivanja događaja A u posmatranih n eksperimenata.

Ako slučajna veličina X ima $B(n_1,p)$ raspodelu, a slučajna veličina Y ima $B(n_2,p)$ raspodelu, i ako su X i Y nezavisne, tada slučajna veličina $Z = X + Y$ ima $B(n_1 + n_2, p)$ raspodelu.

Neka je verovatnoća događaja A u svakom eksperimentu jednaka p i neka je izvedeno n eksperimenata pod istim uslovima, nezavisno jedan od drugog. Neka I_1 predstavlja indikator događaja A u prvom eksperimentu, I_2 u drugom,..., I_n u n-tom eksperimentu. Pri tome su slučajne veličine $I_1, I_2,..., I_n$ nezavisne. Na osnovu definicije binomne raspodele sledi da za slučajnu veličinu X sa binomnom raspodelom $B(n,p)$ važi $X = I_1 + I_2 + ... + I_n$.

Na osnovu osobina očekivanja i disperzije, iz prethodne jednakosti dobija se da je:

$$E(X) = np,\ D(X) = np(1-p),\ \sigma_X=\sqrt{np(1-p)}.$$

→ *Nastavak na str. 256*

Biostatistika

Биостатистика

EN	Biostatistics
FR	Biostatistique
RU	Статистика биологическая
ZH	生物统计学, shēngwù tǒngjìxué
AR	الأحصاء الحيوي
EL	Βιοστατιστική
DE	Biostatistik
ES	Bioestadística
IT	Biostatistica
SK	Bioštatistika
SL	Biostatistika
HU	Biostatisztika

Povezani termini:

Statistika (150)

Etimologija:

Prefiks bio- < βίος^{EL} (život) se koristi u složenicama koje imaju veze sa proučavanjem života i životnih procesa ili, u novijoj terminologiji, organskom proizvodnjom.

Statistika < statisticum^{LA} (ono što ima veze s državom). Termin statistika, sa prvobitnim značenjem analiziranja podataka o državi, uveo je Gotfrid Ahenval (1719–1772), nemački filozof, pravnik i ekonomista.

Biostatistika ili **biometrika** je zajednički naziv za statistiku primenjenu na živi svet. U tom smislu, biostatistika uključuje demografiju[27], epidemiologiju, klinička ispitivanja...

Ponekad takav naziv ima kurs statistike na medicinskom, biološkom, farmaceutskom ili srodnim visokoškolskim ustanovama.

Na primer, primena statistike u medicini podrazumeva dobro poznavanje vrsta obeležja, vrsta grafičkog prikazivanja podataka, poznavanje planiranja eksperimenata, pravilnu primenu parametarskih i neparametarskih testova, analizu preživljavanja,... pri čemu se delovi ili čitave oblasti statistike koje su bitne u medicinskim istraživanjima u klasičnim kursevima eventualno samo spominju.

[27] Demografija proučava ljudske grupe posebno u odnosu na rođenja, umiranja, skalpanje brakova, broj zaposlenih, migracije, itd.

Boks plot dijagram

Бокс плот дијаграм

EN Box plot
FR Diagramme en boîte
RU Диаграмма размаха
ZH 箱形图, xiāngxíngtú
AR مخطط الصندوق
EL Διάγραμμα Box plot
DE Boxplot
ES Diagrama de caja
IT Diagramma a scatola
SK Krabicový graf
SL Škatla ploskvi
HU Doboz ábrázolás

Povezani termini:

Kvartil (140), medijana (154), normalna raspodela (172)

Etimologija:

Box^{EN} (kutija) < buxis^{LA} (kutija) < πυξίς^{EL} (kutija od šimširovog drveta) < πύξος^{EL} (šimšir). U engleskom jeziku se šimšir naziva *box* ili *box wood*.

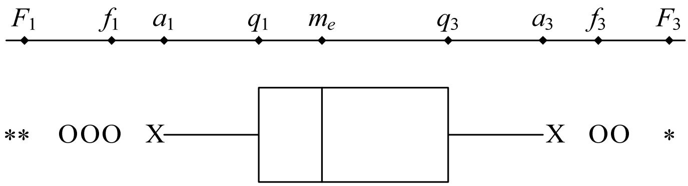
Plot^{EN} (plac). *Plot* potiče od proto-germanskog korena *plataz (parče zemlje), čije je poreklo nepoznato, pa je moguće da ima veze i sa našom rečju *plot* (odn. proto-slovenskim *plet-), kao ogradom kojom se određuje granica između dva placa.

Dijagram < διάγραμμα^{EL} (ono što je ocrtano linijama) < διά^{EL} (kroz, od) + γράμμα^{EL} (slovo, ono što je nacrtano)

Boks plot dijagram se dobija tako što se na izabranoj osi odrede tačke koje odgovaraju <u>uzoračkoj medijani</u> i <u>kvartilima</u> q_1 i q_3. Zatim se računaju unutrašnje f_1 i f_3 i spoljašnje granice F_1 i F_3 na sledeći način:

$$f_1 = q_1 - 1.5(q_3 - q_1), f_3 = q_1 + 1.5(q_3 - q_1),$$
$$F_1 = q_1 - 3(q_3 - q_1), F_3 = q_1 + 3(q_3 - q_1)$$

i onda se određuje a_1 – najmanji među elementima <u>uzorka</u> koji je veći od f_1 i a_3 – najveći među elementima uzorka koji je manji od f_3. Dijagram se sastoji od pravougaonika čija je jedna strana paralelna izabranoj osi i jednaka odsečku (q_1, q_3). Dimenzija druge strane pravougaonika nije od značaja, pa se bira proizvoljno. U pravougaonik se ucrta prava linija koja odgovara uzoračkoj medijani m_e. Ako je ta linija blizu sredine pravougaonika, raspodela obeležja na uzorku bi mogla biti neka simetrična raspodela, u suprotnom je u pitanju asimetrična raspodela.



Na dijagramu na slici su znakom X obeleženi elementi a_1 i a_3, kružićem O svi elementi uzorka koji se nalaze u intervalima $[F_1, f_1]$ i $[f_3, F_3]$, a zvezdicom * svi elementi uzorka manji od F_1 ili veći od F_3.

Boks plot dijagram je predložio Džon Tjuki.

Džon Vajlder Tjuki (1916–2000), američki matematičar. Poznat po stvaranju box-plot dijagrama i FFT algoritma. Radeći sa Džonom fon Nojmanom na razvoju kompjuterske tehnologije, stvorio je termin „bit“ kao skraćenicu za „binary digit“.

Broj stepeni slobode

Број степени слободе

EN Number of degrees of freedom
FR Nombre de degrés de liberté
RU Число степеней свободы
ZH 自由度数, zìyóu dùshu
AR عدد درجات الحرية
EL Αριθμός των βαθμών ελευθερίας
DE Anzahl der Freiheitsgrade
ES Número de grados de libertad
IT Numero di gradi di libertà
SK Počet stupňov voľnosti
SL Število stopinj prostosti
HU Szabadságfok

Povezani termini:

F-raspodela (45), Hi-kvadrat raspodela (102), raspodela Kolmogorova (45), *t*-raspodela (45)

Pojam koji se vezuje za neke raspodele u verovatnoći i statistici, pre svega za χ^2-raspodelu, Studentovu t-raspodelu i Fišerovu F-raspodelu. **Broj stepeni slobode** je jedan od parametara u ovim raspodelama, a zbog značenja koje ima u statistici, uzima se da ima celobrojne vrednosti[28]. Npr. χ^2-raspodela sa n stepeni slobode se dobija kao zbir n nezavisnih slučajnih veličina sa normalnom normiranom raspodelom, a kod χ^2-testa broj stepeni slobode test statistike zavisi i od broja nepoznatih parametara u raspodeli obeležja.

Studentova raspodela sa n stepeni slobode se dobija na osnovu jedne normalne normirane raspodele i, od nje nezavisne, jedne χ^2-raspodele sa n stepeni slobode.

Fišerova raspodela ima dva stepena slobode, n_1 i n_2, jer se dobija kao količnik dve nezavisne χ^2-raspodele, sa n_1 i n_2 stepeni slobode.

Raspodela Kolmogorova se javlja kao raspodela supremuma razlike empirijske i teorijske funkcije raspodele posmatranog obeležja.

Za sve navedene raspodele postoje tablice iz kojih se za dati broj stepena slobode i argument funkcije raspodele, može pročitati (približna, uglavnom na 5 decimala) vrednost funkcije raspodele.

Pojam broja stepeni slobode je uveo ser Ronald Fišer.

[28] Izraz za gustinu raspodele kod ovih raspodela ima smisla i kad je posmatrani parametar realan broj.

Centralna granična teorema

Централна гранична теорема

EN	Central limit theorem
FR	Théorème centrale limite
RU	Центральная предельная теорема
ZH	中心极限定理, zhōngxīn jíxiàn dìnglǐ
AR	نظرية النهاية المركزية
EL	Κεντρικό οριακό θεώρημα
DE	Zentraler Grenzwertsatz
ES	Teorema central del límite
IT	Teorema del limite centrale
SK	Centrálna limitná veta
SL	Centralni limitni izrek
HU	Központi határeloszlás-tétel

Povezani termini:

Binomna raspodela (38), konvergencija u zakonu raspodele, Muavr-Laplasova teorema (47), normalna raspodela (172)

Etimologija:

Centar < centrum[LA] (središte) < κέντρο[EL] < κεντεῖν[EL] (pobosti pritke, ograditi)

Teorema < θεώρημα[EL] (opažanje, tvrđenje, iskaz) < θεωρέω[EL] (posmatrati) < θέα[EL] (pogled) + ὁράω[EL] (gledati)

Neka je dat niz $X_1, X_2,...$ niz slučajnih veličina i neka slučajna veličina X ima $N(0,1)$ raspodelu. Neka je

$$\overline{X}_n = (X_1 + \ldots + X_n)/n.$$

Ako važi

$$\frac{\overline{X}_n - E(\overline{X}_n)}{\sqrt{D(\overline{X}_n)}} \xrightarrow{Z} X, \; n \to \infty,$$

tada se kaže da za niz $X_1, X_2,...$ važi **centralna granična teorema**. Ovo znači da će raspodela za $\overline{X}_n$ konvergirati[29] ka normalnoj raspodeli[30]. Postoji više tvrđenja kojima se, uz izvesne pretpostavke, za određeni niz pokazuje da važi centralna granična teorema. Jedno od njih je sledeća teorema:

Neka su date nezavisne slučajne veličine $X_1, X_2,...$ koje imaju istu raspodelu, sa konačnim matematičkim očekivanjem m i disperzijom σ^2. Tada za posmatrani niz slučajnih veličina važi centralna granična teorema, tj.

$$P\left\{\frac{\overline{X}_n - m}{\sigma / \sqrt{n}} \leq x\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt, \; n \to \infty.$$

Specijalni slučaj ove teoreme je Muavr-Laplasova teorema[31].

---

[29] Ovde se koristi konvergencija u zakonu raspodele.

[30] Gaus je zapravo dobio normalnu raspodelu kao graničnu gustinu raspodele uzoračke sredine na uzorku slučajnih grešaka merenja.

[31] U Muavr-Laplasovoj teoremi su X_j indikatori istog događaja u nizu nezavisnih eksperimenata, tako da će njihov zbir imati binomnu raspodelu. Stoga zapravo Muavr-Laplasova teorema govori o tome da se binomna raspodela može dobro aprokisimirati normalnom raspodelom za dovoljno veliko n.

Centralna teorema matematičke statistike

Централна теорема математичке статистике

EN Glivenko-Cantelli theorem
FR Théorème de Glivenko-Cantelli
RU Теорема Гливенко-Кантелли
ZH 格里文科−坎泰利定理, Gélǐwénkē-Kǎntàilì dìnglǐ
AR نظرية الغاية المركزية في الاحصاء
EL Θεώρημα Glivenko-Cantelli
DE Fundamentalsatz der Statistik
ES Teorema de Glivenko-Cantelli
IT Teorema di Glivenko-Cantelli
SK Glivenkova-Cantelliho veta
SL Glivenko-Cantellijev izrek
HU A matematikai statisztika alaptétele

Povezani termini:

Empirijska funkcija raspodele, test Kolmogorova (49), uzorak (240)

Etimologija:

Centar < centrum[LA] (središte) < κέντρο[EL] < κεντεῖν[EL] (pobosti pritke, ograditi)

Teorema < θεώρημα[EL] (opažanje, tvrđenje, iskaz) < θεωρέω[EL] (posmatrati) < θέα[EL] (pogled) + ὁράω[EL] (gledati)

Matematika < μαθηματικός[EL] (onaj ko voli da uči) < μάθημα[EL] (znanje, učenje, nauka)

Statistika < statisticum[LA] (ono što ima veze s državom). Termin statistika, sa prvobitnim značenjem analiziranja podataka o državi, uveo je Gotfrid Ahenval (1719–1772), nemački filozof, pravnik i ekonomista.

Ako je $F(x)$ teorijska funkcija raspodele obeležja X, a $F_n^*(x)$ empirijska funkcija raspodele dobijena na osnovu prostog slučajnog uzorka obima n, tada, uniformno po x, funkcija $F_n^*(x)$ teži ka $F(x)$ sa verovatnoćom 1, tj.

$$P\left\{\sup_{x\in R}\left|F_n^*(x)-F(x)\right|\longrightarrow 0,\ n\to\infty\right\}=1\,.$$

Ova teorema, koja se naziva i **Glivenko-Kantelijeva**, daje opravdanje za prenošenje zaključaka koji se donose o raspodeli obeležja na uzorku na celu populaciju, pod uslovom da je obim uzorka dovoljno veliki. Brzina konvergencije se razlikuje za različite raspodele $F(x)$.

Valerij Ivanovič Glivenko (1896–1940), sovjetski matematičar. Bavio se matematičkom logikom i teorijom verovatnoće. Predavao je na Moskovskom gradskom pedagoškom institutu.

Frančesko Paolo Kanteli (1875–1966), italijanski matematičar. Doktorirao u Palermu na temu nebeske mehanike i u ranim radovima se pretežno bavio astronomijom. Radeći u Institutu za depozite i kredite, počeo je da se bavi i verovatnoćom i aktuarskom matematikom, poljima u kojima je imao najviše uspeha.

Raspodela maksimalnog odstupanja empirijske funkcije raspodele od funkcije raspodele obeležja je određena (raspodela Kolmogorova) i posebno je važna jer pokazuje da rezultat ne zavisi od raspodele obeležja.

Na osnovu toga je utvrđen postupak testiranja hipoteze (test Kolmogorova) da je raspodela obeležja određena funkcijom raspodele $F_o(x)$. Odredi se maksimum apsolutne vrednosti razlike empirijske funkcije raspodele i pretpostavljene funkcije raspodele. Za dati prag značajnosti u tablicama za raspodelu Kolmogorova se nalazi vrednost na osnovu koje se utvrđuje da li će hipoteza biti prihvaćena.

Černovljeva lica

Черновљева лица

EN Chernoff's faces
FR Visages de Chernoff
RU Лица Чернова
ZH 切儿诺夫的面, Qiē'érnuòfū de miàn
AR وجوه تشرنوف
EL Πρόσωπα του Chernoff
DE Chernoff-Gesichter
ES Caras de Chernoff
IT Fronti di Chernoff
SK Chernoffove tváre
SL Chernoff obrazi
HU Chernoff arca

Povezani termini:

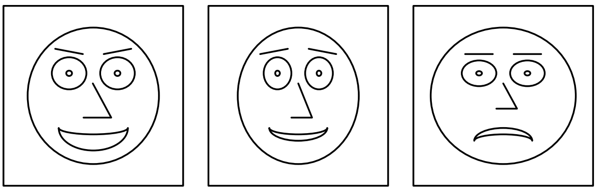
Grafik (94), višedimenzionalna obeležja

Jedan od načina grafičkog prikazivanja višedimenzionalnih obeležja su i **Černovljeva lica**. Standardni grafički prikazi vrednosti jedno- i dvodimenzionalnih obeležja u koordinatnom sistemu nisu primenljivi na višedimenzionalna obeležja (eventualno se razmatraju neke projekcije skupa podataka na jedno- ili dvodimenzionalne prostore – npr. u analizi glavnih komponenata[32]).

Herman Černof (1923), američki matematičar i statističar, koji se bavi prepoznavanjem oblika (*Pattern recognition*), teorijom uzoraka u oblasti sekvencijalnih uzoraka (*Sequential sampling*), i teorijom planiranja eksperimenata (*Experimental design*).

Černovljevo lice predstavlja višedimenzionalno obeležje tako što komponentama obeležja pridružuje: oval lica (mala i velika poluosa te elipse) položaj i veličinu očiju, položaj, veličinu i oblik nosa, obrva i usana...

U zavisnosti od dimenzije obeležja, lice će imati manje ili više detalja, a „portreti" uzoraka za isto obeležje iz različitih populacija ukazaće na sličnosti ili razlike među tim populacijama.



[32] Analiza glavnih komponenata (*Principal Component Analysis)* je jedna od metoda kojom se analiziraju višedimenzionalna obeležja. Postupak se sastoji u tome da se umesto polaznih n veličina posmatraju neke njihove nezavisne linearne kombinacije, konstruisane tako da m ($m < n$) od njih sadrže najveći deo ukupne varijacije polaznog skupa podataka.

Disjunktni događaji

Дисјунктни догађаји

EN Disjoint events
FR Evénements disjoints
RU Несовместные события
ZH 不相交事件, bùxiāngjiāo shìjiàn
AR أحداث منفصلة
EL Διακριτά γεγονότα
DE Disjunkte Ereignisse
ES Sucesos disjuntos
IT Eventi disgiunti
SK Nezlučiteľné udalosti
SL Disjunktni dogodki
HU Kizáró események

Povezani termini:

Potpun sistem događaja (11), presek događaja, slučajni događaj (200)

Etimologija:

Disjunktni < dis-^{LA} (od, odvojeno) + iūnctiō^{LA} (sjedinjenje, jedinstvo) < iungō^{LA} (spojiti)

Slučajni događaji A i B iz istog prostora elementarnih ishoda (vezani za isti eksperiment) su **disjunktni** ako se ne mogu istovremeno realizovati, tj. ako je njihov presek nemoguć događaj:

$$A \cap B = \varnothing.$$

Svaki događaj je disjunktan sa svojim komplementom. Svaki događaj je disjunktan sa nemogućim događajem.

Često se posmatraju skupovi disjunktnih događaja čija je unija siguran događaj. Takve skupove nazivamo potpunim sistemima događaja, jer se pri svakom izvođenju eksperimenta mora desiti jedan od tih događaja. Znači, oni u potpunosti obuhvataju rezultate eksperimenta.

Različite vrednosti slučajne veličine razlažu prostor elementarnih ishoda na potpun sistem događaja.

Posebno treba obratiti pažnju na razliku pojmova: disjunktni događaji i nezavisni događaji, jer se u oba slučaja posmatra presek tih događaja. Ako su događaji disjunktni, oni su obavezno zavisni[33], a ako su događaji nezavisni, oni su i disjunktni samo ako je (bar) jedan od njih nemoguć događaj.

---

[33] *Primer*. Pri bacanju dve numerisane kockice događaji A = „dobijen je zbir 5“ i B = „dobijen je zbir 7“ su disjunktni i zavisni.
Po definiciji, nezavisni događaji A i B imaju osobinu da je uslovna verovatnoća $P(A/B) = P(A)$. Znači, uslovna verovatnoća da se desi događaj A, ako znamo da se desio događaj B, jednaka je verovatnoći da se desi A, bez ikakvih dodatnih informacija o događaju B. Međutim, za događaje A i B iz datog primera je jasno da ako se desio jedan od njih, npr. B, onda se A ne može desiti. Naime, ovde je $P(A/B) = 0 \neq P(A)$, i $P(A) > 0$.

Disperzija
Дисперзија

EN	Variance
FR	Variance
RU	Дисперсия
ZH	方差, fāngchā
AR	التباين
EL	Διασπορά
DE	Dispersion
ES	Dispersión
IT	Dispersione
SK	Rozptyl
SL	Disperzija
HU	Szórás

Povezani termini:

Matematičko očekivanje (152), momenti raspodele

Etimologija:

Disperzija < dis-^{LA} (od, odvojeno) + spargere^{LA} (rasipati)

Drugi naziv:

Varijansa

Uz matematičko očekivanje, jedna od najčešće korišćenih numeričkih karakteristika slučajnih veličina, je **disperzija (varijansa)**. Po definiciji jednaka je $D(X) = E(X - E(X))^2$, pa predstavlja centralni moment drugog reda slučajne veličine X. Disperzija je nenegativna $D(X) \geq 0$, a može se računati i po formuli $D(X) = E(X^2) - (E(X))^2$. Takođe, ako je $P\{X = \mathrm{a}\} = 1$, gde je a neka konstanta, tada je $D(X) = 0$. Ako je k konstanta, tada je $D(kX) = k^2 \cdot D(X)$.

Ako su X i Y nezavisne slučajne veličine koje imaju disperziju i a i b proizvoljne konstante, tada je $D(aX + bY) = a^2D(X) + b^2D(Y)$, a ako su X i Y zavisne slučajne veličine, onda je

$$D(aX+bY) = a^2D(X) + b^2D(Y) + 2ab(E(XY) - E(X)E(Y))$$

Disperzija predstavlja meru odstupanja slučajne veličine od njenog matematičkog očekivanja, ukazujući na raspršenost vrednosti slučajne veličine oko matematičkog očekivanja.

Činjenica da različite raspodele mogu imati isto matematičko očekivanje zahteva da se na neki način meri koliko vrednosti slučajne veličine odstupaju od matematičkog očekivanja[34], a jedan od načina je da se izračuna disperzija.

Činjenica je da je najmanja vrednost izraza $E(X–C)^2$ minimalna ako je $C = E(X)$, što predstavlja razlog za definiciju disperzije kakva je data na početku teksta.

→ *Nastavak na str. 257*

[34] *Primer*. Slučajne veličine

$$X:\begin{pmatrix} -1 & 1 \\ 0.5 & 0.5 \end{pmatrix} \text{ i } Y:\begin{pmatrix} -10 & 10 \\ 0.5 & 0.5 \end{pmatrix}$$

imaju očekivanje 0, a disperzije 1 i 100.

Disperziona analiza

Дисперзиона анализа

EN	Analysis of variance
FR	Analyse de variance
RU	Дисперсионный анализ
ZH	方差分析, fāngchā fēnxī
AR	تحليل التباين
EL	Ανάλυση διακύμανσης
DE	Varianzanalyse
ES	Análisis de variancia
IT	Analisi della varianza
SK	Analýza rozptylu
SL	Analiza variance
HU	Szórásanalízis

Povezani termini:

Disperzija (54), kvadratna forma (57), normalna raspodela (172), testiranje hipoteze (216)

Etimologija:

Disperzija < dis-^{LA} (od, odvojeno) + spargere^{LA} (rasipati)

Analiza < ἀναλύω^{EL} (otkriti) < ἀνά^{EL} (na) + λύω^{EL} (olabaviti).

Disperziona analiza je oblast matematičke statistike koju je uveo ser Ronald Fišer 1918. godine. Uobičajena je oznaka ANOVA (**An**alysis **o**f **va**riance) ili AOV. Može biti jednofaktorska, dvofaktorska i višefaktorska. U jednofaktorskoj disperzionoj analizi se na datom uzorku proverava hipoteza H_0 o odsustvu uticaja različitih nivoa jednog faktora na vrednosti obeležja X. U disperzionoj analizi se takođe pretpostavlja da je uticaj svih ostalih faktora na vrednosti obeležja X zanemarljiv. Testiranje se na ovaj način vrši ukoliko je raspodela obeležja normalna, a disperzije na potpopulacijama iste (uslov homogenosti disperzije). Ako ti uslovi nisu ispunjeni testiranje o odsustvu uticaja nivoa posmatranog faktora na obeležje možemo dobiti korišćenjem Kruskal-Valisovog testa. Tada se ne koriste vrednosti iz uzorka neposredno, nego njihovi rangovi.

ser Ronald Ejlmer Fišer (1890–1962), po mnogima najznačajniji statističar XX veka. Počeo je da se bavi statistikom zbog proučavanja zakona nasleđivanja u genetici. Njegovi prvi radovi su uveli metodu maksimalne verodostojnosti, a nešto kasnije je razvio teoriju planiranja eksperimenata (radeći u Istraživačkoj stanici Rotemsted, gde se bavio utvrđivanjem kvaliteta prinosa različitih poljoprivrednih kultura) i analizu varijanse. Plemićku titulu je dobio 1952. godine zbog doprinosa nauci.

Neka je, na primer, obeležje vek trajanja sijalica iste snage, a faktor – proizvođač. Onda bi nivoi bili različiti proizvođači. Na osnovu uzoraka iz svake fabrike proizvođača sijalica, izračunavale bi se sume kvadrata odstupanja, od ukupne srednje vrednosti (a takav zbir predstavlja jednu kvadratnu formu), kao i od srednjih vrednosti po proizvođačima.

→ *Nastavak na str. 257*

Dovoljna statistika

Довољна статистика

EN	Sufficient statistic
FR	Statistique exhaustive
RU	Достаточная статистика
ZH	充分统计量, chōngfèn tǒngjìliàng
AR	الأحصاءة الكافية
EL	Επαρκής στατιστική
DE	Suffiziente Statistik
ES	Estadístico suficiente
IT	Statistica sufficiente
SK	Postačujúca štatistika
SL	Zadostna statistika
HU	Elégséges statisztika

Povezani termini:

Efikasnost, kvalitet statistika, nepristrasnost, postojanost,

Etimologija:

Statistika < statisticum[LA] (ono što ima veze s državom). Termin statistika, sa prvobitnim značenjem analiziranja podataka o državi, uveo je Gotfrid Ahenval (1719–1772), nemački filozof, pravnik i ekonomista.

U matematičkoj statistici velika pažnja se posvećuje osobinama statistika kojima ocenjujemo nepoznate parametre.

Najčešće pominjani i primenjivani kriterijumi za utvrđivanje kvaliteta statistika su: nepristrasnost, postojanost (stabilnost), efikasnost i dovoljnost.

Statistika $S = f(X_1, X_2,..., X_n)$ je **dovoljna statistika** za parametar θ u raspodeli obeležja X ako je uslovna raspodela prostog slučajnog uzorka $(X_1, X_2,..., X_n)$ pri fiksiranoj vrednosti s statistike S nezavisna od θ.

Navedena definicija nije jednostavna za primenu jer zahteva određivanje uslovne raspodele. Stoga se, ukoliko je to moguće, koristi Fišer-Nejmanova teorema o faktorizaciji.

Dovoljna statistika ne mora postojati, a ako postoji, nije jedinstvena. Naime, ako S jeste dovoljna statistika za θ, a funkcija g bijekcija, tada je i $g(\theta)$ dovoljna statistika za θ.

Dovoljna statistika sadrži svu informaciju o parametru koja je sadržana u uzorku. Ilustrativan primer je obeležje sa uniformnom raspodelom U(0, θ). Statistike $2\overline{X}_n$ i $\max\limits_{1 \le j \le n} X_j$ su ocene za θ. Neka su dobijeni uzorci: 0,4,0,0,2, 0 i 1,1,1,2,1,0. Prva ocena daje rezultat 2 u oba slučaja, ali je to očigledno pogrešno u prvom slučaju. Druga ocena daje rezultate 4 i 2 i nijedan ne protivreči podacima iz uzorka, u odnosu na pretpostavljenu raspodelu.

Termin „dovoljna statistika“ je uveo ser Ronald Fišer 1922. godine.

Dvodimenzionalna raspodela

Дводимензионална расподела

EN Bivariate distribution
FR Distribution bidimensionnelle
RU Двумерное распределение
ZH 二元分布, èryuán fēnbù
AR التوزيع الثنائي
EL Διμεταβλητή κατανομή
DE Zweidimensionale Verteilung
ES Distribución de dos variables
IT Distribuzione bivariata
SK Dvojrozmerné rozdelenie
SL Dvorazsežna porazdelitev
HU Kétváltozós eloszlás

Povezani termini:

Diskretne slučajne veličine (199), jednodimenzionalna raspodela, neprekidne slučajne veličine (199)

Etimologija:

Dimenzija < dīmensio^{LA} (premeravanje) < mētior^{LA} (meriti)

Obeležja u statistici mogu biti jednodimenzionalna i višedimenzionalna. Među višedimenzionalnim obeležjima često razmatramo slučaj dvodimenzionalnih obeležja, na primer, sa ciljem da utvrdimo postojanje i oblik funkcionalne zavisnosti jednog obeležja od drugog – telesne težine od visine, uspeha na ispitu iz statistike i uspeha na ispitu iz analize, i sl.

Dvodimenzionalna obeležja mogu biti takva da su obe komponente diskretne slučajne veličine, obe neprekidne ili jedna diskretna a jedna neprekidna. U svakom slučaju je korisno poznavanje zajedničke raspodele.

Ako su obeležja X i Y diskretna, sa vrednostima iz konačnih ili prebrojivih skupova, zajedničku raspodelu daju verovatnoće $P(X=a, Y=b)$, pri čemu a pripada skupu vrednosti za X, a b skupu vrednosti za Y.

Ako su obeležja neprekidna posmatra se zajednička gustina raspodele $f(x, y)$, x pripada skupu vrednosti za X, a y skupu vrednosti za Y.

Ako je X diskretna, na primer sa vrednostima a_1, a_2, a_3, a Y neprekidna, onda će za svaku vrednost za X biti data raspodela za Y odgovarajućom gustinom: $f_1(y)$, $f_2(y)$, $f_3(y)$.

Pri fiksiranoj vrednosti jedne slučajne veličine, npr. $Y=y_0$ dobijamo uslovnu raspodelu slučajne veličine X kao količnik zajedničke raspodele $f(x, y_0)$ i raspodele za Y u tački y_0.

Iz zajedničke raspodele dvodimenzionalne slučajne veličine ili obeležja može se dobiti raspodela svake od komponenata. Tako dobijene raspodele se nazivaju marginalne raspodele. Marginalna raspodela za X u neprekidnom slučaju je, na primer,

$$\int_{-\infty}^{\infty} f(x, y)dy\,.$$

Efikasna ocena

Ефикасна оцена

EN Efficient estimator
FR Estimateur efficace
RU Эффективная оценка
ZH 有效估计量, yǒuxiào gūjìliàng
AR تقدير الكفاءة
EL Αποτελεσματική εκτιμήτρια
DE Effizienter Schätzer
ES Estimador eficiente
IT Estimatore efficiente
SK Efektívny odhad
SL Učinkovita cenilka
HU Hatásos becslés

Povezani termini:

Disperzija (54), kvalitet ocena, nejednakost Rao-Kramera (21), ocenjivanje parametara (178)

Etimologija:

Efikasan < efficiēns^{LA} (onaj koji izvršava) < ē^{LA} (iz) + faciō^{LA} (raditi). U engleskom se jeziku ova reč u značenju „delotvornog“ prvi put javlja 1787. godine.

Jedan od kriterijuma za upoređivanje različitih nepristrasnih ocena istog parametra je **efikasnost**. Ako su, na osnovu prostog slučajnog uzorka $(X_1,..., X_n)$ dobijene ocene Y_n i Z_n za isti parametar i ako su te ocene nepristrasne, onda je ocena Y_n efikasnija od ocene Z_n ako je disperzija ocene Y_n manja od disperzije ocene Z_n.

U nekim klasama ocena određene su najefikasnije ocene. Na primer, u klasi linearnih nepristrasnih ocena matematičkog očekivanja najefikasnija ocena je uzoračka sredina, tj. njena disperzija je manja ili jednaka od disperzije bilo koje druge linearne ocene matematičkog očekivanja.

Takođe se, u klasi ocena regularnih u smislu Rao-Kramera može da odredi donja granica disperzije nepristrasnih ocena posmatranog parametra, pa se pod efikasnošću ocene smatra količnik disperzije te ocene i dobijene donje granice iz nejednakosti Rao-Kramera[35].

Takođe, može se govoriti o relativnoj efikasnosti različitih ocena istog parametra, kao i o asimptotskoj relativnoj efikasnosti.

[35] *Primer 1.* Neka je raspodela obeležja $N(m,\sigma^2)$, pri čemu je σ^2 poznato. Disperzija uzoračke sredine koja je nepristrasna ocena za matematičko očekivanje je σ^2/n, gde je n obim uzorka, i poklapa se sa donjom granicom disperzija po nejednakosti Rao-Kramera, pa je efikasnost te ocene 1.
Primer 2. Neka je raspodela obeležja $N(m,\sigma^2)$, pri čemu je m poznato. Uzoračka disperzija je tada nepristrasna i efikasna ocena (po Rao-Krameru).
Primer 3. Neka je raspodela obeležja $N(m,\sigma^2)$, pri čemu m nije poznato. Popravljena uzoračka disperzija je nepristrasna ocena disperzije, a njena efikasnost (po Rao-Krameru) je $(n-1)/n$.

Eksponencijalna familija raspodela

Експоненцијална фамилија расподела

EN Exponential family
FR Famille exponentielle
RU Семейство экспоненциальных распределений
ZH 指数族, zhǐshùzú
AR العائلة الأسية
EL Εκθετική οικογένεια
DE Exponentialfamilie
ES Familia exponencial
IT Famiglia esponenziale
SK Exponenciálny systém
SL Eksponentna družina
HU Exponenciális eloszláscsalád

Povezani termini:

Dovoljna statistika (58), normalna raspodela (172), parametri u raspodeli, regularna eksponencijalna familija raspodela

Etimologija:

Eksponent < expōnere[LA] (objasniti) < ex-[LA] (iz) + pōnō [LA] (staviti, dosuditi)

Familija < familia[LA] (domaćinstvo) < famulus[LA] (rob). Smatra se da *famulus* potiče od korena *fāmn̥ (dom) koji je nastao od starijeg proto-indoevropskog korena *d^{h}eh-mn̥ (dom), a koji je u srpskom dao reč *dom*.

Normalna raspodela zauzima posebno mesto u teoriji verovatnoće i u matematičkoj statistici, kako sa istorijske i teorijske tako i sa praktične tačke gledišta, obzirom da veliki broj pojava ima normalnu raspodelu.

U razmatranju nekih osobina obeležja ustanovljeno je da se model normalne raspodele uklapa u širu familiju raspodela – raspodele eksponencijalnog tipa.

Obeležje X ima raspodelu eksponencijalnog tipa ako se gustina ili zakon raspodele za X može napisati u obliku:

$$\exp\{A(\theta)K(x) + S(x) + q(\theta)\}$$

gde je θ nepoznati parametar u raspodeli obeležja X. Navedeni oblik je za jednodimenzionalne parametre[36]. Ako je parametar višedimenzionalni $\theta = (\theta_1, \theta_2, ..., \theta_r)$, uopštenje je prirodno:

$$\exp\{A(\theta_1, \theta_2, ..., \theta_r)K(x) + S(x) + q(\theta_1, \theta_2, ..., \theta_r)\}.$$

Eksponencijalne familije omogućavaju jednostavno određivanje dovoljnih statistika na osnovu Fišer-Nejmanove teoreme. U jednodimenzionalnom slučaju se na osnovu prostog slučajnog uzorka $(X_1, X_2, ..., X_n)$ dobija da je statistika

$$S = \sum_{j=1}^{n} K(X_j)$$

dovoljna za parametar θ.

U višedimenzionalnom slučaju treba obratiti pažnju na činjenicu da koordinate dovoljne statistike za parametar θ, ne moraju biti dovoljne statistike za koordinate θ.

---

[36] Ako raspodela pripada eksponencijalnoj familiji, reprezentacija u datom obliku nije jedinstvena, što je jednostavno zaključiti jer se umesto $A(\theta)K(x)$ može pisati $(cA(\theta))(K(x)/c) = A_1(\theta)K_1(x)$.

Eksponencijalna raspodela

Експоненцијална расподела

EN Exponential distribution
FR Distribution exponentielle
RU Экспоненциальное распределение
ZH 指数分布, zhǐshù fēnbù
AR التوزيع الآسي
EL Εκθετική κατανομή
DE Exponentialverteilung
ES Distribución exponencial
IT Distribuzione esponenziale
SK Exponenciálne rozdelenie
SL Eksponentna porazdelitev
HU Exponenciális eloszlás

Povezani termini:

Erlangova raspodela (76), eksponencijalna familija raspodela (64), neprekidne raspodele

Etimologija:

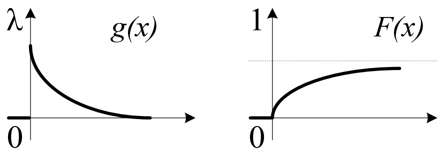
Eksponent < expōnere^{LA} (objasniti) < ex-^{LA} (iz) + pōnō ^{LA} (staviti, dosuditi)

Ako je gustina raspodele slučajne veličine oblika

$$g(x)=\begin{cases}0 & , \quad x\le 0\\ \lambda e^{-\lambda x} & , \quad x>0\end{cases}$$

tada slučajna veličina X ima **eksponencijalnu raspodelu** sa parametrom λ, $\lambda > 0$, što se označava sa $X{:}\varepsilon(\lambda)$. Odgovarajuća funkcija raspodele je

$$F(x) = 1 - e^{-\lambda x},\ x \ge 0.$$



Ako $X{:}\varepsilon(\lambda)$ tada je $E(X)=\dfrac{1}{\lambda}$, $D(X)=\dfrac{1}{\lambda^2}$, dok su koeficijenti asimetrije i spljoštenosti redom jednaki:

$$f_1 = 2,\ f_2 = 6.$$

Eksponencijalna raspodela pripada eksponencijalnoj familiji raspodela, jer se gustina raspodele može napisati u obliku $e^{ln\lambda-\lambda x}$, koji se uklapa u definiciju eksponencijalne klase raspodela.

Eksponencijalna raspodela je jedina neprekidna raspodela sa osobinom odsustva memorije (*lack of memory*)[37]. Naime, za svaka dva pozitivna broja t i s važi:

$$P\{X>t+s \mid X>t\}=P\{X>s\}.$$

Možemo interpretirati ovu jednakost na sledeći način: ako je vreme trajanja pojave prevazišlo trenutak t, onda verovatnoća da prevaziđe trenutak $t + s$ zavisi samo od s, a ne od „celokupne prošlosti".

[37] Koristi se i termin *forgetfulness property*. Od diskretnih slučajnih veličina ovo svojstvo ima samo geometrijska raspodela.

Ekstrapolacija

Екстраполација

EN	Extrapolation
FR	Extrapolation
RU	Экстраполяция
ZH	外推法, wàituīfǎ
AR	الاستكمال
EL	Προέκταση
DE	Extrapolation
ES	Extrapolación
IT	Estrapolazione
SK	Extrapolácia
SL	Ekstrapolacije
HU	Extrapoláció

Povezani termini:

Interpolacija (69), prognoziranje

Etimologija:

Ekstrapolacija < exter^LA (izvan) + polīre^LA (glancati); termin je nastao u drugoj polovini XIX veka, zamenom prefiksa u reči interpolacija < interpolare^LA (izmeniti dodavanjem novog materijala)

U matematičkoj statistici na osnovu skupa podataka (uzorka) donosimo zaključke o populaciji iz koje podaci potiču.

Osnovna ideja je da se formira model posmatrane pojave i na osnovu toga prognozira dalji razvoj posmatrane pojave, ako je to vremenska serija (npr. potrošnja benzina), onda se na mogu prognozirati vrednosti te serije u bližoj ili daljoj budućnosti (npr. planirati snabdevenost gorivom na benzinskoj pumpi).

Donošenje zaključka o vrednosti pojave van skupa podataka koji je na raspolaganju se naziva **ekstrapolacija**.

Najbolji rezultati se dobijaju ukoliko se prognoziraju vrednosti u bliskoj budućnosti. Korišćenjem računara podaci se obrađuju u realnom vremenu i tako ažurira prognoza (npr. vremenska prognoza).

Ukoliko pak neki podaci u uzorku ne postoje (nisu bili prikupljeni ili su izgubljeni) onda se na osnovu dobijenog modela može prognozirati vrednost koja nedostaje – tada je u pitanju interpolacija[38].

[38] Termin interpolacija (*interpolation*) se koristi i u numeričkoj analizi. Tako npr. postoji Lagranžev interpolacioni polinom, kao jedinstveni polinom (n–1)-vog stepena koji prolazi kroz datih n nekolinearnih tačaka.

Ekstremne vrednosti u uzorku

Екстремне вредности у узорку

EN	Outliers
FR	Valeurs aberrantes
RU	Выбросы
ZH	奇异值, qíyìzhí
AR	القيم المتطرفة
EL	Ακραίες τιμές
DE	Ausreißer
ES	Valores extremos
IT	Valori anomali
SK	Odľahlé pozorovania
SL	Osamelci
HU	Szélsőértékek

Povezani termini:

Boks plot dijagram (42), normalna raspodela (172)

Etimologija:

Ekstreman < extremus^{LA} (najdalji) < exter^{LA} (izvan)

Podaci koji su nam na raspolaganju mogu, ponekad sadržati vrednosti koje mnogo odstupaju od glavnine skupa podataka. Postavlja se pitanje da li su takve vrednosti moguće ili su rezultat neke greške (greške merenja ili greške pri beleženju merene vrednosti).

Po analogiji sa osobinama normalne raspodele (a može se oceniti i po nejednakosti Čebiševa) veliki broj vrednosti obeležja mora biti u intervalu između prvog i trećeg kvartila. Pri grafičkom prikazu pomoću tzv. box-plot dijagrama se prvi i treći kvartil i koriste.

U odnosu na pretpostavljenu raspodelu obeležja može se odrediti verovatnoća dobijanja veoma velikih ili veoma malih vrednosti, pa tako možemo testirati hipotezu da li su neke vrednosti ekstremne i neprihvatljive ili pak ekstremne, ali prihvatljive.

Što se tiče numeričkih karakteristika populacije, neke su osetljive na postojanje ekstremnih vrednosti u uzorku (npr. uzoračka sredina), a neke su manje osetljive (npr. medijana).

U nekim oblastima primene **ekstremne vrednosti u uzorku** se koriste za određivanje tzv. raspodele ekstremnih vrednosti (*extreme value distribution*). Interesantno je da se kao raspodele ekstremnih vrednosti dobijaju tri tipa raspodela: Gumbelova, Frešeova i Vejbulova.

Ernest Hjalmar Valodi Vejbul (1887–1979), švedski inženjer koji se, između ostalog, bavio proučavanjem podvodnih eksplozija, kao i vremenom trajanja komponenata nekog proizvoda u teoriji pouzdanosti (*reliability theory*). Njegova raspodela se primenjuje u meteorologiji.

Empirijska funkcija raspodele

Емпиријска функција расподеле

EN Sampling distribution
FR Distribution empirique
RU Выборочная функция распределения
ZH 取样分布, qǔyàng fēnbù
AR توزيع المعاينة
EL Δειγματική κατανομή
DE Stichprobenverteilung
ES Distribución del muestreo
IT Distribuzione campionaria
SK Výberové rozdelenie
SL Vzorčna porazdelitev
HU Empirikus eloszlás

Povezani termini:

Binomna raspodela (38), centralna teorema matematičke statistike (48), funkcija raspodele (88), uzorak (240)

Etimologija:

Empirijski < ἐμπειρικός^{EL} (iskusan) < ἐν^{EL} (u) + πεῖρα^{EL} (proba, pokušaj, probiti, proseći).

Funkcija < fūnctus^{LA} (izvršen) < fungor^{LA} (obaviti, završiti)

Neka je $(X_1, X_2,..., X_n)$ prost slučajni uzorak obima n za posmatrano obeležje. Slučajna veličina

$$F_n^*(x) = \frac{1}{n}\sum_{k=1}^{n} I\left\{X_k \le x\right\}, x \in R,$$

gde I označava indikator događaja, je **empirijska funkcija raspodele**. Pri tom $nF_n^*(x) : B(n, F(x))$, gde je $F(x)$ odgovarajuća teorijska funkcija raspodele za posmatrano obeležje, pa iz Bernulijevog zakona velikih brojeva sledi da empirijska funkcija raspodele konvergira u verovatnoći ka funkciji raspodele obeležja:

$$F_n^*(x) \xrightarrow{P} F(x), n \to \infty.$$

Štaviše, prema centralnoj teoremi matematičke statistike ta konvergencija je uniformna po x.

Sa druge strane ako imamo realizovani uzorak i ako je n_x broj elemenata uzorka za koje je vrednost obeležja X manja ili jednaka od realnog broja x, tada se realizovana vrednost empirijske funkcije raspodele u tački x dobija po formuli: $F_n^*(x) = n_x/n$.[39]

Empirijska funkcija raspodele je jednaka relativnoj učestalosti događaja $\{X \le x\}$. To je jedna stepenasta funkcija koja uzima vrednosti iz segmenta [0,1], neopadajuća je za svako x i neprekidna sa desne strane.

---

[39] *Primer*. Pri bacanju tri numerisane kocke posmatra se obeležje X: broj dobijenih šestica. Rezultati su bili redom: 0,1,0,1,1,0,1,0,1,0,1,0, 2,1,2,2. Empirijska funkcija raspodele se dobija pomoću tabele:

broj šestica	*0*	*1*	*2*	*3*
frekvencije	*6*	*7*	*3*	*0*

i jednaka je:

$$F_{16}^*(x) = \begin{cases} 0 & , \quad x < 0 \\ 6/16 & , \quad 0 \le x < 1 \\ 13/16 & , \quad 1 \le x < 2 \\ 1 & , \quad x \ge 2 \end{cases}$$

Enskombov kvartet

Енскомбов квартет

EN	Anscombe’s quartet
FR	Quartet d’Anscombe
RU	Квартет Энскомба
ZH	安斯库姆四重奏, Ānsīkùmǔ sìchóngzòu
AR	تحليل انحدار انسكوب
EL	Κουαρτέτο του Anscombe
DE	Anscombe-Quartett
ES	Cuarteto de Anscombe
IT	Quartetto di Anscombe
SK	Anscombova regresia
SL	Anscombeov kvartet
HU	Anscombe kvartett

Povezani termini:

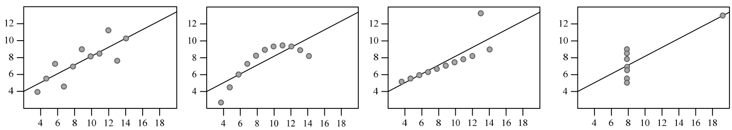
Grafik (94), metoda najmanjih kvadrata (75), regresija (75)

Etimologija:

Kvartet < quattuor[LA] (četiri). Latinska reč *quattuor* i naše *četiri* imaju zajednički koren u proto-indoevropskom *kʷetwóres.

Engleski matematičar Enskomb je konstruisao četiri skupa podataka kod kojih se nakon primene metode najmanjih kvadrata (*least square method*)[40] dobijaju iste jednačine linearne regresije[41], dodatno je ukupna suma kvadrata odstupanja ista u svim slučajevima. Zanimljivost proizilazi iz grafičkog prikaza tih skupova podataka, kad se očigledno vidi da u jednom slučaju dobijena linearna zavisnost odgovara podacima, u drugom – zavisnost (prema grafiku) uopšte nije linearna, u trećem, zavisnost je linearna, ali linija koja se „dobro" slaže sa podacima, nije ista koja se dobija računski. U četvrtom slučaju su sve osim jedne vrednosti nezavisno promenljive *X* jednake, tako da je očigledno da dobijena prava (regresiona linija) nije ni blizu datim podacima.

Frensis Džon Enskomb (1918–2001), engleski statističar. Njegova je izreka: „Bolje je shvatiti šta problem stvarno jeste i rešiti ga makar približno, nego izmišljati neku zamenu koju možemo rešiti tačno".



→ *Nastavak na str. 258*

[40] Način određivanja nepoznatih koeficijenata u funkciji koja je izabrana kao model posmatrane pojave. Koeficijenti se nalaze iz uslova postizanja minimuma sume kvadrata odstupanja dobijenih od, po modelu, pretpostavljenih vrednosti.

[41] Kad je zavisnost veličine Y od veličine X linearna, govorimo o regresiji. Ako su podaci (x_i,y_i), $i = 1, n$, a zavisnost oblika $Y = aX + bm$ onda se po metodi najmanjih kvadrata a i b određuju iz uslova da suma $\sum_{i=1}^{n}(y_i - (ax_i - b))^2$ bude minimalna.

Erlangova raspodela

Ерлангова расподела

EN	Erlang distribution
FR	Distribution d’Erlang
RU	Распределение Эрланга
ZH	厄兰分布, Èlán fēnbù
AR	توزيع ارلانج
EL	Κατανομή Erlang
DE	Erlangsche Verteilung
ES	Distribución de Erlang
IT	Distribuzione di Erlang
SK	Erlangovo rozdelenie
SL	Erlangova porazdelitev
HU	Erlang-eloszlás

Povezani termini:

Eksponencijalna raspodela (66), gama raspodela (92)

Ako su slučajne veličine $X_1, X_2, \ldots, X_n$ nezavisne i sve imaju istu eksponencijalnu raspodelu $\varepsilon(\lambda)$, tada slučajna veličina $Y = X_1 + \ldots + X_n$ ima gustinu raspodele

$$g_Y(y) = \begin{cases} 0 & , \quad y \leq 0 \\ \dfrac{\lambda(\lambda y)^{n-1} e^{-\lambda y}}{(n-1)!} & , \quad y > 0 \end{cases}$$

i kaže se da Y ima **Erlangovu raspodelu** n-tog reda. Erlangova raspodela je specijalni slučaj gama raspodele. Zajedno sa eksponencijalnom raspodelom se često koristi u opisivanju sistema masovnog opsluživanja[42].

Agner Krarup Erlang (1878–1929), danski inženjer koji se bavio i teorijom verovatnoće. Radio je u kopenhagenskoj telefonskoj centrali, a njegov prvi rad se zvao "Teorija verovatnoće i telefonski razgovori".

[42] Sistemi masovnog opsluživanja (SMO) proučavaju raspodelu broja opsluženih klijenata, njihovo vreme čekanja do usluge, vreme trajanja usluživanja, itd. Kako su, dakle, u pitanju nizovi klijenata koji dolaze radi dobijanja određene usluge, to se u engleskom koristi termin *queuing system.* U SMO su bitne: raspodela vremena između uzastopnih dolazaka u sistem, raspodela vremena čekanja, raspodela vremena opsluživanja, kao i discipline opsluživanja (FIFO – *first-in-first-out*) ili FILO – *first-in-last-out*), broj kanala opsluživanja (npr. broj šaltera u pošti). Tako imamo jedno- i višekanalne sisteme opsluživanja, zatim sisteme sa otkazom (ako je klijent došao, a sistem je zauzet, klijent dobija otkaz, tj. ne bude primljen u sistem) i sisteme sa čekanjem (kad je bitno i to da li je vreme čekanja ograničeno ili ne), itd. Pri proučavanju sistema masovnog opsluživanja koriste se statističke metode za ocenjivanje parametara u raspodelama, kao i za opisivanje rada sistema u celosti. SMO se mogu uspešno proučavati i modeliranjem.

Fišer-Nejmanova teorema

Фишер-Нејманова теорема

EN Fisher-Neyman theorem
FR Théorème de Fisher-Neyman
RU Теорема Неймана-Фишера
ZH 费舍尔–奈曼定理, Fèishè'ěr-Nàimàn dìnglǐ
AR نظرية فيشر - نيمان
EL Θεώρημα Neyman-Fisher
DE Satz von Fisher-Neyman
ES Teorema de Fisher-Neyman
IT Teorema di Fisher-Neyman
SK Fisherova-Neymannova veta
SL Fisher-Neymanov izrek
HU Fisher-Neyman-tétel

Povezani termini:

Dovoljna statistika (58), funkcija verodostojnosti (90), raspodela uzorka

Etimologija:

Teorema < θεώρημα[EL] (opažanje, tvrđenje, iskaz) < θεωρέω[EL] (posmatrati) < θέα[EL] (pogled) + ὁράω[EL] (gledati)

Neka je θ nepoznati parametar u raspodeli obeležja X čija je gustina ili zakon raspodele obeležen sa $f(x,\theta)$. Pitanje utvrđivanja dovoljnosti neke statistike kojom se parametar θ ocenjuje primenom definicije dovoljne statistike može da bude vrlo složeno jer je neophodno znati uslovnu raspodelu uzorka pri fiksiranoj vrednosti posmatrane statistike. U nekim slučajevima[43] utvrđivanje dovoljnosti može jednostavno da se ostvari ako važi **Fišer-Nejmanova teorema o faktorizaciji**:

Statistika U je dovoljna statistika za parametar θ na osnovu uzorka obima n, ako i samo ako se raspodela uzorka (tj. funkcija verodostojnosti) može napisati u obliku proizvoda dve nenegativne funkcije, od kojih jedna zavisi od uzorka samo preko statistike U i može zavisiti od θ, a druga funkcija *ne zavisi* od θ.

Jirži Nejman (1894–1981), poljski statističar koji je najveći deo života proveo u Americi. Sarađivao je sa ocem i sinom – Karlom i Egonom Pirsonom i zajedno su izgrađivali teoriju testiranja hipoteza. Mnoga tvrđenja nose njihova imena.

[43] *Primer.* Obeležje X ima raspodelu eksponencijalnog tipa koja je data u obliku: $\exp\{A(\theta)K(x) + S(x) + q(\theta)\}$, gde je θ nepoznati parametar. Funkcija verodostojnosti za uzorak obima n će biti

$$L = \exp(A(\theta)U + nq(\theta)) \cdot \exp(\sum_{j=1}^{n} S(x_j)\text{ , gde je } U = \sum_{j=1}^{n} K(x_j)\text{ ,}$$

tako da je prva od funkcija u izrazu za L funkcija koja zavisi od (dovoljne) statistike U i od parametra θ, a druga funkcija zavisi samo od uzorka.

Frekvencija

Фреквенција

EN	Frequency
FR	Fréquence
RU	Частота
ZH	频率, pínlǜ
AR	التكرار
EL	Συχνότητα
DE	Häufigkeit
ES	Frecuencia
IT	Frequenza
SK	Frekvencia
SL	Frekvenca
HU	Gyakoriság

Povezani termini:

Obeležje (176), uzorak (240)

Etimologija:

Frekvencija < frequēns^{LA} (čest, prepun). Ova reč je, smatra se, nastala od proto-indoevropskog korena *bhrek- (sabiti se, nagurati se), pa je možda srodna i sa našom rečju *zbrka*.

Broj pojavljivanja neke vrednosti u uzorku je (apsolutna) **frekvencija** te vrednosti (tog modaliteta obeležja). Zbir svih apsolutnih frekvencija jednak je obimu uzorka[44] (*sample size*). Relativna frekvencija je količnik broja pojavljivanja neke vrednosti u uzorku i obima uzorka. Zbir relativnih frekvencija je jednak 1. Ako se podaci iz uzorka predstavljaju samo kao elementi pojedinih intervala, onda će se u okviru istog uzorka menjati frekvencije po intervalima kad se promene granice intervala. O ovoj jednostavnoj činjenici treba voditi računa, jer nevelike izmene graničnih tačaka intervala mogu da dovedu[45] do velikih izmena frekvencija po intervalima, a time do različitih ocena parametara i/ili do različitih zaključaka u slučaju testiranja hipoteza.

---

[44] Obim uzorka je broj elemenata u uzorku.

[45] *Primer.* Neka je u uzorku obima 21 dobijeno:

vrednost	0.1	0.6	0.9	1.1	1.6	1.9	2.1	2.6	2.9
frekvencija	1	5	1	1	5	1	1	5	1

Ako su potom formirani intervali:

interval	(0,0.5)	[0.5,1)	[1,1.5)	[1.5,2)	[2,2.5)	[2.5,3)
frekvencija	1	6	1	6	1	6

ali pri drugoj podeli, frekvencije po intervalima se značajno menjaju:

I	(0,0.6)	[0.6,1.2)	[1.2,1.8)	[1.8,2.2)	[2.2,2.6)	[2.6,3)
F	1	7	5	2	0	6

a ako izmenimo da intervali budu zatvoreni s desne strane, tada je:

I	(0,0.6]	(0.6,1.2]	(1.2,1.8]	(1.8,2.2]	(2.2,2.6]	(2.6,3]
F	6	2	5	2	5	1

Jedna od sugestija je i da se za granice intervala uzimaju vrednosti koji imaju veći broj decimala nego podaci iz uzorka. Na taj način podatak iz uzorka ne može biti granična tačka.

Funkcija gubitaka

Функција губитака

EN	Loss function
FR	Fonction de perte
RU	Функция потерь
ZH	损失函数, sǔnshī hánshù
AR	دالة الخساره
EL	Συνάρτηση απώλειας
DE	Verlustfunktion
ES	Función de pérdida
IT	Funzione di perdita
SK	Stratová funkcia
SL	Funkcija izgube
HU	Veszteségfüggvény

Povezani termini:

Bajesovske metode ocenjivanja, teorija odlučivanja

Etimologija:

Funkcija < fūnctus^{LA} (izvršen) < fungor^{LA} (obaviti, završiti)

Funkcija gubitaka je funkcija koju treba minimizirati u postupku određivanja modela na osnovu uzorka. Na primer, u metodi najmanjih kvadrata funkcija gubitka je zbir kvadrata reziduala – razlika između eksperimentalno dobijenih i po modelu pretpostavljenih vrednosti posmatrane pojave.

Funkcija gubitaka se takođe vezuje sa kvalitetom ocene. Uglavnom se posmatraju ograničene funkcije gubitaka. Često se pretpostavlja da je funkcija gubitaka konveksna[46], a radi jednostavnosti i da je to kvadratna funkcija.

U bajesovskom ocenjivanju najjednostavnije i najčešće korišćene funkcije gubitaka su:

a) kvadratna funkcija $(\theta - \hat{\theta})^2$, kada je ocena za θ jednaka uslovnom matematičkom očekivanju $E(\theta/X)$, gde X označava prost slučajni uzorak obima n,

b) apsolutna vrednost razlike parametra i ocene $|\theta - \hat{\theta}|$, kada je ocena za θ jednaka medijani uslovne raspodele θ/X, gde X označava prost slučajni uzorak obima n.

[46] Funkcija f je konveksna (*convex function*) na nekom intervalu ako za svake dve tačke a i b iz tog intervala i realni broj $c \in (0,1)$ važi $f(c \cdot a + (1 - c) \cdot b) \leq c \cdot f(a) + (1 - c) \cdot f(b)$. Drugim rečima, grafik funkcije se nalazi ispod sečice – odsečka koji spaja tačke $(a, f(a))$ i $(b, f(b))$. Ako se na grafik funkcije može konstruisati tangenta u svakoj tački posmatranog intervala, onda će funkcija biti konveksna ako se u svakoj tački tog intervala grafik funkcije nalazi iznad tangente na grafik funkcije u toj tački.

Funkcija gustine

Функција густине

EN	Probability density function
FR	Fonction de densité
RU	Функция плотности распределения
ZH	密度函数, mìdù hánshù
AR	دالة كثافة الاحتمال
EL	Συνάρτηση πυκνότητας πιθανότητας
DE	Wahrscheinlichkeitsdichtefunktion
ES	Función de densidad
IT	Funzione di densità
SK	Hustota pravdepodobnosti
SL	Funkcija gostote
HU	Sűrűségfüggvény

Povezani termini:

Disperzija (54), funkcija raspodele (88), histogram (108), matematičko očekivanje (150), slučajna veličina (198)

Etimologija:

Funkcija < fūnctus^{LA} (izvršen) < fungor^{LA} (obaviti, završiti)

Drugi naziv:

Gustina raspodele

Funkcija gustine ili **gustina raspodele** je funkcija koja se javlja u definiciji apsolutno neprekidne slučajne veličine: Ako je F funkcija raspodele slučajne veličine X i ako postoji funkcija g definisana na R i takva da za svako $x \in R$ važi:

$$1^{\text{o}}\ g(x) \geq 0\text{, i }\ 2^{\text{o}}\ F(x) = \int_{-\infty}^{x} g(t)dt$$

tada je F funkcija raspodele apsolutno neprekidne slučajne veličine X. Funkcija f se naziva gustina raspodele apsolutno neprekidne slučajne veličine X. Gustina raspodele je značajna pri proučavanju slučajnih veličina, a može biti korišćena pri zadavanju i prepoznavanju[47] slučajnih veličina.

Na osnovu osobina funkcije raspodele i definicije apsolutno neprekidne slučajne veličine zaključuje se da funkcija gustine (gustina raspodele) mora imati svojstva[48]: $g(x) \geq 0\ \ i\ \int_{-\infty}^{+\infty} g(t)dt = 1$. Ako je X apsolutno neprekidna slučajna veličina sa gustinom raspodele $g(x)$ i funkcijom raspodele $F(x)$, tada je: $g(x) = F'(x)$, a važi i

$$P\{a < X \leq b\} = \int_{a}^{b} g(x)dx = F(b) - F(a),\ \ a, b \in R,\ \ a \leq b.$$

Diskretna slučajna veličina nema gustinu raspodele.

---

[47] Na osnovu histograma za neko neprekidno obeležje može se najpre vizuelno napraviti poređenje sa nekom od poznatih gustina raspodele, a zatim nekim testom pretpostavku opovrgnuti ili prihvatiti.

[48] Bitno je da mora imati samo ta dva svojstva. Znači, gustina raspodele može biti, na primer, i prekidna funkcija ili imati vertikalnu asimptotu, ali ako zadovoljava navedena svojstva onda postoji slučajna veličina kojoj je posmatrana funkcija gustina raspodele. Stoga i svaki čitalac „Rečnika statističkih termina“ može definisati „svoju“ gustinu raspodele ☺

Funkcija moći testa

Функција моћи теста

EN Power function

FR Fonction puissance

RU Функция мощности критерия

ZH 幂函数, mì hánshù

AR منحني القوي

EL Καμπύλη ισχύος

DE Gütefunktion

ES Función de potencia

IT Funzione di potenza

SK Mocninová funkcia

SL Funkcija moči preizkusa

HU Hatványfüggvény

Povezani termini:

Nulta hipoteza (174), parametar, test

Etimologija:

Funkcija < fūnctus^{LA} (izvršen) < fungor^{LA} (obaviti, završiti)

Test < test^{stFR} (zemljani kotao, obično onaj u kome su isprobavani metali) < testum^{LA} (zemljana posuda)

Statistički testovi za testiranje parametarskih hipoteza se mogu na razne načine međusobno porediti. Jedan od načina za utvrđivanje kvaliteta testa je i njegova moć – pod tim podrazumevamo verovatnoću γ da pri tačnoj alternativnoj hipotezi H_1 vrednost test-statistike T_n bude u kritičnoj oblasti W određenoj pri pretpostavci da je nulta hipoteza tačna:

$$P_{H_1}(T_n \in W) = \gamma$$

Na taj način moć testa pokazuje kolike su šanse donošenja ispravne odluke pri određenoj alternativnoj vrednosti parametra. Veoma je dobro ako je moć testa bliska jedinici. Moć testa je funkcija vrednosti parametra, pa za sve vrednosti parametra koje obuhvata alternativna hipoteza se može posmatrati odgovarajući skup vrednosti i tako dobiti **funkcija moći testa**. Njenom analizom utvrđujemo za koje vrednosti parametra koje alternativna hipoteza dopušta, test ima najveću moć. Jasno je da će, u opštem slučaju, ako su obe hipoteze proste, moć testa biti veća ukoliko je razlika vrednosti parametra pri nultoj i pri alternativnoj hipotezi velika.

Ukoliko izračunavanje moći testa složeno, zbog oblika raspodele obeležja, ili nemoguće u slučaju da je raspodela obeležja nepoznata, onda se test ponavlja veliki broj puta na simuliranim podacima i moć testa se ocenjuje frekvencijom broja pravilnih odluka o odbacivanju nulte hipoteze kad ona nije tačna. Taj pristup se primenjuje i za neparametarske testove. Kako su to uvek obimna izračunavanja, jasno je da bez korišćenja računara ne bi ni mogla biti primenjena.

Funkcija raspodele

Функција расподеле

EN Distribution function
FR Fonction de répartition
RU Функция распределения
ZH 分布函数, fēnbù hánshù
AR دالة التوزيع
EL Συνάρτηση κατανομής
DE Verteilungsfunktion
ES Función de distribución
IT Funzione di ripartizione
SK Distribučná funkcia
SL Porazdelitvena funkcija
HU Eloszlásfüggvény

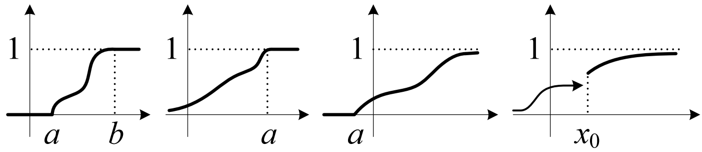
Povezani termini:

Centralna teorema mat. statistike (48), empirijska funkcija raspodele (72), funkcija gustine (82), obeležje (176), slučajna veličina (198)

Etimologija:

Funkcija < fūnctus[LA] (izvršen) < fungor[LA] (obaviti, završiti)

Grafik bilo koje **funkcije raspodele** je neopadajuća funkcija ograničena između 0 i 1, i neprekidna s desne strane. Sa grafika funkcije raspodele može da se ustanovi i koji je skup vrednosti odgovarajuće slučajne veličine.



Sa verovatnoćom 1 vrednosti ovih slučajnih veličina su, redom, u intervalima: (a,b), $(-\infty,a)$, (a,∞) i $(-\infty,\infty)$. Tačka x_0, u kojoj je skok funkcije raspodele, je vrednost koju slučajna veličina ima sa verovatnoćom različitom od nule.

Funkcija raspodele je važna funkcionalna karakteristika koja se odnosi na svaku slučajnu veličinu X, bez obzira na to koliko ona ima vrednosti. Realna funkcija F je funkcija raspodele slučajne veličine X ako je $F(x) = P(X \leq x)$, za svako $x \in R$. Odatle slede osobine:

1° F neopadajuća funkcija

2° $F(-\infty) = \lim\limits_{x\to-\infty} F(x) = 0,\ F(+\infty) = \lim\limits_{x\to\infty} F(x) = 1,$

3° F neprekidna sa desne strane za svako $x \in R$,

4° F ima najviše prebrojivo mnogo tačaka prekida (i to su prekidi prve vrste).

Sa teorijske tačke gledišta veoma je značajna činjenica da za svaku funkciju F koja zadovoljava uslove 1°, 2° i 3°, postoji prostor verovatnoća i na njemu zadana slučajna veličina X, kojoj je F funkcija raspodele. Takođe, ako slučajne veličine X i Y imaju istu funkciju raspodele, onda je $P[X \neq Y] = 0$.

→ *Nastavak na str. 258*

Funkcija verodostojnosti

Функција веродостојности

EN	Likelihood function
FR	Fonction de vraisemblance
RU	Функция правдоподобия
ZH	似然函数, sìrán hánshù
AR	دالة الارجحية
EL	Συνάρτηση πιθανοφάνειας
DE	Likelihood-Funktion
ES	Función de verosimilitud
IT	Funzione di verosimiglianza
SK	Vierohodnostná funkcia
SL	Funkcija verjetja
HU	Likelihood-függvény

Povezani termini:

Fišer-Nejmanova teorema (78), metoda maksimalne verodostojnosti (156), raspodela obeležja, test količnika verodostojnosti, uzorak (240)

Etimologija:

Funkcija < fūnctus^{LA} (izvršen) < fungor^{LA} (obaviti, završiti)

Neka je θ nepoznati parametar u raspodeli obeležja X i neka je Θ skup svih mogućih vrednosti parametra θ, tzv. parametarski prostor. Neka je

$$\vec{X} = (X_1, X_2, ..., X_n)$$

prost slučajni uzorak za neprekidno obeležje X i neka je $f(x,\theta)$ gustina raspodele obeležja X. Funkcija

$$L(\vec{X};\theta) = f(X_1;\theta) \cdot f(X_2;\theta) \cdot ... \cdot f(X_n;\theta)$$

se naziva **funkcija verodostojnosti**.

Ako je obeležje X diskretno sa zakonom raspodele $p(x,\theta) = P(X = x)$, onda je funkcija verodostojnosti

$$L(\vec{X};\theta) = p(X_1;\theta) \cdot p(X_2;\theta) \cdot ... \cdot p(X_n;\theta).$$

Na realizovanom uzorku funkcija verodostojnosti je funkcija parametra θ.

Funkcija verodostojnosti je funkcija raspodele prostog slučajnog uzorka.

Koristi se u metodi maksimalne verodostojnosti, testu količnika verodostojnosti, sekvencijalnom testu količnika verodostojnosti, Fišer-Nejmanovoj teoremi, itd.

Često se umesto funkcije verodostojnosti koristi njen logaritam

$$\ln L(\vec{X},\theta) = \sum_{j=1}^{n} \ln f(X_j;\theta),$$

zbog mogućih jednostavnijih izračunavanja i činjenice da će ekstremne vrednosti funkcija

$$L(\vec{X},\theta) \text{ i } \ln L(\vec{X},\theta)$$

biti u istim tačkama, jer je funkcija *ln* monotono rastuća funkcija.

Funkciju verodostojnosti, kao i metodu maksimalne verodostojnosti, uveo je ser Ronald Fišer.

Gama raspodela

Гама расподела

EN	Gamma distribution
FR	Distribution gamma
RU	Гамма-распределение
ZH	伽马分布, gāmǎ fēnbù
AR	توزيع جاما
EL	Κατανομή Γάμα
DE	Gamma-Verteilung
ES	Distribución gamma
IT	Distribuzione gamma
SK	Gamma rozdelenie
SL	Gamma porazdelitev
HU	Gamma-eloszlás

Povezani termini:

Disperzija (54), eksponencijalna raspodela (66), normalna raspodela (172), očekivanje (152)

Etimologija:

Gama (γάμμα) je treće slovo grčkog alfabeta. Pošto je grčki alfabet nastao iz feničanskog pisma, nazivi slova u grčkom jeziku su takođe preuzeti iz ovog semitskog jezika. Naziv *gama* (hebrejski ekvivalent *gimel*) potiče od proto-kanaanitskog korena koji označava praćku ili štap.

Slučajna veličina *X* ima **gama raspodelu** sa parametrima *a* i *b* ako je njena gustina raspodele oblika

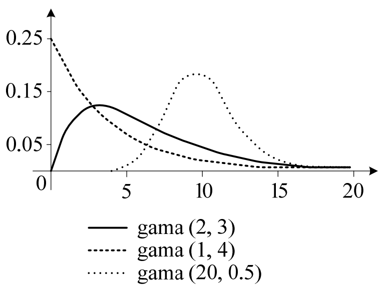
$$f(x) = \frac{1}{b^a \Gamma(a,b)} x^{a-1} \exp(-x/b),\ 0 < x\ .$$

Parametri *a* i *b* su pozitivni, a $\Gamma(a,b)$ označava gama funkciju, koja ovde služi kao normirajući faktor, tj. obezbeđuje da važi

$$\int_0^1 f(x)dx = 1\ .$$

Očekivana vrednost slučajne veličine sa gama raspodelom navedenog oblika je *ab,* dok je disperzija ab^2. Za različite vrednosti parametara se dobijaju razne familije gustina, a kao specijalni slučaj za $a = 1$, $b = 1$ se dobija eksponencijalna raspodela. Ako slučajna veličina *Z* ima normalnu normiranu raspodelu, onda $Z^2/2$ ima gama raspodelu sa $a = 1/2$, $b = 1$.

Gama raspodela spada u eksponencijalnu familiju raspodela, što daje mogućnost jednostavnog određivanja dovoljnih statistika za parametre ove raspodele.



Grafik

График

EN	Chart, diagram
FR	Graphique, diagramme
RU	График, диаграмма
ZH	统计图表, tǒngjì túbiǎo
AR	الرسم البيان
EL	Γράφημα, διάγραμμα
DE	Grafik, Diagramm
ES	Gráfico, diagrama
IT	Grafico, diagramma
SK	Graf, diagram
SL	Grafikon, diagram
HU	Gráf, diagramm

Povezani termini:

Uzorak (240), obeležje (176), populacija (182)

Etimologija:

Grafik < γράφω^{EL} (pisati, crtati)

Drugi naziv:

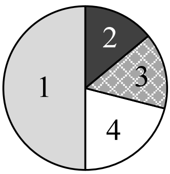
Dijagram

Postoje mnogobrojni i raznovrsni **grafici**. Na računaru se kod grafičkog prikazivanja podataka nudi obilje mogućnosti. Pogodno je i iskoristiti ih, i isprobati više različitih[49] grafikona za iste podatke.

Svaki dijagram mora imati objašnjenje („legendu“) o tome šta i na koji način (u procentima u odnosu na ceo uzorak/populaciju, u apsolutnim/relativnim frekvencijama, u hiljadama, itd.) predstavlja.

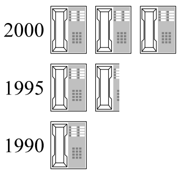
Na dijagramima se mogu porediti vrednosti posmatranog obeležja u raznim populacijama, tokom godina, itd.

Dijagrami su pogodni za predstavljanje jedno- i dvodimenzionalnih obeležja. Za višedimenzionalna obeležja grafičke mogućnosti reprezentacije su manje, ali postoje (npr. Černovljeva lica).

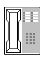


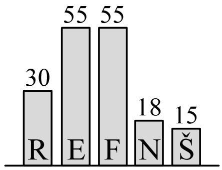
Anketa o omiljenoj boji odeće

1. Svetla: 50%
2. Tamna: 14%
3. Šarena: 15%
4. Svejedno: 21%



Proizvodnja telefona

 = 100 000 kom.



Broj učenika osmog razreda

R – ruski
E – engleski
F – francuski
N – nemački
Š – španski

---

[49] Kružni dijagram (*pie chart*) predstavlja podatke tako da su kružni isečci proporcionalni frekvencijama, piktogram (*pictogram*) predstavlja podatke sličicama koje liče na elemente populacije, dijagram sa dužima (*bar chart*) predstavlja frekvencije vertikalnim odsečcima.

Greška druge vrste

Грешка друге врсте

EN	Type II error
FR	Erreur de seconde espèce
RU	Ошибка второго рода
ZH	二型差误, èrxíng chāwù
AR	الخطأ من النوع الثاني
EL	Σφάλμα δεύτερου τύπου
DE	Fehler zweiter Art
ES	Error de tipo dos
IT	Errore di seconda specie
SK	Chyba druhého typu
SL	Napaka druge vrste
HU	Másodfajú hiba

Povezani termini:

Greška prve vrste (98), kritična oblast (136), test

Pojam koji se pojavljuje pri testiranju statističkih hipoteza, a odnosi se na donošenje odluke koja nije u skladu sa stvarnom, a istraživaču nepoznatom raspodelom obeležja. Hipotezu kojom se precizira raspodela obeležja X, bilo da je u pitanju vrednost nekog parametra ili sama raspodela, nazivamo nultom hipotezom i označavamo sa H_0, i njoj suprotstavljamo alternativnu hipotezu, u oznaci H_1. Moguće je da je alternativna hipoteza zaista tačna, a da u postupku testiranja, na osnovu posmatranog uzorka, odlučimo da prihvatamo nultu hipotezu kao tačnu. U tom slučaju, dakle ako se prihvata nulta hipoteza kada je tačna alternativna hipoteza, čini se **greška druge vrste** (ili drugog tipa). Verovatnoća ove greške treba da bude što manja. Nije moguće istovremeno smanjiti[50] i grešku prve vrste i grešku druge vrste.

Statistički test je u potpunosti određen kritičnom oblašću W_n kao skupom tačaka $(x_1, x_2, ..., x_n)$ u n-dimenzionalnom euklidskom prostoru na sledeći način. Ako realizovani uzorak $(x_1, x_2, ..., x_n) \in W_n$, tada se odbacuje H_0, a ako $(x_1, x_2, ..., x_n) \notin W_n$, onda se prihvata H_0. Prema tome, verovatnoća greške druge vrste je:

$$\beta = P_{H_1}\left\{\left(X_1, X_2, ..., X_n\right) \notin W_n\right\}.$$

[50] *Primer.* Pretpostavimo da je raspodela obeležja normalna, sa poznatom disperzijom i da se testira hipoteza da je očekivanje 5 protiv alternativne da je očekivanje 7. Ako je za neki prag značajnosti dobijena kritična oblast $(c,+\infty)$, onda bi za manji prag značajnosti kritična oblast bila $(c_1,+\infty)$, gde je $c_1 > c$. Ali je tada verovatnoća greške drugog tipa veća nego verovatnoća greške drugog tipa u odnosu na $(c,+\infty)$

Greška prve vrste

Грешка прве врсте

EN	Type I error
FR	Erreur de première espèce
RU	Ошибка первого рода
ZH	一型差误, yīxíng chāwù
AR	الخطأ من النوع الأول
EL	Σφάλμα πρώτου τύπου
DE	Fehler erster Art
ES	Error de tipo uno
IT	Errore di prima specie
SK	Chyba prvého typu
SL	Napaka prve vrste
HU	Elsőfajú hiba

Povezani termini:

Greška druge vrste (96), kritična oblast (136), test

Pojam koji se pojavljuje pri testiranju statističkih hipoteza, a odnosi se na donošenje odluke koja nije u skladu sa stvarnom, a istraživaču nepoznatom raspodelom obeležja. Hipotezu kojom se precizira raspodela obeležja X, bilo da je u pitanju vrednost nekog parametra ili sama raspodela, nazivamo nultom hipotezom i označavamo sa H_0, i njoj suprotstavljamo alternativnu hipotezu, u oznaci H_1. Moguće je da je nulta hipoteza zaista tačna, a da u postupku testiranja, na osnovu posmatranog uzorka, odlučimo da odbacimo nultu hipotezu kao netačnu. U tom slučaju, dakle ako se odbacuje nulta hipoteza kada ona jeste tačna hipoteza, čini se **greška prve vrste** (ili prvog tipa). Verovatnoća ove greške treba da bude što manja. Nije moguće istovremeno smanjivati i grešku prve vrste i grešku druge vrste.

Statistički test je u potpunosti određen kritičnom oblašću W_n kao skupom tačaka $(x_1, x_2, ..., x_n)$ u n-dimenzionalnom euklidskom prostoru na sledeći način. Ako realizovani uzorak $(x_1, x_2, ..., x_n) \in W_n$, tada se odbacuje H_0, a ako $(x_1, x_2, ..., x_n) \notin W_n$, onda se prihvata H_0. Prema tome, verovatnoće greške druge vrste je:

$$\alpha = P_{H_0}\left\{\left(X_1, X_2, ..., X_n\right) \in W_n\right\}.$$

Ova formula je „univerzalna“ u smislu da se sreće kod svih testova, i parametarskih i neparametarskih (slobodnih od raspodele).

Obično se uzima da je α jednako 0.01 ili 0.05. Stoga za neke testove kad se ne opisuje detaljno postupak, srećemo formulaciju „hipoteza se prihvata sa značajnošću 0.01, ako je vrednost statistike (navodi se koje) manja od vrednosti (navodi se konkretna tablična vrednost)“.

Gumbelova raspodela

Гумбелова расподела

EN Gumbel distribution
FR Distribution de Gumbel
RU Распределение Гумбеля
ZH 冈贝儿分布, Gāngbèi'ér fēnbù
AR توزيع جامبل
EL Κατανομή Gumbel
DE Gumbel-Verteilung
ES Distribución de Gumbel
IT Distribuzione di Gumbel
SK Gumbelovo rozdelenie
SL Gumbelova porazdelitev
HU Gumbel-eloszlás

Povezani termini:

Gustina raspodele (84), raspodela ekstremnih vrednosti

Drugi nazivi:

Dvostruko (dvojno) eksponencijalna raspodela, Fišer-Tipetova raspodela, raspodela ekstremnih vrednosti I tipa, Gompercova raspodela, log-Vejbulova raspodela.

Gumbelova raspodela je jedna od raspodela ekstremnih vrednosti[51]. Njena gustina raspodele je

$$g(x) = \frac{1}{\beta}\exp\left\{-\frac{x-\alpha}{\beta} - \exp\left(-\frac{x-\alpha}{\beta}\right)\right\}, -\infty < x < \infty.$$

a funkcija raspodele

$$F(x) = \exp\left\{-\exp\left(-\frac{x-\alpha}{\beta}\right)\right\}.$$

Mod raspodele je α, očekivanje $\alpha + \gamma\beta$, gde je γ Ojlerova konstanta, jednaka 0.5772... Disperzija Gumbelove raspodele je $\beta^2\pi^2/6$.

Emil Julijus Gumbel (1891–1966), nemački Jevrejin koji je zbog pacifističkih ideja od 1932. godine živeo najpre u Francuskoj, gde je na Univerzitetu u Lionu objavio rad o raspodeli koja danas nosi njegovo ime, a kasnije je otišao u Ameriku.

Raspodele ekstremnih vednosti se koriste u meteorologiji, hidrologiji, analizi preživljavanja (*survival analysis*)[52], ekonomiji, i u drugim oblastima.

---

[51] Polazna tačka za proučavanje raspodela ekstremnih vrednosti su raspodele statistika minimalnog i maksimalnog ranga u posmatranom prostom slučajnom uzorku obima n.

Ako je dat uzorak $(X_1, X_2,..., X_n)$, onda je statistika poretka prvog ranga $X_{min} = \min(X_1, X_2,..., X_n)$, a statistika poretka maksimalnog ranga je $X_{max} = \max(X_1, X_2,..., X_n)$. Kako je $\min(-f) = -\max f$, jasno je da je dovoljno proučavati raspodelu samo za minimum, ili samo za maksimum.

[52] Oblast primene koja se koristi, na primer, u medicini, kad se određuju šanse da pacijentov životni vek potraje preko neke granice.

Hi-kvadrat raspodela

Хи-квадрат расподела

EN Chi-squared distribution
FR Distribution de khi carré
RU Хи-квадрат распределение
ZH 卡方分布, kǎfāng fēnbù
AR توزيع كاي تربيع
EL Κατανομή Χι-τετράγωνο
DE Chi-Quadrat-Verteilung
ES Distribución de ji-cuadrado
IT Distribuzione chi-quadrato
SK Chí-kvadrát rozdelenie
SL Hi-kvadrat porazdelitev
HU Khi-négyzet eloszlás

Povezani termini:

Broj stepeni slobode (44), gama raspodela (90), gustina raspodele (84), normalna raspodela (172)

Etimologija:

Kvadrat < quadrō^{LA} (dopuniti, urediti) < quattuor^{LA} (četiri); latinska reč *quattuor* i naše *četiri* imaju zajednički koren u proto-indoevropskom *k^{w}etwóres.

Neka su slučajne veličine X_1, X_2,..., X_n nezavisne i neka sve imaju $N(0,1)$ raspodelu. Za slučajnu veličinu $X = X_1^2 + X_2^2 + ... + X_n^2$ se kaže da ima **χ^2-raspodelu** sa n stepeni slobode, što se označava sa $X : \chi_n^2$, a čita "hi-kvadrat sa n stepeni slobode". Gustina raspodele slučajne veličine je $X : \chi_n^2$.

$$g(x) = \begin{cases} 0 & , \quad x \le 0 \\ \dfrac{x^{n/2-1} e^{-x/2}}{2^{n/2} G\left(\dfrac{n}{2}\right)} & , \quad x > 0 \end{cases} \qquad (*)$$

U izrazu (*) G označava gama funkciju, a broj n može biti proizvoljan pozitivan broj, ali s obzirom na prethodnu relaciju i zbog primena u statistici najčešće se posmatra slučaj kada je n prirodan broj.

Verovatnoće vezane za χ^2-raspodelu se daju u tablicama, jer je neposredno izračunavanje komplikovano.

Ako $X : \chi_n^2$, tada je $E(X) = n$, $D(X) = 2n$. Za dovoljno velike vrednosti n, raspodela χ_n^2 može se aproksimirati normalnom raspodelom $N(n,2n)$. Ova aproksimacija se primenjuje za $n \ge 30$, pa je to ujedno i razlog što u tablicama često ne figurišu vrednosti n veće od 30.

Ako su $X : \chi_n^2$ i $Y : \chi_m^2$ nezavisne slučajne veličine, tada je $Z = X + Y$ slučajna veličina sa χ_{n+m}^2 raspodelom.

Za $n > 2$ karakteristični oblik grafika gustine raspodele χ_n^2 je dat na slici.

χ_n^2-raspodela se javlja kao tačna ili aproksimirajuća raspodela nekih statistika, i stoga se javlja pri određivanju intervala poverenja ili pri testiranju statističkih hipoteza.

Hi-kvadrat test

Хи-квадрат тест

EN Chi-squared test
FR Test du khi carré
RU Критерий хи-квадрат
ZH 卡方检验, kǎfāng jiǎnyàn
AR اختبارمربع كاي
EL Έλεγχος Χι-τετράγωνο
DE Chi-Quadrat-Test
ES Prueba de ji-cuadrado
IT Test chi-quadrato
SK Chí-kvadrát test
SL Hi-kvadrat preizkus
HU Khi-négyzet próba

Povezani termini:

Binomna raspodela (38), neparametarski test (162), test-statistika

Etimologija:

Kvadrat < quadrō^{LA} (dopuniti, urediti) < quattuor^{LA} (četiri); latinska reč *quattuor* i naše *četiri* imaju zajednički koren u proto-indoevropskom *k^{w}etwóres.

Test < test^{stFR} (zemljani kotao, obično onaj u kome su isprobavani metali) < testum^{LA} (zemljana posuda)

Drugi naziv:

U literaturi na engleskom sreće se i termin *Pearson goodness--of-fit test*

Neka je u raspodeli obeležja X nepoznato s parametara. Parametri se ocenjuju na osnovu uzorka, pomoću metode maksimalne verodostojnosti. Ako se parametri ocenjuju po nekoj drugoj metodi, asimptotska raspodela test-statistike (*) ne mora da bude χ^2-tipa. Skup mogućih vrednosti obeležja X se razbija na r disjunktnih podskupova $S_1, S_2, \ldots, S_r$[53]. Neka m_j broj elemenata iz uzorka koji su u skupu S_j, $j = 1, 2, \ldots, r$. Brojevi m_j su realizovane vrednosti slučajnih veličina M_j, koje, pri nultoj hipotezi, imaju binomnu raspodelu $B(n,p_j)$, $j = 1, 2, \ldots, r$, gde je

$$p_j = P_{H_0}\left\{X \in S_j\right\}.$$

Te verovatnoće se nalaze prema postavljenoj hipotezi, tj. na osnovu funkcije raspodele $F_0(x)$. Test-statistika je:

$$\chi_U^2 = \sum_{j=1}^{r} \frac{\left(M_j - np_j\right)^2}{np_j}. \qquad (*)$$

Ako je hipoteza H_0 tačna, pri $n \to \infty$, test-statistika ima asimptotski χ_{r-s-1} raspodelu. Za prag značajnosti α, pomoću tablice za χ^2-raspodelu dobija se $c = \chi^2_{r-s-1;\alpha}$ tako da važi

$$P\left\{\chi_U^2 \geq c\right\} = \alpha.$$

→ *Nastavak na str. 258*

---

[53] Podela skupa dopustivih vrednosti obeležja na podskupove (klase) je, donekle, proizvoljna. Ako su moguće vrednosti obeležja svi realni brojevi, onda skupovi $S_1, S_2, \ldots, S_r$ predstavljaju jednu podelu intervala $(-\infty,\infty)$, ako su moguće vrednosti samo pozitivni realni brojevi, onda skupovi $S_1, S_2, \ldots, S_r$ predstavljaju jednu podelu intervala $(0,\infty)$, a ako obeležje ima samo konačno mnogo vrednosti, onda će skupovi $S_1, S_2, \ldots, S_r$ predstavljati ili pojedinačne vrednosti obeležja ili unije nekoliko vrednosti. Ako obeležje ima beskonačno, ali prebrojivo mnogo vrednosti, onda će skupovi $S_1, S_2, \ldots, S_r$ predstavljati neke pojedinačne vrednosti obeležja kao i uniju vrednosti.

Hi-kvadrat test nezavisnosti

Хи-квадрат тест независности

EN Chi-squared test for independence
FR Test d’indépendance du khi carré
RU Критерий независимости хи-квадрат
ZH 卡方独立性检验, kǎfāng dúlìxìng jiǎnyàn
AR اختبار مربع كاي للأستقلاية
EL Έλεγχος ανεξαρτησίας με το τεστ Χι-τετράγωνο
DE Chi-Quadrat-Unabhängigkeitstest
ES Prueba de independencia de ji-cuadrado
IT Test chi-quadrato di indipendenza
SK Chí-kvadrát test nezávislosti
SL Hi-kvadrat preizkus hipoteze neodvisnosti
HU Khi-négyzet függetlenségi próbá

Povezani termini:

Dvodimenzionalna raspodela (60), hi-kvadrat test (104), neparametarski test (162), nezavisna slučajna veličina, ocenjivanje verovatnoća, zajednička raspodela

Etimologija:

Kvadrat < quadrō$^{\text{LA}}$ (dopuniti, urediti) < quattuor$^{\text{LA}}$ (četiri); latinska reč *quattuor* i naše *četiri* imaju zajednički koren u proto-indoevropskom *k$^{\text{w}}$etwóres.

Test < test$^{\text{stFR}}$ (zemljani kotao, obično onaj u kome su isprobavani metali) < testum$^{\text{LA}}$ (zemljana posuda)

Ispitivanje da li su dva obeležja X i Y nezavisna ili zavisna je čest zadatak u matematičkoj statistici. Hipoteza $H_0(F_{X,Y} = F_X \cdot F_Y)$ je hipoteza nezavisnosti, a testira se protiv alternativne hipoteze $H_1(F_{X,Y} \neq F_X \cdot F_Y)$.

Testiranje se ostvaruje na osnovu dvodimenzionalnog prostog slučajnog uzorka obima n: (X_i, Y_i), $i = 1, 2,..., n$. Neka su podaci iz dvodimenzionalnog uzorka obima n raspoređeni u r kategorija obeležja X, a u s kategorija po vrednostima obeležja Y[54]. Ti podaci se onda daju u tabeli kontingencije (tabeli povezanosti):

$X \setminus Y$	y_1	y_2	…	y_s	zbir
x_1	n_{11}	n_{12}	…	n_{1s}	$n(x_1)$
x_2	n_{21}	n_{22}	…	n_{2s}	$n(x_2)$
…	…	…	…	…	…
x_r	n_{r1}	n_{r2}	…	n_{rs}	$n(x_r)$
zbir	$n(y_1)$	$n(y_1)$	…	$n(x_s)$	N

Brojevi u središnjem delu tabele označavaju da se par (x_i, y_j) pojavio n_{ij} puta u uzorku. Zbir

$$n = \sum_{i=1}^{r} \sum_{j=1}^{s} n_{ij}$$

je jednak obimu uzorka. Neka su marginalni zbirovi po vrstama, tj. po vrednostima obeležja X: $n(x_1)$, $n(x_2)$,..., $n(x_r)$, a po kolonama, po vrednostima obeležja Y: $n(y_1)$, $n(y_2)$,..., $n(y_s)$.

→ *Nastavak na str. 259*

[54] Kako za Y, tako i za X vrednosti mogu biti date i opisno (X – boja očiju, Y – pol osobe), i po intervalima (X – visina, Y – težina jedinke).

Histogram

Хистограм

EN	Histogram
FR	Histogramme
RU	Гистограмма
ZH	柱状图, zhùzhuàngtú
AR	المدرج التكراري
EL	Ιστόγραμμα
DE	Histogramm
ES	Histograma
IT	Istogramma
SK	Histogram
SL	Histogram
HU	Hisztogram

Povezani termini:

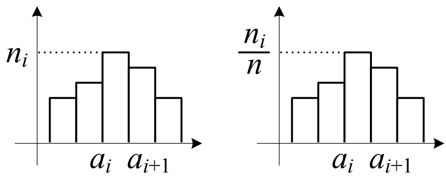
Frekvencija (80), gustina raspodele (84), hi-kvadrat test (104), uzorak (240)

Etimologija:

Histogram < ἱστός^{EL} (jarbol) + γράμμα^{EL} (slovo, znak).

Histogrami se dobijaju kad su podaci iz uzorka dati u obliku tabele u kojoj se nalaze intervali mogućih vrednosti obeležja i odgovarajuće frekvencije.

U koordinatnom sistemu se predstave dati intervali (a_i, a_{i+1}), $i = 1, 2,..., k$ i nad njima nacrtaju pravougaonici visine jednake apsolutnim n_i ili relativnim frekvencijama n_i/n, $i = 1, 2,..., k$. Po izgledu su obe slike iste, ali je različita podela na vertikalnoj osi.



Ukoliko je zbir površinâ svih pravougaonika jednak 1, tada histogram predstavlja uzorački ekvivalent gustini raspodele obeležja. To je moguće, na primer, ako su svi intervali dužine 1 i ako su visine pravougaonika relativne frekvencije n_i/n, $i = 1, 2,..., k$[55].

Poređenje histograma sa gustinom raspodele obeležja koristi se, na primer, u Pirsonovom χ^2-testu.

[55] Ako su dužine intervala jednake onda će histogram sa apsolutnim, histogram sa relativnim frekvencijama i histogram sa ukupnom površinom 1 izgledati isto. To je razlog što se opredeljujemo za jednake intervale. U nekim situacijama se mogu javiti intervali različitih dužina i tada treba vodti računa o tome da će histogrami sa apsolutnim i relativnim frekvencijama izgledati isto, ali da histogram sa ukupnom površinom 1 može izgledati značajno drukčije, jer se tada visine pravougaonika računaju po formulu n_i/nd_i, gde su d_i dužine intervala, $i = 1, 2,..., k$.

→ *Primer na str. 259*

Homogenost

Хомогеност

EN	Homogeneity
FR	Homogénéité
RU	Однородность
ZH	同质性, tóngzhìxìng
AR	التجانس
EL	Ομοιογένεια
DE	Einheitlichkeit
ES	Homogeneidad
IT	Omogeneità
SK	Homogenita
SL	Enovitost
HU	Homogenitás

Povezani termini:

Analiza varijanse (56), disperzija (54), populacija (182)

Etimologija:

Homogenost < ὁμός[EL] (isti) + γένος[EL] (vrsta)

Zahtev da podaci iz različitih potpopulacija[56] posmatrane populacije imaju neku međusobnu sličnost, se naziva **homogenost**. Na primer, u analizi varijanse, to će biti zahtev da disperzije obeležja na potpopulacijama budu jednake, ili, bar, vrlo bliske po vrednosti. Ukoliko to nije ispunjeno, onda nema smisla nastaviti sa izračunavanjima po šemi analize varijanse, jer postupak dovodi do rešenja koje ima smisla interpretirati samo ako je postupak primenjen uz sve potrebne preduslove.

Ako su disperzije obeležja na potpopulacijama jednake, tada se u engleskom jeziku za to koristi termin *homoskedastic*, odnosno *heteroskedastic* ako su disperzije na potpopulacijama različite.

[56] Potpopulacije (*subpopulation*) su delovi populacije izdvojeni po nekom kriterijumu (npr. učenici škole po razredima ili pak podela prema stranom jeziku koji uče).

Indeks krivolinijske korelacije

Индекс криволинијске корелације

EN	Index of non-linear correlation
FR	Indice de corrélation non-linéaire
RU	Индекс нелинейной корреляции
ZH	非线性相关指数, fēixiànxìng xiāngguān zhǐshù
AR	مؤشر الارتباط الابحبائي
EL	Δείκτης μή-γραμμικής συσχέτισης
DE	Nichtlinearer Korrelationskoeffizient
ES	Índice de correlación no lineal
IT	Indice di correlazione non lineare
SK	Index nelineárne korelácie
SL	Indeks nelinearne korelacije
HU	Nemlineáris korrelációs index

Povezani termini:

Korelacija, metoda najmanjih kvadrata (75), ocenjivanje parametara (178)

Etimologija:

Indeks < index[LA] (uhoda; kažiprst) < indicō[LA] (ukazati)

Linija < līnea[LA] (konac, nit) < līnum[LA] (lan)

Korelacija < correlatio[LA] (međusobna povezanost) < co-[LA] (sa) + relatio[LA] (odnos) < referō[LA] (nositi nazad) < re[LA] (ponovo) + ferre[LA] (nositi)

Zavisnost obeležja (zavisno promenljive) Y od obeležja X (nezavisno promenljive) može biti, ukoliko postoji, različitih oblika. Najjednostavnija je linearna zavisnost, ali su mogući i svi drugi (krivolinijski) oblici zavisnosti. Koeficijente funkcije $Y = f(X, a_1, \ldots, a_k)$ odredimo metodom najmanjih kvadrata koristeći podatke (x_j, y_j), $j = 1,2,\ldots,n$. Tada dobijamo vrednosti (ocenjene) za parametre kao i ocenjene vrednosti za Y u tačkama x_j, $j = 1,2,\ldots,n$.

Bliskost raspoloživih podataka i dobijene krive $y = f(x)$ može da se meri i **indeksom krivolinijske korelacije** koji je jednak:

$$R^2 = 1 - \frac{\sum_{j=1}^{n} (y_j - \hat{y}_j)^2}{\sum_{j=1}^{n} (y_j - \overline{y}_n)^2}$$

U brojiocu su razlike stvarnih i ocenjenih vrednosti za Y, a u imeniocu razlike stvarnih vrednosti i uzoračke sredine za Y.

Bitne osobine su da je indeks krivolinijske korelacije uvek između 0 i 1, a u slučaju jednostruke linearne regresije jednak kvadratu koeficijenta korelacije veličina X i Y.

Interval poverenja

Интервал поверења

EN	Confidence interval
FR	Intervalle de confiance
RU	Доверительный интерваль
ZH	置信区间, zhìxìn qūjiān
AR	فترة الثقة
EL	Διάστημα εμπιστοσύνης
DE	Konfidenzintervall
ES	Intervalo de confianza
IT	Intervallo di confidenza
SK	Konfidenčný interval
SL	Interval zaupanja
HU	Konfidenciaintervallum

Povezani termini:

Ocenjivanje parametara (178), statistike (167), tačkaste ocene (179), uzorak (240)

Etimologija:

Interval < intervallum^{LA} (međuprostor, vremenski period) < inter^{LA} (između) + vallum^{LA} (bedem, zid)

Neka je θ nepoznati parametar u raspodeli obeležja X i neka je $(X_1, X_2,..., X_n)$ prost slučajan uzorak obima n za to obeležje. Neka su definisane statistike $f(X_1, X_2,..., X_n)$ i $g(X_1, X_2,..., X_n)$ tako da važe uslovi:

$$P\{f \le g\} = 1,$$
$$P\{f \le \theta \le g\} = \beta, \beta \in [0,1].$$

Tada se $[f, g]$ naziva **interval poverenja** za nepoznati parametar θ sa nivoom poverenja β. Obično se kaže da je to $100\beta\%$ interval poverenja za nepoznati parametar. Realizovana vrednost tačkaste ocene parametra može dosta odstupati od stvarne vrednosti parametra, a da je pri tom nepoznato koliko je to odstupanje. Stoga se, na osnovu prostog slučajnog uzorka i određuje interval koji, sa unapred zadatom verovatnoćom, sadrži nepoznati parametar.

U primenama se uzima da je vrednost nivoa poverenja β bliska jedinici (najčešće je $\beta = 0.9$ ili $\beta = 0.95$).

Smisao ovako određenih intervala poverenja je da očekujemo da od (velikog broja) K uzoraka istog obima njih $K\beta$ sadrži nepoznati parametar.

S obzirom da je suštinski na raspolaganju samo jedna jednačina, a to je $P\{f \le \theta \le g\} = \beta$, u izrazima za statistike f i g može da figuriše samo jedna nepoznata veličina. Ukoliko ih ima više zadatak određivanja granica će imati više rešenja. Da bi se dobilo jedno rešenje, treba postaviti dodatne uslove.

U zavisnosti od prirode parametra ili posmatrane pojave može jedna od granica intervala poverenja biti fiksirani broj (npr. u slučaju intervala poverenja za disperziju često je $f = 0$).

Interval poverenja za očekivanje

Интервал поверења за очекивање

EN Confidence interval for the mean
FR Intervalle de confiance pour l’espérance
RU Доверительный интерваль для ожидания
ZH 期望的置信区间, qīwàng de zhìxìn qūjiān
AR فترة الثقة للمتوسط
EL Διάστημα εμπιστοσύνης για το μέσο όρο
DE Konfidenzintervall für Erwartungswert
ES Intervalo de confianza para la esperanza
IT Intervallo di confidenza per la speranza
SK Konfidenčný interval pre strednú hodnotu
SL Interval zaupanja za upanje
HU A várható érték konfidenciaintervalluma

Povezani termini:

Interval poverenja (114), kontrola kvaliteta (128), matematičko očekivanje (152), uzoračka disperzija (224), uzoračka sredina (230)

Etimologija:

Interval < inter[LA] (između) + vallum[LA] (bedem, zid)

Matematika < μαθηματικός[EL] (onaj ko voli da uči) < μάθημα[EL] (znanje, učenje, nauka)

A) **Interval poverenja** za matematičko očekivanje m obeležja X sa normalnom raspodelom i poznatom disperzijom sa nivoom poverenja β je simetričan u odnosu na $\overline{X}_n$. Iz uslova $P\{\overline{X}_n - \varepsilon \leq m \leq \overline{X}_n + \varepsilon\} = \beta$ se određuje ε. Koristeći tablice za normalnu raspodelu za dato β, nalazi se z_β i tako se dobija interval poverenja za matematičko očekivanje obeležja X:

$$\left[\overline{X}_n - \frac{z_\beta}{\sqrt{n}}\sigma,\ \overline{X}_n + \frac{z_\beta}{\sqrt{n}}\sigma\right].$$

Realizovana vrednost $\overline{x}_n$ daje realizovani interval poverenja. Njegova dužina predstavlja konstantnu veličinu pri datim vrednostima β, σ i n. S druge strane, ako se ne menjaju σ i n, a povećava vrednost β, širi se i interval poverenja. Ako se ne menjaju β i σ, a povećava n, interval se sužava.

B) **Interval poverenja** za matematičko očekivanje m obeležja X sa normalnom raspodelom i nepoznatom disperzijom je oblika

$$\left[\overline{X}_n - \frac{t_{n-1;1-\beta}}{\sqrt{n-1}}\overline{S}_n,\ \overline{X}_n + \frac{t_{n-1;1-\beta}}{\sqrt{n-1}}\overline{S}_n\right]$$

Dužina intervala poverenja u ovom slučaju predstavlja slučajnu veličinu, jer je funkcija slučajne veličine $\overline{S}_n$. Ako se za isti uzorak menja β, interval poverenja će se povećavati sa povećanjem β.

C) Postupak opisan pod B) se primenjuje pri određivanju intervala poverenja za matematičko očekivanje obeležja koje nema normalnu raspodelu, a u pitanju je veliki uzorak, jer na osnovu centralne granične teoreme statistika $\overline{X}_n$ ima približno normalnu raspodelu, bez obzira na raspodelu obeležja X.

→ *Nastavak na str. 260*

Izvestan događaj

Известан догађај

EN	Certain event
FR	Evénement certain
RU	Достоверное событие
ZH	必然事件, bìrán shìjiàn
AR	الحدث المؤكد
EL	Βέβαιο γεγονός
DE	Sicheres Ereignis
ES	Evento deterministico
IT	Evento certo
SK	Istá udalosť
SL	Gotov dogodek
HU	Biztos esemény

Povezani termini:

Populacija (182), prostor elementarnih ishoda (190), slučajni događaj (200), uzorak (240)

Drugi naziv:

Siguran događaj

Slučajni događaj A je svaki podskup skupa ishoda Ω eksperimenta sa slučajnim ishodima, tj. $A \subseteq \Omega$. Stoga i ceo skup elementarnih ishoda Ω predstavlja slučajan događaj. Taj slučajni događaj se sigurno realizuje pri svakom izvođenju eksperimenta, pa se stoga zove **izvestan događaj** ili **siguran događaj**. Komplement sigurnog događaja je nemoguć događaj.

Unija proizvoljnog događaja i njegovog komplementa je siguran događaj. Verovatnoća sigurnog događaja je jednaka 1. Verovatnoća nemogućeg događaja je 0[57].

U matematičkoj statistici skupu Ω odgovara cela populacija, a slučajni izbor jednog elementa populacije je analogon elementarnom ishodu. Stoga je (prost slučajni) uzorak analogon slučajnom događaju, a analogon izvesnom događaju je situacija u kojoj ispitujemo populaciju u celini. Zato se u teoriji uzoraka kada se razmatraju razne vrste uzoraka iz konačnih populacija sreću formule iz kombinatorike, iste one koje se javljaju pri izračunavanju verovatnoća događaja primenom klasične definicije verovatnoće .

---

[57] Skoro nemoguć događaj je događaj A čija je verovatnoća jednaka 0, ali je pri tom $A \neq \Phi$. Za razliku od nemogućeg, koji uvek postoji, skoro nemoguć događaj ne mora postojati u svakom prostoru elementarnih ishoda.

Primeri skoro nemogućih događaja sreću se kad prostor elementarnih ishoda ima neprebrojivo mnogo elemenata. Na primer, u eksperimentu slučajnog izbora tačke u krugu, skoro nemogući događaji su: „izbor centra kruga“, „izbor jedne određene tačke u krugu“, „izbor tačke na nekoj određenoj tetivi kruga“, itd.

Koeficijent asimetrije

Коефицијент асиметрије

EN	Coefficient of skewness
FR	Coefficient d’asymétrie
RU	Коэффициент асимметрии
ZH	偏斜系数, piānxié xìshù
AR	معامل الالتواء
EL	Συντελεστής ασυμμετρίας
DE	Schiefekoeffizient
ES	Coeficiente de asimetría
IT	Coefficiente di asimmetria
SK	Koeficient šikmosti
SL	Koeficient asimetrije
HU	Asszimetria mutató

Povezani termini:

Disperzija (54), matematičko očekivanje (152), raspodela

Etimologija:

Koeficijent < co-^{LA} (sa) + efficiēns^{LA} (onaj koji izvršava) < ē^{LA} (iz) + faciō^{LA} (raditi); ovaj termin je prvi upotrebio francuski matematičar Fransoa Vijet (1540–1603), otac moderne algebre.

Starogrčki prefiks ἀ označava negaciju. U tom smislu ovaj termin možemo prevesti i opisno kao „odsustvo simetrije“.

Simetrija < συμμετρία^{EL} < σύν^{EL} (sa) + μέτρον^{EL} (mera)

Jedna mera asimetrije je **koeficijent asimetrije** koji se računa po formuli

$$f_1 = \frac{E(X-m)^3}{\sigma^3},$$

gde su $E(X) = m$ i $D(X) = \sigma^2$ matematičko očekivanje i disperzija slučajne veličine X. Koeficijent asimetrije je broj koji ne zavisi od mernih jedinica. Ako je raspodela simetrična sa konačnim momentima prvog, drugog i trećeg reda, koeficijent asimetrije, je, prema datoj formuli, jednak 0.

Druge moguće mere asimetrije su:

a) Pirsonov koeficijent asimetrije (*Pearson's coefficient of skewness*) jednak količniku $(E(X)-\mu)/\sigma$, gde je μ mod (ili medijana, ako X nema mod).

b) Kvartilni koeficijent asimetrije (*Quartile coefficient of skewness* ili *Bowley coefficient of skewness*) jednak količniku $(Q_1+Q_3-2Q_2)/(Q_3-Q_1)$, gde je sa Q_j, $j = 1,2,3$ označen j-ti kvartil raspodele X. Ovaj koeficijent ima vrednosti u intervalu $[-1,1]$.

Ukoliko za posmatrano obeležje X imamo uzorak obima n, onda ćemo asimetriju raspodele obeležja oceniti uzoračkim koeficijentom asimetrije:

$$\frac{\frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n)^3}{\left(\sqrt{\frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2}\right)^3}$$

Koeficijent korelacije

Коефицијент корелације

EN Coefficient of correlation
FR Coefficient de corrélation
RU Коэффициент корреляции
ZH 相关系数, xiāngguān xìshù
AR معامل الارتباط
EL Συντελεστής συσχέτισης
DE Korrelationskoeffizient
ES Coeficiente de correlación
IT Coefficiente di correlazione
SK Korelačný koeficient
SL Koeficient korelacije
HU Korrelációs együttható

Povezani termini:

Disperzija (54), matematičko očekivanje (152), parcijalni koeficijent korelacije (123)

Etimologija:

Koeficijent < co-[LA] (sa) + efficiēns[LA] (onaj koji izvršava) < ē[LA] (iz) + faciō[LA] (raditi); ovaj termin je prvi upotrebio francuski matematičar Fransoa Vijet (1540–1603), otac moderne algebre.

Korelacija < correlatio[LA] (međusobna povezanost) < co-[LA] (sa) + relatio[LA] (odnos) < referō[LA] (nositi nazad) < re[LA] (ponovo) + ferre[LA] (nositi)

Koeficijent korelacije meri stepen zavisnosti slučajnih veličina X i Y koje imaju konačna očekivanja i disperzije i jednak je

$$\rho_{X,Y} = \frac{E(X - E(X))(Y - E(Y))}{\sqrt{D(X)D(Y)}}.$$

Ako su X i Y nezavisne slučajne veličine, tada je $\rho_{X,Y} = 0$, dok obrnuto ne važi[58]. Vrednosti koeficijenata korelacije se nalaze u intervalu $[-1,1]$. Ako je $Y = aX + b$ i $a > 0$, tada je $\rho_{X,Y} = 1$, a ako je $a < 0$ tada je $\rho_{X,Y} = -1$. Isto tako, ako je $\rho_{X,Y} = 1$ili $\rho_{X,Y} = -1$, tada je Y linearna funkcija od X (sa verovatnoćom 1). Zato se ovaj koeficijent često naziva **koeficijentom linearne korelacije**.

Koeficijent korelacije može biti blizak 1 (ili –1) ali to ne mora značiti da su X i Y međusobno linearno zavisne, već da postoji treća slučajna veličina Z, od koje zavise i X i Y. Korišćenjem parcijalnog koeficijenta korelacije nalazi se zavisnost X od Y bez uticaja Z.

Parcijalni koeficijent korelacije X i Y bez uticaja Z je

$$\rho_{X,Y} = \frac{\rho_{XY} - \rho_{XZ}\rho_{YZ}}{\sqrt{(1-\rho_{XZ}^2)(1-\rho_{YZ}^2)}}.$$

Odgovarajuća vrednost na osnovu uzorka za sva tri posmatrana obeležja se računa po formulama za uzorački koeficijent korelacije.

[58] *Primer*. Ako je X sa simetričnom raspodelom i očekivanjem 0 i ukoliko postoji $E(X^3)$ biće takođe jednako 0. Neka je $Y = X^2$, onda je $E(XY) = E(X^3) = 0$, pa je koeficijent korelacije jednak 0, iako su slučajne veličine X i Y zavisne.

Koeficijent spljoštenosti

Коефицијент спљоштености

EN	Coefficient of kurtosis
FR	Coefficient d'aplatissement
RU	Коэффициент эксцесса
ZH	峰态系数, fēngtài xìshù
AR	معامل التفرطح
EL	Συντελεστής κύρτωσης
DE	Exzeßkoeffizient
ES	Coeficiente de exceso
IT	Coefficiente di eccesso
SK	Koeficient špicatosti
SL	Koeficient presežka
HU	Laposság

Povezani termini:

Disperzija (54), matematičko očekivanje (152), raspodela

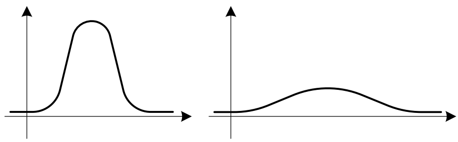
Etimologija:

Koeficijent < co-^{LA} (sa) + efficiēns^{LA} (onaj koji izvršava) < ē^{LA} (iz) + faciō^{LA} (raditi); ovaj termin je prvi upotrebio francuski matematičar Fransoa Vijet (1540–1603), otac moderne algebre.

Koeficijent spljoštenosti, kao i koeficijent asimetrije su mere oblika grafika gustine raspodele. Neka su $E(X) = m$ i $D(X) = \sigma^2$ matematičko očekivanje i disperzija slučajne veličine X. Koeficijent spljoštenosti, ako postoji, je

$$f_2 = \frac{E(X-m)^4}{\sigma^4} - 3.$$

Za normalnu raspodelu je koeficijent spljoštenosti jednak 0. Ako je spljoštenost za neku raspodelu pozitivna, onda je njen grafik kao na slici levo – manja spljoštenost, a ako je spljoštenost negativna, onda je grafik kao na slici desno – veća spljoštenost:



Za svaku slučajnu veličinu koeficijent spljoštenosti je veći ili jednak –2.

Ukoliko imamo podatke iz uzorka obima n za neku populaciju, onda se spljoštenost raspodele obeležja može oceniti uzoračkim koeficijentom spljoštenosti:

$$\frac{\frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n)^4}{\left(\sqrt{\frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2}\right)^4} - 3.$$

Koeficijent varijacije

Коефицијент варијације

EN	Coefficient of variation
FR	Coefficient de variation
RU	Коэффициент вариации
ZH	变异系数, biànyì xìshù
AR	معامل التغير
EL	Συντελεστής διακύμανσης
DE	Variationskoeffizient
ES	Coeficiente de variación
IT	Coefficiente di variazione
SK	Koeficient rozptylu
SL	Koeficient variacije
HU	Relatív szórás

Povezani termini:

Disperzija (54), matematičko očekivanje (152), raspodela

Etimologija:

Koeficijent < co-[LA] (sa) + efficiēns[LA] (onaj koji izvršava) < ē[LA] (iz) + faciō[LA] (raditi); ovaj termin je prvi upotrebio francuski matematičar Fransoa Vijet (1540–1603), otac moderne algebre.

Varijacija < variātiō[LA] (različnost, promena) < varius[LA] (različit, drugačiji)

Koeficijent varijacije je količnik matematičkog očekivanja i standardne devijacije posmatrane slučajne veličine X:

$$\nu = \frac{E(X)}{\sqrt{D(X)}}.$$

Predstavlja neimenovan broj, pa se može koristiti pri poređenju različitih pojava[59]. Ukoliko je očekivana vrednost jednaka 0, koeficijent varijacije ne donosi više informacija o slučajnoj veličini nego samo matematičko očekivanje.

Koeficijent varijacije obeležja može biti ocenjen po analogiji sa datom formulom na sledeći način:

$$\hat{\nu} = \frac{\overline{X}_n}{\sqrt{\overline{S}_n^2}}.$$

Termin je uveo Karl Pirson 1896. godine.

[59] *Primer.* Koeficijent varijacije eksponencijalne raspodele je 1, koeficijent varijacije χ^2-raspodele sa n stepeni slobode je $(n/2)^{1/2}$, koeficijent varijacije gama raspodele sa parametrima a i b je $a^{1/2}$, a koeficijent varijacije Erlangove raspodele reda n je $n^{1/2}$.

Kontrola kvaliteta

Контрола квалитета

EN	Quality control
FR	Contrôle de qualité
RU	Контроль качества
ZH	质量控制, zhìliàng kòngzhì
AR	مراقبة الجودة
EL	Έλεγχος ποιότητας
DE	Qualitätsregelung
ES	Control de la calidad
IT	Controllo di qualità
SK	Kontrola kvality
SL	Kontrola kakovosti
HU	Minőségellenőrzés

Povezani termini:

Normalna raspodela (172), očekivanje (152), standardana devijacija (202)

Etimologija:

Kontrola < contrarotulus[LA] (kontrolna traka u kasi koja služi za proveravanje računa) < contra[LA] (protiv) + rotulla[LA] (točkić)

Kvalitet < quālitas[LA] (svojstvo) < quālis[LA] (kakav). Latinski retor i filozof Marko Tulije Ciceron (106–43) kalkirao je ovu reč po uzoru na grčku reč ποιότης (svojstvo < ποῖος[EL], kakav) koju je prvi skovao Platon (oko 428–348).

Statistička kontrola kvaliteta obuhvata različite postupke kontrolisanja gotovih proizvoda, kao i procesa proizvodnje. Sa teorijske tačke gledišta neki vidovi kontrole kvaliteta su krajnje jednostavni, kao npr. kontrolna karta. Ona predstavlja interval poverenja određenog, unapred fiksiranog, nivoa poverenja za posmatranu karakteristiku proizvoda (težina, dužina, i sl.) gde se jednostavno, beleženjem vrednosti posmatranog obeležja vidi da li se „proizvodnja nalazi u granicama kontrole". Svaki izlazak van tih granica može biti signal za ispitivanje i utvrđivanje razloga koji su doveli do odstupanja, pa se zatim primene odgovarajuće korekcije proizvodnog procesa.

Postoje razne vrste kontrolnih karti. Najjednostavnije su tzv. $\bar{x}$ kontrolne kartice, (*x-bar*) koje je osmislio Šuhart. To je grafik sa ucrtanom željenom (i po standardu proizvodnje predviđenom) vrednošću posmatrane karakteristike, i linijama koje su paralelne toj liniji i nalaze se na rastojanju $\pm 3\sigma$ od nje. Standardom proizvodnje je predviđena vrednost standardne devijacije σ. U jednakim vremenskim razmacima se na slučajana način izdvajaju uzorci određenog, uvek istog, obima i za svaki od njih računa uzoračka srednja vrednost. Sve dok su srednje vrednosti u granicama $\pm 3\sigma$ smatra se da je proizvodnja u granicama kontrole. Obim uzorka i vremena uzimanja uzoraka su veličine koje se određuju na osnovu osobina proizvodnog procesa i cene koštanja uzimanja uzorka.

Valter Endrju Šuhart (1891–1967), američki inženjer, fizičar i statističar. Često se naziva „ocem statističke kontrole kvaliteta". Zaslužan i za stvaranje Šuhartovog, ili PDCA ciklusa (*plan–do–check–act*), metode menadžmenta proizvodnje u četiri koraka.

Konvergencija u verovatnoći

Конвергенција у вероватноћи

EN	Convergence in probability
FR	Convergence de probabilité
RU	Сходимость по вероятности
ZH	依概率收敛, yī gàilǜ shōuliǎn
AR	التقارب الاحتمالي
EL	Σύγκλιση σε πιθανότητες
DE	Stochastische Konvergenz
ES	Convergencia en probabilidad
IT	Convergenza in probabilità
SK	Konvergencia podľa pravdepodobnosti
SL	Konvergenca v verjetnosti
HU	Sztochasztikus konvergencia

Povezani termini:

Granične raspodele, nizovi slučajnih veličina

Etimologija:

Konvergencija < cum^{LA} (sa) + vergere^{LA} (naginjati se)

Pri proučavanju nizova slučajnih veličina često se postavlja pitanje konvergencije. Za razliku od nizova realnih brojeva, nizovi slučajnih veličina mogu na razne načine konvergirati ka nekim brojevima (specijalni slučaj slučajne veličine) ili ka slučajnim veličinama. Kako svaki broj b možemo smatrati za slučajnu veličinu Y za koju je $P(Y=b)=1$, zapravo uvek imamo da niz slučajnih veličina konvergira ka nekoj slučajnoj veličini.

Jedna vrsta konvergencije nizova slučajnih veličina se naziva **konvergencija u verovatnoći**.

Niz slučajnih veličina $X_1, X_2,..., X_n,...$, zadatih na istom prostoru verovatnoća (Ω, F, P) konvergira u verovatnoći ka slučajnoj veličini X, ako za svako važi

$$P\ \{|Xn - X| \geq \varepsilon\} \to 0, n \to \infty.$$

U teoriji verovatnoće definišu se i druge vrste konvergencije: konvergencija u srednjem reda k[60], konvergencija u raspodeli[61], skoro sigurna konvergencija[62] i proučavaju međusobni odnosi tih konvergencija.

---

[60] Ako je za niz slučajnih veličina $X_1, X_2,..., X_n,...$ i sl. veličinu X granična vrednost niza $E(X_j–X)^k$ jednaka 0, tada je u pitanju konvergencija u srednjem reda k.

[61] Ako za niz $X_1, X_2,..., X_n,...$ i sl. veličinu X važi da funkcije raspodele članova niza konvergiraju ka funkciji raspodele veličine X, onda je u pitanju konvergencija u raspodeli (kaže se i „u zakonu raspodele“)

[62] Ako sa verovatnoćom 1 članovi niza $X_1, X_2,..., X_n,...$ konvergiraju ka slučajnoj veličini X, onda je u pitanju skoro sigurna konvergencija.

Kovarijansa

Коваријанса

EN	Covariance
FR	Covariance
RU	Ковариантность
ZH	协方差, xiéfāngchā
AR	التغاير
EL	Συνδιακύμανση
DE	Kovarianz
ES	Covariancia
IT	Covarianza
SK	Kovariancia
SL	Kovarianca
HU	Kovariancia

Povezani termini:

Korelacija (131), matematičko očekivanje (152), nezavisnost

Etimologija:

Kovarijansa < cum^{LA} (sa) + variantia^{LA} (različnost) < varius^{LA} (različit, drugačiji)

Matematičko očekivanje slučajne veličine

$$(X - E(X))(Y - E(Y))$$

je **kovarijansa** slučajnih veličina X i Y, i označava se sa $C_{X,Y}$:

$$C_{X,Y} = E[(X - E(X))(Y - E(Y))].$$

Ako je kovarijansa jednaka nuli, slučajne veličine X i Y su nekorelisane. Ako su slučajne veličine X i Y nezavisne, onda je kovarijansa jednaka nuli. Obrnuto tvrđenje ne važi. Naime, ako je $C_{X,Y} = 0$, slučajne veličine X i Y mogu biti kako zavisne, tako i nezavisne[63].

Kovarijansa slučajnih veličina X i Y može da se računa i po formuli

$$C_{X,Y} = E(XY) - E(X)E(Y).$$

Kovarijansa se koristi pri izračunavanju koeficijenta korelacije slučajnih veličina X i Y.

Ako imamo uzorak (X_j, Y_j), $j = \overline{1, n}$, obima n za obeležja X i Y, onda će uzoračka kovarijansa biti

$$\frac{1}{n}\sum_{j=1}^{n} X_j Y_j - \frac{1}{n}\sum_{j=1}^{n} X_j \cdot \frac{1}{n}\sum_{j=1}^{n} Y_j$$

Termin kovarijansa uveo je ser Ronald Fišer 1930. godine.

[63] Primer za zavisne slučajne veličine čija je kovarijansa jednaka 0 su X i Y, gde je $P(X = 1) = P(X = -1) = 1/2$, $Y = X^2$.

Kriterijumi saglasnosti

Критеријуми сагласности

EN	Test of goodness of fit
FR	Validité de l’ajustement
RU	Критерий согласия
ZH	适合度检测, shìhédù jiǎncè
AR	اختبار حسن المطابقة
EL	Έλεγχος καλής προσαρμογής
DE	Güte der Anpassung
ES	Prueba de bondad de ajuste
IT	Bontà dell’adattamento
SK	Test dobrej zhody
SL	Preizkusi skladnosti
HU	Illeszkedés

Povezani termini:

Empirijska funkcija raspodele (72), histogram (108), nulta hipoteza (174), raspodela obeležja, statistički test

Etimologija:

Kriterijum < κριτήριον^EL (merilo, sud) < κρίνω^EL (suditi)

Ukoliko nulta hipoteza znači da je raspodela obeležja određena raspodela F, onda se postupak testiranja naziva i **kriterijum saglasnosti**.

Među najpoznatijim su Pirsonov χ^2-test, kao i testovi Kolmogorova, Kolmogorova-Smirnova, Kuipera...

Pirsonov test poredi histogram sa pretpostavljenom gustinom ili zakonom raspodele i primenljiv je kako na diskretna tako i na neprekidna obeležja, pri čemu u pretpostavljenoj raspodeli obeležja mogu da figurišu i nepoznati parametri. Kod Pirsonovog testa raspodela test-statistike ne zavisi od raspodele obeležja.

Test Kolmogorova poredi pretpostavljanu raspodelu sa empirijskom funkcijom raspodele i primenljiv je, u osnovnoj varijanti na neprekidna obeležja sa potpuno određenom raspodelom (bez nepoznatih parametara). Kod testa Kolmogorova raspodela test-statistike ne zavisi od raspodele obeležja.

Andrej Nikolajevič Kolmogorov (1903–1987), izuzetno poznat i cenjen ruski (sovjetski) matematičar i statističar. Raspodela koja nosi njegovo ime je raspodela test-statistike u neparametarskom testu kojim se proverava da li su podaci iz uzorka dobijeni iz populacije sa datom neprekidnom raspodelom. Kolmogorov se bavio i primenama, a poznat je i po svom radu na unapređenju nastave matematike u sredjim školama. Aksiomatski je zasnovao teoriju verovatnoće 1933. godine.

Za testiranje saglasnosti podataka sa normalnom raspodelom postoje i specifični testovi zasnovani na poređenju uzoračkih koeficijenata asimetrije i spljoštenosti sa odgovarajućim koeficijentima normalne raspodele. Jedan takav test je test Harke-Bera koji se koristi u analizi vremenskih serija. Kod nekih testova saglasnosti (Anderson-Darling test, Šapiro-Vilksov test) raspodela test-statistike zavisi od pretpostavljene raspodele obeležja.

Kritična oblast
Критична област

EN	Critical region
FR	Région critique
RU	Критическая область
ZH	判别区域, pànbié qūyù
AR	المنطقة الحرجة
EL	Κρίσιμη περιοχή
DE	Ablehnungsbereich
ES	Región crítica
IT	Regione critica
SK	Kritická oblasť
SL	Kritično območje
HU	Kritikus tartomány

Povezani termini:

Realizovani uzorak (189), statistički test, uzorak (240)

Etimologija:

Kritičan < κριτικός^{EL} (razborit) < κρίνω^{EL} (suditi)

Statistički test je u potpunosti određen **kritičnom oblašću** W_n kao skupom tačaka $(x_1, x_2,..., x_n)$ u n-dimenzionalnom euklidskom prostoru. Ako za realizovani uzorak važi $(x_1, x_2,..., x_n) \in W_n$, tada se odbacuje H_0, a ako $(x_1, x_2,..., x_n) \notin W_n$, onda se prihvata H_0.

Testiranje hipoteza o jednodimenzionalnom parametru θ se ostvaruje korišćenjem pogodno odabrane statistike Θ_n i zato se umesto uslova $(x_1, x_2,..., x_n) \in W_n$ dobija uslov $\Theta_n \in W$, gde je W oblast u jednodimenzionalnom prostoru.

U višedimenzionalnom slučaju situacija je analogna.

Oblik kritične oblasti određuje alternativna hipoteza:

- u slučaju alternativne hipoteze $H_1(\theta \neq \theta_0)$, kritična oblast je unija intervala $(-\infty, f)$ i $(g, +\infty)$, gde se statistike f i g određuju u funkciji statistike kojom se ocenjuje parametar θ,
- u slučaju alternativne hipoteze $H_1(\theta < \theta_0)$, kritična oblast je interval $(-\infty, f)$, gde se statistika f određuje u funkciji statistike kojom se ocenjuje parametar θ,
- u slučaju alternativne hipoteze, $H_1(\theta > \theta_0)$, kritična oblast je interval $(g, +\infty)$, gde se statistika g određuje u funkciji statistike kojom se ocenjuje parametar θ.

U prvom slučaju kritična oblast je dvostrana (*two-sided critical region*), a u druga dva slučaja – jednostrana (*one-sided critical region*).

U zavisnosti od dopustivog skupa vrednosti parametra moguće su neke izmene oblika kritične oblasti, tako što se za kritičnu oblast uzima presek skupa dopustivih vrednosti parametra i intervala za kritičnu oblast kao što je navedeno. Sve ovo se odnosi na parametre čiji je dopustivi skup vrednosti neki interval ili cela realna prava.

Kvantil

Квантил

EN	Quantile
FR	Quantile
RU	Квантиль
ZH	分位数, fēnwèishù
AR	العشير
EL	Ποσοστημόριο
DE	Quantil
ES	Cuantilo
IT	Quantile
SK	Kvantil
SL	Kvantil
HU	Kvantilis

Povezani termini:

Funkcija raspodele (88), kvartil (140)

Etimologija:

Kvantil < quantitas^{LA} (količina) < quantus^{LA} (koliki).

Kvantili su vrednosti obeležja kojima odgovara određena vrednost funkcije raspodele obeležja. Dakle, realni broj x je kvantil reda p ako je vrednost funkcije raspodele obeležja u tački x jednaka p: $F(x) = p$. Kvantili su zapravo vezani za inverznu funkciju raspodele $x = F^{-1}(p)$.

Jasno je da kvantili nekog reda ne moraju biti jedinstveni, npr. kod diskretnih raspodela. Uzorački kvantili se po istom principu određuju na osnovu empirijske funkcije raspodele. Uzorački kvantili stoga ne moraju biti jedinstveni, jer je empirijska funkcija raspodele stepenasta funkcija, kao što su i funkcije raspodele diskretnih slučajnih veličina.

Neki kvantili imaju posebna imena:

- medijana ($p = 0.5$),
- prvi ($p = 0.25$) i treći ($p = 0.75$) kvartil,
- decil ($p = 0.1$).

Često u zaglavlju statističkih tablica piše: „Kvantili ... raspodele“, što, dakle, samo znači da su dati brojevi vrednosti obeležja posmatrane raspodele kojima odgovaraju date verovatnoće.

Kvartil
Квартил

EN	Quartile
FR	Quartile
RU	Квартиль
ZH	四分位数, sìfēnwèishù
AR	الربيع
EL	Τεταρτημόριο
DE	Quartil
ES	Cuartilo
IT	Quartile
SK	Kvartil
SL	Kvartil
HU	Kvartilis

Povezani termini:

Funkcija raspodele (88), kvantil (138)

Etimologija:

Kvartil < quartus$^{\mathrm{LA}}$ (četvrti) < quattuor$^{\mathrm{LA}}$ (četiri); latinska reč *quattuor* i naše *četiri* imaju zajednički koren u proto-indoevropskom *k^{w}etwóres.

Kvartili se označavaju (često, mada ne i obavezno) sa q_1, q_2, q_3 i nazivaju prvi, drugi i treći kvartil.

Redom su jednaki vrednostima obeležja za koje je funkcija raspodele jednaka 0.25, 0.5 i 0.75. Dakle, kvartili su specijalni slučajevi kvantila raspodele. Drugi kvartil je zapravo medijana obeležja.

Jasno je da kvartili nekog reda ne moraju biti jedinstveni, npr. kod diskretnih raspodela. Uzorački kvartili se po istom principu određuju na osnovu empirijske funkcije raspodele. Uzorački kvartili stoga ne moraju biti jedinstveni, jer je empirijska funkcija raspodele stepenasta funkcija, kao što su i funkcije raspodele diskretnih slučajnih veličina.

Kvartili i interkvartilni razmak[64] se koriste u grafičkom prikazivanju podataka pomoću boks plot dijagrama.

[64] Interkvartilni razmak (*interquartile range*) je razlika $q_3 – q_1$.

Linearna korelacija

Линеарна корелација

EN	Linear correlation
FR	Corrélation linéaire
RU	Линейная корреляция
ZH	线性相关, xiànxìng xiāngguān
AR	الارتباط الخطي
EL	Γραμμική συσχέτιση
DE	Lineare Korrelation
ES	Correlación lineal
IT	Correlazione lineare
SK	Lineárna korelácia
SL	Linearna korelacija
HU	Lineáris korreláció

Povezani termini:

Koeficijent korelacije (122), uzorak (240), zavisnost

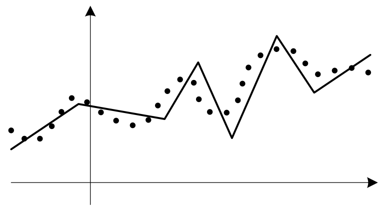
Etimologija:

Linija < līnea$^{\text{LA}}$ (konac, nit) < līnum$^{\text{LA}}$ (lan)

Korelacija < correlatio$^{\text{LA}}$ (međusobna povezanost) < co-$^{\text{LA}}$ (sa) + relatio$^{\text{LA}}$ (odnos) < referō$^{\text{LA}}$ (nositi nazad) < re$^{\text{LA}}$ (ponovo) + ferre$^{\text{LA}}$ (nositi)

Korelacija uopšte označava povezanost dva ili više obeležja, pa u tom smislu **linearna korelacija** označava vezu u obliku linearne funkcije $Y = aX + b$, ili pak $Z = aX + bY + c$ itd. Koeficijenti koji se ovde pojavljuju se ocenjuju na osnovu podataka kojima raspolažemo: (x_i, y_i) odnosno (x_i, y_i, z_i), $i = 1, n$. Kvalitet dobijenih ocena se utvrđuje prema nekim kriterijumima, npr. nepristrasnost, postojanost, efikasnost. Jedna od metoda za dobijanje ocena je metoda najmanjih kvadrata. Ukoliko je poznata raspodela veličina koje posmatramo, može se primeniti i metoda maksimalne verodostojnosti.

Ukoliko zavisnost obeležja nije linearna na celom skupu vrednosti (što se već sa grafika može uočiti), ipak se linearna veza može primeniti po principu „deo po deo“, tj. na manjim intervalima vrednosti obeležja se određuju veze u obliku linearnih funkcija, a onda se svi rezultati objedine (videti sliku) – zavisnost tada (približno) opisuje dobijena izlomljena linija.



Linearna regresija

Линеарна regresija

EN	Linear regression
FR	Régression linéaire
RU	Линейная регрессия
ZH	线性回归, xiànxìng huíguī
AR	الانحدار الخطي
EL	Γραμμική παλινδρόμηση
DE	Lineare Regression
ES	Regresión lineal
IT	Regressione lineare
SK	Lineárna regresia
SL	Linearna regresija
HU	Lineáris regresszió

Povezani termini:

Uzorak (240), zavisnost obeležja

Etimologija:

Linija < līnea$^{\text{LA}}$ (konac, nit) < līnum$^{\text{LA}}$ (lan)

Regresija < regressio$^{\text{LA}}$ < regredior$^{\text{LA}}$ (vratiti se) < re-$^{\text{LA}}$ (ponovo) + gradior$^{\text{LA}}$ (hodati)

Pri proučavanju dva obeležja od interesa je ustanoviti da li i kakva zavisnost među njima postoji. Ukoliko nemamo prethodnih saznanja o proučavanim pojavama podatke ćemo prvo predstaviti grafički. Ako se uočava postojanje neke zakonitosti, npr. povećanje vrednosti obeležja X prati povećanje vrednosti obeležja Y, tada biramo neku funkciju f koja ima sličnosti sa grafikom sa podacima. Takva funkcija će imati neke nepoznate parametre ($Y = aX + b$, $Y = aX^2 + bX + C$,...) i te parametre treba izračunati (tj. „oceniti“) na osnovu podataka (x_i, y_i), $i = 1, n$.

Ukoliko je zavisnost linearna, govori se o **linearnoj regresiji** Y u odnosu na X, a nepoznati koeficijenti se određuju iz uslova da suma kvadrata odstupanja bude minimalna[65].

$$\sum_{j=1}^{n}\left(y_j - (ax_j + b)\right)^2$$

Odatle se dobija da a i b predstavljaju rešenja sistema jednačina[66]

$$\sum_{j=1}^{n}\left(y_j - (ax_j + b)\right)x_j = 0, \quad \sum_{j=1}^{n}\left(y_j - (ax_j + b)\right) = 0.$$

i nalazimo da je ocena za a jednaka Δ_1/Δ, gde je

$$\Delta_1 = \sum_{j=1}^{n} x_j \sum_{j=1}^{n} y_j - n\sum_{j=1}^{n} x_j y_j, \quad \Delta = \left(\sum_{j=1}^{n} x_j\right)^2 - n\sum_{j=1}^{n} x_j^2$$

a da je ocena za b jednaka $\hat{b} = \overline{y}_n - \hat{a}\overline{x}_n$.

[65] Taj zahtev je u osnovi metode najmanjih kvadrata (*least square method*), kad, u opštem slučaju, određujemo nepoznate koeficijente a_1,..., a_k u izrazu funkcije f iz uslova da suma kvadrata odstupanja y_j od $f(x_j, a_1,...,a_k)$ bude minimalna.

[66] Što se naziva i sistem normalnih jednačina.

Linearna transformacija podataka

Линеарна трансформација података

EN	Linear data transformation
FR	Transformation linéaire des données
RU	Линейное преобразование данных
ZH	线性数据转换, xiànxìng shùjù zhuǎnhuàn
AR	التحويل الخطي من البيانات
EL	Γραμμικός μετασχηματισμός δεδομένων
DE	Lineare Datentransformation
ES	Transformación lineal de datos
IT	Trasformazione lineare dei dati
SK	Lineárna transformácia dát
SL	Linearna transformacija podatkov
HU	Lineáris adatalakítás

Povezani termini:

Uzoračka sredina (230), uzorak (240)

Etimologija:

Linija < līnea^{LA} (konac, nit) < līnum^{LA} (lan)

Transformacija < trans^{LA} (preko) + fōrma^{LA} (oblik)

Translacija podataka:

Umesto vrednosti $x_1, x_2, \ldots, x_m$ uzimaju se $x_i^* = x_i - a$, $i = 1, 2, \ldots, m$, gde je a neki realni broj. Realni broj a, koji se koristi u ovoj transformaciji se naziva radna nula. Translacija se ostvaruje tako da radna nula bude, na primer, središnji podatak u nizu vrednosti obeležja jer tada novodobijeni podaci budu delom pozitivni, delom negativni, pa se grupisanjem sabiraka olakšava i ubrzava izračunavanje.

Na primer, uzoračka sredina za novodobijene podatke će biti u vezi sa uzoračkom sredinom polaznih podataka:

$$\bar{x}_n^* = \frac{1}{n}\left(n_1 x_1^* + n_2 x_2^* + \ldots + n_m x_m^*\right) =$$

$$= \frac{1}{n}\left(n_1 x_1 + n_2 x_2 + \ldots + n_m x_m\right) - a = \bar{x}_n - a.$$

Translacija je specijalni slučaj **linearne transformacije podataka**:

Umesto $x_1, x_2, \ldots, x_m$ uzima se $x_i^* = \mathrm{b}x_i - a$, $i = 1, 2, \ldots, m$, gde su b i a neki realni brojevi. Na primer, za uzoračku sredinu za novodobijene podatke važi veza sa uzoračkom sredinom polaznih podataka:

$$\bar{x}_n^* = \frac{1}{n}\left(n_1 x_1^* + n_2 x_2^* + \ldots + n_m x_m^*\right) =$$

$$= \frac{b}{n}\left(n_1 x_1 + n_2 x_2 + \ldots + n_m x_m\right) - a = b\bar{x}_n - a.$$

Ova transformacija podataka se najčešće primenjuje kad su vrednosti u uzorku ili vrlo veliki brojevi ili brojevi vrlo bliski nuli. Ako je $b = 1$, dobija se translacija podataka.

→ *Nastavak na str. 260*

Marginalna raspodela

Маргинална расподела

EN Marginal distribution
FR Distribution marginale
RU Маргинальное распределение
ZH 边际分布, biānjì fēnbù
AR التوزيع الهامشي
EL Περιθώρια κατανομή
DE Randverteilung
ES Distribución marginal
IT Distribuzione marginale
SK Marginálne rozdelenie
SL Marginalna porazdelitev
HU Marginális eloszlás

Povezani termini:

Funkcija raspodele (88), gustina raspodele (84), nezavisnost

Etimologija:

Marginalan < margō^LA (rub, ivica, granica)

Ako se posmatraju višedimenzionalne slučajne veličine, onda se raspodele pojedinačnih komponenata nazivaju **marginalne raspodele**. Na primer, u slučaju dvodimenzionalne slučajne veličine (X, Y) funkcije raspodele slučajnih veličina X i Y se mogu dobiti takođe iz zajedničke funkcije raspodele.

Funkcija

$$F_1(x) = \lim_{y \to +\infty} F(x, y) = F(x, +\infty)$$

je marginalna funkcija raspodele veličine X, a funkcija

$$F_2(y) = \lim_{x \to +\infty} F(x, y) = F(+\infty, y)$$

je marginalna funkcija raspodele veličine Y.

U slučaju neprekidnih slučajnih veličina X i Y se od marginalnih funkcija raspodele dobijaju marginalne gustine raspodele:

$$F_1'(x) = f_1(x) \text{ i } F_2'(y) = f_2(y).$$

Marginalne gustine raspodela se mogu dobiti i iz zajedničke gustine raspodele dvodimenzionalne slučajne veličine:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y)dy, \quad f_2(y) = \int_{-\infty}^{+\infty} f(x, y)dx.$$

Na osnovu rečenog jasno je da se iz zajedničke raspodele dobijaju marginalne raspodele. Obrnuto, da se iz marginalnih raspodela dobije zajednička, je moguće samo u slučaju nezavisnih veličina X i Y, kada je zajednička raspodela jednaka proizvodu marginalnih raspodela.

Matematička statistika

Математичка статистика

EN Mathematical statistics
FR Statistique mathématique
RU Математическая статистика
ZH 数理统计学, shùlǐ tǒngjìxué
AR الأحصاء الرياضي
EL Μαθηματική στατιστική
DE Matematische Statistik
ES Estadística matemática
IT Statistica matematica
SK Matematická štatistika
SL Matematična statistika
HU Matematikai statisztika

Povezani termini:

Teorija verovatnoće (212) i sve iz ovog rečnika (☺)

Etimologija:

Matematika < μαθηματικός^EL (onaj ko voli da uči) < μάθημα^EL (znanje, učenje, nauka)

Statistika < statisticum^LA (ono što ima veze s državom). Termin statistika, sa prvobitnim značenjem analiziranja podataka o državi, uveo je Gotfrid Ahenval (1719–1772), nemački filozof, pravnik i ekonomista.

Statistika se bavi:

1. Prikupljanjem podataka,
2. Prikazivanjem podataka,
3. Analizom podataka,
4. Zaključivanjem na osnovu podataka.

Statistika je danas prisutna i u svakodnevnom životu, a izvesnim aspektima „običnog“ života bavila se i u svojim počecima. Naime, statistika je prvobitno proučavala tzv. masovne pojave u ljudskom društvu kroz prikupljanje, upoređivanje i tumačenje podataka o stanovništvu, imovini, vojnoj snazi, itd.

Još u IV veku pre nove ere, Aristotel (384–322) je uradio „popis“ 158 gradova-država na teritoriji gde su živeli Grci, što bi mogao da bude, na prostorima današnje Evrope, začetak nauke o društvu – demografije.

Kasnije, u XVII veku, u Nemačkoj je takođe bio urađen komparativni opis država, a u Engleskoj, isto u XVII veku napravljen je, u vreme epidemije kuge, popis umrlih i vrsta bolesti od kojih su preminuli, po parohijama. To je, sa današnje tačke gledišta, primena statistike u medicini.

U istoriji statistike važno mesto ima i Edmond Halej (1656–1742) zbog svojih ideja u vezi sa uplatom životnog osiguranja. Svakako, za razvoj statistike su važni i Karl Fridrih Gaus i Pjer-Simon Laplas, naučnici i matematičari koji su poznati i širem krugu, a ne samo onima koji se bave matematikom. Manje je, međutim, poznat doprinos Adolfa Ketlea (1796–1874) iz Belgije, koji je dao sistematizaciju i kategorizaciju kriminalnog ponašanja i uzroka koji do njega dovode.

→ *Nastavak na str. 260*

Matematičko očekivanje

Математичко очекивање

EN Mathematical expectation
FR Espérance mathématique
RU Математическое ожидание
ZH 数学期望, shùxué qīwàng
AR التوقع الرياضي
EL Προσδοκία
DE Erwartungswert
ES Esperanza matemática
IT Speranza matematica
SK Stredná hodnota
SL Matematično upanje
HU Várható érték

Povezani termini:

Diskretne slučajne veličine (199), neprekidne slučajne veličine (199)

Etimologija:

Matematika < μαθηματικός[EL] (onaj ko voli da uči) < μάθημα[EL] (znanje, učenje, nauka)

Matematičko očekivanje predstavlja jednu od najvažnijih numeričkih karakteristika slučajne veličine.

Ako je X diskretna slučajna veličina sa zakonom raspodele: $P\{X = x_k\} = p_k$, $k = 1,..., n$, tada je matematičko očekivanje slučajne veličine X jednako

$$E(X) = x_1p_1 + x_2p_2 + ... + x_np_n.$$

Ako je X elementarna slučajna veličina sa zakonom raspodele $P\{X = x_j\} = pj, j \in J \subseteq N$, gde je J beskonačan podskup skupa prirodnih brojeva. Kaže da slučajna veličina X ima matematičko očekivanje i da je njeno matematičko očekivanje jednako

$$E(X) = \sum_j x_j p_j \ ,$$

ako je red na desnoj strani apsolutno konvergentan.

Ako je X apsolutno neprekidna slučajna veličina sa gustinom raspodele $g(x)$, onda je njeno matematičko očekivanje

$$E(X) = \int_{-\infty}^{+\infty} x \cdot g(x)dx \,,$$

ako je integral na desnoj strani apsolutno konvergentan.

Postoji beskonačno mnogo slučajnih veličina koje imaju isto matematičko očekivanje. Stoga se zaključci o nekoj slučajnoj veličini ne mogu izvoditi samo na osnovu poznavanja njenog matematičkog očekivanja.

Ne mora svaka slučajna veličina da ima matematičko očekivanje. Ako slučajna veličina ima matematičko očekivanje, onda je to neki realan broj.

→ *Nastavak na str. 261*

Medijana

Медијана

EN	Median
FR	Médiane
RU	Медиана
ZH	中线, zhōngxiàn
AR	الوسيط
EL	Διάμεσος
DE	Median
ES	Mediana
IT	Mediana
SK	Medián
SL	Mediana
HU	Medián

Povezani termini:

Funkcija raspodele (88)

Etimologija:

Medijana < mediānus^{LA} (onaj koji pripada sredini) < medius^{LA} (srednji). Latinska reč *medius* i srpska reč *međa* imaju zajednički koren u proto-indoevropskom *med^{h}yos (srednji).

Neka je $F(x)$ funkcija raspodele slučajne veličine X. **Medijana** je vrednost M_e argumenta funkcije raspodele za koju je

$$F(M_e) = 0.5.$$

Na ovaj način definisana medijana ne mora da postoji[67] a ukoliko postoji ne mora biti jedinstvena[68]. U statistici se medijana definiše na precizniji način, tako da je jedinstvena za dati uzorak.

Ako je raspodela simetrična i ima matematičko očekivanje onda se medijana i matematičko očekivanje poklapaju. Kod asimetričnih raspodela medijana može biti manja ili veća od matematičkog očekivanja[69].

---

[67] *Primer.* Neka je X diskretna slučajna veličina sa vrednostima 1, 2 i 3 i odgovarajućim verovatnoćama 0.2, 0.5 i 0.3. Funkcija raspodele posmatrane slučajne veličine nije jednaka 0.5 ni u jednoj tački.

[68] *Primer.* Neka je X diskretna slučajna veličina sa vrednostima 1, 2 i 3, i odgovarajućim verovatnoćama 0.4, 0.1 i 0.5. Tada je za sve vrednosti između 2 i 3 funkcija raspodele jednaka 0.5.

[69] *Primer.* Medijana hi-kvadrat raspodele sa n stepeni slobode je približno $n - 2/3 + 4/27n$, pa je manja od očekivanja. Kod eksponencijalne raspodele sa parametrom λ medijana je $\lambda ln2$.

Metoda maksimalne verodostojnosti

Метода максималне веродостојности

EN	Maximum likelihood method
FR	Méthode du maximum de vraisemblance
RU	Метод максимального правдоподобия
ZH	极大似然法, jídà sìrán fǎ
AR	طريقة الارجحية القصوى
EL	Μέθοδος μέγιστης πιθανοφάνειας
DE	Maximum-Likelihood-Methode
ES	Método del máximo de verosimilitud
IT	Metodo della massima verosimiglianza
SK	Metóda maximálnej vierohodnosti
SL	Metoda največjega verjetja
HU	Maximum-likelihood módszer

Povezani termini:

Obeležje (176), ocenjivanje parametara (178), raspodela obeležja

Etimologija:

Metoda < μέθοδος$^{\mathrm{EL}}$ (način) < μετά$^{\mathrm{EL}}$ (posle) + ὁδός$^{\mathrm{EL}}$ (put)

Maksimalan < maximus$^{\mathrm{LA}}$ (najveći) < magnus$^{\mathrm{LA}}$ (veliki)

Neka je θ nepoznati parametar u raspodeli obeležja X. Neka je $(X_1, X_2,..., X_n)$ prost slučajni uzorak, a $f(x, \theta)$ gustina raspodele za obeležje X.

Za ocenu parametra θ se uzima statistika za koju se dostiže supremum funkcije verodostojnosti

$$L(X; \theta) = f(X_1; \theta) \cdot f(X_2; \theta) \cdot ... \cdot f(X_\mathrm{n}; \theta).$$

U nekim slučajevima to će ujedno biti maksimalna vrednost funkcije verodostojnosti. Navedena metoda dobijanja ocena se naziva **metoda maksimalne verodostojnosti**, a ocene dobijene tom metodom su ocene maksimalne verodostojnosti. Statistika kojom se po metodi maksimalne verodostojnosti ocenjuje neki parametar ne mora da bude jedinstvena.

Ako je funkcija verodostojnosti diferencijabilna, a parametar θ k-dimenzionalni, najpre se određuju vrednosti parametara koje su rešenja sistema jednačina

$$\frac{dL(X;\theta)}{d\theta_j} = 0, j = 1,...,k$$

ili sistema jednačina $\frac{d\ln(L(X;\theta))}{d\theta_j} = 0, j = 1,...,k$.

Među dobijenim rešenjima jednačine mogu se nalaziti vrednosti za koje funkcija L (ili lnL) postiže svoj supremum, odnosno maksimum ako je totalni diferencijal drugog reda L (ili lnL) negativan u tim tačkama.

Ako je posmatrano obeležje diskretnog tipa, tada je funkcija verodostojnosti jednaka proizvodu verovatnoća postizanja pojedinih vrednosti iz uzorka

$$L(X; \theta) = p(X_1; \theta) \cdot p(X_2; \theta) \cdot ... \cdot p(X_\mathrm{n}; \theta),$$

gde je $p(t; \theta) = P\{X = t\}$. Zatim se određuje maksimalna vrednost funkcije L (ili lnL) kao u prethodnom slučaju.

→ *Nastavak na str. 261*

Metoda momenata

Метода момената

EN	Method of moments
FR	Méthode des moments
RU	Метод моментов
ZH	矩法, jǔfǎ
AR	طريقة العزوم
EL	Μέθοδος των ροπών
DE	Momentenmethode
ES	Método de los momentos
IT	Metodo dei momenti
SK	Momentová metóda
SL	Metoda momentov
HU	Momentum módszer

Povezani termini:

Obeležje (176), ocenjivanje parametara (178), raspodela obeležja

Etimologija:

Metoda < μέθοδος^{EL} (način) < μετά^{EL} (posle) + ὁδός^{EL} (put)

Momenat < mōmentum^{LA} (pokret) < movēre^{LA} (kretati se)

Metoda momenata je jedna od opštih metoda za dobijanje ocena parametara u raspodeli obeležja, i primenljiva je i za diskretna, i za neprekidna obeležja. Ako u raspodeli obeležja X ima r, $r \geq 1$, nepoznatih parametara i ako je $(X_1,..., X_n)$ uzorak za posmatrano obeležje, onda se statistike kojima se ti parametri ocenjuju mogu dobiti rešavanjem sistema od r jednačina

$$E(X^j) = \frac{1}{n}\sum_{k=1}^{n} X_k^j,\ j = 1,...,r\ .$$

Ovi sistemi jednačina su uglavnom nelinearni, tako da se ne može dati opšti postupak za njihovo rešavanje[70]. U nekim slučajevima se rešenja određuju približno, metodama numeričke matematike.

Metoda momenata se ne može primeniti ako momenti raspodele obeležja ne postoje.

Statistike koje se dobijaju po metodi momenata su često manje efikasne ocene od statistika koje se dobijaju po metodi maksimalne verodostojnosti (za isti parametar). Za neke raspodele i neke parametre u tim raspodelama poklapaju se ocene dobijene metodom momenata i metodom maksimalne verodostojnosti.

[70] Umesto navedenog, može se koristiti i sistem jednačina koji se dobija izjednačavanjem r centralnim momenata u raspodeli obeležja sa odgovarajućih r centralnih uzoračkih momenata. Takođe, nije obavezno momente uzimati redom, od prvog do r-tog, bitno je samo da bude uzeto nekih r momenata. U zavisnosti od raspodele obeležja X moguće je dobiti jednostavniji sistem jednačina ako se kombinuju opšti i centralni momenti.

Modalna vrednost

Модална вредност

EN	Modal value
FR	Valeur modale
RU	Модальное значение
ZH	众数, zhòngshù
AR	المنوال
EL	Κορυφή
DE	Modalwert
ES	Valor modal
IT	Valore modale
SK	Modus
SL	Modus
HU	Modális érték

Povezani termini:

Asimetrija (18), očekivanje (152), raspodela obeležja

Etimologija:

Modalni < modus^{LA} (mera, ograničenje)

Neka je data diskretna slučajna veličina X svojim zakonom raspodele

$$p_j = P\{X = x_j\}, j \in J \subseteq N.$$

Svaka vrednost slučajne veličine X čije su odgovarajuće verovatnoće veće od susednih je mod raspodele.

Neka je data neprekidna slučajna veličina X svojom gustinom raspodele $f(x)$, $x \in R$. Apscisa svake tačke lokalnog maksimuma funkcije $f(x)$ je **modalna vrednost** ili **moda** raspodele[71].

Na osnovu definicije se zaključuje da raspodela može imati imati jedan mod (ili modus) i tad se kaže da je raspodela unimodalna. Raspodela može imati dva (bimodalna) ili više (polimodalna) moda, a neke raspodele nemaju mod, na primer, ravnomerna raspodela.

Kod simetričnih raspodela koje imaju očekivanje, mod i očekivanje se poklapaju. Kod asimetričnih raspodela mod može biti manji ili veći od očekivanja[72], a kod nekih raspodela ne zavisi od parametara raspodele[73].

---

[71] Ako raspodela ima lokalni minimum u tački a, onda se tačka a naziva antimod.

[72] *Primer.* Za χ^2-raspodelu sa n stepeni slobode, ako je n veće od 2, mod je jednak n–2.

[73] Kod Paretove raspodele, čija je gustina a/x^{a*1}, $x \geq 1$, mod je jednak 1, kod eksponencijalne raspodele mod je jednak 0.

Neparametarski test

Непараметарски тест

EN Nonparametric test
FR Test non parametrique
RU Непараметрический критерий
ZH 非参数检验, fēicānshù jiǎnyàn
AR الاختبار اللابارامتري
EL Μή-παραμετρικό τεστ
DE Nichtparametrische Test
ES Prueba no paramétrica
IT Test non parametrico
SK Neparametrický test
SL Neparametrični preizkus
HU Nemparaméteres próbá

Povezani termini:

Raspodela obeležja, test-statistika, testiranje hipoteza (216)

Etimologija:

Parametar < παρά^{EL} (pored) + μέτρον^{EL} (mera)

Test < test^{stFR} (zemljani kotao, obično onaj u kome su isprobavani metali) < testum^{LA} (zemljana posuda)

Hipoteze o raspodeli obeležja (koje se ne odnose na parametre, već na samu raspodelu obeležja) se nazivaju neparametarske hipoteze, a odgovarajući testovi su testovi saglasnosti (obeležja sa pretpostavljenom raspodelom) ili **neparametarski testovi**.

Neka je dat prost slučajni uzorak (X_1, X_2,..., X_n) za obeležje X. Neka je potrebno sa pragom značajnosti α, testirati hipotezu H_0 da obeležje X ima datu funkciju raspodele $F_0(x)$, što se zapisuje u obliku $H_0(X : F_0(x))$ ili $H_0(F(x) = F_0(x))$. Alternativna hipoteza je da je raspodela $F(x)$ različita od $F_0(x)$, tj. $H_1(F(x) \neq F_0(x))$.

Postoji više neparametarskih testova: Pirsonov test, test Kolmogorova i testovi koji koriste rangove[74] (među kojima je i Vilkokson-Man-Vitni test). Svaki ima neke svoje specifičnosti, ali su svi prilično opšti, u smislu da se nultom hipotezom može pretpostaviti raspodela obeležja iz široke klase raspodela. Pirsonov test poredi histogram i pretpostavljenu raspodelu, a može biti primenjen na testiranje hipoteze H_0 o nezavisnosti dva obeležja, X i Y, jer je ta hipoteza ekvivalentna hipotezi da je zajednička raspodela za (X, Y) proizvod marginalnih raspodela. U slučaju neprekidnog obeležja, test Kolmogorova poredi empirijsku funkciju raspodele i pretpostavljenu raspodelu u kojoj nema nepoznatih parametara. Zbog tog ograničenja su razvijene modifikacije testa Kolmogorova. Test rangova polazi od varijacionog niza na osnovu koga se računa test-statistika.

[74] Ako je dat uzorak: 2, 8, 5, 7, 6, 3 rangove njegovih elemenata određujemo tako što prvo podatke poređamo u neopadajući niz: 2, 3, 5, 6, 7, 8. Redni broj podatka u ovom nizu je njegov rang. Dakle, element 2 ima rang 1, element 8 ima rang 6, element 5 ima rang 3, element 7 ima rang 5, element 6 ima rang 4 i element 3 ima rang 2. Zbir svih rangova u uzorku obima n je $n(n + 1)/2$.

Nepristrasna ocena

Непристрасна оцена

EN	Unbiased estimator
FR	Estimateur sans biais
RU	Несмещенная оценка
ZH	无偏估计量, wúpiān gūjìliàng
AR	تقدير غير متحيز
EL	Αμερόληπτη εκτιμήτρια
DE	Erwartungstreuer Schätzer
ES	Estimador sin sesgo
IT	Stimatore non distorto
SK	Nevychýlený odhad
SL	Nepristranska cenilka
HU	Torzítatlan becslés

Povezani termini:

Očekivanje (152), statistika (167), uzorak (240)

Neka je dat prost slučajni uzorak $(X_1, X_2, ..., X_n)$ obima n za obeležje X i neka je θ nepoznati parametar u raspodeli obeležja X. Statistika $Y = f(X_1, X_2, ..., X_n)$ je **nepristrasna ocena** parametra θ ako je $E(Y) = \theta$.

Razlika između očekivane vrednosti statistike Y i vrednosti parametra koji je ocenjen tom statistikom se naziva pristrasnost (*bias*).

Nepristrasnost ocene je poželjna osobina, jer očekivana vrednost statistike jednaka ocenjivanom parametru. Moguće je odrediti više nepristrasnih ocena za isti parametar. Takve su, na primer, linearne nepristrasne ocene matematičkog očekivanja obeležja. One su oblika:

$$\sum_{j=1}^{n} a_j X_j \text{, gde je } \sum_{j=1}^{n} a_j = 1 .$$

U slučaju kad su svi koeficijenti isti i jednaki $1/n$, dobija se uzoračka sredina, i ona ima najmanju disperziju u klasi linearnih nepristrasnih ocena za matematičko očekivanje obeležja.

Ako statistika $Y = f(X_1, X_2, ..., X_n)$ kojom se ocenjuje neki parametar θ postaje nepristrasna u slučaju kada obim uzorka teži beskonačnosti, tj. ako je $E(Y) \to \theta$, $n \to \infty$, tada je statistika Y asimptotski nepristrasna ocena posmatranog parametra θ.

Nezavisne slučajne promenljive

Независне случајне променљиве

EN Independent random variables
FR Variables aléatoires indépendantes
RU Независимые случайные переменные
ZH 独立随机变量, dúlì suíjī biànliàng
AR المتغيرات العشوائية المستقلة
EL Ανεξάρτητες τυχαίες μεταβλητές
DE Unabhängige Zufallsvariablen
ES Variables aleatorias independientes
IT Variabili casuali indipendenti
SK Nezávislé náhodné premenné
SL Neodvisne slučajne spremenljivke
HU Független valószínűségi változók

Povezani termini:

Funkcija raspodele (88), gustina raspodele (84), verovatnoća (248)

Neka su date jednodimenzionalne slučajne veličine X i Y. Neka su U i V proizvoljni Borelovi skupovi[75] na realnoj pravoj. Ako je

$$P(X \in U, Z \in V) = P(X \in U)P(Z \in V),$$

tada su **slučajne promenljive** X i Y **nezavisne**. Slučajne veličine koje nisu nezavisne su zavisne.

Ukoliko su slučajne veličine X i Y nezavisne, tada je zajednička funkcija raspodele slučajne veličine (X, Y) jednaka proizvodu funkcija raspodele slučajnih veličina X i Y

$$P(X \le x, Y \le y) = P(X \le x)P(Y \le y), \forall x, \forall y.$$

Ukoliko su slučajne veličine X i Y diskretne i nezavisne, onda se zajednička raspodela dobija kao proizvod marginalnih raspodela

$$P(X = x, Y = y) = P(X = x)P(Y = y), \forall x, \forall y,$$

a ukoliko su te slučajne veličine neprekidne, onda je zajednička gustina raspodele $g(x,y)$ jednaka proizvodu marginalnih gustina raspodele $g_X(x)g_Y(y)$.

Ako u zajedničkom zakonu raspodele slučajne veličine (X,Y) postoji bar jedna verovatnoća koja je jednaka 0, onda su slučajne promenljive X i Y zavisne.

Pojam nezavisnosti je važan u statistici jer su elementi prostog slučajnog uzorka nezavisne slučajne promenljive. Nezavisnost omogućava da se odrede osobine statistika[76] formiranih na osnovu prostog slučajnog uzorka.

[75] Borelovi skupovi na realnoj pravoj su intervali oblika (a, b) i sve njihove konačne i prebrojive unije i preseci.

[76] Statistika je svaka (merljiva) funkcija uzorka. Neke statistike su: uzoračka sredina, uzoračka disperzija, uzoračka medijana, statistike poretka, itd.

Nezavisni događaji

Независни догађаји

EN	Independent events
FR	Evénements indépendants
RU	Независимые события
ZH	独立事件, dúlì shìjiàn
AR	الأحداث المستقلة
EL	Ανεξάρτητα γεγονότα
DE	Unabhängige Ereignisse
ES	Eventos independientes
IT	Eventi indipendenti
SK	Nezávislé udalosti
SL	Neodvisni dogodki
HU	Független események

Povezani termini:

Uslovna verovatnoća (222), verovatnoća (248)

Nezavisnost događaja je jedan od bitnih pojmova u teoriji verovatnoće, a samim tim i u matematičkoj statistici. Ako su događaji A i B iz istog prostora verovatnoća i ako važi $P(AB) = P(A)P(B)$, tada su događaji A i B nezavisni. Na osnovu toga, zaključuje se da za nezavisne događaje A i B važi:

$$P(A \mid B) = P(A) \text{ i } P(B \mid A) = P(B).$$

Ako se posmatra više događaja iz istog prostora elementarnih ishoda, onda se govori da su nezavisni u ukupnosti (ili se samo kaže nezavisni) ako je verovatnoća preseka bilo kojih događaja iz tog skupa događaja jednaka proizvodu verovatnoćâ izdvojenih događaja[77].

Ako su A, B, C i D nezavisni u ukupnosti, onda su nezavisni, na primer, i događaji A, B, C i $\overline{D}$ itd. Suština je u tome da će novoformirani događaji biti nezavisni ako su njihove komponente iz različitih događaja polaznog skupa.

Na osnovu nezavisnosti događaja definiše se nezavisnost slučajnih promenljivih.

Treba razlikovati pojmove: disjunktni događaji i nezavisni događaji: ako su događaji disjunktni, onda su oni zavisni, osim u slučaju da je (bar) jedan od njih nemoguć događaj.

---

[77] Da bi se ustanovila nezavisnost u ukupnosti za tri događaja terba proveriti da li važe *sve* jednakosti:

$$P(AB) = P(A)P(B),$$
$$P(AC) = P(A)P(C),$$
$$P(BC) = P(B)P(C) \text{ i}$$
$$P(ABC) = P(A)P(B)P(C).$$

U slučaju kad se proverava nezavisnost u ukupnosti za n događaja treba proveriti 2^n–n–1 jednakost i samo ukoliko su sve ispunjene, događaji će biti nezavisni u ukupnosti..

Nivo poverenja

Ниво поверења

EN Confidence level
FR Niveau de confiance
RU Уровень доверия
ZH 置信系数, zhìxìn xìshù
AR مستوى الثقة
EL Επίπεδο εμπιστοσύνης
DE Konfidenzniveau
ES Nivel de confianza
IT Livello di confidenza
SK Konfidenčná hladina
SL Raven zaupanja
HU Konfidencia szint

Povezani termini:

Interval poverenja (114), ocenjivanje parametara (178), statistike (167), testiranje hipoteza (216), verovatnoća (248)

Etimologija:

Nivo < niveau[FR] (ravan) < lībella[LA] (libela) < libra[LA] (vaga)

Kada se na osnovu prostog slučajnog uzorka određuje intervalna ocena nepoznatog parametra, tj. interval koji, sa unapred zadatom verovatnoćom β, sadrži nepoznati parametar, onda se ta verovatnoća naziva **nivo poverenja**. Zbog značenja te verovatnoće prirodno je da ona bude bliska jedinici, a u primenama se uzima da je vrednost nivoa poverenja β najčešće je $\beta = 0.9$ ili $\beta = 0.95$ ili $\beta = 0.99$.

Veza nivoa poverenja i statistika koje određuju granice intervala poverenja je data u definiciji: ako je θ nepoznati parametar u raspodeli obeležja X ($X_1, X_2,..., X_n$) prost slučajan uzorak obima n za posmatrano obeležje i statistike $f(X_1, X_2,..., X_n)$ i $g(X_1, X_2,..., X_n)$ takve da važi:

$$P\{f \le g\} = 1 \text{ i } P\{f \le \theta \le g\} = \beta, \beta \in [0,1].$$

Tada se $[f, g]$ naziva interval poverenja za nepoznati parametar θ sa nivoom poverenja β. Obično se kaže da je to 100β% interval poverenja za nepoznati parametar (tj. 90% interval poverenja ili 95% interval poverenja...).

U statističkoj kontroli kvaliteta korišćenjem $\overline{\text{x}}$-bar kartica se, najčešće sa 99% nivoom poverenja, daju granice u kojima posmatrana karakteristika proizvoda treba da bude da bi se smatralo da je „proces pod kontrolom“.

U slučaju testiranja hipoteza o nepoznatom parametru θ, ako je alternativna hipoteza dvostrana, onda je interval poverenja komplement kritične oblasti testa, a samim tim za nivo poverenja β i prag značajnosti α važi $\beta = 1-\alpha$.

Normalna raspodela

Нормална расподела

EN Normal distribution
FR Distribution normale
RU Нормальное распределение
ZH 正态分布, zhèngtài fēnbù
AR التوزيع الطبيعي
EL Κανονική κατανομή
DE Normalverteilung
ES Distribución normal
IT Distribuzione normale
SK Normálne rozdelenie
SL Normalna porazdelitev
HU Normális eloszlás

Povezani termini:

Centralna granična teorema (46), disperzija (54), funkcija raspodele (88), gustina raspodele (84), momenti raspodele (158), očekivanje (152)

Etimologija:

Normalan < normālis^{LA} (uspravan) < norma^{LA} (stolarski ugaonik, deo stolarskog alata)

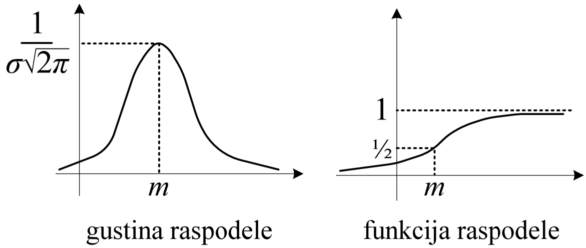
Drugi naziv:

Gausova raspodela

U teoriji verovatnoće i matematičkoj statistici posebno značajno mesto zauzima **normalna** ili **Gausova raspodela**. Ako je gustina raspodele slučajne veličine X:

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \ x \in R, \sigma > 0, m \in R,$$

tada se kaže da X ima normalnu raspodelu sa parametrima m i σ^2 što se označava sa $X{:}N(m, \sigma^2)$.



Gustina raspodele g i funkcija raspodele F slučajne veličine $X{:}N(m, \sigma^2)$

Na grafiku gustine raspodele $g(x)$ slučajne veličine koja ima normalnu raspodelu $N(m, \sigma^2)$ uočavaju se: lokalni maksimum u tački $x = m$, simetrija grafika u odnosu na pravu $x = m$, dve prevojne tačke $x = m \pm \sigma$, horizontalna asimptota $y = 0$, i pri $x \to -\infty$, i pri $x \to +\infty$, maksimalna vrednost jednaka $1/(\sigma\sqrt{2\pi})$.

Funkcija raspodele $F(x)$ slučajne veličine koja ima normalnu raspodelu $N(m, \sigma^2)$ ima prevojnu tačku $x = m$, i njena vrednost u toj tački je jednaka ½, dok su $y = 0$ i $y = 1$ horizontalne asimptote pri $x \to -\infty$, odnosno $x \to +\infty$.

Svi centralni momenti neparnog reda su jednaki nuli $\mu_{2k+1} = 0$, $k = 0, 1, 2, \ldots$, a centralni momenti parnog reda su $\mu_{2k} = 1 \cdot 3 \cdot \ldots \cdot (2k-1)\sigma^{2k} = (2k-1)!!\sigma^{2k}$, $k = 1, 2, \ldots$

→ *Nastavak na str. 262*

Nulta hipoteza
Нулта хипотеза

EN	Null hypothesis
FR	Hypothèse nulle
RU	Нулевая гипотеза
ZH	解消假设, jiěxiāo jiǎshè
AR	الفرضية المبدئية
EL	Μηδενική υπόθεση
DE	Nullhypothese
ES	Hipótesis de nulidad
IT	Ipotesi nulla
SK	Nulová hypotéza
SL	Ničelna domneva
HU	Nullhipotézis

Povezani termini:

Greška druge vrste (96), greška prve vrste (98), kritična oblast testa (136), moć testa (86), prag značajnosti (186), testiranje hipoteze (216)

Etimologija:

Nula < nūllus^LA (nijedan)

Hipoteza < ὑπόθεση^EL (staviti ispod, pretpostaviti) < ὑπό^EL (ispod) + τίθημι^EL (postaviti)

Hipoteza koja se testira se zove **nulta hipoteza** i obično označava sa H_0. Nulta hipoteza je obično ona koja je na neki način važnija u konkretnom zadatku, tj. hipoteza čije bi odbacivanje moglo da ima negativne posledice. Jasan primer za to je nulta hipoteza da je neki kozmetički preparat štetan. Odbacivanje te hipoteze bi značilo da preparat nije štetan, iako zapravo jeste, pa bi to moglo da ima nesagledive negativne posledice.

U paru s nultom hipotezom uvek se javlja hipoteza H_1 ili alternativna hipoteza koja na neki način protivureči hipotezi H_0. Obe hipoteze mogu biti proste ili složene. Za hipotezu se kaže da je prosta, ako se odnosi na jednu vrednost parametra kojom je raspodela obeležja potpuno određena, npr. $H_0(\theta = \theta_0)$. Ako hipoteza nije prosta, onda je složena, kao što su hipoteze:

$$H_0(\theta \neq \theta_0),\ H_0(\theta < \theta_0),\ H_0(\theta > \theta_0).$$

Nulta hipoteza se može odnositi i na raspodelu obeležja, npr. $H_0(F_X(x) = F_0(x))$. Na osnovu nulte hipoteze se često može ustanoviti i koja test-statistika će se koristiti. Na primer, ako se nulta hipoteza odnosi na srednju vrednost obeležja, tada uzoračka sredina može biti upotrebljena kao test-statistika.

Ako se odbacuje nulta hipoteza kada je ona tačna, čini se greška prvog tipa, a ako se prihvata nulta hipoteza kada je tačna alternativna hipoteza, čini se greška drugog tipa.

Verovatnoća donošenja pravilne odluke: hipoteza H_0 nije tačna i odbacuje se, naziva se moć testa. Moć testa treba da bude veća od verovatnoće greške prve vrste.

Obeležje

Обележје

EN	Characteristic
FR	Caractère
RU	Признак
ZH	特征, tèzhēng
AR	الخاصية
EL	Χαρακτηριστικό
DE	Merkmal
ES	Característica
IT	Carattere
SK	Charakteristika
SL	Lastnost
HU	Jellemző

Povezani termini:

Populacija (182), uzorak (240)

Elemente populacije posmatramo u odnosu na izvesnu varijabilnu kvantitativnu ili kvalitativnu osobinu koja se naziva **obeležje**[78]. Obeležje koje se posmatra može biti jednodimenzionalno, dvodimenzionalno ili višedimenzionalno.

Kvalitativna karakteristika se često može prevesti u kvantitativnu (numeričku). Prevođenje kvalitativne karakteristike u kvantitativnu nije jednoznačno i predstavlja dogovor među osobama koje sprovode određeno istraživanje. Prevođenje kvalitativnih u kvantitativne karakteristike se koji put naziva šifriranje podataka. Tako se dobijaju kompaktnije baze podataka i omogućava obrada podataka statističkim metodama. U nekim statističkim ispitivanjima nije neophodno prevođenje kvalitativnih karakteristika u kvantitativne. Jedan od takvih primera je χ^2-test.

Ako se na slučajan način izabere jedan element populacije ne zna se unapred vrednost obeležja (modalitet) koju taj element ima. To znači da se vrednost obeležja na slučajno izabranom elementu populacije može shvatiti kao vrednost slučajne veličine, koja može biti jednodimenzionalna, dvodimenzionalna ili višedimenzionalna. Raspodela verovatnoća te slučajne veličine se naziva raspodela obeležja.

Zadatak statistike je određivanje raspodele obeležja, ili, bar nekih parametara te raspodele. Radi toga se prikupljaju podaci o populaciji (merenjem, anketama, popisima, itd.).

[78] *Primeri obeležja:* krvna grupa (jednodimenzionalno), krvna grupa i Rh faktor (dvodimenzionalno), temperatura vazduha, atmosferski pritisak i vlažnost vazduha (trodimenzionalno), boja očiju (kvalitativno) i sl.

Ocenjivanje parametara

Оцењивање параметара

EN Estimation of parameters
FR Estimation de paramètres
RU Оценка параметров
ZH 参数估计, cānshù gūjì
AR تقدير المعلمات
EL Εκτίμηση παραμέτρων
DE Parameterschätzung
ES Estimación de parámetros
IT Stima dei parametri
SK Odhad parametrov
SL Ocenjevanje parametrov
HU Paraméterbecslés

Povezani termini:

Obeležje (176), raspodela obeležja, uzorak (240)

Etimologija:

Parametar < παρά^{EL} (pored) + μέτρον^{EL} (mera)

Neka je dat prost slučajni uzorak ($X_1, X_2,..., X_n$) obima n za posmatrano obeležje X.

Statistika $Y = f(X_1, X_2,..., X_n)$ je slučajna veličina koja implicitno zavisi od parametara u raspodeli obeležja X.

Ako se statistikom Y ocenjuje parametar θ, tada se statistika Y naziva **ocena parametra** θ i označava se sa $\hat{\theta}$ ili $\hat{\theta}_n$.

Nepoznati parametri raspodele posmatranog obeležja se ocenjuju na osnovu uzorka, kao realizovane vrednosti pogodno odabranih statistika. Taj postupak se naziva **ocenjivanje parametara**. Pošto je realizovana vrednost statistike neki realni broj, tj. neka tačka na realnoj pravoj, ovakve ocene parametara se nazivaju tačkaste ocene.

Pošto se realizovana vrednost tačkaste ocene razlikuje od jednog uzorka do drugog (iako su istog obima), moguć je i drugi pristup ocenjivanju – tzv. intervalno ocenjivanje. Tada se, sa određenim, unapred zadatim, nivoom poverenja β utvrđuje interval u kome će se nepoznati parametar nalaziti sa verovatnoćom β.

Statistike koje se koriste pri ocenjivanju parametara treba da imaju određene osobine kojima se opravdava njihova primena u ocenjivanju parametara. Jedna od tih osobina je nepristrasnost, a tu su takođe i osobine: postojanost (stabilnost), efikasnost, dovoljnost, i druge.

Pošto je moguće da se različitim statistikama ocenjuje isti parametar[79], bitno je da postoji način za poređenje ocena. Jedan od načina je poređenje ocena po efikasnosti.

[79] Neka je raspodela obeležja $U(0,\theta)$. Tada su npr. statistika poretka maksimalnog ranga X_{max} i $2\overline{X}_n$ moguće ocene za θ.

Poligon frekvencija

Полигон фреквенција

EN	Frequency polygon
FR	Polygone de fréquences
RU	Полигон частот
ZH	频率多边形, pínlǜ duōbiānxíng
AR	المضلع التكراري
EL	Πολύγωνο συχνοτήτων
DE	Häufigkeitspolygon
ES	Polígono de frecuencias
IT	Poligono di frequenza
SK	Frekvenčný polygón
SL	Frekvenčni poligon
HU	Gyakorisági poligon

Povezani termini:

Apsolutna frekvencija (211), grafik (94), realizovani uzorak (189), relativna frekvencija (211), uzorak (240)

Etimologija:

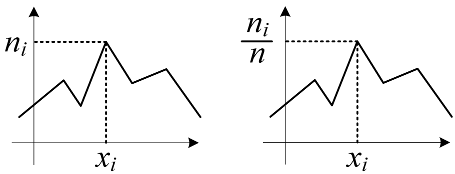
Poligon < πολύ^EL (mnogo) + γωνία^EL (ugao).

Frekvencija < frequēns^LA (čest, prepun). Ova reč je, smatra se, nastala od proto-indoevropskog korena *bhrek- (sabiti se, nagurati se), pa je možda srodna i sa našom rečju *zbrka*.

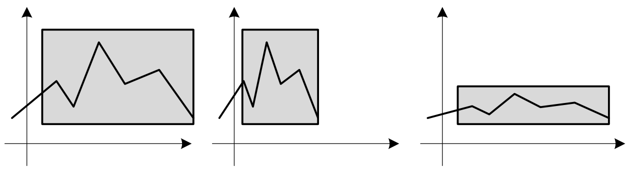
Poligoni frekvencija se dobijaju kad su podaci dati u obliku tabele:

vrednost obeležja	x_1	x_2	...	x_k
frekvencija	n_1	n_2	...	n_k

Tada se u koordinatnom sistemu predstave tačke (x_i, n_i), ili tačke (x_i, n_i/n), $i = 1, 2,..., k$ i spoje izlomljenom linijom. Po izgledu su obe slike iste, ali je različita podela na vertikalnoj osi.



Na sledećoj slici su dati poligoni frekvencija za iste podatke, ali je izbor mernih jedinica na koordinatnim osama različit.



Prethodna ilustracija ukazuje na značaj pravilnog izbora dimenzija grafika. Smatra se da se najbolje uočava međusobni odnos elemenata na grafiku ako se njegova maksimalna visina i njegova maksimalna širina (misli se na deo gde su podaci, što je predstavljeno osenčenim pravougaonikom na slici) odnose kao 1:√2, a to je odnos tzv. zlatnog preseka, poznat još u staroj Grčkoj.

Populacija
Популација

EN	Population
FR	Population
RU	Генеральная совокупность
ZH	总体, zǒngtǐ
AR	السكان
EL	Πληθυσμός
DE	Population
ES	Población
IT	Popolazione
SK	Populácia
SL	Populacija
HU	Sokaság

Povezani termini:

Obeležje (176), statistika (150)

Etimologija:

Populacija < populātiō^{LA} (naseljavanje) < populus^{LA} (narod)

Izučavani skup objekata se naziva **populacija**[80] (osnovni skup ili generalna kolekcija). Elemente tog skupa posmatramo u odnosu na izvesnu varijabilnu kvantitativnu ili kvalitativnu osobinu koja se naziva obeležje.

Kvalitativna karakteristika se često prevodi u kvantitativnu (numeričku), ali to prevođenje nije jednoznačno i predstavlja dogovor među osobama koje sprovode određeno istraživanje.

Populacije mogu imati konačno ili beskonačno mnogo elemenata.

Ako se na slučajan način izabere jedan element populacije ne zna se unapred vrednost obeležja (modalitet) koju taj element ima. To znači da se vrednost obeležja na slučajno izabranom elementu populacije može shvatiti kao vrednost slučajne veličine. Raspodela verovatnoća te slučajne veličine se naziva raspodela obeležja, i predstavlja bitnu karakteristiku populacije.

Zadatak statistike je određivanje raspodele obeležja, ili, bar nekih parametara te raspodele. Radi toga se prikupljaju podaci o populaciji: merenjima, anketama, popisima, itd. Takođe, podaci o populaciji mogu biti dobijeni proučavanjem nekog njenog dela (uzorak), pod uslovom da je pravilno sproveden postupak izbora elemenata populacije u uzorak (pitanje reprezentativnosti uzorka).

[80] *Primeri populacija:* skup formula u ovoj knjizi, skup reči u ovoj knjizi, skup stanovnika jednog grada ili neke oblasti, skup učenika jedne škole, skup vrednosti temperature vazduha na jednom mernom mestu, vodostaj reke tokom godine, itd.

Postojana ocena

Постојана оцена

EN	Consistent estimator
FR	Estimateur consistant
RU	Состоятельная оценка
ZH	相容估计量, xiāngróng gūjìliàng
AR	تقدير متناسق
EL	Συνεπής εκτιμήτρια
DE	Konsistenter Schätzer
ES	Estimador consistente
IT	Stimatore consistente
SK	Konzistentný odhad
SL	Dosledna cenilka
HU	Konzisztens becslés

Povezani termini:

Ocenjivanje parametra (178), nepristrasnost (164), statistika (150), uzorak (240)

Drugi naziv:

Konzistentna ocena

Za realizovani uzorak $(x_1, x_2,..., x_n)$ se računa realizovana vrednost statistike $Y = f(X_1,..., X_n)$, tj. računa se $y = f(x_1, x_2,..., x_n)$. S obzirom da je Y slučajna veličina, ta realizovana vrednost će u izvesnoj meri odstupati od $E(Y)$. Poželjno je da se to odstupanje smanjuje kako se povećava obim uzorka. Stoga se, u kategoriji nepristrasnih ocena, izdvajaju one koje imaju dole navedeno svojstvo, tzv. postojanost (stabilnost).

Neka je θ nepoznati parametar u raspodeli obeležja X i neka je dat prost slučajni uzorak $(X_1, X_2,..., X_n)$ obima n za obeležje X. Statistika $Y = f(X_1, X_2,..., X_n)$ je **postojana ocena** (u verovatnoći)[81] parametra θ ako

$$Y \xrightarrow{P} \theta,\ n \to \infty .$$

Prilikom utvrđivanja postojanosti ocena često se koristi tvrđenje[82]:

Ako je statistika $Y = f(X_1, X_2,..., X_n)$ nepristrasna ocena za nepoznati parametar θ i ako $D(Y) \to 0$, $n \to \infty$, tada je Y postojana ocena za θ[83].

[81] Osim postojane ocene u verovatnoći, definišu se i jako postojana ocena (kada je u pitanju skoro sigurna konvergencija), kao i postojanost u srednje kvadratnom smislu (kada je u pitanju konvergencija u srednje kvadratnom).

[82] Tvrđenje se dokazuje primenom nejednakosti Čebiševa.

[83] Na osnovu tog tvrđenja dobijamo da je uzoračka sredina postojana ocena matematičkog očekivanja obeležja X koje ima konačnu disperziju σ^2, jer je disperzija uzoračke sredine jednaka σ^2/n. Uzoračka sredina pripada klasi linearnih nepristrasnih ocena matematičkog očekivanja, ali ocene koje pripadaju toj klasi nisu obavezno postojane.

Prag značajnosti

Праг значајности

EN	Level of significance
FR	Niveau de signification
RU	Уровень значимости
ZH	显著性水平, xiănzhùxìng shuĭpíng
AR	مستوى الدلالة
EL	Επίπεδο σημαντικότητας
DE	Signifikanzniveau
ES	Nivel de significación
IT	Livello di significatività
SK	Hladina významnosti
SL	Stopnja značilnosti
HU	Szignifikanciaszint

Povezani termini:

Greška prve vrste (98), kritična oblast (136), testiranje hipoteze (216)

Verovatnoća α odbacivanja hipoteze H_0, ako je ona tačna treba da bude mala. Broj α se još naziva **prag značajnosti** ili **nivo značajnosti**, a takođe i verovatnoća greške prve vrste. U praksi se najčešće uzima da je α jednako 0.01, 0.05 ili 0.1. Prag značajnosti određuje veličinu kritične oblasti.

Iskaz „Verovatnoća da pri tačnoj hipotezi H_0 uzorak $(X_1,..., X_n)$ bude u kritičnoj oblasti W_n jednaka je pragu značajnosti", odnosno

$$P_{H_0}\left\{\left(X_1, X_2, ..., X_n\right) \in W_n\right\} = \alpha$$

je karakterističan za sve testove. Može se javiti i u obliku „Verovatnoća da pri tačnoj hipotezi H_0 test-statistika Θ_n bude u kritičnoj oblasti W jednaka je pragu značajnosti"

$$P_{H_0}\left\{\Theta_n \in W\right\} = \alpha .$$

Pri testiranju parametarskih hipoteza ukoliko je kritična oblast dvostrana, prag značajnosti je komplementarna verovatnoća nivou poverenja za intervalnu ocenu posmatranog parametra, jer je interval poverenja tada komplement kritične oblasti.

Često se sreće formulacija da je neki parametar ili neka numerička karakteristika obeležja „statistički značajna veličina". To zapravo znači je hipoteza da je vrednost te veličine različita od nule prihvaćena posle testiranja sa nekim pragom značajnosti[84].

[84] Na primer, ako je model zavisnosti $Y = a_1X_1 + a_2X_2$ i za a_1 se kaže da nije statistički značajna veličina, proizilazi da Y zavisi samo od X_2.

Prost slučajan uzorak

Прост случајан узорак

EN Simple random sample
FR Echantillon aléatoire simple
RU Простая случайная выборка
ZH 简单随机样本, jiǎndān suíjī yàngběn
AR العينة العشوائية البسيطة
EL Απλό τυχαίο δείγμα
DE Einfache Zufallsstichprobe
ES Muestra aleatoria simple
IT Campione casuale semplice
SK Jednoduchý náhodný výber
SL Enostavni slučajni vzorec
HU Egyszerű véletlen minta

Povezani termini:

Nezavisnost slučajnih veličina, populacija (182), reprezentativni uzorak (196), uzorak (240)

Neka se u populaciji posmatra obeležje X. **Prost slučajni uzorak** obima n za posmatrano obeležje je n-dimenzionalna slučajna veličina $(X_1, X_2, \ldots, X_n)$ pri čemu su slučajne veličine $X_1, X_2, \ldots, X_n$ nezavisne i sve imaju istu raspodelu kao posmatrano obeležje X.

Prost slučajni uzorak je teorijska kategorija koja obezbeđuje reprezentativnost uzorka.

Realizovani uzorak predstavlja konkretan niz vrednosti obeležja dobijenih na elementima populacije koji su izabrani u uzorak. To je, dakle, jedna konkretna n-torka brojeva. I pri tom se unapred ne zna koje vrednosti obeležja će biti u konkretnom uzorku.

Metode matematičke statistike mogu dati odgovore na pitanja:

- koji elementi u uzorku će se češće javljati,
- u kojim granicama će, i sa kojom verovatnoćom, broj pojavljivanja jednog modaliteta u uzorku odstupati od očekivanog broja pojavljivanja tog modaliteta,
- da li je raspodela obeležja na populaciji (poznata raspodela ili neka pretpostavljena raspodela) u skladu (u saglasnosti) sa raspodelom obeležja koja je dobijena na uzorku, itd.

Uglavnom je iz konteksta jasno da li se govori o prostom slučajnom uzorku kao n-dimenzionalnoj slučajnoj veličini ili o relizovanom uzorku – uređenoj n-torki brojeva.

Prostor elementarnih ishoda

Простор елементарних исхода

EN	Sample space
FR	Espace échantillonnal
RU	Пространство элементарных событий
ZH	样本空间, yàngběn kōngjiān
AR	فضاء العينة
EL	Δειγματοχώρος
DE	Ergebnisraum
ES	Espacio muestral
IT	Spazio campionario
SK	Priestor elementárnych udalostí
SL	Prostor vzorcev
HU	Eseménytér

Povezani termini:

Uzorački prostor (191), verovatnoća (248)

Etimologija:

Elementaran < elementum[LA] (pramaterija, osnovni princip)

Prilikom proučavanja raznih problema uočavaju se pojave koje se ostvaruju pri realizaciji nekog kompleksa uslova. Jedna realizacija posmatranog kompleksa uslova naziva se eksperimentom. Posmatraju se eksperimenti koje je moguće ponoviti neograničen broj puta pri istim uslovima. Pre izvođenja eksperimenta se tačno zna šta se smatra ishodom eksperimenta i poznat je skup mogućih ishoda, ali ishod svakog pojedinačnog eksperimenta nije unapred poznat. Takvi eksperimenti se često nazivaju stohastički eksperimenti, za razliku od determinističkih eksperimenata, u kojima su ishodi poznati i pre izvođenja eksperimenta.

U teoriji verovatnoće je elementaran događaj ili elementaran ishod osnovni pojam i ne definiše se. Skup svih elementarnih ishoda jednog eksperimenta je prostor elementarnih ishoda.

U zavisnosti od rezultata koji se posmatraju, formira se odgovarajući **prostor elementarnih ishoda**. Prostor elementarnih ishoda može imati konačno mnogo, prebrojivo mnogo ili neprebrojivo mnogo elementarnih ishoda.

U statistici prostoru elementarnih ishoda odgovara uzorački prostor. Uzorački prostor za uzorak obima n je skup svih tačaka $(x_1, x_2,\ldots, x_n)$ n-dimenzionalnog prostora koje mogu biti realizovane vrednosti prostog slučajnog uzorka obima n. Uzorački prostor može imati konačno, prebrojivo mnogo ili neprebrojivo mnogo tačaka.[85]

---

[85] Ako je obeležje X sa uniformnom raspodelom na intervalu (0,1), onda je uzorački prostor za uzorak obima 2 – *kvadrat*, za uzorak obima 3 – *kocka*, a za uzorak obima većeg od 3 – *hiperkocka*. Ako je obeležje X broj na gornjoj strani numerisane kockice, i kockicu bacamo 5 puta (ili, što je isto, bacamo 5 kockica kao u igri Jamb), tada uzorački prostor ima $6^5 = 7776$ elemenata: (1,1,1,1,1), (1,1,1,1,2),... , (6,6,6,6,6).

Puasonova raspodela

Пуасонова расподела

EN	Poisson distribution
FR	Distribution de Poisson
RU	Распределение Пуассона
ZH	泊松分布, Pōsōng fēnbù
AR	توزيع بواسون
EL	Κατανομή Poisson
DE	Poissonsche Verteilung
ES	Distribución de Poisson
IT	Distribuzione di Poisson
SK	Poissonovo rozdelenie
SL	Poissonova porazdelitev
HU	Poisson-eloszlás

Povezani termini:

Binomna raspodela (38), disperzija (54), moda (228), momenti (238), očekivanje (152)

Ako je $\lambda > 0$ i

$$X:\begin{pmatrix} 0 & 1 & \dots & n & \dots \\ p_0 & p_1 & \dots & p_n & \dots \end{pmatrix},\ p_j = \frac{\lambda^j}{j!}e^{-\lambda}, j = 0,1,2...$$

tada se kaže da X ima **Puasonovu raspodelu** sa parametrom λ, što se zapisuje sa $X{:}P(\lambda)$. Ova slučajna veličina ima prebrojivo mnogo vrednosti, a uspešno se koristi pri opisivanju slučajnih pojava kod kojih se najverovatnije javlja mali broj realizacija posmatranog događaja, dok se verovatnoće da se javi veliki broj realizacija smanjuju sa povećanjem broja realizacija[86].

Ako se posmatra binomna raspodela $S_n : B(n, p_n)$ kod koje verovatnoća p_n realizacije događaja A u n-tom opitu zavisi od n, tako da $np_n \to \lambda$, $n \to \infty$, tada važi tzv. **Puasonova aproksimacija** binomne raspodele, tj.

$$P\{S_n = j\} \to \frac{\lambda^j}{j!}e^{-\lambda}, n \to \infty, j = 0,1,2...$$

Ovo znači da će se aproksimacija koristiti ako je

$$p_\lambda \approx \lambda / n.$$

Aproksimacija verovatnoćâ iz binomnog zakona

$$P\{S_n = j\} \approx \frac{\lambda^j}{j!}e^{-\lambda}$$

se obično koristi za velike vrednosti n (n veće od 30) i ako je $\lambda = np < 10$.

Ako parametar λ nije prirodan broj, tada raspodela ima jedan mod jednak celom delu broja λ, a ako je λ prirodan broj, tada Puasonova raspodela ima maksimum u dvema tačkama $(\lambda - 1)$ i λ. Matematičko očekivanje i disperzija slučajne veličine X koja ima Puasonovu raspodelu sa parametrom λ su jednaki λ.

→ *Nastavak na str. 262*

---

[86] Na primer, broj rođenih trojki ima Puasonovu raspodelu.

Ravnomerna raspodela

Равномерна расподела

EN	Uniform distribution
FR	Distribution uniforme
RU	Равномерное распределение
ZH	均匀分布, jūnyún fēnbù
AR	التوزيع المنتظم
EL	Ομοιόμορφη κατανομή
DE	Gleichverteilung
ES	Distribución uniforme
IT	Distribuzione uniforme
SK	Rovnomerné rozdelenie
SL	Enakomerna porazdelitev
HU	Egyenletes eloszlás

Povezani termini:

Asimetrija (18), disperzija (54), gustina raspodele (84), matematičko očekivanje (152), spljoštenost (124)

Drugi naziv:

Uniformna raspodela

Ako je gustina raspodele slučajne veličine X jednaka

$$g(x)=\begin{cases}\dfrac{1}{b-a} & , \quad x\in(a,b),\\ 0 & , \quad x\notin(a,b),\end{cases}$$

tada se kaže da je X slučajna veličina sa **ravnomernom raspodelom na intervalu (*a*,*b*)**, a koristi se i termin **uniformna raspodela**.

Uobičajena oznaka ravnomerne raspodele je

$$X:\ U(a,\ b).$$

Matematičko očekivanje i disperzija ove slučajne veličine su:

$$E(X)=\frac{a+b}{2}\ ,\quad D(X)=\frac{1}{12}(b-a)^2\,.$$

Sem toga, centralni momenti trećeg i četvrtog reda i koeficijenti asimetrije i spljoštenosti su redom jednaki:

$$\mu_3=0\,,\quad \mu_4=\frac{(b-a)^4}{80}\,,\quad f_1=\frac{\mu_3}{\sigma^3}=0\,,$$

$$f_2=\frac{\mu_4}{\sigma^4}-3=-1.2\,.$$

Ravnomernu raspodelu imaju greške grubih merenja pomoću aparata sa krupnom podelom, kada se izmerena vrednost zaokružuje na najbliži ceo broj (na najbliži manji ili na najbliži veći ceo broj). Ako se tačka M slučajno bira na realnoj pravoj na intervalu dužine d, tada je rastojanje tačke M od levog kraja intervala slučajna veličina sa uniformnom raspodelom $U(0,d)$.

Ravnomerna raspodela je od posebnog značaja u modeliranju slučajnih veličina (metode Monte-Karlo)

Reprezentativni uzorak

Репрезентативни узорак

EN Representative sample
FR Echantillon représentatif
RU Репрезентативная выборка
ZH 具有代表性的样本, jùyǒu dàibiǎoxìng de yàngběn
AR عينة تمثيلية
EL Αντιπροσωπευτικό δείγμα
DE Repräsentative Stichprobe
ES Muestra representativa
IT Campione rappresentativo
SK Reprezentatívny výber
SL Reprezentativni vzorec
HU Reprezentatív minta

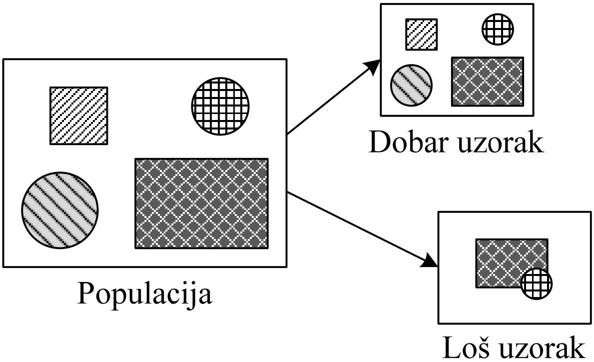
Povezani termini:

Obeležje (176), populacija (182), prost slučajni uzorak (188)

Etimologija:

Reprezentativan < repræsentāre[LA] (predstavljati) < re-[LA] (opet) + præsentāre[LA] (staviti ispred) < præ-[LA] (ispred) + esse[LA] (biti)

Kako je zadatak statistike određivanje raspodele obeležja na populaciji na osnovu obeležja na uzorku, od velike je važnosti obezbeđivanje **reprezentativnosti uzorka**. Neophodno je obezbediti uzorak koji dobro reprezentuje populaciju, tj. predstavlja „umanjenu“, a nikako „iskrivljenu“, niti „uveličanu“ sliku jednog dela populacije, kao što je prikazano na sledećoj slici:



Metode izbora uzorka treba da isključuju sistematske greške. Dopustive su samo slučajne greške, čiji se uticaj može procenjivati na osnovu teorije verovatnoće. Postoje metode eliminisanja grubih grešaka.

Jedan od načina dobijanja reprezentativnog uzorka je uzorak sa vraćanjem. Ako je populacija veoma velika, a uzorak veoma malog obima u odnosu na obim populacije, tada i uzorak dobijen po šemi izbora bez vraćanja ima dobru reprezentativnost.

Slučajna promenljiva

Случајна променљива

EN Random variable
FR Variable aléatoire
RU Случайная переменная
ZH 随机变量, suíjī biànliàng
AR المتغير العشوائي
EL Τυχαία μεταβλητή
DE Zufallsvariable
ES Variable aleatoria
IT Variabile aleatoria
SK Náhodná premenná
SL Slučajna spremenljivka
HU Valószínűségi változó

Povezani termini:

Preslikavanje, prostor elementarnih ishoda (190), verovatnoća (248)

Drugi naziv:

Slučajna veličina

Slučajna promenljiva ili **slučajna veličina** X je preslikavanje koje elementarnim ishodima dodeljuje realne brojeve. Domen tog preslikavanja je prostor elementarnih ishoda, a skup vrednosti neki podskup skupa R^n. Ako je $n = 1$, slučajna promenljiva X je jednodimenzionalna.

Ako je n veće od 1, slučajna promenljiva je višedimenzionalna. Tada je svaka njena „koordinata" jedna jednodimenzionalna slučajna veličina.

Preslikavanje X mora biti finitno, tj.

$$P(X = -\infty) = P(X = \infty) = 0$$

i mora biti merljivo, tj. mora biti moguće računanje verovatnoća događaja oblika $(X \leq x)$ za svaki realni broj x. To je važno da bi se mogla odrediti funkcija raspodele slučajne veličine.

Ako je skup vrednosti slučajne veličine konačan ili prebrojiv, kaže se da je slučajna veličina diskretna[87], sa konačno ili prebrojivo mnogo vrednosti. Ukoliko ima prebrojivo mnogo vrednosti naziva se i elementarna slučajna veličina. Takve slučajne veličine su u potpunosti određene svojim zakonom raspodele, tj. skupom vrednosti koje mogu imati i verovatnoćama sa kojima imaju te vrednosti. Ako je skup vrednosti neki interval, onda je slučajna veličina neprekidna, a koristi se i termin apsolutno neprekidna. Takve slučajne veličine su u potpunosti određene svojom gustinom raspodele. Bilo da su dikretne ili neprekidne slučajne veličine su jednoznačno (do na događaj verovatnoće 0) određene svojim funkcijama raspodele.

Slučajnoj veličini u statistici odgovara obeležje.

[87] Diskretne slučajne veličine su slučajne veličine sa binomnom raspodelom, sa Puasonovom raspodelom, sa geometrijskom raspodelom, sa dikretnom uniformnom raspodelom... Neprekidne slučajne veličine su slučajne veličine sa normalnom, eksponencijalnom, uniformnom raspodelom...

Slučajni događaj

Случајни догађај

EN Random event
FR Evénement aléatoire
RU Случайное событие
ZH 随机事件, suíjī shìjiàn
AR الحدث العشوائي
EL Τυχαίο γεγονός
DE Zufallsereignis
ES Evento aleatorio
IT Evento accidentale
SK Náhodná udalosť
SL Slučajni dogodek
HU Véletlen esemény

Povezani termini:

Prostor elemetarnih ishoda (190), slučajna veličina (198), verovatnoća (248)

Skup svih elementarnih ishoda jednog eksperimenta je prostor elementarnih ishoda Ω. Ako je *A* podskup skupa Ω, tada je *A* **slučajni događaj** i on može imati konačno mnogo, prebrojivo mnogo ili neprebrojivo mnogo elementarnih ishoda, zavisno od prirode Ω.

Pošto su prostor elementarnih ishoda i slučajni događaji skupovi, to se za relacije i operacije sa slučajnim događajima koriste termini kao u teoriji skupova, uz neke specifične termine, kao što je – realizovanje događaja: ako je *A* slučajni događaj i ako je rezultat eksperimenta jedan od elementarnih događaja koji pripada *A*, tj. elementarni događaj koji ima osobinu kojom se događaj *A* definiše, tada se kaže: *događaj A se ostvario (realizovao).*

Skup *A* može biti jednočlan, pa je svaki elementarni ishod – slučajni događaj. Ceo skup Ω i prazan skup Φ su takođe slučajni događaji: Ω je siguran (izvestan, pouzdan) događaj, a prazan skup Φ je nemoguć događaj.

Slučajni događaji se označavaju velikim slovima abecede *A*, *B*, *C*,... ili, po potrebi, sa indeksima: A_1, A_2, A_3,... Pošto se u teoriji verovatnoće i statistici razmatraju samo slučajni događaji, često će, umesto slučajni događaj, pisati samo – događaj.

U teoriji verovatnoće se verovatnoće događaja računaju po određenim pravilima, u matematičkoj statistici se verovatnoće događaja ocenjuju svojim relativnim frekvencijama.

U statistici slučajnom događaju odgovara neki podskup uzoračkog prostora.[88]

[88] Na primer, događaju da je uzoračka sredina manja od nekog broja *c* odgovara skup paralelnih hiperravni, sa slobodnim članom manjim od posmatraog broja *c*.

Standardna devijacija

Стандардна девијација

EN	Standard deviation
FR	Ecart type
RU	Среднеквадратическое отклонение
ZH	标准偏差, biāozhǔn piānchā
AR	الانحراف المعياري
EL	Τυπική απόκλιση
DE	Standardabweichung
ES	Desviación estándar
IT	Scarto quadratico medio
SK	Štandartná odchýlka
SL	Standardni odklon
HU	Standardeltérés

Povezani termini:

Disperzija (54), koeficijent varijacije (126), kontrola kvaliteta (128), očekivanje (152)

Etimologija:

Standardan < estandart^{stFR} (ratni steg, zborno mesto); ili extendere^{LA} (istegnuti, produžiti) < ex-^{LA} (iz) + tendere^{LA} (istegnuti)

Devijacija < dēviāre^{LA} (skrenuti) < de^{LA} (od) + via^{LA} (put)

Pozitivna vrednost kvadratnog korena disperzije slučajne veličine X se naziva **standardna devijacija** i često označava sa σ. Znači, $\sigma^2 = D(X)$.

Ako slučajna veličina X ima matematičko očekivanje i disperziju, onda je standardizovana veličina X:

$$X^* = \frac{X - E(X)}{\sqrt{D(X)}} = \frac{X - E(X)}{\sigma},$$

i njeno očekivanje i disperzija će redom biti jednaki 0 i 1.

Jedna od važnih primena standardne devijacije je u statističkoj kontroli kvaliteta. Ukoliko je dimenzija D proizvoda njegova bitna karakteristika, a D_0 vrednost koju bi po propisima trebalo da ima, tada na osnovu osobina normalne raspodele i pravila 3-sigma, proizilazi da će u intervalu (D_0–3σ, D_0+3σ) biti 99,7% proizvoda. Stoga, ukoliko se za uzorak iz proizvodnje dobije da uzoračka sredina ne pripada navedenom intervalu, onda se smatra da je proces van kontrole i da je neophodna neka korekcija proizvodnog procesa. Postoje i razni drugi principi za statističku kontrolu kvaliteta.

Standardna devijacija u uzorku (x_1, x_2, ..., x_n) je kvadratni koren uzoračke disperzije, i njeno računanje zavisi od toga kako su podaci dati.

Ukoliko slučajna veličina (obeležje) ima dimenziju d, tada je i dimenzija standardne devijacije jednaka d, dok je dimenzija disperzije $D(X)$ jednaka d^2. Zato je X^* bez fizičke dimenzije

Standardna devijacija se koristi pri definisanju koeficijenta varijacije.

Statistike poretka

Статистике поретка

EN Order statistics
FR Statistiques d’ordre
RU Порядковые статистики
ZH 次序统计量, cìxù tǒngjìliàng
AR الأحصاءات المرتبة
EL Στατιστικές τάξης μεγέθους
DE Ordnungsstatistik
ES Estadísticas de orden
IT Statistiche d’ordine
SK Poriadkové štatistiky
SL Vrstilne statistike
HU Rendstatisztikák

Povezani termini:

Empirijska funkcija raspodele (72), medijana (154), uzorak (240)

Etimologija:

Statistika < statisticum^{LA} (ono što ima veze s državom). Termin statistika, sa prvobitnim značenjem analiziranja podataka o državi, uveo je Gotfrid Ahenval (1719–1772), nemački filozof, pravnik i ekonomista.

Statistike poretka $Y_1, Y_2,..., Y_n$ prvog (minimalnog), drugog,..., n-tog (maksimalnog) ranga su redom prvi, drugi,..., n-ti element varijacionog niza. Varijacioni niz je niz koji se dobija kad se podaci iz uzorka urede u neopadajućem poretku. Varijacioni niz se koristi i pri formiranju empirijske funkcije raspodele.

Posebno se izdvajaju:

statistika poretka prvog ranga $Y_1 = \min_{1 \le i \le n} X_i$,

statistika poretka *n*-tog ranga $Y_n = \max_{1 \le i \le n} X_i$,

a pomoću njih se definiše raspon uzorka $R = Y_n - Y_1$.

Statistike poretka se koriste i za definisanje medijane uzorka. Naime, uzoračka medijana je statistika

$$M_e = \begin{cases} Y_{k+1}, & ako \quad je \quad n = 2k+1 \\ \frac{1}{2}\left(Y_k + Y_{k+1}\right), & ako \quad je \quad n = 2k \end{cases}$$

Raspodele statistika poretka mogu se izraziti u funkciji raspodele obeležja[89].

[89] Raspodela statistike poretka minimalnog ranga za obeležje X: $P(Y_1 < x) = 1 - P(Y_1 \ge x) = 1 - P(X_1 \ge x) \ldots P(X_n \ge x) = 1 - (P(X \ge x))^n = 1 - (1 - P(X < x))^n$, za svaki realni broj x.
Raspodela statistike poretka maksimalnog ranga za obeležje X: $P(Y_n < x) = P(X_1 < x) \ldots P(X_n < x) = (P(X < x))^n$, za svaki realni broj x.

Stem end lif dijagram

Стем енд лиф дијаграм

EN	Stem and leaf diagram
FR	Diagramme branche-et-feuilles
RU	Диаграмма “ствол с листьями”
ZH	枝叶图, zhīyètú
AR	الرسم بطريقة الساق والورقة
EL	Διάγραμμα μίσχου-φύλλου
DE	Stamm-und-Blätter-Diagramm
ES	Diagrama de tallos y hojas
IT	Diagramma ramo-foglia
SK	Číslicový dendrogram
SL	Histogram s številkami
HU	Leveles ág bemutatás

Povezani termini:

Raspodela obeležja, grafik (94), uzorak (240)

Etimologija:

Dijagram < διάγραμμα^{EL} (ono što je ocrtano linijama) < διά^{EL} (kroz, od) + γράμμα^{EL} (slovo, ono što je nacrtano)

Dijagram „grana i lišće“ je jedan od jednostavnih načina sortiranja i istovremenog grafičkog prikazivanja podataka. Postupak dobijanja je najjednostavnije objasniti na primeru. Neka su dati podaci:

33.1, 33.4, 34.8, 33.8, 34.7, 34.3, 34.5, 34.6, 34.6, 34.9, 33.3, 34.2, 35.2, 35.6, 36.0, 36.1, 35.2, 35.1, 34.0, 33.9, 33.6, 33.4, 33.7, 34.2, 35.1, 34.2, 33.7, 34.2, 35.8, 35.3

Povučemo vertikalnu liniju, „granu“ i sa jedne strane te vertikalne linije se piše deo svakog podatka sem poslednje cifre, i tako se formira „grana“, a desno od linije se piše redom, na odgovarajućim mestima, preostala cifra od svakog podatka i tako se dobija „lišće“. Za date podatke četiri broja, 33, 34, 35 i 36 čine „granu“, a prvih nekoliko „listova“ dobijaju se od podataka 33.1, 33.4, 34.8, 33.8 i 34.7

33	1 4 8
34	8 7
35	
36	

Podaci se čitaju s leva na desno i odozgo prema dole.

Ako bi se postupak nastavio na isti način, dobilo bi se previše „listova“ za ovako malu „granu“. Stoga se u ovakvim slučajevima dijagram formira „udvajanjem“ podataka na levoj strani od povučene linije i to na sledeći način: prva vrsta se formira za podatke od 33.0 do 33.4, druga vrsta za podatke od 33.5 do 33.9, itd.

Prvih pet „listova“ su dati na donjoj slici levo, a kompletan dijagram na slici desno. Na tom dijagramu se zatim slobodnom rukom nacrta glatka kriva koja obuhvata sve „lišće“. Ako se takav crtež rotira za 90 stepeni, dobija se (približno) oblik raspodele obeležja na uzorku.

→ *Nastavak na str. 263*

Tabela kontingencije

Табела контингенције

EN	Contingency table
FR	Table de contingence
RU	Таблица сопряженности
ZH	列联表, lièlián biǎo
AR	جدول المطابقة
EL	Πίνακας συνάφειας
DE	Kontingenztafel
ES	Tabla de contingencia
IT	Tavola di contingenza
SK	Kontingenčná tabuľka
SL	Kontingenčna tabela
HU	Kontingencia táblázat

Povezani termini:

Hi-kvadrat test (104), koeficijent korelacije (122), uzorak (240)

Etimologija:

Tabela < tabula[LA] (sto, ploča)

Kontingencija < contingere[LA] (dodirnuti, sresti) < co-[LA] (sa) + tangere[LA] (dodirnuti)

Neka se istovremeno posmatraju dva obeležja X i Y na elementima uzorka obima n. Kad se prebroji koliko puta se koji par vrednosti pojavio u uzorku, podaci iz uzorka se mogu dati u obliku sledeće tabele koja se naziva **tabela kontingencije** (vrednosti posmatranih obeležja mogu biti date i intervalno).

$X \setminus Y$	y_1	y_2	…	y_s
x_1	n_{11}	n_{12}	…	n_{1s}
x_2	n_{21}	n_{22}	…	n_{2s}
…	…	…	…	…
x_r	n_{r1}	n_{r2}	…	n_{rs}

Broj n_{ij} označava da se par vrednosti (x_i, y_j) pojavio n_{ij} puta u uzorku i pri tom je

$$n_{ij} \geq 0,\ i = 1, 2,..., r,\ j = 1, 2,..., s \text{ i važi } \sum_{i=1}^{r}\sum_{j=1}^{s} n_{ij} = n.$$

Vrednosti obeležja se, prema uobičajenom načinu pisanja, pišu u rastućem poretku, od najmanjih ka najvećim.

Na osnovu ovakve tabele se može uočiti da li postoji neki oblik zavisnosti posmatranih veličina (na primer, ako se duž dijagonale javljaju najveće vrednosti, onda to znači da sa porastom vrednosti jednog obeležja, dolazi do porasta vrednosti i drugog, i sl.)

Prilikom testiranja nezavisnosti obeležja X i Y primenom χ^2-testa iz tabele kontigencije se podaci koriste za izračunavanje realizovane vrednosti test-statistike.

Takođe, pomoću tabele kontingencije se izračunava koeficijent korelacije.

Tablično prikazivanje podataka

Таблично приказивање података

EN	Tabulation of data
FR	Tabulation des données
RU	Табуляция данных
ZH	资料制表, zīliào zhìbiǎo
AR	تبويب البيانات
EL	Πινακοποίηση δεδομένων
DE	Auftabellierung von Daten
ES	Tabulación de datos
IT	Tabulazione dei dati
SK	Tabelovanie údajov
SL	Tabelovanje podatkov
HU	Adatok táblázata

Povezani termini:

Frekvencija (80), uzorak (240)

Etimologija:

Tablica < tabula^{LA} (sto, ploča)

Postoje dva osnovna načina za **tablično predstavljanje podataka** dobijenih iz uzorka:

Podaci dati pojedinačnim brojnim vrednostima

Neka je dat realizovani uzorak obima n. Neka su, među dobijenim vrednostima, sa $x_1, x_2, \ldots, x_k$ označene različite vrednosti u rastućem poretku. Neka je n_1 puta dobijena vrednost x_1 u uzorku, n_2 puta dobijena vrednost $x_2, \ldots, n_k$ puta vrednost x_k, pri čemu je $n_1 + n_2 + \ldots + n_k = n$. Brojevi $n_1, n_2, \ldots, n_k$ se nazivaju (apsolutne) frekvencije ili *učestalosti*. Podatke iz uzorke unosimo u tabelu:

vrednost obeležja	$x_1 \quad x_2 \quad \ldots \quad x_k$
frekvencija	$n_1 \quad n_2 \quad \ldots \quad n_k$

U tabeli se mogu dati relativne frekvencije n_j/n, $j = 1, 2, \ldots, k$ umesto apsolutnih frekvencija.

Podaci dati po intervalima

Ako je n veliki broj, često se vrednosti obeležja zadaju po intervalima (klasama), mada se tako gubi jedan (znatan) deo informacije o posmatranom obeležju. Prvo se odrede najmanja i najveća vrednost u uzorku. Zatim se odaberu brojevi a_m (broj nešto manji od najmanje vrednosti u uzorku) i a_M (broj nešto veći od najveće vrednosti u uzorku). Zatim se interval $[a_m, a_M]$ podeli na k disjunktnih podintervala, čije dužine mogu, ali ne moraju, biti jednake: $[a_1, a_2), [a_2, a_3), \ldots, [a_k, a_{k+1})$, pri čemu je $a_1 = a_m$ i $a_{k+1} = a_M$.

Najčešće se uzima između 6 i 20 intervala (klasa). Neki autori predlažu da broj k klasa zadovoljava nejednačinu $k \leq 5 \log_{10} n$, a neki, pak, $k \approx 1 + 3.32 \log_{10} n$.

→ *Nastavak na str. 263*

Teorija verovatnoće

Теорија вероватноће

EN	Probability theory
FR	Théorie des probabilités
RU	Теория вероятностей
ZH	概率论, gàilǜlùn
AR	نظرية الاحتمالات
EL	Θεωρία πιθανοτήτων
DE	Wahrscheinlichkeitstheorie
ES	Teoría de la probabilidad
IT	Teoria della probabilità
SK	Teória pravdepodobnosti
SL	Verjetnostni račun
HU	Valószínűségszámítás

Povezani termini:

Matematička statistika (150)

Etimologija:

Teorija < θεωρία[EL] (razmišljanje) < θεωρέω[EL] (posmatrati) < θέα[EL] (pogled) + ὁράω[EL] (gledati)

Teorija verovatnoće proučava i objašnjava zakonitosti koje nastaju pri istovremenom uticaju velikog broja slučajnih faktora. Ova matematička disciplina je osnova matematičke statistike, teorije slučajnih procesa, teorije masovnog opsluživanja, teorije pouzdanosti, itd. Primenjuje se u raznim oblastima, kao što su: statistička fizika, geodezija (račun izravnanja), stohastička hidrologija, biologija (zakoni nasleđivanja), medicina, meteorologija (prognoziranje vremena), astronomija, demografija, ekonomija, itd.

Prvi radovi evropskih matematičara koji su sadržali osnovne ideje teorije verovatnoće pojavili su se u XVI i XVII veku i bili u vezi sa određivanjem šansi dobitka u igrama na sreću. Takve radove možemo naći kod Kardana, Paskala i Fermaa.

Đirolamo Kardano (1501–1576), italijanski matematičar, filozof i lekar. U mladosti se bavio isključivo medicinom, da bi kasnije postao profesor matematike u Bolonji i Milanu. Za Kardana se obično vezuju formule za određivanje korena jednačine trećeg stepena, koje je objavio 1545. godine, mada je taj rezultat prvi dobio Nikolo Tartalja, takođe italijanski matematičar tog doba. Kardano je napisao knjigu o igrama sa kockicama i sistematski obradio računanje verovatnoća događaja u tim igrama.

Blez Paskal (1623–1662), francuski matematičar, fizičar i filozof. U domu njegovih roditelja su se redovno sastajali matematičari i fizičari, a iz tih sastanaka je kasnije nastala Pariska akademija nauka. Pre Paskala niko od matematičara nije računao verovatnoće događaja na način na koji se to i sada radi. Među mnogim rezultatima koji se vezuju za Paskalovo ime je i poznati „Paskalov trougao“, šema u kojoj su zapisani binomni koeficijenti. U okviru evropske matematike to je bila značajna novost i važan rezultat. Interesantno je, međutim, da su Kinezi u jednoj knjizi objavljenoj prvih godina XIV veka (1303) već imali takvu šemu.

Pjer de Ferma (1601–1665), francuski pravnik i matematičar. Bavio se teorijom brojeva, geometrijom, algebrom i teorijom verovatnoće. Njegova prepiska sa Blezom Paskalom je osnov teorije verovatnoće.

Sledeća etapa se vezuje za Jakoba Bernulija. Teorema koju je on dokazao, a koja se danas naziva Bernulijev zakon velikih brojeva, predstavljala je teoretsku osnovu tada razmatranih činjenica. Značajan doprinos razvoju teorije verovatnoće dali su takođe Muavr, Laplas, Gaus, Puason, kao i Čebišev, Markov i mnogi drugi naučnici. Aksiomatsko zasnivanje teorije verovatnoće je rezultat radova Kolomogorova iz tridesetih godina XX veka.

Abraham de Muavr (1667–1754), engleski matematičar, rođen u Francuskoj. Dokazao je važnu teoremu iz teorije verovatnoće. U teoriju verovatnoće je uveo pojam slučajnog događaja. Osim teorije verovatnoće bavio se teorijom redova i teorijom kompleksnih brojeva.

Pjer-Simon Laplas (1749–1827), francuski matematičar, fizičar i astronom. Učio je u školi benediktinskog monaškog reda, ali je kasnije postao ateista. Njegovi rezultati su fundamantalni u matematici, eksperimentalnoj i matematičkoj fizici i astronomiji. Razvio je i sistematizovao rezultate Paskala, Fermaa, Bernulija, dokazao teoremu koja se naziva Muavr-Laplasova teorema. Uveo je pojam matematičkog očekivanja. Njegova knjiga „Analitička teorija verovatnoće“ je još za njegovog života imala tri izdanja.

Karl Fridrih Gaus (1777–1855), nemački matematičar i astronom. U radu iz 1809. godine razvio je metodu najmanjih kvadrata. Bio je direktor opservatorije u Getingenu. U svom radu je ustanovio da zbir velikog broja malih grešaka ima specifičnu raspodelu, koja se danas naziva Gausova raspodela.

Simeon Deni Puason (1781–1840), francuski matematičar koji je objavljivao radove iz raznih oblasti matematike. Bio je profesor matematike na visokim školama u Parizu. U knjizi koja se odnosila na analiziranje sudskih presuda uveo je raspodelu koja danas nosi njegovo ime. Inače, pripisuje mu se izreka „Život je vredan samo zbog dve stvari: učenja i podučavanja matematike“.

U teoriji verovatnoće polazna tačka je prostor elementarnih ishoda, kao što je to u statistici populacija. Elementarnim ishodima odgovaraju elementi populacije. Verovatnoći događaja odgovara, ako se eksperimenti izvode nezavisno jedan od drugog i pod isptim uslovima, relativna frekvencija. Raspodelama slučajnih veličina odgovaraju raspodele obeležja, numeričkim karakteristikama slučajnih veličina analogne su odgovarajuće karakteristike u raspodeli obeležja koje posmatramo na datoj populaciji.

Nezavisnost, korelacija, granične teoreme, sve to i mnogo drugog što se proučava u teoriji verovatnoće ima svoj analogon u matematičkoj statistici.

Dokazi ispravnosti statističkih procedura se daju metodama teorije verovatnoće.

Testiranje hipoteze

Тестирање хипотезе

EN Testing of hypothesis
FR Test d’hypothèse
RU Проверка гипотезы
ZH 假设检验, jiǎshè jiǎnyàn
AR اختبار الفرضيات
EL Έλεγχος υπόθεσης
DE Hypothesentest
ES Test de hipótesis
IT Test di verifica d’ipotesi
SK Testovanie hypotéz
SL Preizkušanje domnev
HU Hipotézisvizsgálat

Povezani termini:

Interval poverenja (114), prag značajnosti (186), statistički test

Etimologija:

Test < test[stFR] (zemljani kotao, obično onaj u kome su isprobavani metali) < testum[LA] (zemljana posuda)

Hipoteza < ὑπόθεση[EL] (staviti ispod, pretpostaviti) < ὑπό[EL] (ispod) + τίθημι[EL] (postaviti)

Testiranje statističkih hipoteza znači da se primenom statističkih metoda utvrđuje da li se, na osnovu uzorka, može, i sa kojom verovatnoćom, prihvatiti/odbaciti pretpostavka o konkretnoj brojčanoj vrednosti nekog parametra u raspodeli posmatranog obeležja ili pak o konkretnoj raspodeli obeležja.

Ako je hipoteza o vrednosti parametra, onda se postupak njenog potvrđivanja ili odbacivanja na osnovu podataka iz uzorka naziva parametarski test. Statistika koja se koristi u tom postupku se naziva test-statistika. Testiranje hipoteza o parametrima i određivanje intervala poverenja za te parametre su zadaci u kojima se koriste iste statistike.

Pri testiranju hipoteza o očekivanju često se koristi uzoračka sredina, pri testiranju hipoteza o disperziji – uzoračka disperzija... tj. koriste se statistike kojima se posmatrani parametar može ocenjivati. Raspodela statistika koje se koriste ovde zavisi od raspodele posmatrnog obeležja.

S druge strane postoje testovi (npr. hi-kvadrat test) u kojima test-statistike imaju raspodelu koja ne zavisi od raspodele obeležja, pa se za te testove koristi termin testovi slobodni od raspodele.

Hipoteza se može prihvatiti ili odbaciti. Umesto *hipoteza se odbacuje* često se kaže *nema razloga za prihvatanje hipoteze*, a umesto *hipoteza se prihvata* kaže se *nema razloga da se hipoteza odbaci*. U svakom slučaju važno je da se vodi računa o tome da je odluka doneta na osnovu datog uzorka i za dati prag značajnosti. Za drugi uzorak i/ili drugi prag značajnosti, odluka može biti i suprotna. Stoga uvek podrazumevamo da iskaz hipoteza *se odbacuje* znači da se hipoteza odbacuje za dati uzorak i dati prag značajnosti.

Testiranje hipoteze o očekivanju

Тестирање хипотезе о очекивању

EN	Hypothesis test for mathematical expectation
FR	Test d'hypothèse de l'espérance mathématique
RU	Проверка гипотезы о математическом ожидании
ZH	数学期望的假设检验, shùxué qīwàng de jiǎshè jiǎnyàn
AR	إختبار الفرضيات لتوقع الرياضي
EL	Έλεγχος υπόθεσης της μαθηματικής προσδοκίας
DE	Hypothesentest für den Erwartungswert
ES	Test de hipótesis sobre de la esperanza matemática
IT	Test di verifica d'ipotesi sulla speranza matematica
SK	Testovanie hypotézy o strednej hodnote
SL	Preizkušanje domneve o pričakovani vrednosti
HU	A várható értékre vonatkozó hipotézisvizsgálat

Povezani termini:

Centralna granična teorema (46), disperzija (54), matematičko očekivanje (152), normalna raspodela (172), statistički test, Studentova raspodela (45)

Etimologija:

Test < test^{stFR} (zemljani kotao, obično onaj u kome su isprobavani metali) < testum^{LA} (zemljana posuda)

Hipoteza < ὑπόθεση^{EL} (staviti ispod, pretpostaviti) < ὑπό^{EL} (ispod) + τίθημι^{EL} (postaviti)

Matematika < μαθηματικός^{EL} (onaj ko voli da uči) < μάθημα^{EL} (znanje, učenje, nauka)

A) **Testiranje hipoteze** $H_0(m = m_0)$ o matematičkom očekivanju obeležja X koje ima normalnu raspodelu $N(m, \sigma^2)$, ako je disperzija σ^2 poznata i ako je alternativna hipoteza $H_1(m \neq m_0)$ podrazumeva da se pomoću tablica za normalnu raspodelu nađe kritična vrednost

$$c = F^{-1}\left(\frac{1-\alpha}{2}\right),$$

gde F označava funkciju raspodele slučajne veličine sa raspodelom $N(0,1)$. Zatim se računa ε iz $c = (\varepsilon/\sigma)\sqrt{n}$. Ako je $|\bar{x}_n - m_0| \geq \varepsilon$, odbacuje se nulta hipoteza. Tu je $\bar{x}_n$ je realizovana vrednost statistike $\bar{X}_n$. Znači, ako $\bar{x}_n \in W$, gde je kritična oblast oblika

$$W = (-\infty, m_0 - \varepsilon] \cup [m_0 + \varepsilon, +\infty),$$

nulta hipoteza se odbacuje, a prihvata alternativna hipoteza. Ako $\bar{x}_n \notin W$, nulta hipoteza se prihvata.

B) **Testiranje hipoteze** $H_0(m = m_0)$ o matematičkom očekivanju obeležja X koje ima normalnu raspodelu $N(m, \sigma^2)$, ako disperzija σ^2 nije poznata u slučaju da je $H_1(m \neq m_0)$ podrazumeva da se nulta hipoteza odbacuje ako je $|\bar{x}_n - m_0| \geq \varepsilon$. Ako je $|\bar{x}_n - m_0| < \varepsilon$, nulta hipoteza se prihvata.

Veličina ε se određuje iz

$$P_{H_0}\left\{ \left|\bar{X}_n - m_0\right| \geq \varepsilon \right\} = P_{H_0}\left\{ \frac{\left|\bar{X}_n - m_0\right|}{\bar{S}_n}\sqrt{n-1} \geq \frac{\varepsilon}{\bar{S}_n}\sqrt{n-1} \right\} = P\left\{ |T| \geq c \right\} = \alpha$$

Koristeći tablice za Studentovu raspodelu sa n–1 stepeni slobode, dobija se $c = t_{n-1;\alpha}$, a zatim se računa ε.

C) Ako raspodela obeležja nije normalna, ali je uzorak velikog obima, može se primeniti postupak pod B) na osnovu centralne granične teoreme.

Uniformno najmoćniji test

Униформно најмоћнији тест

EN	Uniformly most powerful test
FR	Test uniformément le plus puissant
RU	Равномерно найболее мощный критерий
ZH	一致最大功效检验, yízhì zuìdà gōngxiào jiǎnyàn
AR	اختبار الأكثر قوي بأنتظام
EL	Ομοιόμορφα ισχυρότερος έλεγχος
DE	Gleichmäßig trennschärfster Test
ES	Prueba uniformemente más poderosa
IT	Test uniformemente il più potente
SK	Rovnomerne najsilnejší test
SL	Enakomerno najmočnejši preizkus
HU	Egyenletesen legerősebb próba

Povezani termini:

Moć testa (86), testiranje hipoteze (216)

Etimologija:

Uniforman < ūnus^{LA} (jedan) + fōrma^{LA} (oblik)

Test < test^{stFR} (zemljani kotao, obično onaj u kome su isprobavani metali) < testum^{LA} (zemljana posuda)

Veliki broj različitih testova nameće potrebu za njihovim poređenjem. Jedan od mogućih pokazatelja kvaliteta statističkog testa se naziva moć testa i predstavlja verovatnoću donošenja pravilne odluke o odbacivanju nulte hipoteze, kad ona zaista nije tačna. Poželjno je da ta verovatnoća bude što veća, a svakako veća od praga značajnosti – verovatnoće donošenja odluke o odbacivanju nulte hipoteze kad ona jeste tačna.

U slučaju testiranja proste parametarske hipoteze H_0 protiv proste alternativne hipoteze H_1, Nejman-Pirsonova lema utvrđuje postojanje najmoćnijeg testa sa pragom značajnosti α. Kritična oblast tog testa se dobija na osnovu količnika verodostojnosti L_1/L_0, gde indeks označava da li se koristi parametar prema nultoj ili prema alternativnoj hipotezi. Test će biti najmoćniji u smislu da će njegova moć biti veća ili jednaka moći svakog drugog testa kod koga je veličina kritične oblasti manja ili jednaka α.

Uniformno najmoćniji test je pojam koji se javlja u slučaju složene alternativne hipoteze, i predstavlja test moćniji od testova pri bilo kojoj vrednosti parametra koje su dopustive u okviru alternativne hipoteze. Ako je pri tom nulta hipoteza prosta, uniformno najmoćniji test je onaj za koji oblik kritične oblasti ne zavisi od vrednosti parametra u alternativnoj hipotezi, pa se kaže da je test nepristrasan. Ako je nulta hipoteza složena, postojanje uniformno najmoćnijeg testa se utvrđuje za klasu raspodela koje imaju monoton količnik verodostojnosti u odnosu na neku funkciju uzorka. Ukoliko je raspodela obeležja iz klase eksponencijalne familije raspodela i funkcija $A(\theta)$ u toj reprezentaciji monotona funkcija, onda će količnik verodostojnosti biti monotona funkcija u odnosu na dovoljnu statistiku $\Sigma K(X_j)$, i postojaće uniformno najmoćniji test.

Uslovna verovatnoća

Условна вероватноћа

EN Conditional probability
FR Probabilité conditionnelle
RU Условная вероятность
ZH 条件概率, tiáojiàn gàilǜ
AR الاحتمال الشرطي
EL Δεσμευμένη πιθανότητα
DE Bedingte Wahrscheinlichkeit
ES Probabilidad condicional
IT Probabilità condizionale
SK Podmienená pravdepodobnosť
SL Pogojna verjetnost
HU Feltételes valószínűség

Povezani termini:

Obim populacije (182), relativna frekvencija (80), uzorak (240), verovatnoća (248)

Neka su A i B događaji iz istog prostora verovatnoća (Ω, F, P) i neka je $P(B) > 0$. Tada je **uslovna verovatnoća** događaja A, ako se ostvario događaj B, jednaka

$$P(A|B) = P(AB)/P(B).$$

Za uslovnu verovatnoću koristi se i oznaka $P_B(A)$.

Uslovne verovatnoće događaja iz istog prostora verovatnoća, u odnosu na neki događaj iz tog prostora, imaju sve osobine verovatnoće: uslovna verovatnoća je nenegativna (za svaki događaj A je $P(A|B) \geq 0$), normirana (za siguran događaj je $P(\Omega|B) = 1$) i aditivna funkcija, tj. za disjunktne događaje važi:

$$P\left[\left(\sum_j A_j\right) | B\right] = \sum_j P(A_j | B).$$

I ostale osobine verovatnoće se prenose na uslovne verovatnoće (npr. zbir uslovne verovatnoće nekog događaja i uslovne verovatnoće njegovog komplementa je jednak 1).

Kako verovatnoće događaja ocenjujemo njihovim relativnim frekvencijama, to će ocena verovatnoće $P(A|B)$ biti $(n_{AB}/N)/(n_B/N) = n_{AB}/n_B$, gde je N ukupan broj izvršenih eksperimenata (ili obim (konačne) populacije), n_B broj eksperimenata u kojima se desio događaj B (obim podpopulacije koja ima svojstvo „B“), n_{AB} broj eksperimenata u kojima su se desili i A i B (obim potpopulacije čiji elementi imaju i svojstvo „A“ i svojstvo „B“).

Po analogiji sa formulom uslovne verovatnoće se definiše i uslovna raspodela Y, $f(y|x)$, u odnosu na X kao količnik zajedničke i marginalne raspodele, a pomoću uslovne raspodele se definiše uslovno matematičko očekivanje:

$$E(Y | X = x) = \int_{-\infty}^{\infty} y f(y | x) dy,$$

u neprekidnom slučaju.

Uzoračka disperzija

Узорачка дисперзија

EN Sample variance
FR Variance de l’échantillon
RU Выборочная дисперсия
ZH 样本方差, yàngběn fāngchā
AR تباين العينة
EL Διασπορά δείγματος
DE Streuung einer Stichprobe
ES Dispersión de la muestra
IT Dispersione campionaria
SK Výberový rozptyl
SL Vzorčna disperzija
HU Tapasztalati szórás

Povezani termini:

Tabelarno prikazivanje podataka (210), uzoračka sredina (230), uzorak (240)

Etimologija:

Disperzija < dis-^{LA} (od, odvojeno) + spargere^{LA} (rasipati)

Neka je $(X_1, X_2,..., X_n)$ prost slučajni uzorak obima n za posmatrano obeležje X. Ukoliko je poznato matematičko očekivanje za obeležje X, $E(X) = m$, tada je **uzoračka disperzija** statistika

$$\hat{S}_n^2 = \frac{1}{n}\left(\left(X_1 - m\right)^2 + \ldots + \left(X_n - m\right)^2\right) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - m\right)^2 \qquad (1)$$

Ako matematičko očekivanje obeležja nije poznato, tada je

$$\bar{S}_n^2 = \frac{1}{n}\left(\left(X_1 - \bar{X}_n\right)^2 + \ldots + \left(X_n - \bar{X}_n\right)^2\right) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \qquad (2)$$

uzoračka disperzija, a popravljena uzoračka disperzija je

$$\tilde{S}_n^2 = \frac{1}{n-1}\left(\left(X_1 - \bar{X}_n\right)^2 + \ldots + \left(X_n - \bar{X}_n\right)^2\right) = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \qquad (3)$$

Važno: Ako matematičko očekivanje nije poznato, tada izraz u jednakosti (1) nije statistika i ne može se koristiti.

Uzoračka disperzija iz (2) se može računati i po formuli:

$$\bar{S}_n^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}_n^2$$

Veza uzoračke disperzije i popravljene uzoračke disperzije je

$$\tilde{S}_n^2 = \frac{n}{n-1}\bar{S}_n^2 .$$

Pri izračunavanju realizovane vrednosti ($\hat{s}_n^2$, $\bar{s}_n^2$, $\tilde{s}_n^2$) uzoračke disperzije postupa se slično kao pri izračunavanju realizovane vrednosti uzoračke sredine, a u zavisnosti od načina na koji su podaci zadati.

→ *Nastavak na str. 264*

Uzoračka medijana

Узорачка медијана

EN	Sample median
FR	Médiane de l’échantillon
RU	Выборочная медиана
ZH	样本中位数, yàngběn zhōngwèishù
AR	وسيط العينة
EL	Διάμεσος δείγματος
DE	Median einer Stichprobe
ES	Mediana de la muestra
IT	Mediana campionaria
SK	Výberový medián
SL	Vzorčna mediana
HU	Minta médian

Povezani termini:

Histogram (108), uzorak (240)

Etimologija:

Medijana < mediānus^{LA} (onaj koji pripada sredini) < medius^{LA} (srednji). Latinska reč *medius* i srpska reč *međa* imaju zajednički koren u proto-indoevropskom *med^{h}yos (srednji).

Medijana uzorka se ako su podaci dati kao različite pojedinačne vrednosti obeležja x_i sa frekvencijama n_i, $i = 1, 2, ..., m$ puta, dobija tako što se prvo napiše varijacioni niz tj. elementi uzorka se poređaju u neopadajući niz. Varijacioni niz je onaj u kome se svaka vrednost x_i obeležja X ponavlja onoliko puta kolika je odgovarajuća frekvencija n_i. Neka su y_1, y_2,..., y_n elementi dobijenog varijacionog niza. Važi:

$$y_1 \leq y_2 \leq ... \leq y_n .$$

Tada je uzoračka medijana m_e jednaka

$$m_e = \begin{cases} y_{k+1}, & ako \quad je \quad n = 2k+1 \\ \frac{1}{2}(y_k + y_{k+1}), & ako \quad je \quad n = 2k \end{cases} .$$

Ako su podaci iz uzorka sređeni po intervalima, tako da se vrednosti obeležja u intervalu $[a_i, a_{i+1})$ pojavljuju n_i, $i = 1, 2,..., m$ puta u uzorku, pri čemu su dužine intervala jednake c, a medijana se nalazi u intervalu $[a_i, a_{i+1})$, tada je uzoračka medijana jednaka

$$m_e = a_i + \left(\frac{n}{2} - \sum_{j=1}^{i-1} n_j \right) \cdot \frac{c}{n_i} .$$

Ovim je dat izraz za realizovanu vrednost medijane uzorka.

Na histogramu je medijana ona tačka m_e za koju prava $x = m_e$ deli histogram na delove jednakih površina.

Takođe, važi da je medijana uvek između srednje vrednosti i mode.

Bitna osobina medijane je njena robusnost. Pod ovim se podrazumeva da izmena ekstremnih vrednosti u uzorku ne utiče na vrednost medijane, za razliku od uzoračke sredine koja je osetljiva na takve izmene[90].

[90] U uzorcima: 1,1,2,2,4,5,6,7,8 i 1,2,2,3,4,5,6,8,17 medijana je ista i jednaka 4, dok je uzoračka sredina u prvom slučaju 4, u drugom 5.

Uzoračka modalna vrednost

Узорачка модална вредност

EN Sample mode
FR Mode de l'échantillon
RU Мода выборки
ZH 样本众数, yàngběn zhòngshù
AR منوال العينة
EL Δειγματική κορυφή
DE Modus einer Stichprobe
ES Moda de la muestra
IT Moda campionaria
SK Výberový modus
SL Vzorčni modus
HU Minta módusz

Povezani termini:

Medijana (154), očekivanje (152), tabelarno predstavljanje podataka (210), uzorak (240)

Etimologija:

Modalni < modus[LA] (mera, ograničenje)

Drugi naziv:

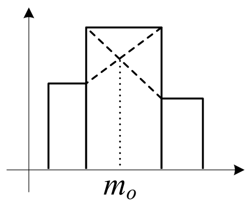
Moda uzorka

Moda uzorka je, ako su podaci dati kao različite pojedinačne vrednosti obeležja x_i sa frekvencijama n_i, $i = 1, 2,..., m$ puta, svaka vrednost x_i obeležja X za čiju odgovarajuću frekvenciju n_i važi: $n_i > n_{i-1}$ i $n_i > n_{i+1}$.

Ako su podaci iz uzorka sređeni po intervalima, tako da se vrednosti obeležja u intervalu $[a_i, a_{i+1})$ pojavljuju n_i, $i = 1, 2,..., m$ puta u uzorku, moda se nalazi u intervalu $[a_i, a_{i+1})$ koji se naziva modalna klasa, i za koji važi: $n_i > n_{i-1}$ i $n_i > n_{i+1}$. Moda uzorka m_o je, uz oznake $\delta = n_i - n_{i-1}$, $\Delta = n_i - n_{i+1}$, $c = a_{i+1} - a_i$:

$$m_o = a_i + (\partial c/\partial+\Delta).$$

Ovim je dat izraz za realizovanu vrednost ove statistike. Grafički je jednostavno dobiti modu uzorka u okviru modalne klase (pravougaonik čija je visina veća od visine susednih pravougaonika na histogramu), kao projekciju presečne tačke duži (isprekidana linija) na osu Ox:



Ako postoji samo jedna moda, tada se radi o unimodalnoj raspodeli obeležja, a ako ima više od jedne mode reč je o polimodalnoj raspodeli. Specijalno, raspodela sa dve mode se naziva bimodalna raspodela.

Kod simetričnih unimodalnih jednodimenzionalnih raspodela se poklapaju matematičko očekivanje m, moda m_o i medijana m_e. Za asimetrične (tj. nesimetrične) unimodalne raspodele je, približno, rastojanje mode i medijane dvostruko veće od rastojanja očekivane vrednosti i medijane, pa važi $|m-m_o| = 3|m-m_e|$, što može da posluži kao brza, gruba procena jedne od odgovarajućih uzoračkih veličina kada se znaju preostale dve.

Uzoračka sredina

Узорачка средина

EN	Sample mean
FR	Moyenne de l’échantillon
RU	Выборочное среднее
ZH	样本平均数, yàngběn píngjūnshù
AR	متوسط العينة
EL	Δειγματικός μέσος
DE	Stichprobenmittelwert
ES	Media de la muestra
IT	Media campionaria
SK	Výberový priemer
SL	Vzorčna sredina
HU	Mintaközép

Povezani termini:

Aritmetička sredina (16), nepristrasnost (164), postojanost, uzorak (240), Neka je $(X_1, X_2,..., X_n)$ prost slučajni uzorak obima n za posmatrano obeležje X. **Uzoračka sredina** je statistika

$$\overline{X}_n = \frac{1}{n}\left(X_1 + X_2 + ... + X_n\right) = \frac{1}{n}\sum_{i=1}^{n} X_i .$$

Ako su podaci u uzorku, tzv. realizovani uzorak, dati kao niz vrednosti $x_1, x_2,..., x_n$, tada je realizovana vrednost uzoračke sredine jednaka

$$\overline{x}_n = \frac{1}{n}\left(x_1 + x_2 + ... + x_n\right) = \frac{1}{n}\sum_{i=1}^{n} x_i . \quad (1)$$

Na osnovu definicije prostog slučajnog uzorka i osobina matematičkog očekivanja sledi da je matematičko očekivanje uzoračke sredine jednako matematičkom očekivanju obeležja, tj. $E(\overline{X}_n) = E(X) = m$.

Veza disperzije uzoračke sredine i disperzije σ^2 obeležja X u populaciji je

$$D\left(\overline{X}_n\right) = \frac{1}{n}D\left(X\right) = \frac{\sigma^2}{n} ,$$

što znači da je uzoračka sredina postojana ocena matematičkog očekivanja obeležja.

Ako su podaci dati kao različite pojedinačne vrednosti obeležja x_i sa frekvencijama n_i, $i = 1, 2,...,m$ puta, tada se realizovana vrednost statistike $\overline{X}_n$ računa po formuli:

$$\overline{x}_n = \frac{1}{n}\left(n_1 x_1 + n_2 x_2 + ... + n_m x_m\right) = \frac{1}{n}\sum_{i=1}^{m} n_i x_i . \quad (2)$$

Ako su podaci iz uzorka sređeni po intervalima, tako da se vrednosti obeležja u intervalu $[a_i, a_{i+1})$ pojavljuju n_i, $i = 1, 2, ..., m$ puta u uzorku.

→ *Nastavak na str. 264*

Uzorački koeficijent korelacije

Узорачки коефицијент корелације

EN Sample coefficient of correlation
FR Coefficient de corrélation de l’échantillon
RU Выборочный коэффициент корреляции
ZH 样本相关系数, yàngběn xiāngguān xìshù
AR معامل الارتباط العينة
EL Δειγματικός συντελεστής συσχέτισης
DE Korrelationskoeffizient einer Stichprobe
ES Coeficiente de correlación de la muestra
IT Coefficiente di correlazione campionaria
SK Výberový korelačný koeficient
SL Vzorčni koeficient korelacije
HU Mintabeli korrelációs együttható

Povezani termini:

Korelacija, uzoračka disperzija (224), uzoračka srednja vrednost

Etimologija:

Koeficijent < co-^{LA} (sa) + efficiēns^{LA} (onaj koji izvršava) < ē^{LA} (iz) + faciō^{LA} (raditi); ovaj termin je prvi upotrebio francuski matematičar Fransoa Vijet (1540–1603), otac moderne algebre.

Korelacija < correlatio^{LA} (međusobna povezanost) < co-^{LA} (sa) + relatio^{LA} (odnos) < referō^{LA} (nositi nazad) < re^{LA} (ponovo) + ferre^{LA} (nositi)

Drugi nazivi:

U literaturi na engleskom sreću se i nazivi *Pearson’s correlation coefficient* i *product-moment correlation coefficient*

Neka je dat prost slučajni uzorak obima n pri čemu se istovremeno posmatraju dva obeležja X i Y na elementima uzorka. Posle sređivanja, podaci iz uzorka se mogu dati u obliku sledeće tabele koja se naziva korelaciona tabela. Vrednosti posmatranih obeležja mogu biti date i intervalno, a mogu biti date i opisno.

$X \setminus Y$	y_1	y_2	…	y_s	*zbir*
x_1	n_{11}	n_{12}	…	y_{1s}	$n(x_1)$
x_2	n_{21}	n_{22}	…	n_{2s}	$n(x_2)$
…	…	…	…	…	…
x_r	n_{r1}	n_{r2}	…	n_{rs}	$n(x_r)$
zbir	$n(y_1)$	$n(y_2)$	…	$n(y_s)$	n

Broj n_{ij} označava da se par vrednosti (x_i, y_j) pojavio n_{ij} puta u uzorku, i pri tom je $n_{ij} \geq 0$, $i = 1, 2,..., r$, $j = 1, 2,..., s$ i važi $\sum_{i=1}^{r}\sum_{j=1}^{s} n_{ij} = n$.

Neka su marginalni zbirovi po vrstama (po vrednostima obeležja X) jednaki $n(x_1)$, $n(x_2)$, ..., $n(x_r)$, a po kolonama (po vrednostima obeležja Y) $n(y_1)$, $n(y_2)$, ..., $n(y_s)$.

Kada su podaci dati u obliku korelacione tabele, uzoračke sredine obeležja X i Y se računaju na osnovu sledećih formula:

$$\overline{x}_n = \frac{1}{n}\sum_{i=1}^{r} n(x_i)\cdot x_i\ ,\ \overline{y}_n = \frac{1}{n}\sum_{j=1}^{s} n(y_j)\cdot y_j\ .$$

Uzoračke disperzije za obeležja X i Y se računaju po formulama:

$$\overline{s}_X^2 = \frac{1}{n}\sum_{i=1}^{r} n(x_i)\left(x_i - \overline{x}_n\right)^2\ ,\ \overline{s}_Y^2 = \frac{1}{n}\sum_{i=1}^{s} n(y_i)\left(y_i - \overline{y}_n\right)^2\ .$$

U slučaju kad se posmatraju dva obeležja X i Y, jedan od važnih podataka je stepen njihove međusobne zavisnosti. Ta zavisnost se „meri“ uzoračkim koeficijentom korelacije. Kada su podaci dati u obliku korelacione tabele, tada se uzorački koeficijent korelacije računa po formuli:

$$r_{X,Y} = \frac{\frac{1}{n}\sum_{i=1}^{r}\sum_{j=1}^{s} n_{ij} x_i y_j - \overline{x}_n \overline{y}_n}{\sqrt{\overline{s}_X^2 \cdot \overline{s}_Y^2}}. \qquad (1)$$

Ako se posmatraju dva obeležja X i Y, i njihove vrednosti redom beleže na elementima uzorka, dobija se sledeća tabela njihovih vrednosti:

vrednosti za X	x_1	x_2	…	x_n
vrednosti za Y	y_1	y_2	…	y_n

U slučaju kada su podaci dati u obliku takve tabele, uzoračke sredine obeležja X i Y se računaju na osnovu uobičajenih formula:

$$\overline{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \overline{y}_n = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

Uzoračke disperzije za obeležja X i Y se tada računaju takođe po uobičajenim formulama:

$$\overline{s}_X^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}_n\right)^2, \quad \overline{s}_Y^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \overline{y}_n\right)^2.$$

Uzorački koeficijent korelacije je

$$r_{X,Y} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x}_n \overline{y}_n}{\sqrt{\overline{s}_X^2 \cdot \overline{s}_Y^2}}. \qquad (2)$$

Formulama (1) i (2) date su realizovane vrednosti uzoračkog koeficijenta korelacije koji se, na osnovu dvodimenzionalnog prostog slučajnog uzorka (X_i, Y_i), $i = 1,\ldots, n$, računa po formuli

$$R_{X,Y} = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i Y_i - \overline{X}_n \overline{Y}_n}{\sqrt{\overline{S}_X^2 \cdot \overline{S}_Y^2}} = \frac{\frac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X}_n)(Y_i - \overline{Y}_n)}{\sqrt{\overline{S}_X^2 \cdot \overline{S}_Y^2}}. \quad (3)$$

Brojioci u izrazima (1), (2) i (3) predstavljaju uzoračku kovarijansu.

Osobine uzoračkog koeficijenta korelacije su slične osobinama koeficijenta korelacije dve slučajne veličine. Naime, i za uzorački koeficijent korelacije važi da su njegove vrednosti u intervalu [–1,1]. Što je koeficijent korelacije po apsolutnoj vrednosti bliži jedinici, to je veća linearna povezanost posmatranih obeležja.

Ukoliko povećanje vrednosti jednog obeležja prati povećanje vrednosti drugog, kaže se da su obeležja pozitivno korelisana (*positive correlation*), a tada je koeficijent korelacije pozitivan. Ukoliko pak povećanje vrednosti jednog obeležja prati smanjivanje vrednosti drugog, tada su obeležja negativno korelisana (*negative correlation*), a koeficijent korelacije negativan. Formule za uzorački koeficijent korelacije „imitiraju" formule za koeficijent korelacije, koje su, dajemo ih radi podsećanja:

$$\rho_{X,Y} = \frac{E(X-EX)(Y-EY)}{\sqrt{DX \cdot DY}}, \text{ ili } \rho_{X,Y} = \frac{E(XY) - EX \cdot EY}{\sqrt{DX \cdot DY}}.$$

Uzorački kvantili

Узорачки квантили

EN	Sample quantiles
FR	Quantiles de l’échantillon
RU	Выборочные квантили
ZH	样本分位数, yàngběn fēnwèishù
AR	تجزئ العينة
EL	Ποσοστημόρια του δείγματος
DE	Stichprobenquantilen
ES	Cuantiles de una muestra
IT	Quantili campionari
SK	Výberové kvantily
SL	Vzorčni kvantili
HU	Minta kvantilis

Povezani termini:

Empirijska funkcija raspodele (72), medijana (154), uzorak (240)

Etimologija:

Kvantil < quantitas$^{\mathrm{LA}}$ (količina) < quantus$^{\mathrm{LA}}$ (koliko).

Uzorački *p*-procentni kvantil (kvantil reda p) je vrednost obeležja (broj) koja je veća ili jednaka od p% elemenata iz uzorka.

Specijalno, uzorački 25%-ni kvantil se naziva prvi uzorački kvartil (*lower quartile*) i označava se sa q_1, uzorački 50%-ni kvantil se poklapa sa uzoračkom medijanom, a uzorački 75%-ni kvantil se naziva treći uzorački kvartil (*upper quartile*) i označava se sa q_3. Razlika trećeg i prvog kvartila se naziva kvartilni raspon ili interkvartilni raspon.

Prvi kvartil predstavlja medijanu donje, a treći kvartil medijanu gornje polovine podataka iz uzorka kad se oni poređaju u varijacioni niz.

Ako je broj podataka u uzorku mali neposredno se određuju kvantili i/ili kvartili posle formiranja varijacionog niza. Ako je uzorak većeg obima mogu se kvantili i/ili kvartili i približno odrediti (očitati) sa grafika empirijske funkcije raspodele.

Uzorački momenti

Узорачки моменти

EN	Sample moments
FR	Moments de l’échantillon
RU	Выборочные моменты
ZH	样本矩, yàngběn jǔ
AR	عزوم العينة
EL	Ροπές του δείγματος
DE	Momenten der Stichprobe
ES	Momentos de la muestra
IT	Momenti campionari
SK	Výberové momenty
SL	Vzorčni momenti
HU	A minta momentumai

Povezani termini:

Koeficijent asimetrije (120), koeficijent korelacije (122), koeficijent spljoštenosti (124), nepristrasnost (164), tablično prikazivanje podataka (210)

Etimologija:

Momenat < mōmentum^{LA} (pokret) < movēre^{LA} (kretati se)

Obični uzorački moment reda *k* je statistika

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k ,$$

a njena realizovana vrednost jednaka je

$$a_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k \quad \text{(kad su podaci dati bez sređivanja), odn.}$$

$$a_k = \frac{1}{n}\sum_{i=1}^{m} n_i x_i^k \quad \text{(kad su podaci dati tabelarno).}$$

Centralni uzorački moment reda *k* je statistika

$$C_k = \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \overline{X}_n \right)^k ,$$

a njena realizovana vrednost u slučaju kada su podaci dati bez sređivanja, jednaka

$$c_k = \frac{1}{n}\sum_{i=1}^{n} \left(x_i - \overline{x}_n \right)^k ,$$

a u slučaju kada su podaci dati tabelarno je

$$c_k = \frac{1}{n}\sum_{i=1}^{m} n_i \left(x_i - \overline{x}_n \right)^k .$$

Prema tome, uzoračka sredina predstavlja uzorački moment prvog reda, a uzoračka disperzija centralni uzorački moment drugog reda.

Obični i centralni uzorački momenti prvog, drugog, trećeg i četvrtog reda se koriste i prilikom izračunavanja sledećih veličina: uzoračka kovarijansa, uzorački koeficijent korelacije, uzorački koeficijent asimetrije, uzorački koeficijent spljoštenosti, i dr.

Obični uzorački momenti su nepristrasne ocene odgovarajućih momenata u raspodeli posmatranog obeležja, dok su centralni momenti asimptotski nepristrasne ocene odgovarajućih centralnih momenata u raspodeli posmatranog obeležja.

Uzorak
Узорак

EN	Sample
FR	Echantillon
RU	Выборка
ZH	样本, yàngběn
AR	العينة
EL	Δείγμα
DE	Stichprobe
ES	Muestra
IT	Campione
SK	Výber
SL	Vzorec
HU	Minta

Povezani termini:

Grupni uzorak (243), izbor bez vraćanja, izbor sa vraćanjem, periodični uzorak (244), populacija (182), reprezentativni uzorak (196), stratifikovani uzorak (242)

Radi rešavanja zadataka u statistici se prikupljaju podaci o populaciji. Podaci se mogu prikupljati iz cele populacije i tako izučavati populacija u celini. Međutim, ako populacija sadrži veliki broj elemenata, tada izučavanje cele populacije može dugo trajati ili prouzrokovati veće materijalne troškove. A u nekim je slučajevima proučavanje cele populacije principijelno nemoguće. Stoga se iz cele populacije (na slučajan način) uzima jedan deo koji se dalje proučava. Taj deo se naziva **uzorak**. Broj elemenata u uzorku je obim uzorka. Uzorak mora biti reprezentativan[91], a njegov obim se određuje po izvesnim pravilima.

Postupci za obradu podataka iz uzorka su povezani sa vrstom uzorka i načinom njegovog dobijanja. Postoje razne vrste uzoraka: stratifikovani, periodični, grupni... i razne metode dobijanja uzoraka: uzorak sa vraćanjem, uzorak bez vraćanja, i dr.

Razlikuju se metode koje se primenjuju na tzv. male uzorke (*small sample,* čiji je obim najviše 30) i za velike uzorke (*large sample*, obima većeg od 30). Za male uzorke je potrebno znati tačne raspodele slučajnih veličina koje se koriste, dok za uzorke velikog obima tačne raspodele se mogu aproksimirati drugim raspodelama (na primer, normalnom raspodelom – na osnovu centralne granične teoreme).

Oblast statistike koja se bavi uzorkom se naziva teorija uzoraka (*sampling theory*).

U nastavku teksta ukratko će biti opisane osnovne vrste uzoraka:

[91] Uzorak dobro reprezentuje populaciju ukoliko je u odnosu na posmatrano obeležje struktura uzorka vrlo bliska strukturi populacije.

Stratifikovani uzorak

Stratified sampling · Échantillonnage stratifié
Районированная выборка · 分层抽样

U nekim slučajevima se posmatrana populacija može podeliti na disjunktne podgrupe. Svaka takva podela se naziva stratifikacija (raslojavanje), a dobijene podgrupe stratumi (slojevi).

Neka je populacija obima N podeljena na L stratuma obima $N_1, N_2, \ldots, N_L$. Tada je

$$\sum_{j=1}^{L} N_j = N .$$

Uz oznaku $\omega_j = N_j/N$, $j = 1, 2, \ldots L$, važi i

$$\sum_{j=1}^{N} \omega_j = 1 .$$

Ako je populacija beskonačna, tada bi se sa ω_j označavala verovatnoća da slučajno izabrani element populacije bude iz j-tog stratuma.

Izbor elemenata po stratumima može biti sa vraćanjem ili bez vraćanja. Izbori elemenata iz različitih stratuma treba da budu međusobno nezavisni. Iz populacije se dobija uzorak obima n tako što se iz stratuma j dobija uzorak obima n_j, $j = 1, 2, \ldots, L$, pri čemu je

$$n_1 + n_2 + \ldots + n_L = n.$$

Postoje dva osnovna pristupa izbora elemenata po stratumima:

a) Ravnomerni, kad je obim uzorka n_j iz svakog stratuma isti i jednak:

$$n_j = \frac{n}{L}, j = 1, 2, \ldots, L, \text{ i}$$

b) Proporcionalni, kad je obim uzorka n_j iz svakog stratuma proporcionalan obimu stratuma. Tada je:

$$n_j = n \cdot \omega_j, j = 1, 2, \ldots, L.$$

Pristup b) je povoljniji za dalje etape statističkog istraživanja.

Kod stratifikovanog uzorka veliki uticaj na dalje etape statističkog istraživanja ima izbor stratuma. Treba se rukovoditi principom da disperzija obeležja po stratumima bude što manja (tzv. homogenost).

Postoje, dakle, razni načini da se ostvari slučajni izbor elemenata populacije. Bez obzira na to kojim je načinom dobijen, uzorak treba da reprezentuje datu populaciju. To znači da se struktura uzorka, u što većoj meri, poklapa sa strukturom populacije. Ovo svojstvo izabranog uzorka (reprezentativnost) se u praksi može ostvariti sa manjom ili većom preciznošću.

Evo nekoliko primera formiranja stratuma: podela stanovništva po zanimanju, podela privrede po granama, podela preduzeća po broju zaposlenih, podela teritorije po reljefu u nekom ispitivanju biljnog sveta, itd.

Grupni uzorak

Cluster sampling · Échantillonnage en grappes
Выборочное обследование · 整群抽样

Neka se posmatrana populacija može podeliti na disjunktne podgrupe. Od svih grupa na slučajan način se izabere nekoliko, a svi elementi izabranih grupa čine uzorak. Stratifikovani uzorak intuitivno više odgovara ideji reprezentativnog uzorka, ali se grupni uzorak može brže dobiti. Osim toga, podela na grupe ne mora biti ostvarena na isti način kao kod stratifikacije.

Dvoetapni uzorak

Two-phase sampling · Echantillonnage à deux degrés
Двухфазная выборка · 两相抽样

Neka je posmatrana populacija podeljena na grupe. Na slučajan način se izabere nekoliko grupa, a zatim iz izabranih grupa bira određeni broj elemenata. Dakle, dvoetapni uzorak je izvesna kombinacija grupnog i stratifikovanog uzorka. Primeri ovog načina dobijanja uzorka su prilikom primene statistike u lingvistici. Recimo, utvrđivanje raspodele raznih vrsta reči u knjizi nekog pisca: biraju se stranice, a sa svake stranice nekoliko redova i u njima se prebroji frekvencija određenih vrsta reči.

Ukoliko se populacija podeli na grupe velikog obima, tada se grupe izabrane u prvoj etapi mogu podeliti na podgrupe, pa od tih podgrupa birati neke i iz njih po nekoliko elemenata. Onda više nije u pitanju dvoetapni, već troetapni uzorak. Analogno se formiraju i višeetapni uzorci.

Periodični uzorak

Systematic sampling · Échantillonnage systématique
Систематическая выборка · 系统抽样

Neka su elementi populacije na neki način poređani u niz (npr. automobili koji prolaze jednom trakom puta). Elementi za uzorak se izdvajaju periodično: npr. prvi, peti, deveti, trinaesti,... (tj. svaki četvrti).

Nešto opštiji pristup bi bio za period dužine k, da se od prvih k elemenata na slučajan način izabere jedan, a zatim, počevši od tog elementa, bira se svaki k-ti u nizu koji čine ostali elementi populacije. Uzorak se na ovaj način lako dobija, a i zadovoljava neku ideju ravnopravnosti za izbor elemenata u uzorak. Naravno, može se postaviti pitanje međusobne zavisnosti elemenata populacije ovako poređanih u niz (recimo, automobili u koloni na svadbi ☺).

Periodični uzorak se može primeniti kod stratifikovanog uzorka pri izboru po stratumima, može se primeniti kod grupnog uzorka za izbor grupa, kao i kod dvoetapnog uzorka za izbor grupa ili elemenata iz izabranih grupa.

Uzorak sa vraćanjem

Узорак са враћањем

EN Sampling with replacement
FR Sondage avec remise
RU Повторная выборка
ZH 放回抽样, fànghuí chōuyàng
AR المعاينة مع الارجاع
EL Δειγματοληψία με επανατοποθέτηση
DE Stichprobe mit Zurücklegen
ES Muestreo con reposición
IT Campionamento con ripetizione
SK Výber s opakovaním
SL Vzorčenje s ponavljanjem
HU Visszatevéses mintavétel

Povezani termini:

Populacija (182), prost slučajni uzorak (188), ravnomerna raspodela (194)

S obzirom da zaključke o obeležju koje posmatramo na nekoj populaciji donosimo na osnovu uzorka, jasno je da treba posvetiti posebnu pažnju metodologiji dobijanja uzoraka, jer od loših ili nepotpunih podataka nije moguće dobiti kvalitetne zaključke.

U teoriji koristimo prost slučajni uzorak, što obezbeđuje jednakost raspodela i nezavisnost slučajnih veličina koje su komponente tog uzorka.

U praksi pojmu prostog slučajnog uzorka odgovara **uzorak sa vraćanjem**. Pod tim podrazumevamo da se na slučajan način izabere element populacije, da se zabeleži vrednost obeležja na tom elementu i da se pre izbora sledećeg elementa posmatrani „vrati“ u populaciju. Na taj način u svakom trenutku su na raspolaganju svi elementi populacije, i svaki od njih ima jednaku šansu da bude izabran u uzorak. Takođe, elemente biramo na slučajan način i nezavisno jedan od drugog i time obezbeđujemo da su podaci koje dobijemo realizacija prostog slučajnog uzorka, te da se na njih mogu primeniti teorijske metode matematičke statistike zasnovane na prostom slučajnom uzorku.

Za konačne populacije izbor sa vraćanjem može se ostvariti numerisanjem elemenata populacije od 1 do N, a onda se korišćenjem slučajnih brojeva ili RAND funkcije dobija željenih n elemenata koji će biti u uzorku. RAND ili RND funkcija daje nezavisne realizacije slučajne veličine sa uniformnom raspodelom U(0,1), tako da se indeks elementa koji je „izabran“ u uzorak dobija kao celobrojni deo izraza N·RND+1.

Uzorak bez vraćanja (*sampling without replacement*) podrazumeva da se na slučajan način izabere element populacije, da se zabeleži vrednost obeležja na tom elementu i da se taj element više ne može „birati“ u uzorak.

Verovatnoća

Вероватноћа

EN	Probability
FR	Probabilité
RU	Вероятность
ZH	概率, gàilǜ
AR	الاحتمال
EL	Πιθανότητα
DE	Wahrscheinlichkeit
ES	Probabilidad
IT	Probabilità
SK	Pravdepodobnosť
SL	Verjetnost
HU	Valószínűség

Povezani termini:

Frekvencija (80), slučajni događaj (200)

Etimologija:

Aksiom/aksioma < ἀξίωμα^{EL} (preduslov, ono što se smatra prigodnim, očigledna istina) < ἄξιος^{EL} (vredan, pogodan, dostojan) < ἄγω^{EL} (voditi, voziti).

Ako realna funkcija P definisana na σ-polju F[92] događaja iz Ω ima osobine:

$B1^{\circ}$ Za svaki događaj A koji pripada F važi $P(A) \geq 0$.

$B2^{\circ}$ Za siguran događaj Ω važi $P(\Omega) = 1$.

$B3^{\circ}$ Za svaki niz (konačan ili prebrojiv) događaja koji pripadaju F i koji su međusobno disjunktni u parovima važi:

$$P\left(\bigcup_j A_j\right) = \sum_j P(A_j).$$

Tada je funkcija P **verovatnoća** na σ-polju F.

Osobine $B1^{\circ}$, $B2^{\circ}$ i $B3^{\circ}$ su **aksiome** i nazivaju se, redom: nenegativnost, normiranost i aditivnost.

Aksiome verovatnoće $B1^{\circ}$, $B2^{\circ}$ i $B3^{\circ}$ ne daju način za određivanje verovatnoće u konkretnim zadacima, ali se, polazeći od definicije verovatnoće mogu izvesti mnoge opšte osobine verovatnoće.

Teorija verovatnoće je neophodna za bolje razumevanje ideja i metoda matematičke statistike, zato što se mnogi pojmovi definišu i mnoge osobine dokazuju korišćenjem pojmova teorije verovatnoće, kao i same verovatnoće čiji analogon u realnim (eksperimentalnim) situacijama predstavlja frekvencija.

Zanimljivo je da data aksiomatika nije potpuna u sledećem smislu: pri bacanju novčića, i kad definišemo da je verovatnoća dobijanja „pisma“ ½, i kad je, na primer, ⅓, aksiome verovatnoće su zadovoljene. U praksi biramo onaj model koji je bliži realnosti, tj. verovatnoća dobijanja pisma je ½.

[92] Sigma-polje F je familija događaja iz jednog prostora elementarnih ishoda koja sadrži siguran događaj, uz svaki događaj sadrži i njegov komplement, i uz svaki konačan ili prebrojiv niz događaja sadrži i njihovu uniju.

Višestruka regresija

Вишеструка регресија

EN	Multiple regression
FR	Régression multiple
RU	Множественная регрессия
ZH	多重回归, duōchóng huíguī
AR	الانحدار المتعدد
EL	Πολλαπλή παλινδρόμηση
DE	Mehrfachregression
ES	Regresión múltiple
IT	Regressione multipla
SK	Viacrozmerná regresia
SL	Multipla regresija
HU	Többváltozós regresszió

Povezani termini:

Metoda najmanjih kvadrata (75), regresija (144), teorema Gausa-Markova (251)

Etimologija:

Regresija < regressio^{LA} < regredior^{LA} (vratiti se) < re-^{LA} (ponovo) + gradior^{LA} (hodati)

Ukoliko posmatrano obeležje[93] Y zavisi od više tzv. objašnjavajućih promenljivih[94] $X_1, X_2, \ldots, X_n$, model kojim se opisuje ta zavisnost $Y = f(X_1, X_2, \ldots, X_n)$ može se označiti terminom **višestruka regresija**, po analogiji sa (jednostrukom linearnom) regresijom, ako je:

$$Y = a_0 + a_1X_1 + a_2X_2 + \ldots + a_nX_n + E, \qquad (1)$$

gde je E slučajna greška koju uključujemo u model da bismo obuhvatili sve ono što ne možemo „objasniti".

Mogući su i drugi višestruki modeli, kada zavisnost nije linearna.

Neki modeli u kojima Y zavisi samo od jedne veličine, označimo je sa X_1, npr. polinomijalni model:

$$Y = a_0 + a_1X_1 + a_2X_1^2 + \ldots + a_nX_1^n + E,$$

gde je E slučajna greška, sa smenom promenljvih $Xj = X_1^j$, svodi na prethodni model (1).

Određivanje ocena nepoznatih parametara u ovakvim modelima moguće je raznim metodama, među kojima je i metoda najmanjih kvadrata, za koju je karakteristična teorema Gausa-Markova prema kojoj od svih linearnih ocena nepoznatih parametara, najmanju disperziju ima linearna ocena dobijena metodom najmanjih kvadrata.

[93] Za Y se koriste razni termini: odgovor (*response*), zavisna promenljiva (*dependent variable*), ishod (*outcome*), i dr. a moguće je da neke oblasti primene imaju i neke druge specifične termine.

[94] Za objašnjavajuće promenljive (*explanatory variables*) se koriste i termini: kontrolisane promenljive (*controlled variables*), veličine koje služe za prognozu (*predictors*), pozadinske veličine (*background variables*), i dr. a moguće je da neke oblasti primene imaju i neke druge specifične termine.

Zakon velikih brojeva

Закон великих бројева

EN Law of large numbers
FR Loi des grandes nombres
RU Закон больших чисел
ZH 大数定律, dàshù dìnglǜ
AR قانون الأعداد الكبيرة
EL Νόμος των μεγάλων αριθμών
DE Gesetz der großen Zahlen
ES Ley de los grandes números
IT Legge dei grandi numeri
SK Zákon veľkých čísel
SL Zakon velikih števil
HU A nagy számok törvénye

Povezani termini:

Konvergencija u verovatnoći (130), nezavisne slučajne veličine (166)

Neka je X_1, X_2,... niz nezavisnih slučajnih veličina definisanih nad istim prostorom verovatnoća. Ako za svaki pozitivan broj ε važi:

$$P\left\{\left|\frac{1}{n}\sum_{j=1}^{n} X_j - \frac{1}{n}\sum_{j=1}^{n} EX_j\right| \geq \varepsilon\right\} \to 0, n \to \infty \qquad (1)$$

tada se kaže da za posmatrani niz važi slabi zakon velikih brojeva. Iz (1) proizilazi da je smisao zakona velikih brojeva da uzoračka sredina konvergira ka svom očekivanju.

Ako se pretpostavi da za niz slučajnih veličina važe neki dodatni uslovi, onda će uzoračka sredina da konvergira u verovatnoći ka svom matematičkom očekivanju

$$\overline{X}_n \xrightarrow{P} E(\overline{X}_n),\ n \to \infty,\ \overline{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j\,.$$

i važiće **zakon velikih brojeva**. Tako imamo Bernulijev zakon velikih brojeva, a takođe i Hinčinov zakon velikih brojeva, Čebiševljev zakon velikih brojeva, Markovljev zakon velikih brojeva, i dr.

Aleksandar Jakovljevič Hinčin (1894–1959), sovjetski matematičar i pedagog, jedan od osnivača ruske škole Teorije verovatnoće. Bavio se i Teorijom slučajnih procesa, Statističkom fizikom, Teorijom masovnog opsluživanja, a napisao je i veliki broj udžbenika za srednje škole i fakultete. Bio je član sovjetske Akademije nauka od 1939. godine.

Andrej Andrejevič Markov (1856–1922), ruski matematičar kome je jedan od profesora bio i Čebišev. Među njegovim doprinosima matematici i njenim primenama su Markovljevi procesi i teorema Gausa-Markova.

→ *Nastavak na str. 265*

Zvonasta kriva

Звонаста крива

EN	Bell-shaped curve
FR	Courbe en cloche
RU	Колоколообразная кривая
ZH	钟形曲线, zhōngxíng qūxiàn
AR	المنحني الجرسي
EL	Κωδωνοειδής καμπύλη
DE	Glockenkurve
ES	Curva campanular
IT	Curva a campana
SK	Zvonovitá krivka
SL	Krivulja v obliki zvona
HU	Haranggörbe

Povezani termini:

Asimetrija (18), hi-kvadrat raspodela (102), histogram (108), moda (160), medijana (154), matematičko očekivanje (152), normalna raspodela (172), Studentova raspodela (45)

Zvonasta kriva je termin koji se koristi za opisivanje grafika gustine raspodele koji je sličan ravnom preseku zvona. Osim normalne raspodele, za koju se termin najčešće vezuje kao sinonim, i druge raspodele imaju taj oblik. Na primer, Studentova t-raspodela takođe ima zvonast oblik, bez obzira na broj stepeni slobode. Košijeva raspodela[95] je takođe sličnog oblika. χ^2-raspodela sa velikim (većim od 30) brojem stepeni slobode takođe pokazuje sličnost sa zvonastom krivom.

Ogisten-Luj Koši (1789–1857), francuski inženjer i matematičar. Njegov doprinos je značajan u mnogim matematičkim disciplinama, pa i u teoriji verovatnoće gde jedna raspodela nosi njegovo ime. Košijevi rezultati se sreću i u matematičkoj analizi, teoriji determinanata, parcijalnim diferencijalnim jednačinama, teoriji funkcija kompleksne promenljive. Zbog političkih opredeljenja 18 godina nije bio unapređen u zvanje redovnog profesora.

Kod svih ovih raspodela je karakteristično da su unimodalne, da je mod jednak ili vrlo blizak očekivanju, kao i da su simetrične ili vrlo blago asimetrične.

Ukoliko histogram pokazuje sličnost sa raspodelama ovakvog tipa, i za njega se može reći da predstavlja zvonastu krivu.

[95] Neprekidna raspodela sa gustinom, u opštem slučaju oblika $f(x) = k/(\pi\{k^2 + (x - m)^2\})$, $k > 0$. Standardni oblik je sa $k = 1$, $m = 0$, i predstavlja Studentovu raspodelu sa jednim stepenom slobode. Dakle, ako su X i Y nezavisne normalne normirane slučajne veličine, onda je X/Y slučajna veličina sa standardnom Košijevom raspodelom. Košijeva raspodela *nema* ni matematičko očekivanje ni disperziju. Zbir nezavisnih slučajnih veličina sa Košijevom raspodelom, kao i svaka linearna kombinacija Košijevih raspodela je takođe Košijeva raspodela. Znači da centralna granična teorema na ovu raspodelu nije primenljiva. I mod i medijana ove raspodele su jednaki m.

Nastavak teksta

Asimetrija (*nastavak sa str. 19*)

U statistici asimetriju u skupu podataka uočavamo i pri raznim grafičkim prikazima. A asimetriju raspodele, na primer, na histogramu. Tada je važno voditi računa o dužinama intervala, jer za neki skup podataka mala izmena granica može da dovede do značajne izmene histograma, a time i do pogrešnih prvih zaključaka o populaciji.

Binomna raspodela (*nastavak sa str. 39*)

U zakonu raspodele slučajne veličine X sa binomnom raspodelom $B(n,p)$ najveća je verovatnoća $p_{n,k}$ za onu vrednost k koja zadovoljava dvostruku nejednakost:

$$(np + p) - 1 \leq k \leq np + p,$$

a kako je matematičko očekivanje $E(X) = np$, biće najveća verovatnoća u zakonu raspodele za onu vrednost k koja zadovoljava dvostruku nejednakost

$$E(X) + p - 1 \leq k \leq E(X) + p,$$

tako da je, za binomnu raspodelu, najverovatnija vrednost slučajne veličine bliska njenom matematičkom očekivanju. Koeficijent asimetrije za binomnu raspodelu je jednak

$$f_1 = \frac{1-2p}{\sqrt{np(1-p)}},$$

pa se i ovako potvrđuje činjenica da će raspodela biti simetrična pri $p = 1/2$.

Koeficijent spljoštenosti je

$$f_2 = \frac{1-6p(1-p)}{np(1-p)}.$$

Za velike vrednosti n, npr. n veće od 30, binomna raspodela se može aproksimirati Puasonovom (ako je $np \leq 10$) ili normalnom raspodelom (ako je $np \geq 10$).

Boks plot dijagram (*nastavak sa str. 43*)

Dijagram se interpretira u skladu sa osobinom normalne raspodele kod koje se van intervala (f_1, f_3) nalazi samo 0.7% vrednosti obeležja. Stoga se, po analogiji sa normalnom raspodelom, sve vrednosti iz uzorka koje se nalaze van intervala (f_1, f_3) smatraju za vrednosti obeležja koje neuobičajeno mnogo odstupaju od očekivane vrednosti, tzv. ekstremne vrednosti, koje se nazivaju i autlajeri (*outlier* – onaj koji je van granica, što bi se u našem jeziku moglo reći „leživan“ ☺).

Disperzija (*nastavak sa str. 55*)

U slučaju normalne raspodele matematičko očekivanje i disperzija su ujedno i parametri koji tu gustinu raspodele jednoznačno određuju. Sa raspodelama drugih slučajnih veličina ne mora biti tako. Videti, na primer, beta raspodelu.

Disperziona analiza (*nastavak sa str. 57*)

Računaju se:

$$S_A^2 = \sum_{i=1}^{k} n_i\left(\overline{X}_i - \overline{X}\right)^2, \qquad S_R^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}_i\right)^2,$$

gde n_i predstavlja obim uzorka i-tog proizvođača.

Za nultu hipotezu H_0 su kritične velike vrednosti količnika $\frac{(n-k)S_A^2}{(k-1)S_R^2}$, pa se, za dati prag značajnosti α, kritična oblast određuje iz tablica za Fišerovu raspodelu, na osnovu

$$P_{H_0}\left\{F_{k-1,n-k} \geq c\right\} = \alpha,$$

gde je c tablična vrednost za Fišerovu raspodelu za odgovarajući broj stepeni slobode $c = F_{k-1,n-1;\alpha}$.

Enskombov kvartet (*nastavak sa str. 75*)

Značaj ovog primera je da treba prvo pogledati kakvi su podaci na osnovu grafika, pa onda pristupiti izračunavanjima, a ne upuštati se u izračunavanja po nekom modelu bez prethodne provere da li taj model uopšte odgovara podacima i proučavanoj pojavi.

Funkcija raspodele (*nastavak sa str. 89*)

U praksi se funkcija raspodele koristi za računanje verovatnoća[96] u vezi sa odgovarajućom slučajnom veličinom. Za neke najvažnije slučajne veličine postoje tablice vrednosti funkcije raspodele.

Takođe je veoma bitna veza empirijske funkcije raspodele i funkcije raspodele obeležja, data centralnom teoremom matematičke statistike.

Hi-kvadrat test (*nastavak sa str. 105*)

Ako je hipoteza H_0 tačna, pri $n \to \infty$, test-statistika ima asimptotski χ_{r-s-1} raspodelu. Za prag značajnosti α, pomoću tablice za χ^2-raspodelu dobija se $c = \chi^2_{r-s-1;\alpha}$ tako da važi

$$P\left\{\chi_U^2 \geq c\right\} = \alpha .$$

Ako je realizovana vrednost test-statistike veća od kritične vrednosti c, nulta hipoteza se odbacuje. U suprotnom se, na osnovu dobijenih podataka i za dati prag značajnosti, nulta hipoteza se prihvata.

---

[96] Ako je F funkcija raspodele slučajne veličine X i a i b proizvoljni realni brojevi, tada je:

$$P\{a < X \leq b\} = F(b) - F(a).$$

Takođe važe i jednakosti:

$$P\{X > b\} = 1 - F(b),\ P\{X \geq b\} = 1 - F(b) - P(X = b),$$
$$P\{a \leq X \leq b\} = F(b) - F(a) + P(X = a),$$
$$P\{X \leq a \cup X > b\} = F(a) + (1 - F(b)),\ a < b.$$

Napomena: Ukoliko se dobije u k-tom podskupu (klasi) S_k da je $np_k < 5$, taj podskup se spaja sa nekim drugim podskupom, tako da u novoj podeli za sve podskupove važi uslov da je očekivani broj elemenata uzorka u tom podskupu (klasi) veći ili jednak 5.

Hi-kvadrat test nezavisnosti (*nastavak sa str. 107*)

Ako je tačna nulta hipoteza, test-statistika

$$\chi_U^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{\left(n \cdot n_{ij} - n(x_i)n(y_j)\right)^2}{n \cdot n(x_i)n(y_j)} \qquad (*)$$

pri $n \to \infty$, ima asimptotski $\chi^2_{(r-1)(s-1)}$ raspodelu. Za dati prag značajnosti α, pomoću tablice za $\chi^2_{(r-1)(s-1)}$ raspodelu dobija se $c = \chi^2_{(r-1)(s-1);\alpha}$ tako da važi:

$$P\left\{\chi_U^2 \geq c\right\} = \alpha .$$

Hipoteza o nezavisnosti se odbacuje ako je realizovana vrednost test-statistike veća od c, a u suprotnom se prihvata.

Formula (*) i na prvi pogled ima sličnosti sa formulom za test-statistiku u osnovnom χ^2-testu. Međutim, one jesu istovetne, i ovaj test je samo varijanta χ^2-testa.

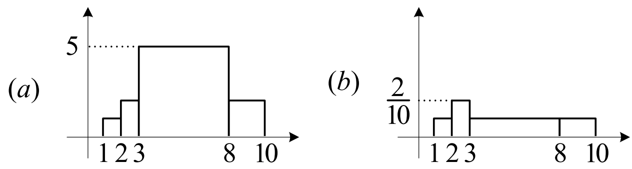
U Pirsonovom χ^2-testu (i svim njegovim varijantama) obim uzorka treba da bude veliki, bar toliko da se obezbedi da očekivane frekvencije u svim klasama budu veće od 5.

Histogram (*nastavak sa str. 109*)

Primer. Dati su podaci:

vrednost	[1,2)	[2,3)	[3,8)	[8,10)
frekvencija	1	2	5	2

Histogrami za date podatke su: sa apsolutnim frekvencijama na slici (a), sa ukupnom površinom 1 na slici (b):



Interval poverenja za očekivanje (*nast. sa str. 117*)

Znači, koristi se činjenica da je raspodela statistike

$$\frac{\overline{X}_n - m}{\overline{S}_n}\sqrt{n-1}$$

približno raspodela t_{n-1}, a postupak će biti kao prethodno opisani.

Interval poverenja za matematičko očekivanje se koristi i u statističkoj kontroli kvaliteta. Tada se obično koristi 90%-tni interval poverenja, ili 95%-tni ili pak 99%-tni, pa se u formulama eksplicitno i daje odgovarajuća tablična vrednost z_β.

Linearna transformacija podataka (*nast. sa str. 147*)

Transformacije podataka su vrlo često korišćene kad nije bilo računara, jer su predstavljale način da se izračunavanja olakšaju, posebno kad se radi sa većim skupom podataka. Danas se koriste samo u specijalnim situacijama, kad su vrednosti podataka ili veoma male ili veoma velike po apsolutnoj vrednosti, jer računari „ne osećaju" potrebu za takvim olakšicama.

Matematička statistika (*nastavak sa str. 151*)

Kasnije, svakako nezaobilazna imena u statistici su ser Ronald Fišer i Čarls Pirson, a takođe i mnogi naučnici iz Rusije i kasnije Sovjetskog Saveza, ali i iz Indije, dok se danas u statistici, i matematici uopšte, sve češće sreću imena kineskih naučnika.

Oblast primene statistike se mnogo proširila, i praktično nema domena ljudske delatnosti bez manjeg ili većeg primenjivanja statističkih metoda. Razvoj računarske tehnike značajno je doprineo brzom razvoju statistike zahvaljujući, između ostalog, mogućnostima bržeg izračunavanja potrebnih veličina na osnovu prikupljenih podataka. Postoje mnogi specifični paketi programa namenjenih za statističku obradu podataka: R, SPSS, Statgraphics, i dr. ali i Excel i Matlab,... takođe imaju široke palete statističkih metoda.

Matematičko očekivanje (*nastavak sa str. 153*)

Neke osobine matematičkog očekivanja:

- $E(X)$ *i* $E(|X|)$ istovremeno postoje ili ne postoje.
- Ako je $P\{X = a\} = 1$, gde je a neka konstanta, tada je $E(X) = a$.
- Ako je $P\{X \geq Y\} = 1$, tada je $E(X) \geq E(Y)$.
- Ako je k konstanta, tada je $E(kX) = k \cdot E(X)$.
- Ako su X i Y dve slučajne veličine, koje imaju matematička očekivanja, tada je $E(X + Y) = E(X) + E(Y)$.
- Ako su X i Y nezavisne slučajne veličine sa matematičkim očekivanjima, tada je $E(X \cdot Y) = E(X) \cdot E(Y)$.

Metoda maksimalne verodostojnosti (*nast. sa str. 157*)

Ocene koje se dobijaju postupkom metode maksimalne verodostojnosti imaju, često, dobre osobine, kao što su nepristrasnost i/ili postojanost, a takođe su često efikasnije od ocena dobijenih nekom drugom metodom. Pod određenim uslovima raspodele ocena dobijenih metodom maksimalne verodostojnosti imaju asimptotski normalnu raspodelu. Ocene dobijene metodom maksimalne verodostojnosti su simetrične funkcije uzorka (ne zavise od redosleda podataka u uzorku). Takođe, ocene po metodi maksimalne verodostojnosti imaju osobinu invarijantnosti:

Neka je g(x) neprekidna funkcija koja ima neprekidnu inverznu funkciju i neka je Y statistika kojom se po metodi maksimalne verodostojnosti ocenjuje parametar θ. Tada je g(Y) statistika kojom se po metodi maksimalne verodostojnosti ocenjuje g(θ).

Normalna raspodela (*nastavak sa str. 173*)

Odatle specijalno sledi da je $\mu_3 = 0$, $\mu_4 = 3\sigma^4$ i zato su koeficijenti asimetrije i spljoštenosti jednaki 0.

$P\{m - 3\sigma \le X \le m + 3\sigma\} = 0.997 \approx 1$, (tzv. pravilo 3-sigma)

Specijalni slučaj predstavlja normalna raspodela sa parametrima $m = 0$ i $\sigma = 1$. Za takvu slučajnu veličinu se kaže da ima normalnu normiranu raspodelu $N(0,1)$. Značaj ove raspodele proizilazi iz sledećeg tvrđenja:

Ako slučajna veličina X ima $N(m, \sigma^2)$ *raspodelu, tada slučajna veličina* $Y = (X - m)/\sigma$ *ima* $N(0,1)$ *raspodelu.*

Ova teorema se primenjuje pri izračunavanju verovatnoća vezanih za slučajne veličine sa normalnom raspodelom, s obzirom da se moraju koristiti tablice sa unapred određenim vrednostima funkcije raspodele. Teorema pokazuje da je za računanje vrednosti funkcije raspodele bilo koje slučajne veličine sa normalnom raspodelom, dovoljno znati vrednosti funkcije raspodele slučajne veličine sa normalnom normiranom raspodelom.

Normalna raspodela je jedna neprekidna raspodela, koja je povezana sa mnogim drugim diskretnim i neprekidnim raspodelama.

Značaj normalne raspodele proizilazi iz centralne granične teoreme.

Puasonova raspodela (*nastavak sa str. 193*)

Momenti trećeg i četvrtog reda su jednaki

$$E(X^3) = \lambda + 3\lambda^2 + \lambda^3,$$
$$E(X^4) = \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4,$$

a treći i četvrti centralni moment su

$$\mu_3 = \lambda,\ \mu_4 = \lambda + 3\lambda^2.$$

Koeficijenti asimetrije i spljoštenosti za ovu raspodelu su

$$f_1 = \frac{1}{\sqrt{\lambda}}, \quad f_2 = \frac{1}{\lambda}.$$

Neka su slučajne veličine $X : P(\lambda)$ i $Y : P(\mu)$ nezavisne. Tada je $Z = X + Y$ slučajna veličina sa Puasonovom raspodelom $P(\lambda + \mu)$.

Stem end lif dijagram (*nastavak sa str. 207*)

33	1 4	33	1 4 3 4
33	8	33	8 9 6 7 7
34	2	34	3 2 0 2 2 2
34	8 7 5 6 9	34	8 7 5 6 9 7
35		35	2 2 1 1 3
35		35	6 8
36		36	0 1

Tablično prikazivanje podataka (*nast. sa str. 211*)

Često se prihvataju i sledeća pravila: a) da sredine intervala budu jednake nekim podacima iz uzorka, b) da granice intervala budu različite od podataka iz uzorka.

Na primer, ako su podaci dati sa d decimala, onda se granice intervala daju sa d+1 decimala, i time ostvaruje ovo pravilo.

Kad su definisani intervali, određuje se broj elemenata uzorka koji se nalazi u svakom od tih intervala i ti podaci se unose u tabelu:

vrednost obeležja	$[a_1, a_2)$	$[a_2, a_3)$	...	$[a_k, a_{k+1})$
frekvencija	n_1	n_2	...	n_k

U tabeli se mogu dati relativne frekvencije n_j/n, $j = 1, 2, ..., k$ umesto apsolutnih frekvencija.

U gornjoj tabeli je $-\infty \leq a_1 < a_2 < ... < a_{k+1} \leq +\infty$ i $n_1 + n_2 + ... + n_k = n$. Ako je $a_1 = -\infty$, onda je prvi interval oblika $(-\infty, a_2)$, a ako je $a_{k+1} = +\infty$, tada je poslednji interval oblika $[a_k, +\infty)$.

Podaci dobijeni iz uzorka se moraju prvenstveno srediti u tabelu kao što je opisano u delu (a) odnosno (b) da bi mogli biti dati odgovori na pitanja o obliku ili karakteristikama raspodele obeležja. Grafičko prikazivanje daje, u izvesnoj meri, početne (delimične) odgovore na pitanja o raspodeli obeležja.

Uzoračka disperzija (*nastavak sa str. 225*)

Ako su podaci dati po intervalima, onda dolazi do razlike između stvarne uzoračke disperzije i one koja je izračunata na osnovu predstavnika intervala. Jedan od načina da se to ispravi su Šepardove korekcije koje se primenjuju ako su intervali jednake dužine d; tada od izračunate vrednosti oduzimamo $d^2/12$ i dobijamo vrednost uzoračke disperzije koja je bliža stvarnoj vrednosti.

Vilijem Flitvud Šepard (1863–1936), australijski matematičar koji je diplomirao na Kembridžu. Bavio se, između ostalog, problemima korelacije.

Uzoračka sredina (*nastavak sa str. 231*)

Onda je prvo je potrebno odrediti predstavnike datih intervala, na primer sredine intervala x_i' (pri čemu se pretpostavlja da su a_1 i a_{m+1} konačni), pa se realizovana vrednost statistike $\overline{X}_n$ računa po formuli:

$$\overline{x}_n = \frac{1}{n}\left(n_1 x'_1 + n_2 x'_2 + \ldots + n_m x'_m\right) = \frac{1}{n}\sum_{i=1}^{m} n_i x'_i . \quad (3)$$

Oblici (1) i (2) daju istu vrednost, u opštem slučaju različitu od vrednosti (3). Realizovani uzorak (uređena n-torka brojeva) daje više informacija o posmatranom obeležju nego uzoračka sredina (jedan broj). Jasno je da uzorci iz različitih populacija mogu da imaju iste uzoračke sredine, a takođe i da više uzoraka istog obima iz jedne populacije može imati (i vrlo) različite vrednosti uzoračke sredine.

Informacija o obeležju na uzorku, a samim tim ni na populaciji, ne može se svesti na navođenje realizovane vrednosti uzoračke sredine.

Raspodela uzoračke sredine je, takođe, jedna važna informacija o posmatranom obeležju. Za uzorke velikog obima može se aproksimirati normalnom raspodelom.

Zakon velikih brojeva (*nastavak sa str. 253*)

Hinčinov zakon velikih brojeva: Neka je X_1, X_2,... niz nezavisnih slučajnih veličina sa istom raspodelom sa konačnim matematičkim očekivanjem jednakim m. Tada za posmatrani niz slučajnih veličina važi zakon velikih brojeva.

Čebiševljev zakon velikih brojeva: Neka je X_1, X_2,... niz nezavisnih slučajnih veličina sa konačnim matematičkim očekivanjima i neka postoji konstanta c tako da važi: $DX_n \leq c$, n = 1, 2,... Tada za posmatrani niz slučajnih veličina važi zakon velikih brojeva.

Markovljev zakon velikih brojeva: Neka je X_1, X_2,... niz nezavisnih slučajnih veličina sa konačnim matematičkim očekivanjima i neka važi:

$$\lim_{n\to\infty} D(\overline{X_n}) = 0 \cdot$$

Tada za posmatrani niz slučajnih veličina važi zakon velikih brojeva.

Oznake korišćene u knjizi

∞	beskonačno *infinite*
$B(a,b)$	beta funkcija *beta function*
$\binom{n}{k}$	binomni koeficijent (čita se „n nad k“) *binomial coefficient, n over k*
$B(n,p)$	binomna raspodela *binomial distribution: n trials, probability of success p*
σ^2, Σ^2	disperzija (kao numerička karakteristika) *variance*
D	disperzija (kao operator) *variance*
$n!!$	dvostruki faktorijel *double factorial*
exp	eksponencijalna funkcija ($exp(x)=e^x$) *exponential function*
x_j, y_j, ...	elementi uzorka *sample elements*
$F_n^*(x)$	empirijska funkcija raspodele *empirical distribution function*

$F_{k,n}$	Fišerova raspodela sa k i n stepeni slobode *Fisher distribution with k and n degrees of freedom*
f_j	frekvencija *frequency*
$f(x)$, $g(x)$,...	funkcija *function*
$F(x)$	funkcija raspodele *distribution function*
L, $L(\mathbf{x},\theta)$, $L_n(\theta)$	funkcija verodostojnosti *likelihood function*
$\Gamma(a,b)$, $G(h)$	gama funkcija *gamma function*
lim, $\lim\limits_{j\to\infty} x_j$	granična vrednost *limit*
α	greška I vrste, prag značajnosti, veličina kritične oblasti *size of the critical region*
β	greška II vrste (u vezi sa testovima), nivo poverenja (u vezi sa intervalima poverenja) *level of confidence*
$g(x)$, $f(x)$	gustina raspodele *density function*

χ_n^2 hi-kvadrat raspodela sa *n* stepeni slobode
chi-squared distribution with n degrees of freedom

$I, I_k, \ldots$ indikator (slučajna veličina)
dummy variable, indicator variable

$\int$ integral
integral

d izvod
derivative

ρ koeficijent korelacije
correlation coefficient

$\xrightarrow{P}$ konvergencija u verovatnoći
convergence in probability

$\xrightarrow{Z}$ konvergencija u zakonu raspodele
convergence in distribution

$\bar{D}$ komplement događaja
complement

Cov, $C_{X,Y}$ kovarijansa
covariance

W kritična oblast testa
critical region, rejection region of a test

q_1, q_3 kvartili
quartiles

ln	logaritam sa osnovom e *natural logarithm*
$\max\limits_{1 \le j \le n} x_j$	maksimalni podatak *maximal value*
$\le$	manje ili jednako *less than or equal to*
E	matematičko očekivanje (operator) *mathematical expectation, expected value*
m, μ	matematičko očekivanje (parametar u raspodeli) *mathematical expectation, expected value*
m_e	medijana *median*
$\min\limits_{1 \le j \le n} x_j$	minimalni podatak *minimal value*
M	moć testa *power, power curve*
$n!$	n-faktorijel *n factorial*
Φ, Ø	nemoguć događaj, prazan skup *impossible event, empty set*
$\notin$	ne pripada *doesn't belong*

$N(0,1)$	normalna normirana raspodela *standard normal distribution*
$N(m,\sigma^2)$	normalna raspodela (sa parametrima m i σ^2) *normal distribution with parameters m and* σ^2
$\int_a^b$	određeni integral u granicama od a do b *definite integral on the interval a to b*
θ	parametar (nepoznat) *parameter (unknown)*
∂	parcijalni izvod *partial derivative*
$\subseteq$	podskup *subset*
F	polje događaja *field of events*
$\approx$	približno jednako *approximately equal to*
$\in$	pripada *belongs*
$\cap$	presek događaja *intersection*
(Ω, F, P)	prostor verovatnoća *probability space*

$P(\lambda)$	Puasonova raspodela sa parametrom λ *Poisson distribution with rate* λ
$\neq$	različito od *not equal to*
$\mathbf{x}$, $\vec{x}$, $(x_1,..., x_n)$	realizovani uzorak *sample*
A, B, C,...	slučajni događaji *random events*
X, Y,...	slučajne veličine *random variables*
U, V,...	statistike (slučajne veličine) *statistics*
H_0, H_1	statističke hipoteze *statistical hypotheses*
t_n	Studentova raspodela sa n stepeni slobode *Student's t-distribution with n degrees of freedom*
Σ	suma (zbir), konačan ili beskonačan (red) *sum, series*
$\sup_x \varphi(x)$	supremum funkcije φ po argumentu x *supremum of function* φ *for argument x*
z_β	tablična vrednost za normalnu raspodelu *table value for normal distribution*

$t_{n,\beta}$ tablična vrednost za Studentovu raspodelu
table value for Student's distribution

$\rightarrow$ teži ka
tends to

$U(a, b)$ uniformna raspodela na intervalu (a, b)
uniform distribution on the interval (a, b)

$P(A|B)$ uslovna verovatnoća događaja A, ako se realizovao događaj B
probability of B conditional on A

$\boldsymbol{X}$, $(X_1,..., X_n)$ uzorak (kao slučajna veličina)
sample

$\overline{S}_n$ uzoračka devijacija
sample deviation

$\overline{S}_n^2$, $\overline{S}^2$, $\hat{S}_n^2, \ldots$ uzoračka disperzija
sample variance

r, $r_{X,Y}$ uzorački koeficijent korelacije
sample correlation coefficient

$\overline{x}$, $\overline{x}_n$ uzoračka sredina (realizovana)
sample mean

$\overline{X}_n$ uzoračka sredina (slučajna veličina)
sample mean

$\geq$ veće ili jednako
greater than or equal to

$P(A)$	verovatnoća događaja A *probability of A*
$P_H(A)$, $P_H\{A\}$	verovatnoća događaja A pri hipotezi H *probability of A under the hypothesis H*
$F(x, y)$	zajednička funkcija raspodele (dvodimenzionalna) *bivariate cumulative density function*
$g(x, y)$	zajednička gustina raspodele (dvodimenzionalna) *bivariate probability density function*

Sadržaj

Predgovor 4

Alternativna hipoteza 8
Aposteriorna verovatnoća...10
Apriorna verovatnoća 12
Apsolutno srednje odstupanje 14
Aritmetička sredina 16
Asimetrija (→256) 18
Asimptotski efikasna ocena 20
Asimptotska raspodela 22
Asimptotska relativna efikasnost....................................24
Autokorelacija 26
Bajesova formula..28
Bajesovsko zaključivanje 30
Basuova teorema 32
Bernulijev zakon velikih brojeva 34
Beta raspodela 36
Binomna raspodela (→256) 38
Biostatistika...40
Boks plot dijagram (→257)...42
Broj stepeni slobode 44
Centralna granična teorema..46
Centralna teorema matematičke statistike 48
Černovljeva lica...50
Disjunktni događaji 52
Disperzija (→257)...54
Disperziona analiza (→257)..56
Dovoljna statistika..58
Dvodimenzionalna raspodela 60
Efikasna ocena...62
Eksponencijalna familija raspodela.................................64
Eksponencijalna raspodela 66
Ekstrapolacija 68
Ekstremne vrednosti u uzorku...70
Empirijska funkcija raspodele 72

Enskombov kvartet (→258) 74
Erlangova raspodela 76
Fišer-Nejmanova teorema 78
Frekvencija 80
Funkcija gubitaka 82
Funkcija gustine 84
Funkcija moći testa 86
Funkcija raspodele (→258) 88
Funkcija verodostojnosti 90
Gama raspodela 92
Grafik 94
Greška druge vrste 96
Greška prve vrste 98
Gumbelova raspodela 100
Hi-kvadrat raspodela 102
Hi-kvadrat test (→258) 104
Hi-kvadrat test nezavisnosti (→259) 106
Histogram (→259) 108
Homogenost 110
Indeks krivolinijske korelacije 112
Interval poverenja 114
Interval poverenja za očekivanje (→260) 116
Izvestan događaj 118
Koeficijent asimetrije 120
Koeficijent korelacije 122
Koeficijent spljoštenosti 124
Koeficijent varijacije 126
Kontrola kvaliteta 128
Konvergencija u verovatnoći 130
Kovarijansa 132
Kriterijumi saglasnosti 134
Kritična oblast 136
Kvantil 138
Kvartil 140
Linearna korelacija 142
Linearna regresija 144
Linearna transformacija podataka (→260) 146

Marginalna raspodela 148
Matematička statistika (→260) 150
Matematičko očekivanje (→261) 152
Medijana 154
Metoda maksimalne verodostojnosti (→261) 156
Metoda momenata 158
Modalna vrednost 160
Neparametarski test 162
Nepristrasna ocena 164
Nezavisne slučajne promenljive 166
Nezavisni događaji 168
Nivo poverenja 170
Normalna raspodela (→262) 172
Nulta hipoteza 174
Obeležje 176
Ocenjivanje parametara 178
Poligon frekvencija 180
Populacija 182
Postojana ocena 184
Prag značajnosti 186
Prost slučajan uzorak 188
Prostor elementarnih ishoda 190
Puasonova raspodela (→262) 192
Ravnomerna raspodela 194
Reprezentativni uzorak 196
Slučajna promenljiva 198
Slučajni događaj 200
Standardna devijacija 202
Statistike poretka 204
Stem end lif dijagram (→263) 206
Tabela kontingencije 208
Tablično prikazivanje podataka (→263) 210
Teorija verovatnoće 212
Testiranje hipoteze 216
Testiranje hipoteze o očekivanju 218
Uniformno najmoćniji test 220
Uslovna verovatnoća 222

Uzoračka disperzija (→264) 224
Uzoračka medijana..226
Uzoračka moda...228
Uzoračka sredina (→264) 230
Uzorački koeficijent korelacije...232
Uzorački kvantili 236
Uzorački momenti 238
Uzorak 240
Uzorak sa vraćanjem 246
Verovatnoća 248
Višestruka regresija 250
Zakon velikih brojeva (→265) 252
Zvonasta kriva 254

Oznake korišćene u knjizi 266
Sadržaj 274
Kratki biografski podaci 277
Ostali pojmovi objašnjeni u knjizi 278
Sadržaj na stranim jezicima 282

Kratki biografski podaci

Bajes, Tomas 29
Basu, Debabrata 33
Bernuli, Jakob..35
Černof, Herman 51
De Muavr, Abraham...214
Enskomb, Džon Frensis..75
Erlang, Agner Krarup 77
Ferma, Pjer 214
Fišer, Ronald Ejlmer 57
Gaus, Karl Fridrih 214
Glivenko, Valerij Ivanovič 49
Gumbel, Emil Julijus 101
Hinčin, Aleksandar Jakovljevič..253
Kanteli, Frančesko Paolo..49
Kardano, Đirolamo 213

Kolmogorov, Andrej Nikolajevič....................................135
Koši, Ogisten-Luj...255
Kramer, Karl Harald...21
Laplas, Pjer-Simon...214
Markov, Andrej Andrejevič..253
Nejman, Jirži...79
Paskal, Blez..213
Puason, Simeon Deni..215
Rao, Kaljampadi Radakrišna..21
Šepard, Vilijem Flitvud..257
Šuhart, Valter Endrju..129
Tjuki, Džon Vajlder...43
Vejbul, Ernest Hjalmar Valodi..71

Ostali pojmovi objašnjeni u knjizi

Aditivnost...249
Analiza glavnih komponenata.......................................51
Analiza preživljavanja..101
Antimod...161
Apriorna raspodela...13
Apsolutna frekvencija...211
Asimptotski nepristrasna ocena.....................................239
Bimodalna raspodela...161
Borelovi skupovi...167
Čebiševljev zakon velikih brojeva.................................265
Demografija..41
Deterministički eksperiment...191
Dijagram sa dužima..95
Diskretna slučajna veličina...199
Dvoetapni uzorak..244
Dvostrana kritična oblast..137
Eksperiment..191
Elementaran ishod..191
Finitno preslikavanje..199
Fišerova raspodela..45
Formula potpune verovatnoće......................................11
Funkcija odluke..31
Grupni uzorak...243

Hi-kvadrat raspodela sa n stepeni slobode 45
Hinčinov zakon velikih brojeva 265
Hipoteza nezavisnosti 107
Informaciona funkcija Fišera 21
Interkvartilni razmak 141
Interpolacija 69
Intervalno ocenjivanje 179
Invarijantnost 261
Izravnanje 17
Jednostrana kritična oblast 137
Koeficijent asimetrije 19
Konvergencija u raspodeli 131
Konvergencija u srednjem reda k 131
Korelaciona tabela 233
Košijeva raspodela 255
Kružni dijagram 95
Kvadratna forma 57
Kvartilni koeficijent asimetrije 121
Linearna nepristrasna ocena 165
Linearna regresija 75
Mali uzorak 241
Marginalna gustina raspodele 149
Markovljev zakon velikih brojeva 265
Merljivo preslikavanje 199
Metoda najmanjih kvadrata 75, 145
Moć testa 175
Modalitet 183
Modalitet obeležja 81
Modalna klasa 229
Moguća ocena 31
Monoton količnik verodostojnosti 221
Muavr-Laplasova teorema 47
Negativna korelacija 235
Negativna spljoštenost 125
Negativno asimetrična raspodela 19
Nejednakost Čebiševa 35
Nejednakost Rao-Kramera 21
Nekorelisane slučajne veličine 133
Nemoguć događaj 201
Nenegativnost 249
Neparametarska hipoteza 163

Neprekidna slučajna veličina 199
Niz pokretnih proseka 17
Normalna normirana raspodela 262
Normiranost 249
Obim uzorka 241
Objašnjavajuća promenljiva 251
Odsustvo memorije 67
Parametarski prostor 31
Parcijalni koeficijent korelacije 123
Paretova raspodela 161
Periodični uzorak 244
Piktogram 95
Pirsonov koeficijent asimetrije 121
Pirsonov test 135
Polimodalna raspodela 161
Polinomijalni model 251
Ponder 17
Ponderisana sredina 17
Potpopulacija 111
Potpun sistem događaja 11, 29
Pozitivna korelacija 235
Pozitivna spljoštenost 125
Pozitivno asimetrična raspodela 19
Prosta hipoteza 9
Prostor verovatnoća 131
Prvi uzorački kvartil 237
Rang 163
Raspodela ekstremnih vrednosti 23
Raspodela Kolmogorova 45
Raspodela obeležja 177, 183
Raspon uzorka 205
Realizovani uzorak 189
Relativna frekvencija 211
Robusnost 227
Sigma-polje F 249
Simetrična funkcija uzorka 261
Sistem masovnog opsluživanja 77
Sistem normalnih jednačina 145
Slabi zakon velikih brojeva 253
Složena hipoteza 9
Skoro nemoguć događaj 119

Skoro sigurna konvergencija..131
Srednji rizik...31
Statistički značajna veličina 187
Statistika..167
Stohastički eksperiment 191
Stratifikacija 242
Stratifikovani uzorak 242
Stratum..242
Studentova raspodela 45
Šepardova korekcija...264
Šifriranje podataka 177
Tabela kontingencije 107
Tačkaste ocene 179
Teorema Gausa-Markova...251
Teorija uzoraka 241
Test Kolmogorova 135
Test saglasnosti 163
Test slobodan od raspodele 217
Test-statistika 217
Translacija podataka 147
Treći uzorački kvartil 237
Unimodalna raspodela..161
Uslovna raspodela 61
Uslovno matematičko očekivanje 223
Uzoračka kovarijansa 133
Uzorački koeficijent asimetrije 121
Uzorački koeficijent spljoštenosti 125
Uzorački prostor...191
Uzorak bez vraćanja..247
Varijacioni niz..205, 227
Veliki uzorak..241
Višeetapni uzorak..244
Vremenska serija...27
$\bar{X}$-kontrolne kartice 129
Zlatni presek..181

Table of Contents

Alternative hypothesis.. 8
Analysis of variance.. 56
Anscombe's quartet.. 74
Arithmetic mean.. 16
Asymptotic distribution.. 22
Asymptotic relative efficiency.. 24
Asymptotically efficient estimator.. 20
Autocorrelation.. 26
Basu's theorem.. 32
Bayes' formula.. 28
Bayesian inference.. 30
Bell-shaped curve.. 254
Bernoulli law of large numbers.. 34
Beta distribution.. 36
Binomial distribution.. 38
Biostatistics.. 40
Bivariate distribution.. 60
Box plot.. 42
Central limit theorem.. 46
Certain event.. 118
Characteristic.. 176
Chart.. 94
Chernoff's faces.. 50
Chi-squared distribution.. 102
Chi-squared test.. 104
Chi-squared test for independence.. 106
Coefficient of correlation.. 122
Coefficient of kurtosis.. 124
Coefficient of skewness.. 120
Coefficient of variation.. 126
Conditional probability.. 222
Confidence interval.. 114
Confidence interval for the mean.. 116
Confidence level.. 170
Consistent estimator.. 184
Contingency table.. 208
Convergence in probability.. 130
Covariance.. 132

Critical region 136
Disjoint events 52
Distribution function 88
Efficient estimator 62
Erlang distribution..76
Estimation of parameters 178
Exponential distribution 66
Exponential family 64
Extrapolation 68
Fisher-Neyman theorem 78
Frequency 80
Frequency polygon 180
Gamma distribution 92
Glivenko-Cantelli theorem 48
Gumbel distribution 100
Histogram 108
Homogeneity 110
Hypothesis test for mathematical expectation 218
Independent events 168
Independent random variables 166
Index of non-linear correlation 112
Law of large numbers 252
Level of significance 186
Likelihood function 90
Linear correlation 142
Linear data transformation 146
Linear regression 144
Loss function 82
Marginal distribution 148
Mathematical expectation 152
Mathematical statistics 150
Maximum likelihood method 156
Mean absolute deviation 14
Median 154
Method of moments 158
Modal value 160
Multiple regression 250
Nonparametric test 162
Normal distribution 172
Null hypothesis 174
Number of degrees of freedom 44

Order statistics 204
Outliers........ 70
Poisson distribution........ 192
Population 182
Posterior probability........ 10
Power function 86
Prior probability 12
Probability........ 248
Probability density function 84
Probability theory........ 212
Quality control 128
Quantile........ 138
Quartile 140
Random event 200
Random variable 198
Representative sample........ 196
Sample 240
Sample coefficient of correlation 232
Sample mean 230
Sample median........ 226
Sample mode........ 228
Sample moments 238
Sample quantiles 236
Sample space........ 190
Sample variance 224
Sampling distribution 72
Sampling with replacement........ 246
Simple random sample........ 188
Skewness........ 18
Standard deviation........ 202
Stem and leaf diagram........ 206
Sufficient statistic 58
Tabulation of data 210
Test of goodness of fit........ 134
Testing of hypothesis 216
Type I error 98
Type II error........ 96
Unbiased estimator........ 164
Uniform distribution 194
Uniformly most powerful test 220
Variance 54

Table des matières

Analyse de variance 56
Asymétrie 18
Autocorrélation 26
Biostatistique 40
Caractère 176
Coefficient d'aplatissement 124
Coefficient d'asymétrie 120
Coefficient de corrélation 122
Coefficient de corrélation de l'échantillon 232
Coefficient de variation 126
Contrôle de qualité 128
Convergence de probabilité 130
Corrélation linéaire 142
Courbe en cloche 254
Covariance 132
Diagramme branche-et-feuilles 206
Diagramme en boîte 42
Distribution asymptotique 22
Distribution bêta 36
Distribution bidimensionnelle 60
Distribution binômiale 38
Distribution d'Erlang 76
Distribution de Gumbel 100
Distribution de khi carré 102
Distribution de Poisson 192
Distribution empirique 72
Distribution exponentielle 66
Distribution gamma 92
Distribution marginale 148
Distribution normale 172
Distribution uniforme 194
Ecart type 202
Echantillon 240
Echantillon aléatoire simple 188
Echantillon représentatif 196
Efficacité relative asymptotique 24
Erreur absolue moyenne 14
Erreur de première espèce 98

Erreur de seconde espèce 96
Espace échantillonnal..190
Espérance mathématique...152
Estimateur asymptotiquement efficace......................................20
Estimateur consistant 184
Estimateur efficace..62
Estimateur sans biais...164
Estimation de paramètres 178
Evénement aléatoire..200
Evénement certain...118
Evénements disjoints..52
Evénements indépendants 168
Extrapolation...68
Famille exponentielle...64
Fonction de densité 84
Fonction de perte..82
Fonction de répartition...88
Fonction de vraisemblance...90
Fonction puissance...86
Formule de Bayes 28
Fréquence..80
Graphique..94
Histogramme...108
Homogénéité 110
Hypothèse alternative..8
Hypothèse nulle 174
Indice de corrélation non-linéaire 112
Inférence de Bayes...30
Intervalle de confiance 114
Intervalle de confiance pour l'espérance..................................116
Loi des grandes nombres...252
Loi des grandes nombres de Bernoulli.......................................34
Médiane 154
Médiane de l'échantillon..226
Méthode des moments 158
Méthode du maximum de vraisemblance.................................156
Mode de l'échantillon 228
Moments de l'échantillon...238
Moyenne arithmetique 16
Moyenne de l'échantillon...230
Niveau de confiance..170

Niveau de signification..186
Nombre de degrés de liberté 44
Polygone de fréquences 180
Population 182
Probabilité 248
Probabilité à posteriori 10
Probabilité à priori 12
Probabilité conditionnelle 222
Région critique 136
Régression linéaire 144
Régression multiple...250
Quantile 138
Quantiles de l'échantillon 236
Quartet d'Anscombe 74
Quartile 140
Sondage avec remise 246
Statistique exhaustive...58
Statistique mathématique 150
Statistiques d'ordre 204
Table de contingence 208
Tabulation des données 210
Théorème centrale limite 46
Théorème de Basu..32
Théorème de Fisher-Neyman 78
Théorème de Glivenko-Cantelli 48
Théorie des probabilités 212
Test d'hypothèse 216
Test d'hypothèse de l'espérance mathématique 218
Test d'indépendance du khi carré 106
Test du khi carré...104
Test non parametrique...162
Test uniformément le plus puissant...220
Transformation linéaire des données.......................................146
Valeur modale 160
Valeurs aberrantes 70
Validité de l'ajustement 134
Variable aléatoire 198
Variables aléatoires indépendantes 166
Variance 54
Variance de l'échantillon 224
Visages de Chernoff...50

Содержание

Автокорреляция 26
Альтернативная гипотеза 8
Апостериорная вероятность 10
Априорная вероятность 12
Арифметическое среднее 16
Асимметрия 18
Асимптотическая относительная эффективность 24
Асимптотически эффективная оценка 20
Асимптотическое распределение 22
Бейесовский вывод 30
Бета-распределение 36
Биномиальное распределение 38
Вероятность 248
Выборка 240
Выборочная дисперсия 224
Выборочная медиана 226
Выборочная функция распределения 72
Выборочное среднее 230
Выборочные квантили 236
Выборочные моменты 238
Выборочный коэффициент корреляции 232
Выбросы 70
Гамма-распределение 92
Генеральная совокупность 182
Гистограмма 108
График 94
Двумерное распределение 60
Диаграмма размаха 42
Диаграмма “ствол с листьями” 206
Дисперсионный анализ 56
Дисперсия 54
Доверительный интерваль 114
Доверительный интерваль для ожидания 116
Достаточная статистика 58
Достоверное событие 118
Закон больших чисел 252
Закон больших чисел Бернулли 34
Индекс нелинейной корреляции 112

Квантиль 138
Квартет Энскомба 74
Квартиль 140
Ковариантность 132
Колоколообразная кривая 254
Контроль качества 128
Коэффициент асимметрии 120
Коэффициент вариации 126
Коэффициент корреляции 122
Коэффициент эксцесса 124
Критерий независимости хи-квадрат 106
Критерий согласия 134
Критерий хи-квадрат 104
Критическая область 136
Линейная корреляция 142
Линейная регрессия 144
Линейное преобразование данных 146
Лица Чернова 50
Маргинальное распределение 148
Математическая статистика 150
Математическое ожидание 152
Медиана 154
Метод максимального правдоподобия 156
Метод моментов 158
Множественная регрессия 250
Мода выборки 228
Модальное значение 160
Независимые случайные переменные 166
Независимые события 168
Непараметрический критерий 162
Несмещенная оценка 164
Несовместные события 52
Нормальное распределение 172
Нулевая гипотеза 174
Однородность 110
Оценка параметров 178
Ошибка второго рода 96
Ошибка первого рода 98
Повторная выборка 246
Полигон частот 180
Порядковые статистики 204

Признак 176
Проверка гипотезы 216
Проверка гипотезы о математическом ожидании 218
Простая случайная выборка 188
Пространство элементарных событий 190
Равномерно найболее мощный критерий 220
Равномерное распределение 194
Распределение Гумбеля 100
Распределение Пуассона 192
Распределение Эрланга 76
Репрезентативная выборка 196
Семейство экспоненциальных распределений 64
Случайная переменная 198
Случайное событие 200
Состоятельная оценка 184
Среднее абсолютное отклонение 14
Среднеквадратическое отклонение 202
Статистика биологическая 40
Сходимость по вероятности 130
Таблица сопряженности 208
Табуляция данных 210
Теорема Басу 32
Теорема Гливенко-Кантелли 48
Теорема Неймана-Фишера 78
Теория вероятностей 212
Уровень доверия 170
Уровень значимости 186
Условная вероятность 222
Формула Бейеса 28
Функция мощности критерия 86
Функция плотности распределения 84
Функция потерь 82
Функция правдоподобия 90
Функция распределения 88
Хи-квадрат распределение 102
Центральная предельная теорема 46
Частота 80
Число степеней свободы 44
Экспоненциальное распределение 66
Экстраполяция 68
Эффективная оценка 62

目录

安斯库姆四重奏……74
巴苏定理……32
贝它分布……36
贝叶斯公式……28
贝叶斯推论……30
备择假设……8
必然事件……118
边际分布……148
变异系数……126
标准偏差……202
伯努利大数定律……34
不相交事件……52
参数估计……178
充分统计量……58
次序统计量……204
大数定律……252
独立事件……168
独立随机变量……166
多重回归……250
厄兰分布……76
二项式分布……38
二型差误……96
二元分布……60
方差……54
方差分析……56
放回抽样……246
非参数检验……162
非线性相关指数……112
费舍尔-奈曼定理……78
分布函数……88
分层抽样……242
分位数……138
峰态系数……124
伽马分布……92

概率 248
概率论 212
冈贝儿分布..100
格里文科-坎泰利定理...48
后验概率 10
极大似然法..156
假设检验 216
简单随机样本..188
解消假设 174
矩法 158
具有代表性的样本..196
均匀分布 194
卡方独立性检验..106
卡方检验 104
卡方分布 102
两相抽样 244
列联表 208
密度函数 84
幂函数 86
判别区域 136
偏斜度 18
偏斜系数 120
频率 80
频率多边形..180
平均绝对误差..14
泊松分布 192
期望的假设检验..218
期望的置信区间..116
奇异值 70
切尔诺夫的面..50
取样分布 72
生物统计学..40
适合度检测..134
数理统计学..150
数学期望 152
四分位数 140

似然函数 90
算术中项 16
随机变量 198
随机事件 200
损失函数 82
特征 176
条件概率 222
同质性 110
统计图表 94
外推法 68
无偏估计量..164
渐近分布 22
渐近相对效率..24
渐近有效估计量..20
系统抽样 244
先验概率 12
显著性水平..186
线性回归 144
线性数据转换..146
线性相关 142
相关系数 122
相容估计量..184
箱形图 42
协方差 132
样本 240
样本方差 224
样本分位数..236
样本矩 238
样本空间 190
样本平均数..230
样本相关系数..232
样本中位数..226
样本众数 228
依概率收敛..130
一型差误 98
一致最大功效检验..220

有效估计量......62
整群抽样......243
正态分布......172
枝叶图......206
指数分布......66
指数族......64
质量控制......128
置信区间......114
置信系数......170
中线......154
中心极限定理......46
钟形曲线......254
众数......160
柱状图......108
资料制表......210
自相关......26
自由度数......44
总体......182

فهرس المحتويات

الاحتمال الأولي.......... 12
الاحتمال الشرطي.......... 222
الاحتمال اللاحق.......... 10
الاحتمال 248
الأحداث المستقلة 168
أحداث منفصلة.......... 52
الأحصاء الحيوي.......... 40
الأحصاء الرياضي.......... 150
الأحصاءات المرتبة.......... 204
الأحصاءة الكافية 58
اختبار الأكثر قوي بأنتظام.......... 220
إختبار الفرضيات لتوقع الرياضي.......... 218
اختبار الفرضيات 216
الاختبار اللابارامتري 162
اختبار حسن المطابقة.......... 134
اختبارمربع كاي.......... 104
اخنبارمربع كاي للأستقلاية.......... 106
الارتباط الخطي 142
الاستكمال.......... 68
الالتواء.......... 18
تقدير المعلمات.......... 178
الانحدار الخطي 144
الانحدار المتعدد 250
الانحراف المعياري.......... 202
برنولي قانون الأعداد الكبيرة.......... 34
تباين العينة 224
التباين 54
تبويب البيانات 210
التجانس.......... 110
تجزئ العينة.......... 236
تحليل التباين 56
تحليل انحدار انسكوب.......... 74
التحويل الخطي من البيانات 146
التغاير.......... 132
التقارب الاحتمالي.......... 130
تقدير الكفاءة.......... 62
تقدير غير متحيز 164
تقدير متناسق.......... 184

التكرار 80
عينة تمثيلية 196
توزيع ارلانج 76
التوزيع الآسي 66
التوزيع التقاربي 22
التوزيع الثنائي 60
التوزيع الطبيعي 172
توزيع المعاينة 72
التوزيع المنتظم 194
التوزيع الهامشي 148
توزيع بواسون 192
توزيع بيتا 36
توزيع جاما 92
توزيع جامبل 100
توزيع ذي الحدين 38
توزيع كاي تربيع 102
التوقع الرياضي 152
جدول المطابقة 208
الحدث العشوائي 200
الحدث المؤكد 118
الخاصية 176
الخطأ من النوع الأول 98
الخطأ من النوع الثاني 96
دالة الارجحية 90
دالة التوزيع 88
دالة الخساره 82
دالة كثافة الاحتمال 84
الربيع 140
الرسم البيان 94
الرسم بطريقة الساق والورقة 206
السكان 182
صيغة بايز 28
طريقة الارجحية القصوى 156
طريقة العزوم 158
العائلة الأسية 64
عدد درجات الحرية 44
عزوم العينة 238
العشير 138
العينة العشوائية البسيطة 188
العينة 240
فترة الثقة للمتوسط 116

فترة الثقة... 114
الفرضية البديلة 8
الفرضية المبدئية... 174
فضاء العينة 190
قانون الأعداد الكبيرة 252
القيم المتطرفة.. 70
الكفاءة النسبية التقربية... 24
لترابط التلقائي 26
المتغير العشوائي 198
المتغيرات العشوائية المستقلة.. 166
متوسط الانحرافات المطلق 14
المتوسط الحسابي.. 16
متوسط العينة.. 230
مخطط الصندوق 42
المدرج التكراري... 108
مراقبة الجودة 128
مستوى الثقة... 170
مستوى الدلالة... 186
المضلع التكراري 180
معادلة بيز 30
معامل الارتباط العينة.. 232
معامل الارتباط 122
معامل الالتواء 120
معامل التغير 126
معامل التفرطح 124
المعاينة مع الارجاع 246
المقدر الكافئ المتقارب 20
المنحني الجرسي 254
منحني القوي.. 86
المنطقة الحرجة.. 136
منوال العينة 228
المنوال 160
مؤشر الارتباط الابحبائي 112
نظرية الاحتمالات.. 212
نظرية الغاية المركزية في الاحصاء 48
نظرية النهاية المركزية.. 46
نظرية باسو.. 32
نظرية فيشر- نيمان 78
وجوه تشرنوف 50
وسيط العينة 226
الوسيط 154

Πίνακας περιεχομένων

A posteriori πιθανότητα 10
A priori πιθανότητα 12
Ακραίες τιμές 70
Αμερόληπτη εκτιμήτρια 164
Ανάλυση διακύμανσης 56
Ανεξάρτητα γεγονότα 168
Ανεξάρτητες τυχαίες μεταβλητές 166
Αντιπροσωπευτικό δείγμα 196
Απλό τυχαίο δείγμα 188
Αποτελεσματική εκτιμήτρια 62
Αριθμητικός μέσος 16
Αριθμός των βαθμών ελευθερίας 44
Ασυμμετρία 18
Ασυμπτωτικά αποτελεσματική εκτιμήτρια 20
Ασυμπτωτική κατανομή 22
Ασυμπτωτική σχετική αποτελεσματικότητα 24
Αυτοσυσχέτιση 26
Βέβαιο γεγονός 118
Βιοστατιστική 40
Γραμμική παλινδρόμηση 144
Γραμμική συσχέτιση 142
Γραμμικός μετασχηματισμός δεδομένων 146
Γράφημα 94
Δείγμα 240
Δειγματική κατανομή 72
Δειγματική κορυφή 228
Δειγματικός μέσος 230
Δειγματικός συντελεστής συσχέτισης 232
Δειγματοληψία με επανατοποθέτηση 246
Δειγματοχώρος 190
Δείκτης μή-γραμμικής συσχέτισης 112
Δεσμευμένη πιθανότητα 222
Διάγραμμα Box plot 42
Διάγραμμα μίσχου-φύλλου 206
Διακριτά γεγονότα 52
Διάμεσος 154
Διάμεσος δείγματος 226
Διασπορά 54

Διασπορά δείγματος...224
Διάστημα εμπιστοσύνης..114
Διάστημα εμπιστοσύνης για το μέσο όρο 116
Διμεταβλητή κατανομή..60
Διωνυμική κατανομή 38
Εκθετική κατανομή...66
Εκθετική οικογένεια...64
Εκτίμηση παραμέτρων...178
Έλεγχος ανεξαρτησίας με το τεστ Χι-τετράγωνο 106
Έλεγχος καλής προσαρμογής..134
Έλεγχος ποιότητας...128
Έλεγχος υπόθεσης..216
Έλεγχος υπόθεσης της μαθηματικής προσδοκίας.............................218
Έλεγχος Χι-τετράγωνο..104
Εναλλακτική υπόθεση...8
Επαρκής στατιστική..58
Επίπεδο εμπιστοσύνης..170
Επίπεδο σημαντικότητας..186
Θεώρημα Basu...32
Θεώρημα Glivenko-Cantelli 48
Θεώρημα Neyman-Fisher 78
Θεωρία πιθανοτήτων..212
Ιστόγραμμα 108
Καμπύλη ισχύος...86
Κανόνας του Bayes 28
Κανονική κατανομή..172
Κατανομή Βήτα 36
Κατανομή Γάμα 92
Κατανομή Erlang 76
Κατανομή Gumbel..100
Κατανομή Poisson 192
Κατανομή Χι-τετράγωνο 102
Κεντρικό οριακό θεώρημα..46
Κορυφή 160
Κουαρτέτο του Anscombe 74
Κρίσιμη περιοχή..136
Κωδωνοειδής καμπύλη 254
Μαθηματική στατιστική 150
Μέθοδος μέγιστης πιθανοφάνειας...156
Μέθοδος των ροπών 158
Μέση απόλυτη απόκλιση 14

Μή-παραμετρικό τεστ..162
Μηδενική υπόθεση..174
Νόμος των μεγάλων αριθμών..252
Νόμος των μεγάλων αριθμών του Bernoulli..34
Ομοιογένεια..110
Ομοιόμορφα ισχυρότερος έλεγχος..220
Ομοιόμορφη κατανομή..194
Περιθώρια κατανομή..148
Πιθανότητα..248
Πίνακας συνάφειας..208
Πινακοποίηση δεδομένων..210
Πληθυσμός..182
Πολλαπλή παλινδρόμηση..250
Πολύγωνο συχνοτήτων..180
Ποσοστημόρια του δείγματος..236
Ποσοστημόριο..138
Προέκταση..68
Προσδοκία..152
Πρόσωπα του Chernoff..50
Ροπές του δείγματος..238
Σφάλμα δεύτερου τύπου..96
Σφάλμα πρώτου τύπου..98
Στατιστικές τάξης μεγέθους..204
Σύγκλιση σε πιθανότητες..130
Συμπερασματολογία κατά Bayes..30
Συνάρτηση απώλειας..82
Συνάρτηση κατανομής..88
Συνάρτηση πιθανοφάνειας..90
Συνάρτηση πυκνότητας πιθανότητας..84
Συνδιακύμανση..132
Συνεπής εκτιμήτρια..184
Συντελεστής ασυμμετρίας..120
Συντελεστής διακύμανσης..126
Συντελεστής κύρτωσης..124
Συντελεστής συσχέτισης..122
Συχνότητα..80
Τεταρτημόριο..140
Τυπική απόκλιση..202
Τυχαία μεταβλητή..198
Τυχαίο γεγονός..200
Χαρακτηριστικό..176

Inhaltsverzeichnis

A-posteriori-Wahrscheinlichkeit 10
A-priori-Wahrscheinlichkeit 12
Ablehnungsbereich 136
Anscombe-Quartett 74
Anzahl der Freiheitsgrade 44
Arithmetisches Mittel 16
Asymptotisch effizienter Schätzer 20
Asymptotische relative Effizienz 24
Auftabellierung von Daten 210
Ausreißer 70
Autokorrelation 26
Basuscher Lehrsatz 32
Bayes-Formel 28
Bayessche Schlußweise 30
Bedingte Wahrscheinlichkeit 222
Bernoullisches Gesetz der großen Zahlen 34
Betaverteilung 36
Binomialverteilung 38
Biostatistik 40
Boxplot 42
Chernoff-Gesichter 50
Chi-Quadrat-Test 104
Chi-Quadrat-Unabhängigkeitstest 106
Chi-Quadrat-Verteilung 102
Disjunkte Ereignisse 52
Dispersion 54
Durchschnittliche Abweichung 14
Effizienter Schätzer 62
Einfache Zufallsstichprobe 188
Einheitlichkeit 110
Ergebnisraum 190
Erlangsche Verteilung 76
Erwartungstreuer Schätzer 164
Erwartungswert 152
Exponentialfamilie 64
Exponentialverteilung 66
Extrapolation 68
Exzeßkoeffizient 124

Fehler erster Art 98
Fehler zweiter Art 96
Fundamentalsatz der Statistik 48
Gamma-Verteilung 92
Gegenhypothese 8
Gesetz der großen Zahlen 252
Gleichmäßig trennschärfster Test 220
Gleichverteilung 194
Glockenkurve 254
Grafik 94
Grenzverteilung 22
Gumbel-Verteilung 100
Güte der Anpassung 134
Gütefunktion 86
Häufigkeit 80
Häufigkeitspolygon 180
Histogramm 108
Hypothesentest 216
Hypothesentest für den Erwartungswert 218
Konfidenzintervall 114
Konfidenzintervall für Erwartungswert 116
Konfidenzniveau 170
Konsistenter Schätzer 184
Kontingenztafel 208
Korrelationskoeffizient 122
Korrelationskoeffizient einer Stichprobe 232
Kovarianz 132
Likelihood-Funktion 90
Lineare Datentransformation 146
Lineare Korrelation 142
Lineare Regression 144
Matematische Statistik 150
Maximum-Likelihood-Methode 156
Median 154
Median einer Stichprobe 226
Mehrfachregression 250
Merkmal 176
Modalwert 160
Modus einer Stichprobe 228
Momenten der Stichprobe 238
Momentenmethode 158

Nichtlinearer Korrelationskoeffizient..112
Nichtparametrische Test 162
Normalverteilung 172
Nullhypothese 174
Ordnungsstatistik 204
Parameterschätzung 178
Poissonsche Verteilung 192
Population 182
Qualitätsregelung 128
Quantil 138
Quartil 140
Randverteilung 148
Repräsentative Stichprobe...196
Satz von Fisher-Neyman 78
Schiefe 18
Schiefekoeffizient 120
Sicheres Ereignis...118
Signifikanzniveau..186
Stamm-und-Blätter-Diagramm 206
Standardabweichung 202
Stichprobe 240
Stichprobe mit Zurücklegen..246
Stichprobenmittelwert...230
Stichprobenquantilen 236
Stichprobenverteilung 72
Stochastische Konvergenz 130
Streuung einer Stichprobe...224
Suffiziente Statistik 58
Unabhängige Ereignisse..168
Unabhängige Zufallsvariablen 166
Varianzanalyse 56
Variationskoeffizient...126
Verlustfunktion 82
Verteilungsfunktion 88
Wahrscheinlichkeit..248
Wahrscheinlichkeitsdichtefunktion..84
Wahrscheinlichkeitstheorie...212
Zentraler Grenzwertsatz...46
Zufallsereignis...200
Zufallsvariable 198
Zweidimensionale Verteilung 60

Tabla de contenido

Análisis de varianza 56
Asimetría 18
Autocorrelación 26
Bioestadística 40
Característica 176
Caras de Chernoff 50
Coeficiente de asimetría 120
Coeficiente de correlación 122
Coeficiente de correlación de la muestra 232
Coeficiente de exceso 124
Coeficiente de variación 126
Control de la calidad 128
Convergencia en probabilidad 130
Correlación lineal 142
Covariancia 132
Cuantiles de una muestra 236
Cuantilo 138
Cuarteto de Anscombe 74
Cuartilo 140
Curva campanular 254
Desviación estándar 202
Desviación promedia 14
Diagrama de caja 42
Diagrama de tallos y hojas 206
Dispersión 54
Dispersión de la muestra 224
Distribución asintótica 22
Distribución beta 36
Distribución binomial 38
Distribución de dos variables 60
Distribución de Erlang 76
Distribución de Gumbel 100
Distribución de ji-cuadrado 102
Distribución de Poisson 192
Distribución del muestreo 72
Distribución exponencial 66
Distribución gamma 92
Distribución marginal 148

Distribución normal 172
Distribución uniforme 194
Eficiencia relativa asintótica 24
Error de tipo dos 96
Error de tipo uno 98
Espacio muestral 190
Esperanza matemática 152
Estadística matemática 150
Estadísticas de orden 204
Estadístico suficiente 58
Estimación de parámetros 178
Estimador asintóticamente eficiente 20
Estimador consistente 184
Estimador eficiente 62
Estimador sin sesgo 164
Evento aleatorio 200
Evento deterministico 118
Eventos independientes 168
Extrapolación 68
Familia exponencial 64
Formula de Bayes 28
Frecuencia 80
Función de densidad 84
Función de distribución 88
Función de pérdida 82
Función de potencia 86
Función de verosimilitud 90
Gráfico 94
Hipótesis alternativa 8
Hipótesis de nulidad 174
Histograma 108
Homogeneidad 110
Índice de correlación no lineal 112
Inferencia de Bayes 30
Intervalo de confianza 114
Intervalo de confianza para la esperanza 116
Ley de los grandes números 252
Ley de los grandes números de Bernoulli 34
Media aritmética 16
Media de la muestra 230
Mediana 154

Mediana de la muestra 226
Método de los momentos 158
Método del máximo de verosimilitud 156
Moda de la muestra 228
Momentos de la muestra 238
Muestra 240
Muestra aleatoria simple 188
Muestra representativa 196
Muestreo con reposición 246
Nivel de confianza 170
Nivel de significación 186
Número de grados de libertad 44
Población 182
Polígono de frecuencias 180
Probabilidad 248
Probabilidad a posteriori 10
Probabilidad a priori 12
Probabilidad condicional 222
Prueba de bondad de ajuste 134
Prueba de independencia de ji-cuadrado 106
Prueba de ji-cuadrado 104
Prueba no paramétrica 162
Prueba uniformemente más poderosa 220
Región crítica 136
Regresión lineal 144
Regresión múltiple 250
Sucesos disjuntos 52
Tabla de contingencia 208
Tabulación de datos 210
Teorema central del límite 46
Teorema de Basu 32
Teorema de Glivenko-Cantelli 48
Teorema de Fisher-Neyman 78
Teoría de la probabilidad 212
Test de hipótesis 216
Test de hipótesis sobre de la esperanza matemática 218
Transformación lineal de datos 146
Valor modal 160
Valores extremos 70
Variable aleatoria 198
Variables aleatorias independientes 166

Indice dei contenuti

Analisi della varianza 56
Asimmetria 18
Autocorrelazione 26
Biostatistica 40
Bontà dell'adattamento 134
Campionamento con ripetizione 246
Campione 240
Campione casuale semplice 188
Campione rappresentativo 196
Carattere 176
Coefficiente di asimmetria 120
Coefficiente di correlazione 122
Coefficiente di correlazione campionaria 232
Coefficiente di eccesso 124
Coefficiente di variazione 126
Controllo di qualità 128
Convergenza in probabilità 130
Correlazione lineare 142
Covarianza 132
Curva a campana 254
Diagramma a scatola 42
Diagramma ramo-foglia 206
Dispersione 54
Dispersione campionaria 224
Distribuzione asintotica 22
Distribuzione beta 36
Distribuzione binomiale 38
Distribuzione bivariata 60
Distribuzione campionaria 72
Distribuzione chi-quadrato 102
Distribuzione di Erlang 76
Distribuzione di Gumbel 100
Distribuzione di Poisson 192
Distribuzione esponenziale 66
Distribuzione gamma 92
Distribuzione marginale 148
Distribuzione normale 172
Distribuzione uniforme 194

Efficienza relativa asintotica 24
Errore di prima specie 98
Errore di seconda specie 96
Estimatore efficiente 62
Estrapolazione 68
Eventi disgiunti 52
Eventi indipendenti 168
Evento accidentale 200
Evento certo 118
Famiglia esponenziale 64
Formula di Bayes 28
Frequenza 80
Fronti di Chernoff 50
Funzione di densità 84
Funzione di perdita 82
Funzione di potenza 86
Funzione di ripartizione 88
Funzione di verosimiglianza 90
Grafico 94
Indice di correlazione non lineare 112
Inferenza Bayesiana 30
Intervallo di confidenza 114
Intervallo di confidenza per la speranza 116
Ipotesi alternativa 8
Ipotesi nulla 174
Istogramma 108
Legge dei grandi numeri 252
Legge dei grandi numeri di Bernoulli 34
Livello di confidenza 170
Livello di significatività 186
Media aritmetica 16
Media campionaria 230
Mediana 154
Mediana campionaria 226
Metodo dei momenti 158
Metodo della massima verosimiglianza 156
Moda campionaria 228
Momenti campionari 238
Numero di gradi di libertà 44
Omogeneità 110
Poligono di frequenza 180

Popolazione 182
Probabilità 248
Probabilità a posteriori 10
Probabilità a priori 12
Probabilità condizionale 222
Quantile 138
Quantili campionari 236
Quartetto di Anscombe 74
Quartile 140
Regione critica 136
Regressione lineare 144
Regressione multipla 250
Scarto medio assoluto 14
Scarto quadratico medio 202
Spazio campionario 190
Speranza matematica 152
Statistica matematica 150
Statistica sufficiente 58
Statistiche d'ordine 204
Stima dei parametri 178
Stimatore asintoticamente efficiente 20
Stimatore consistente 184
Stimatore non distorto 164
Tabulazione dei dati 210
Tavola di contingenza 208
Teorema del limite centrale 46
Teorema di Basu 32
Teorema di Fisher-Neyman 78
Teorema di Glivenko-Cantelli 48
Teoria della probabilità 212
Test chi-quadrato 104
Test chi-quadrato di indipendenza 106
Test di verifica d'ipotesi 216
Test di verifica d'ipotesi sulla speranza matematica 218
Test non parametrico 162
Test uniformemente il più potente 220
Trasformazione lineare dei dati 146
Valore modale 160
Valori anomali 70
Variabile aleatoria 198
Variabili casuali indipendenti 166

Obsah

Alternatívna hypotéza 8
Analýza rozptylu 56
Anscombova regresia 74
Aposteriórna pravdepodobnosť 10
Apriórna pravdepodobnosť 12
Aritmetický priemer 16
Asymptotická relatívna efektívnosť 24
Asymptotické rozdelenie 22
Asymptoticky efektívny odhad 20
Autokorelácia 26
Basuova veta 32
Bayesova formula 28
Bayesovská inferencia 30
Bernoulliho zákon veľkých čísel 34
Beta rozdelenie 36
Binomické rozdelenie 38
Bioštatistika 40
Centrálna limitná veta 46
Charakteristika 176
Chernoffove tváre 50
Chí-kvadrát rozdelenie 102
Chí-kvadrát test 104
Chí-kvadrát test nezávislosti 106
Chyba druhého typu 96
Chyba prvého typu 98
Číslicový dendrogram 206
Distribučná funkcia 88
Dvojrozmerné rozdelenie 60
Efektívny odhad 62
Erlangovo rozdelenie 76
Exponenciálne rozdelenie 66
Exponenciálny systém 64
Extrapolácia 68
Fisherova-Neymannova veta 78
Frekvencia 80
Frekvenčný polygón 180
Gamma rozdelenie 92
Glivenkova-Cantelliho veta 48

Graf 94
Gumbelovo rozdelenie 100
Histogram 108
Hladina významnosti 186
Homogenita 110
Hustota pravdepodobnosti 84
Index nelineárne korelácie 112
Istá udalosť 118
Jednoduchý náhodný výber 188
Koeficient rozptylu 126
Koeficient šikmosti 120
Koeficient špicatosti 124
Konfidenčná hladina 170
Konfidenčný interval 114
Konfidenčný interval pre strednú hodnotu 116
Kontingenčná tabuľka 208
Kontrola kvality 128
Konvergencia podľa pravdepodobnosti 130
Konzistentný odhad 184
Korelačný koeficient 122
Kovariancia 132
Krabicový graf 42
Kritická oblasť 136
Kvantil 138
Kvartil 140
Lineárna korelácia 142
Lineárna regresia 144
Lineárna transformácia dát 146
Marginálne rozdelenie 148
Matematická štatistika 150
Medián 154
Metóda maximálnej vierohodnosti 156
Mocninová funkcia 86
Modus 160
Momentová metóda 158
Náhodná premenná 198
Náhodná udalosť 200
Neparametrický test 162
Nevychýlený odhad 164
Nezávislé náhodné premenné 166
Nezávislé udalosti 168

Nezlučiteľné udalosti 52
Normálne rozdelenie 172
Nulová hypotéza 174
Odhad parametrov 178
Odľahlé pozorovania 70
Počet stupňov voľnosti 44
Podmienená pravdepodobnosť 222
Poissonovo rozdelenie 192
Populácia 182
Poriadkové štatistiky 204
Postačujúca štatistika 58
Pravdepodobnosť 248
Priestor elementárnych udalostí 190
Reprezentatívny výber 196
Rovnomerne najsilnejší test 220
Rovnomerné rozdelenie 194
Rozptyl 54
Stratová funkcia 82
Stredná absolútna odchýlka 14
Stredná hodnota 152
Šikmosť 18
Štandartná odchýlka 202
Tabelovanie údajov 210
Teória pravdepodobnosti 212
Test dobrej zhody 134
Testovanie hypotéz 216
Testovanie hypotézy o strednej hodnote 218
Viacrozmerná regresia 250
Vierohodnostná funkcia 90
Výber 240
Výber s opakovaním 246
Výberové kvantily 236
Výberové momenty 238
Výberové rozdelenie 72
Výberový korelačný koeficient 232
Výberový medián 226
Výberový modus 228
Výberový priemer 230
Výberový rozptyl 224
Zákon veľkých čísel 252
Zvonovitá krivka 254

Kazalo

Alternativna domneva 8
Analiza variance 56
Anscombeov kvartet 74
Apriorna verjetnost 12
Aritmetična sredina 16
Asimetrija 18
Asimptotičnim distribucija 22
Asimptotičnim relativna učinkovitost 24
Asimptotično učinkovita cenilka 20
Avtokorelacija 26
Basu izrek 32
Bayesova formula 28
Bayesovo sklepanje 30
Bernoullijev zakon velikih števil 34
Beta porazdelitev 36
Binomska porazdelitev 38
Biostatistika 40
Centralni limitni izrek 46
Chernoff obrazi 50
Disjunktni dogodki 52
Disperzija 54
Dosledna cenilka 184
Dvorazsežna porazdelitev 60
Eksponentna družina 64
Eksponentna porazdelitev 66
Ekstrapolacija 68
Enakomerna porazdelitev 194
Enakomerno najmočnejši preizkus 220
Enostavni slučajni vzorec 188
Enovitost 110
Erlangova porazdelitev 76
Fisher-Neymanov izrek 78
Frekvenca 80
Frekvenčni poligon 180
Funkcija gostote 84
Funkcija izgube 82
Funkcija moči preizkusa 86
Funkcija verjetja 90

Gamma porazdelitev 92
Glivenko-Cantellijev izrek 48
Gotov dogodek 118
Grafikon 94
Gumbelova porazdelitev 100
Hi-kvadrat porazdelitev 102
Hi-kvadrat preizkus 104
Hi-kvadrat preizkus hipoteze neodvisnosti 106
Histogram 108
Histogram s števílkami 206
Indeks nelinearne korelacije 112
Interval zaupanja 114
Interval zaupanja za upanje 116
Koeficient asimetrije 120
Koeficient korelacije 122
Koeficient presežka 124
Koeficient variacije 126
Kontingenčna tabela 208
Kontrola kakovosti 128
Konvergenca v verjetnosti 130
Kovarianca 132
Kritično območje 136
Krivulja v obliki zvona 254
Kvantil 138
Kvartil 140
Lastnost 176
Linearna korelacija 142
Linearna regresija 144
Linearna transformacija podatkov 146
Marginalna porazdelitev 148
Matematična statistika 150
Matematično upanje 152
Mediana 154
Metoda momentov 158
Metoda največjega verjetja 156
Modus 160
Multipla regresija 250
Napaka druge vrste 96
Napaka prve vrste 98
Neodvisne slučajne spremenljivke 166
Neodvisni dogodki 168

Neparametrični preizkus 162
Nepristranska cenilka 164
Ničelna domneva 174
Normalna porazdelitev 172
Ocenjevanje parametrov 178
Osamelci 70
Pogojna verjetnost 222
Poissonova porazdelitev 192
Populacija 182
Porazdelitvena funkcija 88
Posteriorna verjetnost 10
Povprečni absolutni odklon 14
Preizkusi skladnosti 134
Preizkušanje domnev 216
Preizkušanje domneve o pričakovani vrednosti 218
Prostor vzorcev 190
Raven zaupanja 170
Reprezentativni vzorec 196
Slučajna spremenljivka 198
Slučajni dogodek 200
Standardni odklon 202
Stopnja značilnosti 186
Škatla ploskvi 42
Število stopinj prostosti 44
Tabelovanje podatkov 210
Učinkovita cenilka 62
Verjetnost 248
Verjetnostni račun 212
Vrstilne statistike 204
Vzorčna disperzija 224
Vzorčna mediana 226
Vzorčna porazdelitev 72
Vzorčna sredina 230
Vzorčenje s ponavljanjem 246
Vzorčni koeficient korelacije 232
Vzorčni kvantili 236
Vzorčni modus 228
Vzorčni momenti 238
Vzorec 240
Zadostna statistika 58
Zakon velikih števil 252

Tartalomjegyzék

A matematikai statisztika alaptétele 48
A minta momentumai 238
A nagy számok Bernoulli-féle törvénye 34
A nagy számok törvénye 252
A posteriori valószinűség 10
A priori valószínűség 12
A várható érték konfidenciaintervalluma 116
A várható értékre vonatkozó hipotézisvizsgálat 218
Adatok táblázata 210
Anscombe kvartett 74
Asszimetria mutató 120
Aszimptotikus eloszlás 22
Aszimptotikus relatív hatékonyság 24
Aszimptotikusan hatékony becslés 20
Átlagos abszolút hiba 14
Autokorreláció 26
Basu-tétel 32
Bayes-féle következtetés 30
Bayes-tétel 28
Béta-eloszlás 36
Binomiális eloszlás 38
Biostatisztika 40
Biztos esemény 118
Chernoff arca 50
Doboz ábrázolás 42
Egyenletes eloszlás 194
Egyenletesen legerősebb próba 220
Egyszerű véletlen minta 188
Elégséges statisztika 58
Ellenhipotézis 8
Eloszlásfüggvény 88
Elsőfajú hiba 98
Empirikus eloszlás 72
Erlang-eloszlás 76
Eseménytér 190
Exponenciális eloszlás 66
Exponenciális eloszláscsalád 64
Extrapoláció 68

Feltételes valószínűség ... 222
Ferdeség ... 18
Fisher-Neyman-tétel ... 78
Független események ... 168
Független valószínűségi változók ... 166
Gamma-eloszlás ... 92
Gráf ... 94
Gumbel-eloszlás ... 100
Gyakoriság ... 80
Gyakorisági poligon ... 180
Haranggörbe ... 254
Hatásos becslés ... 62
Hatványfüggvény ... 86
Hipotézisvizsgálat ... 216
Hisztogram ... 108
Homogenitás ... 110
Illeszkedés ... 134
Jellemző ... 176
Kétváltozós eloszlás ... 60
Khi-négyzet eloszlás ... 102
Khi-négyzet függetlenségi próbá ... 106
Khi-négyzet próba ... 104
Kizáró események ... 52
Konfidencia szint ... 170
Konfidenciaintervallum ... 114
Kontingencia táblázat ... 208
Konzisztens becslés ... 184
Korrelációs együttható ... 122
Kovariancia ... 132
Központi határeloszlás-tétel ... 46
Kritikus tartomány ... 136
Kvantilis ... 138
Kvartilis ... 140
Laposság ... 124
Leveles ág bemutatás ... 206
Likelihood-függvény ... 90
Lineáris adatalakítás ... 146
Lineáris korreláció ... 142
Lineáris regresszió ... 144
Marginális eloszlás ... 148
Másodfajú hiba ... 96

Matematikai statisztika 150
Maximum-likelihood módszer 156
Medián 154
Minta 240
Minta kvantilis 236
Minta médian 226
Minta módusz 228
Mintabeli korrelációs együttható 232
Mintaközép 230
Minőségellenőrzés 128
Modális érték 160
Momentum módszer 158
Nemlineáris korrelációs index 112
Nemparaméteres próbá 162
Normális eloszlás 172
Nullhipotézis 174
Paraméterbecslés 178
Poisson-eloszlás 192
Relatív szórás 126
Rendstatisztikák 204
Reprezentatív minta 196
Sokaság 182
Standardeltérés 202
Sűrűségfüggvény 84
Szabadságfok 44
Számtani közép 16
Szélsőértékek 70
Szignifikanciaszint 186
Szórás 54
Szórásanalízis 56
Sztochasztikus konvergencia 130
Tapasztalati szórás 224
Torzítatlan becslés 164
Többváltozós regresszió 250
Valószínűség 248
Valószínűségi változó 198
Valószínűségszámítás 212
Várható érték 152
Véletlen esemény 200
Veszteségfüggvény 82
Visszatevéses mintavétel 246

Izabrani statistički termini su prevedeni na dvanaest jezika – pet velikih svetskih jezika (engleski, francuski, ruski, kineski i arapski) grčki kao jedan od jezika koji su postavili temelje evropske civilizacije i šest jezika zemalja partnera Srbije u okviru Tempus projekta MAS (nemački, španski, italijanski, slovački, slovenački i mađarski).

Korist od ovog rečnika, koji je nešto između rečnika, priručnika i udžbenika će, nadam se, imati ne samo oni koji se prvi put sreću sa statistikom, nego i oni koji u izvesnoj meri poznaju materiju o kojoj je reč.

iz Predgovora

ISBN 978-86-6161-188-9