PROIZVODNJA VITAMINA U SVIJETU I U HRVATSKOJ

Tržište vitamina jedno je od najpropulzivnijih svjetskih tržišta. Već prema potrebama za pojedinim vitaminima, oko 20···60% svjetske proizvodnje vitamina jesu farmaceutski proizvodi, 10···30% čine dodatci prehrambenim proizvodima, a 10···80% dodatci stočnoj hrani (krmivima). Današnja godišnja proizvodnja vitamina u svijetu procjenjuje se na ~100 000 t. Od toga oko 40% otpada na biotehnološku proizvodnju (vitamini B_2 i B_{12} potpuno, a vitamini C i E te provitamin ergosterol djelomično). Među najvažnijim su svjetskim proizvođačima Hoffmann-Laroche (Švicarska), BASF (Njemačka), Takeda (Japan), Sumitomo (Japan), Riken Vitamin (Japan) i Merck (SAD).

U Hrvatskoj vitamine proizvodi samo tvornica PLIVA iz Zagreba, koja pripada među najveće proizvođače vitamina C i vitamina B_6 u svijetu.

Proizvodnja vitamina C započela je u tvornici PLIVA u Zagrebu 1953. u poluindustrijskim količinama od nekoliko desetaka kilograma. Već 1956. proizvodilo se nekoliko tona, a godišnji kapacitet novoizgrađenog postrojenja u 1961. iznosio je 100 t. Proširivanjem postrojenja i poboljšavanjem proizvodnog postupka dostignut je proizvodni kapacitet s više od 1300 t godišnje. Glavnina vitamina C prodaje se na zapadnoeuropskom tržištu i u SAD jer zadovoljava vrhunskim zahtjevima kvalitete.

Industrijska proizvodnja vitamina B_6 započela je u novootvorenom pogonu tvornice PLIVA u Zagrebu 1959. godine. Početni kapacitet proizvodnje dosizao je tek 300 kg godišnje. Tijekom proteklih godina postupak je znatno poboljšan pa je nakon nekoliko rekonstrukcija proizvodnog postrojenja godišnji kapacitet proizvodnje povećan na ~70 t. Vitamin B_6 također udovoljava najstrožim svjetskim propisima o kvaliteti, pa se veći dio proizvodnje izvozi na zapadnoeuropsko i američko tržište.

LIT.: *P. Gyorgy, W. N. Pearson*, The Vitamins, Vol. 6–7. Academic Press, New York [2]1967. – *W. H. Sebrell, Jr., R. S. Harris*, The Vitamins, Vol. 1–5. Academic Press, New York 1967–1972. – *O. Isler, G. Brubacher*, Vitamine I, Fettlösliche Vitamine. Thieme, Stuttgart 1982. – *W. Friedrich*, Vitamins. Walter deGruyter, Berlin –New York 1988. – *E. J. Vandamme*, Biotechnology of Vitamins, Pigments and Growth Factors. Elsevier Applied Science LTD., London–New York 1989. – *Z. Kniewald*, Vitamini i hormoni: Proizvodnja i primjena. Hrvatska sveučilišna naklada, Zagreb 1992.

J. Vorkapić-Furač

VJEROJATNOST, matematički pojam kojim se kvantitativno (brojčano) opisuje slučajnost pojavljivanja uočenog događaja.

Teorija vjerojatnosti kao znanstvena disciplina objašnjava tzv. slučajne pojave. Povijesno gledajući, slučajne su pojave najprije uočene i ozbiljnije analizirane u igrama na sreću (P. Fermat, B. Pascal, XVII. st.), kao što su npr. različite igre s kartama, igraćim kockama i sl. No, ubrzo se uvidjelo da se slučajnost pojavljuje i u drugim ljudskim djelatnostima, pa i u mnogim prirodnim pojavama.

Nakon velikih otkrića u prirodnim znanostima u XVII. st. mnogi znanstvenici nisu uopće priznavali stvarno postojanje slučajnih pojava, već su smatrali da se, u načelu, svaka pojava može objasniti uzročnom vezom između početnih i konačnih stanja. Promatranje neke prirodne pojave kao slučajne objašnjavalo se samo nedovoljnim poznavanjem važnih činjenica o vezama između sadašnjeg i budućeg stanja te pojave. Očekivalo se da će znanost omogućiti spoznavanje tih veza, tako da je promatranje neke pojave kao slučajne relativan pojam, koji je uvjetovan stupnjem razvoja znanosti.

Za razliku od determinističkog stajališta, probabilističko stajalište priznaje stvarno postojanje pojava i procesa koji se podvrgavaju statističkim zakonitostima. To je stajalište omogućilo i poticalo razvoj matematičke teorije, koja će u najopćenitijem obliku izraziti zakone slučajnosti.

Nakon P. Fermata (1601–1665), B. Pascala (1623–1662) i C. Huygensa (1629–1695), koji su prvi počeli matematički obrađivati problem slučajnosti, pa se stoga i smatraju osnivačima teorije vjerojatnosti, razvoju teorije vjerojatnosti uvelike je pridonio J. Bernoulli (1654–1705), koji je napisao knjigu *Umijeće pogađanja*, lat. *Ars conjectandi*, gdje je formulirao i dokazao jedan od prvih graničnih teorema, tzv. *Bernoullijev zakon velikih brojeva*. Zatim slijede radovi A. De Moivrea (1667–1754), P. S. Laplacea (1749–1827), K. F. Gaussa (1777–1855) i S. D. Poissona (1781–1840), kojima se počinje izgrađivati teorija vjerojatnosti kao posebna znanstvena disciplina. Njezinu širenju i produbljivanju znatno pridonose i znanstvenici tzv. *ruske škole teorije vjerojatnosti*: P. L. Čebišev (1821–1894), A. A. Markov (1856–1922) i A. M. Ljapunov (1857–1918), te franc. matematičar E. Borel (1871–1952), koji je dokazao tzv. *jaki zakon velikih brojeva*.

Suvremeni razvoj teorije vjerojatnosti počinje s prvim pokušajima njezina aksiomatiziranja (S. N. Bernstein (1880–1968), R. Mises (1883–1953) i E. Borel). Konačno je A. N. Kolmogorov (1903–1987) dao danas općeprihvaćenu aksiomatiku teorije vjerojatnosti (1933), koja omogućuje njezinu izgradnju kao apstraktne matematičke discipline, tijesno povezane s drugim granama matematike. Kolmogorovljeva aksiomatika dopušta da se teoriji vjerojatnosti, kao i svakoj drugoj aksiomatiziranoj teoriji, daju različite interpretacije. Ta je aksiomatika nastala apstrahiranjem pojma relativne frekvencije slučajnog događaja, tako da se i svaka izjava teorije vjerojatnosti može interpretirati u terminima relativne frekvencije, što omogućuje da se iz apstraktnih shema prijeđe na stvarne pojave. Nagli suvremeni razvoj znanosti nije mimoišao i teoriju vjerojatnosti. Pojavili su se specijalizirani časopisi, publicirana je golema količina radova, napisano je mnogo knjiga, izvršena je podjela na određena specijalna područja u okviru same teorije vjerojatnosti, a razvile su se i mnoge znanstvene discipline u kojima se pri razmatranju glavnih problema pretežno primjenjuje teorija vjerojatnosti (*matematička statistika, teorija informacije, teorija pouzdanosti, teorija repova* i sl.).

DOGAĐAJ I VJEROJATNOST DOGAĐAJA

U suvremenom se pristupu matematizaciji pojma slučajnosti najprije morao razjasniti pojam *događaja* i *vjerojatnosti događaja*. U početku razvoja teorije vjerojatnosti nastojalo se odgovoriti na pitanja o vjerojatnosti konkretnog događaja kao broja koji kvantitativno izražava mogućnost nastupanja uočenog događaja preko određenih definicija vjerojatnosti događaja (klasična, geometrijska ili statistička definicija vjerojatnosti). Međutim, ni jedna od tih definicija nije omogućila izgradnju konzistentne matematičke teorije koja bi poslužila kao model za sve one stvarne situacije za koje se intuitivno očekivalo da budu obuhvaćene jednom takvom teorijom. Rješenje se problema pojavilo u onom trenutku kada je napuštena ideja da se u okviru teorije vjerojatnosti mogu dobiti odgovori na pitanja kolika je vjerojatnost pojedinih događaja, a spoznalo se da treba istaknuti samo bitna svojstva pojma događaja i pojma vjerojatnosti kao matematičkih pojmova koji se na određeni način dovode u vezu s drugim matematičkim pojmovima (skup, broj, funkcija i sl.). Pri tome se na događaj više ne gleda izolirano, već se on razmatra kao element u skupu svih događaja koji se razmatraju u uočenoj slučajnoj pojavi. Taj skup ima određenu strukturu, što omogućuje da se s događajima operira po utvrđenim pravilima i da se dobiju odgovarajuće formule kojima se izražavaju veze između događaja i vjerojatnosti događaja.

Događaji kao skupovi. Temeljna je pretpostavka pri matematičkom razmatranju slučajnih pojava da se može, pri svakom stvarnom eksperimentu ili opažanju neke stvarne pojave, uočiti i definirati određeni neprazni skup Ω svih mogućih ishoda ili rezultata promatrane pojave. Skup Ω obično se naziva *skup svih mogućih ishoda* ili *skup elementarnih događaja* promatrane slučajne pojave $\mathcal{E}$. Zapis $\omega \in \Omega$ označuje da je ω *ishod* uočene slučajne pojave $\mathcal{E}$.

Bacanje igraće kocke tipičan je primjer slučajne pojave u kojemu se uzima da je $\Omega = \{1, 2, 3, 4, 5, 6\}$, tj. brojevi 1, 2, 3, 4, 5 i 6 mogući su ishodi pri bacanju igraće kocke.

Podskup A od Ω ($A \subseteq \Omega$), tj. određeni skup ishoda, naziva se *događaj* u slučajnoj pojavi $\mathcal{E}$. Kaže se da je *nastupio događaj A* ako je u pojavi $\mathcal{E}$ ostvareno $\omega \in A$.

Tako je npr. pojava parnog broja određeni događaj pri bacanju igraće kocke, koji se opisuje skupom $A = \{2, 4, 6\} \subseteq \Omega$. Događaj A ostvaruje se jednim od ishoda 2 ili 4 ili 6.

Međusobni odnosi događaja. Događaji A i B su *jednaki* ako se sastoje od istih ishoda, tj. ako su skupovi A i B ($A \subseteq \Omega$, $B \subseteq \Omega$) jednaki. Piše se $A = B$.

Prazan skup $\emptyset$ predstavlja *nemoguć događaj*, a skup Ω *siguran događaj*.

Skup A^c (komplement skupa $A \subseteq \Omega$ u odnosu na Ω) naziva se *protivan događaj* od A. Događaj A^c je nastupio ako nije nastupio događaj A, tj. ako je u $\mathcal{E}$ ostvaren ishod koji ne pripada skupu A.

Događaju pojavljivanja parnog broja pri bacanju igraće kocke, dakle $A=\{2,4,6\}$, protivan je događaj $A^c=\{1,3,5\}$, tj. pojava neparnog broja.

Skup $A\cup B$ naziva se *unija događaja* A i B. To je događaj koji nastupa onda kada nastupi bar jedan od događaja A, B.

Uoči li se pri bacanju igraće kocke događaj $B=\{1,2\}$, tj. pojava broja manjeg od 3, događaj $A\cup B=\{1,2,4,6\}$ opisuje pojavu parnog broja ili broja manjeg od 3.

Skup $A\cap B$ naziva se *presjek događaja* A i B. To je događaj koji nastupa kada nastupe oba događaja A, B.

U primjeru bacanja igraće kocke događaj $A\cap B=\{2\}$ opisuje pojavu parnog broja manjeg od 4, tj. pojavu broja 2.

Skup $A\setminus B=A\cap B^c$ naziva se *razlika događaja* A i B.

U promotrenom primjeru $A\setminus B=\{4,6\}$ i to je događaj koji opisuje pojavu parnog broja, ali ne manjeg od 3.

$A\subseteq B$ označuje da nastupanje događaja A *implicira* nastupanje događaja B.

$A\cap B=\emptyset$ označuje da se događaji A i B *međusobno isključuju*. Kaže se još da su A i B *disjunktni događaji*.

Razmotreni događaji A i B nisu disjunktni ($A\cap B=\{2\}\neq\emptyset$). Događaji A i A^c primjer su događaja koji se međusobno isključuju, jer je $A\cap A^c=\emptyset$, što znači da je nemoguće istodobno pojavljivanje parnog i neparnog broja.

Algebra događaja. Skup $\mathscr{P}(\Omega)$ svih podskupova od Ω zove se *skup svih mogućih događaja* slučajne pojave $\mathcal{E}$.

Algebra događaja je skup događaja ($\mathscr{A}\subseteq\mathscr{P}(\Omega)$), koji ima svojstvo da je protivan događaj svakog događaja iz $\mathscr{A}$ također događaj iz $\mathscr{A}$:

$$A\in\mathscr{A}\Rightarrow A^c\in\mathscr{A} \tag{1}$$

te da je unija svaka dva događaja iz $\mathscr{A}$ također događaj iz $\mathscr{A}$:

$$A,B\in\mathscr{A}\Rightarrow A\cup B\in\mathscr{A}. \tag{2}$$

Skup $\mathscr{P}(\Omega)$ primjer je algebre događaja, a također i skup $\{\emptyset, \Omega\}$, te skup $\{\emptyset, A, A^c, \Omega\}$, gdje je $A\subseteq\Omega$, zatim $A\neq\emptyset$, te $A\neq\Omega$.

Iz definicije algebre događaja proizlazi da je unija i presjek bilo kojeg konačnog broja događaja zadane algebre opet događaj iz te algebre:

$$A_i\in\mathscr{A}\ (i=1,\ldots,n)\Rightarrow\bigcup_{i=1}^{n}A_i\in\mathscr{A},\ \bigcap_{i=1}^{n}A_i\in\mathscr{A}. \tag{3}$$

Ako je $B_i\in\mathscr{A}\ (i=1,\ldots,n,\ B_i\neq\emptyset)$, te $B_i\cap B_j=\emptyset$, za $i\neq j$ i $\bigcup_{i=1}^{n}B_i=\Omega$, tj. ako je uočeno n disjunktnih događaja od kojih jedan sigurno nastupa u promatranoj pojavi, onda događaji $B_1,\ldots,B_n$ čine *potpunu familiju* (*klasu*) događaja. Tako npr. događaji B i B^c ($B\subseteq\Omega$, $B\neq\emptyset$, $B\neq\Omega$) čine dvočlanu potpunu familiju događaja.

Definicije vjerojatnosti događaja. Pojam vjerojatnosti događaja nastao je apstrahiranjem pojma relativne frekvencije događaja. Ključna je pretpostavka da se slučajni pokus, odnosno opažanje neke slučajne pojave, može neograničeno ponavljati uz iste uvjete. Ako se, dakle, slučajni pokus $\mathcal{E}$ ponovi n puta, pa ako je uočeni događaj A pri tome nastupio m puta $(0\leq m\leq n)$, onda se veličina

$$Q_n(A)=\frac{m}{n} \tag{4}$$

zove *relativna frekvencija* događaja A. Relativna frekvencija ima svojstva:

$$Q_n(\emptyset)=0,\quad Q_n(\Omega)=1, \tag{5}$$

$$0\leq Q_n(A)\leq 1, \tag{6}$$

$$A\cap B=\emptyset\Rightarrow Q_n(A\cup B)=Q_n(A)+Q_n(B), \tag{7}$$

tj. relativna frekvencija nemogućeg događaja ($\emptyset$) jednaka je nuli, sigurnog događaja (Ω) jedan, dok je relativna frekvencija svakog događaja (A) između nule i jedan, a relacijom (7) izrečeno je da je relativna frekvencija unije disjunktnih događaja jednaka zbroju relativnih frekvencija uočenih događaja.

Iskustvena je spoznaja da uz veliki broj (n) ponavljanja pokusa relativna frekvencija $Q_n(A)$ ne ovisi više o n, već samo o uočenom događaju A, pa se može govoriti o broju $\mathrm{P}(A)$ kao određenoj karakteristici događaja A u smislu mjere za mogućnost nastupanja događaja A. Simbolički to se izražava relacijom

$$\lim_{n\to\infty}Q_n(A)=\mathrm{P}(A). \tag{8}$$

Vjerojatnost na algebri događaja $\mathscr{A}$ je funkcija P koja događajima algebre $\mathscr{A}$ pridružuje realne brojeve ($\mathrm{P}:\mathscr{A}\to\mathbf{R}$), tako da vrijedi

$$\mathrm{P}(A)\geq 0, \tag{9}$$

za svaki $A\in\mathscr{A}$, što se naziva *nenegativnost*,

$$\mathrm{P}(\Omega)=1, \tag{10}$$

što se naziva *normiranost*, te

$$A,B\in\mathscr{A}\ \text{ i }\ A\cap B=\emptyset\Rightarrow\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B), \tag{11}$$

što se naziva *aditivnost*.

Broj $\mathrm{P}(A)$ zove se *vjerojatnost događaja* A. To je tzv. *aksiomatska definicija vjerojatnosti.*

Za izgradnju teorije vjerojatnosti u najopćenitijem obliku potrebno je definirati vjerojatnost na tzv. *sigma-algebri* događaja, pri čemu se zahtijeva da funkcija P ima i svojstvo neprekidnosti. U nastavku će se pretpostavljati da događaji o kojima će biti riječ pripadaju određenoj algebri (sigma-algebri) događaja uočenoj u vezi s promatranom slučajnom pojavom i to se neće posebno isticati.

Ostala su svojstva vjerojatnosti:

$$\mathrm{P}(A^c)=1-\mathrm{P}(A), \tag{12}$$

$$\mathrm{P}(\emptyset)=0, \tag{13}$$

$$A\subseteq B\Rightarrow\mathrm{P}(B\setminus A)=\mathrm{P}(B)-\mathrm{P}(A), \tag{14}$$

$$\mathrm{P}(A\cap B)=\mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cup B), \tag{15}$$

$$A_i\cap A_j=\emptyset\ (i\neq j;\ i,j=1,\ldots,n)\Rightarrow\mathrm{P}\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}\mathrm{P}(A_i). \tag{16}$$

U povijesnom se razvoju prva pojavila tzv. *klasična definicija vjerojatnosti* ili *vjerojatnost a priori.* Ona se može primijeniti samo na one slučajne pojave gdje je Ω konačan skup s $v(\Omega)$ elemenata i gdje se može pretpostaviti da su svi ishodi jednako vjerojatni, što je redovit slučaj u igara na sreću. Tada se vjerojatnost $\mathrm{P}(A)$ događaja $A\subseteq\Omega$ definira kao

$$\mathrm{P}(A)=\frac{v(A)}{v(\Omega)}, \tag{17}$$

gdje je $v(A)$ broj elemenata skupa A, odnosno broj *povoljnih ishoda* za događaj A, dok je $v(\Omega)$, broj svih *mogućih ishoda* promatrane slučajne pojave.

Tako je npr. pri bacanju igraće kocke broj svih mogućih ishoda $v(\Omega)=6$. Uzme li se pojava parnog broja kao uočeni događaj A, broj povoljnih ishoda bit će $v(A)=3$, pa je $\mathrm{P}(A)=3/6=0{,}5$.

Vjerojatnost $\mathrm{P}(A)$ definirana u (17) zadovoljava uvjete (9) do (16), tj. klasična se definicija može shvatiti kao specijalni slučaj aksiomatske definicije vjerojatnosti.

Tipičan primjer slučajnog pokusa u kojem je Ω beskonačan skup izbor je točke iz zadanoga beskonačnog skupa točaka (pravca, ravnine, k-dimenzijskog prostora). Tu se ne može primijeniti klasična definicija vjerojatnosti, pa se uvodi tzv. *geometrijska vjerojatnost.* Ako je, dakle, Ω skup koji se sastoji od točaka i koji ima *mjeru* (npr. duljinu, ploštinu ili obujam) $\mu(\Omega)>0$ i ako se pokus sastoji u slučajnom izboru točke iz skupa Ω, onda se vjerojatnost događaja A, tj. vjerojatnost da se izabere točka koja pripada skupu $A\subseteq\Omega$, definira ovako:

$$\mathrm{P}(A)=\frac{\mu(A)}{\mu(\Omega)}, \tag{18}$$

gdje je $\mu(A)$ odgovarajuća mjera skupa A. Može se provjeriti da $\mathrm{P}(A)$, definirano u (18), također ima svojstva navedena u (9) do (16), tj. da se i geometrijska vjerojatnost može shvatiti kao poseban slučaj aksiomatske definicije vjerojatnosti.

Zanimljivo je da geometrijska vjerojatnost dopušta da postoje događaji koji nisu nemogući ($A\neq\emptyset$), a da je njihova vjerojatnost nula, što se ne može postići s klasičnom definicijom vjerojatnosti. Tako, npr., ako je Ω skup točaka određenog kruga, onda vjerojatnost da se pri slučajnom izboru točke dobije središte kruga iznosi nula. Također je nula i vjerojatnost da se dobije točka na uočenoj tetivi toga kruga, iako navedeni događaji nisu nemogući, tj. skup ishoda kojima se može ostvariti navedeni događaj nije prazan skup.

Uvjetna vjerojatnost. Ako je $P(B)>0$, tada je formulom

$$P(A/B)=\frac{P(A\cap B)}{P(B)} \tag{19}$$

definirana *uvjetna vjerojatnost događaja A* u odnosu na događaj B. Broj $P(A/B)$ interpretira se kao vjerojatnost da nastupi događaj A ako se zna da je nastupio događaj B. Uvjetna vjerojatnost također ima svojstva istaknuta u aksiomatskoj definiciji vjerojatnosti, tj. vrijedi

$$P(A/B)\geqq 0, \tag{20}$$

$$P(\Omega/B)=1, \quad P(\emptyset/B)=0, \tag{21}$$

$$A_1\cap A_2=\emptyset \Rightarrow P(A_1\cup A_2/B)=P(A_1/B)+P(A_2/B). \tag{22}$$

Iz (19) još proizlazi:

$$B\subseteq A\Rightarrow P(A/B)=1, \tag{23}$$

$$A\cap B=\emptyset\Rightarrow P(A/B)=0, \tag{24}$$

$$P(A\cap B)=P(B)P(A/B). \tag{25}$$

Ako je i $P(A)>0$, onda vrijedi tzv. *Bayesova formula*:

$$P(B/A)=\frac{P(B)P(A/B)}{P(A)}. \tag{26}$$

Za $n\geqq 2$ vrijedi

$$P(A_1\cap\cdots\cap A_n)=P(A_1)P(A_2/A_1)\cdots P(A_n/A_1\cap\cdots\cap A_{n-1}). \tag{27}$$

Ako događaji $B_1,\ldots,B_n$ $(n\geqq 2)$ čine potpunu familiju događaja, te ako je $P(B_i)>0$ $(i=1,\ldots,n)$, tada za događaj A vrijedi tzv. *formula totalne vjerojatnosti*:

$$P(A)=\sum_{i=1}^{n}P(B_i)P(A/B_i). \tag{28}$$

Događaji A i B su *stohastički nezavisni* ako vrijedi

$$P(A\cap B)=P(A)P(B). \tag{29}$$

Pojam stohastičke nezavisnosti definira se i za skup od n $(n\geqq 2)$ događaja $A_1,\ldots,A_n$. Ako za svaki njegov r-člani $(2\leqq r\leqq n)$ podskup $B_1,\ldots,B_r$ vrijedi

$$P(B_1\cap\cdots\cap B_r)=P(B_1)\cdots P(B_r), \tag{30}$$

onda se kaže da su događaji $A_1,\ldots,A_n$ stohastički nezavisni.

SLUČAJNE VARIJABLE, RAZDIOBE VJEROJATNOSTI

Slučajna je varijabla matematičko teorijski model za pojavu slučajnih rezultata mjerenja pri proučavanju određene mjerljive slučajne pojave. Slučajnost rezultata mjerenja može biti posljedica prirode promatrane pojave, nedovoljne preciznosti instrumenta, subjektivnog faktora (mjeriteljeve nepreciznosti), a također i svega zajedno. Iskustvo je, međutim, pokazalo da i uz prisutnost različitih vrsta slučajnosti postoje određene zakonitosti u ukupnom ponašanju rezultata mjerenja, koje se mogu matematički izraziti odgovarajućim teorijskim razdiobama (distribucijama) vjerojatnosti.

Funkcija razdiobe. Ako je uz svaki ishod određene slučajne pojave $\mathcal{E}$ vezan određeni broj, odnosno ako se slučajni pokus sastoji od mjerenja neke slučajne veličine X, onda se govori o *slučajnoj varijabli* ili *slučajnoj veličini X*. Apstraktno se slučajna varijabla X definira kao određena funkcija koja ishodima slučajne pojave pridružuje realne brojeve. Piše se $\omega\mapsto X(\omega)$ i govori da je broj $x=X(\omega)$ *vrijednost slučajne varijable X* na ishodu ω. Ako je S $(S\subseteq\mathbf{R})$ zadani skup brojeva, onda $\{X\in S\}$ označuje događaj da je slučajna varijabla X poprimila vrijednost iz skupa S, tj.

$$\{X\in S\}=\{\omega\in\Omega\colon X(\omega)\in S\}\subseteq\Omega. \tag{31}$$

Ako je $S=[a,b]$ $(a,b\in\mathbf{R},\ a<b)$, onda je

$$\{X\in[a,b]\}=\{\omega\in\Omega\colon a\leqq X(\omega)\leqq b\}, \tag{32}$$

pa se govori o događaju da je slučajna varijabla X poprimila vrijednost iz segmenta $[a,b]$, odnosno vrijednost koja nije manja od a i nije veća od b. Vjerojatnost događaja definiranog u (31) piše se kao $P(\{X\in S\})$ ili, kraće, $P(S)$, a vjerojatnost događaja definiranog u (32) obično se piše u skraćenu obliku $P(a\leqq X\leqq b)$ i govori se o vjerojatnosti da slučajna varijabla X poprimi vrijednost iz segmenta $[a,b]$. Analogno tome, $P(a<X<b)$ označuje vjerojatnost da slučajna varijabla X poprimi vrijednost iz intervala $\langle a,b\rangle$, tj. vrijednost veću od a i manju od b.

U upotrebi su još i ove tipične oznake:

$$P(X\leqq b)=P(-\infty<X\leqq b), \tag{33 a}$$

$$P(X<b)=P(-\infty<X<b), \tag{33 b}$$

$$P(X\geqq a)=P(a\leqq X<\infty), \tag{33 c}$$

$$P(X>a)=P(a<X<\infty), \tag{33 d}$$

$$P(X=a)=P(a\leqq X\leqq a), \tag{33 e}$$

koje imaju odgovarajuće značenje, što se vidi iz pripadnog zapisa. Pri definiranju slučajne varijable u vezi s određenom slučajnom pojavom kojoj pripada skup Ω svih mogućih ishoda pretpostavlja se da je uočena ona algebra događaja $\mathcal{A}$ koja sadrži sve događaje navedene u (33 a) do (33 e), tako da se može govoriti o njihovim vjerojatnostima. Budući da se svaka od spomenutih vjerojatnosti može izraziti pomoću vjerojatnosti oblika (33 a), definira se funkcija $x\mapsto F(x)$, $x\in\mathbf{R}$ ovako:

$$F(x)=P(X\leqq x) \tag{34}$$

i naziva se *funkcija razdiobe* ili *funkcija distribucije* vjerojatnosti slučajne varijable X. Broj $F(x)$ označuje, dakle, vjerojatnost da slučajna varijabla X poprimi vrijednost koja nije veća od broja x.

Funkcija razdiobe ima ova svojstva:

$$x_1<x_2\Rightarrow F(x_1)\leqq F(x_2), \tag{35}$$

što znači da je funkcija F monotono rastuća, te

$$\lim_{x\to-\infty}F(x)=F(-\infty)=0, \tag{36 a}$$

$$\lim_{x\to\infty}F(x)=F(\infty)=1. \tag{36 b}$$

Nadalje je

$$\lim_{\substack{\varepsilon\to 0\\ \varepsilon>0}}F(x+\varepsilon)=F(x+0)=F(x), \tag{37 a}$$

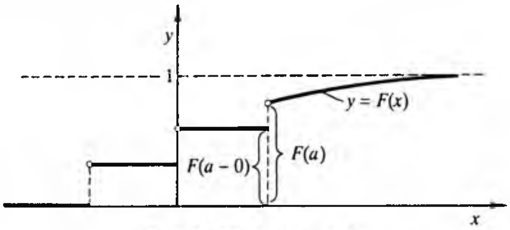
što pokazuje da je funkcija F neprekinuta zdesna u svakoj točki $x\in\mathbf{R}$.

Funkcija razdiobe općenito ne mora biti neprekidna. Ako je, naime, $\lim\limits_{\substack{\varepsilon\to 0\\ \varepsilon>0}}F(x-\varepsilon)=F(x-0)\neq F(x)$, onda funkcija F u točki $x\in\mathbf{R}$ ima prekid (skok veličine $F(x)-F(x-0)$). U skladu s time vrijedi

$$P(X=a)=F(a)-F(a-0). \tag{37 b}$$

Ako je F neprekidna funkcija u točki $a\in\mathbf{R}$, onda je $F(a-0)=F(a)$ i tada je $P(X=a)=0$, tj. vjerojatnost da slučajna varijabla X poprimi vrijednost a jednaka je nuli. Ako F ima skok u točki a, onda je $P(X=a)=F(a)-F(a-0)>0$.

Funkcija razdiobe može imati najviše prebrojivo mnogo skokova, odnosno pozitivnu vjerojatnost može imati konačno ili najviše prebrojivo mnogo točaka apscisne osi (sl. 1). Iz toga slijedi da se vjerojatnosti gotovo svih, za praksu važnih, događaja u vezi sa slučajnom varijablom X mogu izraziti pomoću njezine funkcije razdiobe.



Sl. 1. Graf funkcije razdiobe

Diskretna razdioba vjerojatnosti. Neka je $\mathscr{R}(X)=\{x_1, x_2,\ldots\}$ diskretan (konačan ili prebrojiv) skup realnih brojeva i $p_k \geqq 0\left(k=1,2,\ldots,\sum_k p_k=1\right)$ zadani realni brojevi. Ako se funkcija razdiobe slučajne varijable X može napisati u obliku

$$F(x)=\sum_{\substack{k \\ x_k \leqq x}} p_k, \tag{38}$$

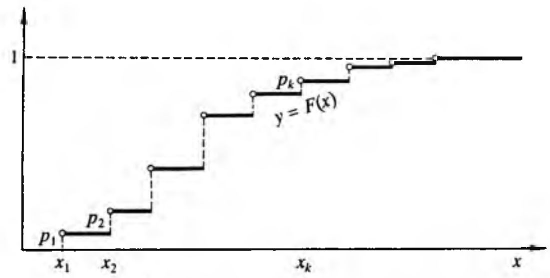
onda se kaže da je X *diskretna slučajna varijabla* sa skupom vrijednosti $\mathscr{R}(X)$ i pripadnim vjerojatnostima $p_k=\mathrm{P}(X=x_k)$, $x_k \in \mathscr{R}(X)$. Kaže se još da slučajnoj varijabli X pripada diskretna razdioba vjerojatnosti (sl. 2).

Sl. 2. Shematski prikaz diskretne razdiobe vjerojatnosti

Diskretna slučajna varijabla matematički je model za one stvarne pojave u kojima se kao rezultati mjerenja mogu pojaviti samo elementi određenoga diskretnog skupa brojeva. Najčešće se radi o cijelim brojevima.

Bacanje igraće kocke može se shvatiti i kao slučajni pokus u kojem se registrira (mjeri) vrijednost slučajne varijable X (broj točkica na gornjoj plohi kocke), za koju je $\mathscr{R}(X)=\{1,2,3,4,5,6\}$ i $p_k=\mathrm{P}(X=k)=1/6$, $k\in\mathscr{R}(X)$.

Sva se važnija svojstva diskretne slučajne varijable mogu izraziti pomoću njezinih vrijednosti x_k i pripadnih vjerojatnosti p_k ($k=1,2,\ldots$), tako da pripadna funkcija razdiobe nema veliko praktično značenje (sl. 3).



Sl. 3. Graf funkcije razdiobe diskretne slučajne varijable

Primjeri diskretnih razdioba. Za slučajnu varijablu X kaže se da ima *binomnu razdiobu* s parametrima m i p ($m\in\mathbf{N}$, $0<p<1$) i piše $X\sim\mathrm{B}(m,p)$ ako je $\mathscr{R}(X)=\{0,1,\ldots,m\}$ i

$$p_k=\mathrm{P}(X=k)=\binom{m}{k}p^k(1-p)^{m-k},\quad k\in\mathscr{R}(X). \tag{39}$$

Binomna razdioba najčešće se u praksi javlja pri ponavljanju pokusa. Ako se slučajni pokus $\mathcal{E}$ u kojem je uočen događaj A vjerojatnosti $\mathrm{P}(A)=p>0$ ponovi m puta, tada je broj nastupa događaja A slučajna varijabla X binomne razdiobe $\mathrm{B}(m,p)$.

Uzme li se, primjerice, bacanje novčića kao slučajni eksperiment $\mathcal{E}$, a pojava grba kao uočeni događaj A za koji je $\mathrm{P}(A)=0{,}5$, te ako se pokus ponovi npr. 10 puta, tada se broj dobivenih grbova može shvatiti kao vrijednost slučajne varijable X binomne razdiobe s parametrima $m=10$ i $p=0{,}5$.

Za slučajnu varijablu X kaže se da ima *hipergeometrijsku razdiobu* s parametrima m, n i r ($m,n,r,\in\mathbf{N}$, $m\leqq r<n$) i piše se $X\sim\mathrm{H}(m,n,r)$ ako je $\mathscr{R}(X)=\{0,1,\ldots,m\}$ i

$$p_k=\mathrm{P}(X=k)=\frac{\binom{r}{k}\binom{n-r}{m-k}}{\binom{n}{m}},\quad k\in\mathscr{R}(X). \tag{40}$$

Glavna interpretacija: Iz n-članog skupa, u kojem ima r elemenata tipa I i $n-r$ elemenata tipa II, slučajno se uzima m-člani podskup. Broj elemenata tipa I u uzetom podskupu slučajna je varijabla $X\sim\mathrm{H}(m,n,r)$. Budući da vrijedi

$$\lim_{\substack{n\to\infty \\ r\to\infty \\ \frac{r}{n}=p}}\frac{\binom{r}{k}\binom{n-r}{m-k}}{\binom{n}{m}}=\binom{m}{k}p^k(1-p)^{m-k},\quad k=0,1,\ldots,m, \tag{41}$$

hipergeometrijska razdioba $\mathrm{H}(m,n,r)$ može se, za velike n i r, aproksimirati binomnom razdiobom $\mathrm{B}(m,p)$ ($p=r/n$).

Slučajna varijabla X ima *Poissonovu razdiobu* parametra a i piše se $X\sim\mathrm{Po}(a)$ ako je $\mathscr{R}(X)=\{0,1,2,\ldots\}$ i

$$p_k=\mathrm{P}(X=k)=\frac{a^k}{k!}\exp(-a),\quad (a>0),\quad k\in\mathscr{R}(X). \tag{42}$$

Budući da je

$$\lim_{\substack{m\to\infty \\ p\to 0 \\ mp=a}}\binom{m}{k}p^k(1-p)^{m-k}=\frac{a^k}{k!}\exp(-a),\quad k=0,1,2,\ldots, \tag{43}$$

binomna se razdioba $\mathrm{B}(m,p)$ može, za velike m i malene p, aproksimirati Poissonovom razdiobom $\mathrm{Po}(a)$ ($a=mp$).

Slučajna varijabla X ima *geometrijsku razdiobu* parametra $p(0<p<1)$ ako je $\mathscr{R}(X)=\{1,2,3,\ldots\}$ i

$$p_k=\mathrm{P}(X=k)=p(1-p)^{k-1},\quad k\in\mathscr{R}(X). \tag{44}$$

Glavna interpretacija: U slučajnom pokusu $\mathcal{E}$ uočen je događaj A vjerojatnosti $\mathrm{P}(A)=p>0$. Pokus $\mathcal{E}$ ponavlja se sve dok ne nastupi događaj A. Broj potrebnih ponavljanja slučajna je varijabla X geometrijske razdiobe parametra p.

Kontinuirana razdioba vjerojatnosti. Kontinuirana slučajna varijabla matematički je model za one stvarne pojave u kojima se kao slučajni rezultati mjerenja mogu dobiti brojevi iz nekog intervala realnih brojeva ili iz čitava skupa realnih brojeva.

Istražuje li se, npr. vijek trajanja određenog proizvoda (žarulje, otpornika i sl.), izrađenog po ustaljenoj tehnologiji, uočit će se da je to varijabilna veličina X koja može poprimiti brojčane vrijednosti iz intervala od nula do beskonačnosti. Statistička će se zakonitost odraziti u različitoj gustoći rezultata mjerenja na pojedinim dijelovima (podintervalima) skupa $\mathbf{R}$. Zato apstraktna definicija glasi: Ako se funkcija razdiobe slučajne varijable X može prikazati u obliku

$$F(x)=\mathrm{P}(X\leqq x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t,\quad x\in\mathbf{R}, \tag{45}$$

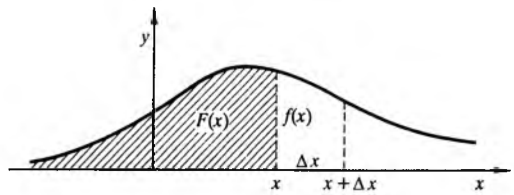
gdje je $t\mapsto f(t)$, $t\in\mathbf{R}$, nenegativna realna funkcija, onda je X *kontinuirana slučajna varijabla* s *funkcijom gustoće vjerojatnosti* f. Kaže se da slučajnoj varijabli X pripada kontinuirana razdioba vjerojatnosti karakterizirana funkcijom gustoće vjerojatnosti f (sl. 4).

Deriviranjem funkcije razdiobe dobiva se funkcija gustoće vjerojatnosti:

$$f(x)=F'(x)=\lim_{\Delta x\to 0}\frac{F(x+\Delta x)-F(x)}{\Delta x}=$$

$$= \lim_{\Delta x \to 0} \frac{P(x \leq X \leq x + \Delta x)}{\Delta x}, \tag{46}$$

pa se $f(x)$ naziva *gustoća vjerojatnosti* u točki $x \in \mathbf{R}$.



Sl. 4. Graf funkcije gustoće vjerojatnosti

Graf funkcije gustoće vjerojatnosti naziva se *krivulja razdiobe* i ona zorno pokazuje kako bi teorijski trebala izgledati razdioba rezultata mjerenja po apscisnoj osi. Površina ispod krivulje razdiobe od $-\infty$ do x brojčano je jednaka vrijednosti $F(x)$ funkcije razdiobe u točki x. Ukupna površina ispod krivulje razdiobe od $-\infty$ do ∞ iznosi jedan, tj. vrijedi

$$\int_{-\infty}^{\infty} f(t)\,dt = F(\infty) = 1. \tag{47}$$

Teorijski udio onih rezultata mjerenja koji pripadaju segmentu $[x, x+\Delta x]$, odnosno vjerojatnost da kontinuirana slučajna varijabla X poprimi vrijednost iz segmenta $[x, x+\Delta x]$, iznosi

$$\int_{x}^{x+\Delta x} f(t)\,dt = P(x \leq X \leq x + \Delta x), \tag{48}$$

tj. brojčano je to jednako površini ispod krivulje razdiobe, a iznad segmenta $[x, x+\Delta x]$. Ako je Δx maleno, onda se može pisati

$$P(x \leq X \leq x + \Delta x) \approx f(x)\Delta x. \tag{49}$$

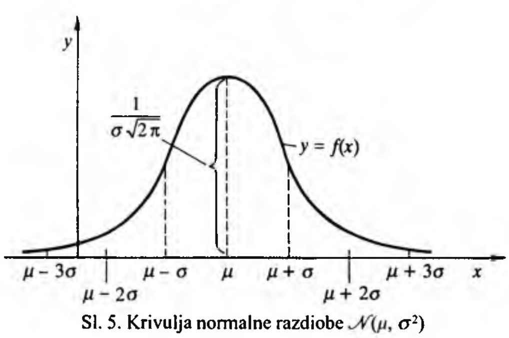
Konkretna kontinuirana razdioba vjerojatnosti obično se zadaje svojom funkcijom gustoće vjerojatnosti.

Primjeri kontinuiranih razdioba. Slučajna varijabla X ima *normalnu* ili *Gaussovu razdiobu* s parametrima μ i σ ($\sigma > 0$) i piše se $X \sim \mathcal{N}(\mu, \sigma^2)$ ako je njezina funkcija gustoće vjerojatnosti zadana izrazom

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]. \tag{50}$$

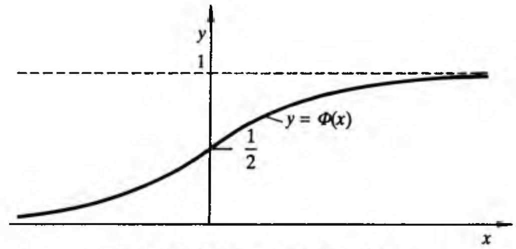
Krivulja normalne razdiobe (sl. 5) simetrična je s obzirom na pravac $y=\mu$, ima maksimum $1/(\sigma\sqrt{2\pi})$ za $x=\mu$, dok za $x=\mu-\sigma$ i $x=\mu+\sigma$ ima točke infleksije. Normalna razdioba $\mathcal{N}(\mu, \sigma^2)$ matematički je model za simetričnu razdiobu rezultata mjerenja kojima gustoća zvonoliko opada udaljavanjem od središta simetrije μ. Pripadna je funkcija razdiobe

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{1}{2}\left(\frac{x-t}{\sigma}\right)^2\right] dt. \tag{51}$$



Sl. 5. Krivulja normalne razdiobe $\mathcal{N}(\mu, \sigma^2)$

Integral na desnoj strani u (51) nije elementarno rješiv i zato se u računima s normalnom razdiobom upotrebljavaju tablice gdje su



Sl. 6. Graf funkcije razdiobe $\mathcal{N}(0, 1)$

navedene vrijednosti funkcije razdiobe, tzv. *standardne* ili *jedinične normalne razdiobe* $\mathcal{N}(0,1)$, koja se dobiva za $\mu=0$ i $\sigma=1$ (sl. 6.). Pripadna se funkcija gustoće vjerojatnosti obično označuje s φ, a funkcija razdiobe s Φ, tako da je

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \tag{52a}$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt, \tag{52b}$$

dok je

$$f(x) = \frac{1}{\sigma}\varphi(x), \tag{53a}$$

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right), \tag{53b}$$

što omogućuje primjenu tablica i na normalnu razdiobu $\mathcal{N}(\mu, \sigma^2)$.

Funkcija φ je parna, tj. vrijedi $\varphi(-x) = \varphi(x)$, dok za funkciju Φ vrijedi

$$\Phi(-x) = 1 - \Phi(x), \quad \Phi(0) = \frac{1}{2}, \tag{54}$$

tako da se obično tabliciraju samo vrijednosti funkcije Φ za $x \geq 0$.

Ako je $X \sim \mathcal{N}(\mu, \sigma^2)$ i $a < b$, onda je

$$P(a < X < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \tag{55}$$

Posebno, za proizvoljno $\lambda > 0$ vrijedi

$$P(\mu - \lambda\sigma < X < \mu + \lambda\sigma) = 2\Phi(\lambda) - 1. \tag{56}$$

Uzme li se da je $\lambda = 3$, dobiva se

$$P(\mu - 3\sigma < X < \mu + 3\sigma) = 0{,}99730, \tag{57}$$

iz čega se vidi da, iako teorijski normalna razdioba svakom realnom broju pridružuje pozitivnu gustoću vjerojatnosti, interval $\langle \mu - 3\sigma, \mu + 3\sigma \rangle$ obuhvaća praktički gotovo sve moguće (99,73%) rezultate mjerenja.

Kaže se da slučajna varijabla X ima *gama-razdiobu* s parametrima α i β ($\alpha > 0$, $\beta > 0$) i piše se $X \sim G(\alpha, \beta)$ ako njezina funkcija gustoće vjerojatnosti glasi

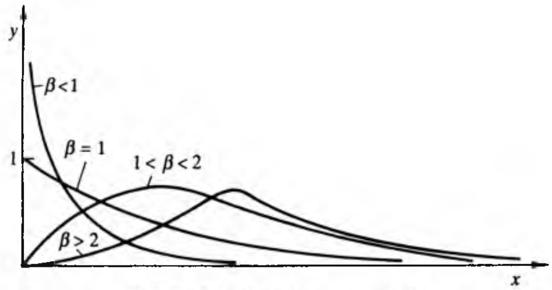
$$f(x) = \begin{cases} 0, & \text{za } x \leq 0 \\ \dfrac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1} \exp(-\alpha x), & \text{za } x > 0, \end{cases} \tag{58}$$

gdje je $\Gamma(\beta) = \int_0^\infty t^{\beta-1} \exp(-t)\,dt$ (*gama-funkcija*).

Ako je $\beta = 1$, tada je to *eksponencijalna razdioba parametra* α i piše se $X \sim \mathrm{E}x(\alpha)$, a pripadna gustoća vjerojatnosti glasi

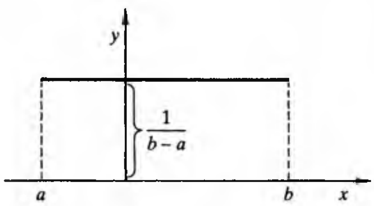
$$f(x) = \begin{cases} 0, & \text{za } x \leq 0 \\ \alpha \exp(-\alpha x), & \text{za } x > 0, \end{cases} \tag{59}$$

Ako je $\alpha = 1/2$ i $\beta = n/2$, $n \in \mathbf{N}$, onda je to *hikvadratna razdioba s n stupnjeva slobode* i piše se $X \sim \chi^2(n)$.



Sl. 7. Različite krivulje gama-razdiobe

Neki se tipovi gama-razdiobe (sl. 7), posebno eksponencijalna razdioba, upotrebljavaju kao matematički model za tumačenje slučajnosti rezultata mjerenja vijeka trajanja određenog tehničkog uređaja ili elementa (žarulja, otpornik, tranzistor i sl.).



Sl. 8. Krivulja jednolike razdiobe

Slučajna varijabla X ima *jednoliku* ili *uniformnu razdiobu* (sl. 8) na segmentu $[a,b]$ $(a<b)$ i piše se $X \sim \mathrm{U}(a,b)$, ako njezina funkcija gustoće vjerojatnosti glasi

$$f(x) = \begin{cases} 0, & \text{za } x<a \text{ i } x>b \\ \dfrac{1}{b-a}, & \text{za } a \leqq x \leqq b \end{cases} \tag{60}$$

Jednolika razdioba teorijski opisuje pojavu slučajnog mjerenja gdje su rezultati jednoliko raspodijeljeni po segmentu $[a,b]$, a izvan tog segmenta nema rezultata mjerenja.

Očekivanje i disperzija. Ako je $x \mapsto h(x)$, $x \in \mathbf{R}$, po dijelovima neprekidna realna funkcija i ako je X zadana slučajna varijabla, onda je $Y = h(X)$ također slučajna varijabla. Dakle, funkcija slučajne varijable opet je slučajna varijabla.

Ako je X diskretna slučajna varijabla sa skupom vrijednosti $\mathcal{R}(X) = \{x_1, x_2, \ldots\}$ i pripadnim vjerojatnostima $p_i = \mathrm{P}(X = x_i)$, $i = 1, 2, \ldots$, onda je $Y = h(X)$ također diskretna slučajna varijabla.

Broj $\mathrm{E}[Y] = \mathrm{E}[h(X)] = \sum_i h(x_i)\, p_i$ je *matematičko očekivanje* (kraće: *očekivanje*) ili *sredina* slučajne varijable Y.

Za $h(x) = x$ dobiva se $\mathrm{E}[X] = \sum_i x_i\, p_i$ i tada je broj $\mathrm{E}[X]$ *očekivanje diskretne slučajne varijable* X.

Uzme li se, npr., $X \sim \mathrm{B}(m,p)$, pokazuje se da je $\mathrm{E}[X] = \sum_{k=0}^{m} k \binom{m}{k} p^k (1-p)^{m-k} = mp$.

Ako je X kontinuirana slučajna varijabla i $x \mapsto h(x)$ strogo monotona i derivabilna funkcija, onda je $Y = h(X)$ također kontinuirana slučajna varijabla. Broj $\mathrm{E}[Y] = \mathrm{E}[h(X)] = \int_{-\infty}^{\infty} h(x) f(x)\,\mathrm{d}x$ naziva se očekivanje slučajne varijable Y. Ako je $h(x) = x$, broj $\mathrm{E}[X] = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x$ naziva se *očekivanje kontinuirane slučajne varijable* X.

Uzme li se npr. $X \sim \mathcal{N}(\mu, \sigma^2)$ i $h(x) = ax + b$, gdje su $a \neq 0$ i b određene konstante, pokazuje se da je $Y = aX + b$ kontinuirana slučajna varijabla normalne distribucije $\mathcal{N}(a\mu + b, a^2\sigma^2)$.

Općenito, ako se slučajna varijabla X kojoj pripada funkcija distribucije $F(x) = \mathrm{P}(X \leqq x)$ podvrgne *afinoj transformaciji*, tj. ako je $h(x) = ax + b$ $(a \neq 0)$, onda slučajnoj varijabli $Y = aX + b$ pripada funkcija razdiobe:

$$G(y) = \mathrm{P}(Y \leqq y) = F\left(\frac{y-b}{a}\right) \tag{61}$$

i tada je

$$\mathrm{E}[Y] = \mathrm{E}[aX + b] = a\,\mathrm{E}[X] + b. \tag{62}$$

Stavi li se u (62) $a = 1$ i $b = -\mathrm{E}[X]$, dobiva se

$$\mathrm{E}[X - \mathrm{E}[X]] = 0, \tag{63}$$

tj. očekivanje razlike slučajne varijable i njezina očekivanja jest nula.

Broj $\mathrm{D}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2]$ je *disperzija* ili *varijanca*, a broj $\sigma = \sqrt{\mathrm{D}[X]}$ je *standardna devijacija* slučajne varijable X. Broj σ^2, odnosno σ, pokazuje veličinu rasipanja vrijednosti slučajne varijable oko matematičkog očekivanja. Očito je $\mathrm{D}[X] \geqq 0$, i $\mathrm{D}[X] = 0$ pokazuje da se radi o tzv. *degeneriranoj slučajnoj varijabli*, tj. vrijedi $\mathrm{P}(X = \mathrm{E}[X]) = 1$.

Vrijede još i ove formule:

$$\mathrm{D}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2, \tag{64}$$

$$\mathrm{D}[aX + b] = a^2\,\mathrm{D}[X], \quad a, b \in \mathbf{R}, \tag{65}$$

$$\min_{a \in \mathbf{R}} \mathrm{E}[(X-a)^2] = \mathrm{E}[(X - \mathrm{E}[X])^2] = \mathrm{D}[X]. \tag{66}$$

Izraz (66) pokazuje da je rasipanje vrijednosti neke slučajne varijable oko njezina očekivanja manje od rasipanja te slučajne varijable oko bilo kojeg drugog realnog broja.

Ako su $\mathrm{E}[X] = \mu$ i $\mathrm{D}[X] = \sigma^2 > 0$ konačni brojevi, tada za svaki $\lambda > 0$ vrijedi *Čebiševljeva nejednakost*:

$$\mathrm{P}(\mu - \lambda\sigma < X < \mu + \lambda\sigma) > 1 - \frac{1}{\lambda^2}, \tag{67}$$

koja se često piše i u obliku

$$\mathrm{P}(|X - \mu| \geqq \lambda\sigma) \leqq \frac{1}{\lambda^2}. \tag{68}$$

Broj $\mathrm{E}[(X-a)^r]$, gdje je $a \in \mathbf{R}$, $r = 0, 1, 2, \ldots$, naziva se *r-ti moment* slučajne varijable X oko broja a. Za $a = \mathrm{E}[X] = \mu$ to su *glavni* ili *centralni momenti*, a za $a = 0$ *pomoćni* ili *ishodišni momenti*. Glavni se momenti obično označuju kao

$$\mu_r = \mathrm{E}[(X - \mu)^r], \tag{69}$$

pa se vidi da je $\mu_0 = 1$, $\mu_1 = 0$ i $\mu_2 = \mathrm{D}[X] = \sigma^2$.

Ako je $\sigma > 0$, definira se tzv. *koeficijent asimetrije*:

$$\mathcal{K} = \frac{\mu_3}{\sigma^3} \tag{70}$$

i *koeficijent spljoštenosti* ili *eksces*:

$$\mathcal{E} = \frac{\mu_4}{\sigma^4} - 3. \tag{71}$$

Za simetrične razdiobe, tj. takve u kojima funkcija razdiobe ima svojstvo $F(\mu - x) = 1 - F(\mu + x)$, koeficijent je asimetrije $\mathcal{K} = 0$, dok je za normalnu razdiobu, koja je, dakako, simetrična, i $\mathcal{E} = 0$.

Za binomnu razdiobu $\mathrm{B}(m,p)$ je $\mathcal{K}=\frac{1-2p}{mp(1-p)}$ i $\mathcal{E}=\frac{1-6p(1-p)}{mp(1-p)}$, pa se vidi da je ona simetrična za $p=\frac{1}{2}$, a inače može biti pozitivno ili negativno asimetrična.

Za Poissonovu razdiobu $\mathrm{Po}(a)$ je $\mathrm{E}[X]=a$, $\mathrm{D}[X]=a$, $\mathcal{K}=\frac{1}{\sqrt{a}}$ i $\mathcal{E}=\frac{1}{a}$, za gama razdiobu $\mathrm{G}(\alpha,\beta)$ je $\mathrm{E}[X]=\frac{\beta}{\alpha}$, $\mathrm{D}[X]=\frac{\beta}{\alpha^2}$, $\mathcal{K}=\frac{2}{\sqrt{\beta}}$ i $\mathcal{E}=\frac{6}{\beta}$, a za jednoliku razdiobu $U=(a,b)$ je $\mathrm{E}[X]=\frac{a+b}{2}$, $\mathrm{D}[X]=\frac{(b-a)^2}{12}$, $\mathcal{K}=0$ i $\mathcal{E}=-1{,}2$.

SLUČAJNI VEKTORI I VIŠEDIMENZIJSKE RAZDIOBE VJEROJATNOSTI

Mnoge praktične situacije nameću potrebu da se istodobno mjeri dvije ili više različitih veličina, pri čemu se rezultati mjerenja podvrgavaju statističkim zakonitostima. Matematički model za apstraktno teorijsko opisivanje takve slučajne pojave naziva se slučajni vektor, kojemu su komponente slučajne varijable. Ovdje se kao rezultat jednoga mjerenja pojavljuje uređeni par, odnosno uređena n-torka brojeva, pa će se statističke zakonitosti matematički izražavati preko pripadnih dvodimenzijskih, odnosno n-dimenzijskih teorijskih razdioba vjerojatnosti.

Vodostaj Save izražava se nizom mjerenja na pojedinim mjernim postajama (Zagreb, Sisak, Jasenovac itd.). Rezultat jednog mjerenja (određenog dana) je niz brojeva $(x_1, x_2, \ldots, x_n)$, koji se može shvatiti kao vrijednost slučajnog vektora $(X_1, X_2, \ldots, X_n)$. Matematički model za opisivanje slučajnog kolebanja vodostaja Save na n mjernih postaja određeni je n-dimenzijski slučajni vektor. Statističke zakonitosti u mnoštvu izvedenih mjerenja teorijski će se izraziti pomoću odgovarajuće n-dimenzijske razdiobe vjerojatnosti.

Dvodimenzijska razdioba vjerojatnosti. Ako se u slučajnom pokusu mjere dvije varijabilne veličine X i Y, onda je uz svaki ishod tog pokusa vezan uređeni par brojeva (x,y) i tada je to dvodimenzijski slučajni vektor (X,Y) s komponentama X i Y. Formalno govoreći, slučajni vektor (X,Y) određena je funkcija koja svakom ishodu $\omega\in\Omega$ pridružuje odgovarajući uređeni par $(x,y)\in\mathbf{R}^2$. Sa $\mathrm{P}(X\leqq b, Y\leqq d)$ označuje se vjerojatnost da X poprimi vrijednost koja nije veća od b, a Y vrijednost koja nije veća od d.

Realna funkcija F dviju realnih varijabli, definirana ovako:

$$F(x,y)=\mathrm{P}(X\leqq x, Y\leqq y),\quad (x,y)\in\mathbf{R}^2, \tag{72}$$

naziva se *funkcija razdiobe slučajnog vektora* (X,Y). Ona ima svojstva slična onima koje ima funkcija razdiobe slučajne varijable (v. (35), (36) i (37)). Također vrijedi

$$\mathrm{P}(x_1<X\leqq x_2,\ y_1<Y\leqq y_2)=$$
$$=F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1)+F(x_1,y_1). \tag{73}$$

Funkcija $x\mapsto F_1(x)=F(x,\infty)$, $x\in\mathbf{R}$, naziva se *marginalna funkcija razdiobe komponente* X, a funkcija $y\mapsto F_2(y)=F(\infty,y)$, $y\in\mathbf{R}$, naziva se *marginalna funkcija razdiobe komponente* Y. Marginalne razdiobe opisuju statističko ponašanje svake slučajne varijable X i Y posebno.

Ako vrijedi $F(x,y)=F_1(x)F_2(y)$, onda su X i Y *stohastički nezavisne slučajne varijable.* Ako su X i Y stohastički nezavisne slučajne varijable, onda njihovo istodobno proučavanje ne daje nikakve nove informacije u odnosu na posebno proučavanje svake od njih.

Diskretna dvodimenzijska razdioba vjerojatnosti. Neka je $\mathcal{R}(X,Y)=\{(x_i,y_j)\in\mathbf{R}^2 : i,j=1,2,\ldots\}$ zadani diskretni skup uređenih parova realnih brojeva i $p_{ij}\geqq 0\left(\sum_i\sum_j p_{ij}=1\right)$ zadani brojevi.

Ako se funkcija razdiobe slučajnog vektora (X,Y) može napisati u obliku

$$F(x,y)=\sum_{x_i\leq x}\sum_{y_j\leq y}p_{ij}, \tag{74}$$

kaže se da je (X,Y) *diskretni slučajni vektor* sa skupom vrijednosti $\mathcal{R}(X,Y)$ i pripadnim vjerojatnostima $p_{ij}=\mathrm{P}(X=x_i,\ Y=y_j)$. Kaže se još da slučajnom vektoru (X,Y) pripada *diskretna dvodimenzijska razdioba vjerojatnosti.*

Ako se, npr., m puta ponavlja slučajni pokus u kojem su uočeni disjunktni događaji A_1 i A_2 ($A_1\cap A_2=\emptyset$, $\mathrm{P}(A_1)=p_1$, $\mathrm{P}(A_2)=p_2$, $p_1+p_2<1$) i pri tome se registrira (mjeri) broj X nastupa događaja A_1 i broj Y nastupa događaja A_2, onda je (X,Y) diskretni slučajni vektor sa skupom vrijednosti $\mathcal{R}(X,Y)=\{(i,j): i,j=0,1,\ldots,m,\ i+j\leqq m\}$ i pripadnim vjerojatnostima:

$$p_{ij}=\mathrm{P}(X=i,Y=j)=\frac{m!}{i!\,j!(m-i-j)!}p_1^i p_2^j(1-p_1-p_2)^{m-i-j}. \tag{75}$$

Pripadna diskretna dvodimenzijska razdioba je *trinomna razdioba* $\mathrm{B}(m,p_1,p_2)$ s parametrima m, p_1 i p_2 ($m\in\mathrm{N}$, $0<p_1<1$; $0<p_2<1$, $p_1+p_2<1$). Piše se $(X,Y)\sim\mathrm{B}(m,p_1,p_2)$.

Marginalne razdiobe diskretnog slučajnog vektora (X,Y) također su diskretne razdiobe i vrijedi

$$\begin{aligned}&\mathcal{R}(X)=\{x_1,x_2,\ldots\},\\ &p_i=\mathrm{P}(X=x_i)=\sum_j p_{ij},\quad i=1,2,\ldots,\end{aligned} \tag{76}$$

$$\begin{aligned}&\mathcal{R}(Y)=\{y_1,y_2,\ldots\},\\ &q_j=\mathrm{P}(Y=y_j)=\sum_i p_{ij},\quad j=1,2,\ldots.\end{aligned} \tag{77}$$

Diskretne slučajne varijable X i Y su nezavisne slučajne varijable ako vrijedi

$$p_{ij}=p_i q_j,\quad i,j=1,2,\ldots. \tag{78}$$

Zatim je

$$\mathrm{E}[X]=m_1=\sum_i x_i\,p_i,\quad \mathrm{D}[X]=s_1^2=\sum_i(x_i-m_1)^2p_i, \tag{79}$$

$$\mathrm{E}[Y]=m_2=\sum_j y_j\,q_j,\quad \mathrm{D}[Y]=s_2^2=\sum_j(y_j-m_2)^2q_j. \tag{80}$$

Za trinomnu razdiobu $\mathrm{B}(m,p_1,p_2)$ marginalne razdiobe su binomne razdiobe, tj. $X\sim\mathrm{B}(m,p_1)$ i $Y\sim\mathrm{B}(m,p_2)$, tako da je $\mathrm{E}[X]=mp_1$, $\mathrm{E}[Y]=mp_2$, $\mathrm{D}[X]=mp_1(1-p_1)$ i $\mathrm{D}[Y]=mp_2(1-p_2)$.

Ako je $q_j>0$, onda je diskretna razdioba sa skupom vrijednosti $\mathcal{R}(X/y_j)=\{x_1,x_2,\ldots\}$ i pripadnim vjerojatnostima $p_{i/j}=\mathrm{P}(X=x_i/Y=y_j)=\frac{p_{ij}}{q_j}$, $i=1,2,\ldots$ *uvjetna razdioba komponente X za fiksirano y_j*. Broj $\mathrm{E}[X/y_j]=\sum_i x_i\,p_{i/j}$ je *uvjetno očekivanje komponente X za fiksirano y_j*. Analogno, ako je $p_i>0$, onda je diskretna razdioba sa skupom vrijednosti $\mathcal{R}(Y/x_i)=\{y_1,y_2,\ldots\}$ i pripadnim vjerojatnostima $q_{j/i}=\mathrm{P}(Y=y_j/X=x_i)=\frac{p_{ij}}{p_i}$, $j=1,2,\ldots$ *uvjetna razdioba komponente Y za fiksirano x_i*. Broj $\mathrm{E}[Y/x_i]=\sum_j y_j\,q_{j/i}$ je *uvjetno očekivanje komponente Y za fiksirano x_i*.

Za trinomnu razdiobu $\mathrm{B}(m,p_1,p_2)$ uvjetne razdiobe su binomne razdiobe. Tako je uvjetna razdioba komponente X za fiksirano j binomna razdioba $\mathrm{B}\left(m-j,\frac{p_1}{1-p_2}\right)$, iz čega proizlazi da je odgovarajuće uvjetno očekivanje $\mathrm{E}[X/j]=(m-j)\frac{p_1}{1-p_2}$. Analogno je $\mathrm{B}\left(m-i,\frac{p_2}{1-p_1}\right)$ uvjetna razdioba komponente Y za fiksirano i, pa je $\mathrm{E}[Y/i]=(m-i)\frac{p_2}{1-p_1}$.

Kontinuirana dvodimenzijska razdioba vjerojatnosti. Ako se funkcija razdiobe slučajnog vektora (X, Y) može napisati u obliku

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,\mathrm{d}u\,\mathrm{d}v, \tag{81}$$

gdje je $(u, v) \mapsto f(u, v) \geq 0$, $(u, v) \in \mathbf{R}^2$, onda je (X, Y) *kontinuirani slučajni vektor* s funkcijom gustoće vjerojatnosti f. Govori se još i da slučajnom vektoru (X, Y) pripada *kontinuirana dvodimenzijska razdioba vjerojatnosti.* Budući da je

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} =$$

$$= \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{\mathrm{P}(x < X \leq x + \Delta x,\, y < Y \leq y + \Delta y)}{\Delta x \Delta y}, \tag{82}$$

opravdan je naziv *gustoća vjerojatnosti* u točki (x,y) za broj $f(x,y)$.

Očigledno je

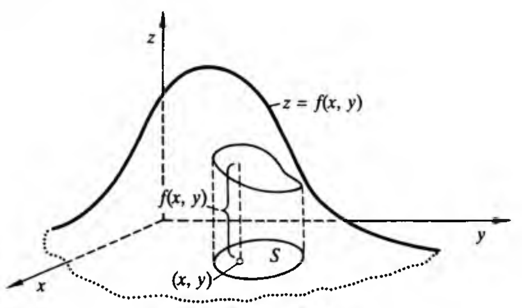
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1. \tag{83}$$

Ako je $S \subseteq \mathbf{R}^2$, onda se vjerojatnost da slučajni vektor (X, Y) poprimi vrijednost iz skupa S može izraziti ovako:

$$\mathrm{P}((X,Y) \in S) = \iint_{(S)} f(x,y)\,\mathrm{d}x\,\mathrm{d}y, \tag{84}$$

tj. ta se vjerojatnost dobiva integriranjem gustoća vjerojatnosti po skupu S.

Graf funkcije gustoće vjerojatnosti, tj. skup $\{(x,y,z) \in \mathbf{R}^3\colon z = f(x,y),\ (x,y) \in \mathbf{R}^2\}$ općenito je neka ploha u prostoru. Ona se naziva *ploha razdiobe* (sl. 9). Iz (83) je vidljivo da je obujam ispod čitave plohe razdiobe, a iznad ravnine xy, jednak jedan, dok se iz (84) vidi da su vjerojatnosti određenih događaja jednake odgovarajućim obujmima ispod plohe razdiobe, a iznad pripadnog skupa S u ravnini xy.



Sl. 9. Ploha razdiobe

Najpoznatija teorijska kontinuirana dvodimenzijska razdioba jest *normalna* ili *Gaussova razdioba* s parametrima $\mu_1, \mu_2, \sigma_1, \sigma_2$ i ρ ($\sigma_1 > 0$, $\sigma_2 > 0$, $0 \leq \rho^2 < 1$). Piše se $(X, Y) \sim \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Slučajni vektor (X, Y) ima *dvodimenzijsku normalnu razdiobu* s navedenim parametrima ako pripadna funkcija gustoće vjerojatnosti glasi

$$f(x,y) = K\exp[-Q(x,y)], \tag{85}$$

gdje je

$$K = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho}},$$

$$Q(x,y) = -\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right].$$

Graf te funkcije zvonolika je ploha, kojoj su presjeci s ravninama $z = c\left(0 < c < \dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\right)$ elipse, s jednadžbom oblika

$$\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 = c. \tag{86}$$

Vidi se da elipse imaju središte u točki $(\mu_1, \mu_2) \in \mathbf{R}^2$ i toj točki pripada najveća gustoća vjerojatnosti $f(\mu_1, \mu_2) = \dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$, dok se udaljavanjem od te točke gustoća vjerojatnosti smanjuje.

Kut α između osi elipse i koordinatne osi dan je izrazom

$$\alpha = \left\{\begin{array}{ll} \dfrac{1}{2}\arctan\dfrac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}, & \text{za } \sigma_1 \neq \sigma_2 \\ \dfrac{\pi}{4}, & \text{za } \sigma_1 = \sigma_2 \end{array}\right\}. \tag{87}$$

Osi elipse bit će usporedne s koordinatnim osima za $\rho = 0$.

Ako je (X, Y) kontinuirani slučajni vektor, onda je

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y, \quad x \in \mathbf{R} \tag{88}$$

marginalna funkcija gustoće vjerojatnosti komponente X, a

$$f_2(y) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x, \quad y \in \mathbf{R}, \tag{89}$$

marginalna funkcija gustoće vjerojatnosti komponente Y.

Zatim je

$$\mathrm{E}[X] = m_1 = \int_{-\infty}^{\infty} x f_1(x)\,\mathrm{d}x, \tag{90 a}$$

$$\mathrm{D}[X] = s_1^2 = \int_{-\infty}^{\infty} (x - m_1)^2 f_1(x)\,\mathrm{d}x, \tag{90 b}$$

$$\mathrm{E}[Y] = m_2 = \int_{-\infty}^{\infty} y f_2(y)\,\mathrm{d}y, \tag{90 c}$$

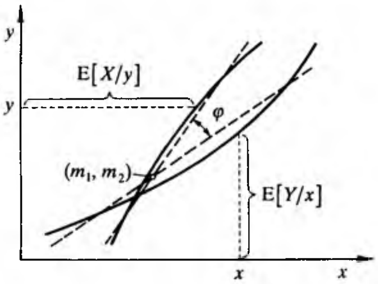
$$\mathrm{D}[Y] = s_2^2 = \int_{-\infty}^{\infty} (y - m_2) f_2(y)\,\mathrm{d}y. \tag{90 d}$$

Za dvodimenzijsku normalnu razdiobu $\mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ marginalne razdiobe su normalne razdiobe, tj. $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ i $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$, tako da je $\mathrm{E}[X] = \mu_1$, $\mathrm{E}[Y] = \mu_2$, $\mathrm{D}[X] = \sigma_1^2$ i $\mathrm{D}[Y] = \sigma_2^2$. Za $\rho = 0$ vrijedi $f(x,y) = f_1(x) \cdot f_2(y)$, tj. X i Y tada su nezavisne slučajne varijable.

Ako je $f_2(y) > 0$, onda je kontinuirana razdioba s funkcijom gustoće vjerojatnosti $p_y(x) = \dfrac{f(x,y)}{f_2(y)}$, ($x \in \mathbf{R}$), *uvjetna razdioba komponente* X za fiksirano y, a broj $\mathrm{E}[X/y] = \int_{-\infty}^{\infty} x\, p_y(x)\,\mathrm{d}x$ *uvjetno očekivanje komponente* X za fiksirano y. Analogno, ako je $f_1(x) > 0$, onda je kontinuirana razdioba s funkcijom gustoće vjerojatnosti $q_x(y) = \dfrac{f(x,y)}{f_1(x)}$, ($x \in \mathbf{R}$), *uvjetna razdioba komponente* Y za fiksirano x, a broj $\mathrm{E}[Y/x] = \int_{-\infty}^{\infty} y\, q_x(y)\,\mathrm{d}y$ *uvjetno očekivanje komponente* Y za fiksirano x.

Ako je $(X, Y) \sim \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, onda je uvjetna razdioba komponente X za fiksirano y normalna razdioba $\mathcal{N}\left(\mu_1 + \rho\dfrac{\sigma_1}{\sigma_2}(y - \mu_2), \sigma_1^2(1-\rho^2)\right)$, tako da je $\mathrm{E}[X/y] = \mu_1 + \rho\dfrac{\sigma_1}{\sigma_2}(y - \mu_2)$. Uvjetna je pak razdioba komponente Y za fiksirano x normalna razdioba $\mathcal{N}\left(\mu_2 + \rho\dfrac{\sigma_2}{\sigma_1}(x - \mu_1), \sigma_2^2(1-\rho^2)\right)$, pa je $\mathrm{E}[Y/x] = \mu_2 + \rho\dfrac{\sigma_1}{\sigma_2}(y - \mu_2)$.

Korelacija. Općenito se funkcije $x \mapsto \mathrm{E}[Y/x]$ i $y \mapsto \mathrm{E}[X/y]$ nazivaju *funkcije regresije*, a njihovi grafovi *krivulje regresije* (sl. 10) zadanoga slučajnog vektora (X, Y).



Sl. 10. Krivulje regresije i pravci regresije

Ako su X i Y nezavisne slučajne varijable, onda je za diskretni slučajni vektor (X, Y)

$$p_{i/j} = p_i, \quad i = 1,2,\ldots, \qquad \mathrm{E}[X/y_j] = \mathrm{E}[X], \quad j = 1,2,\ldots, \tag{91 a}$$

$$q_{j/i} = q_j, \quad j = 1,2,\ldots, \qquad \mathrm{E}[Y/x_i] = \mathrm{E}[Y], \quad i = 1,2,\ldots, \tag{91 b}$$

dok je za kontinuirani slučajni vektor (X, Y)

$$p_y(x) = f_1(x), \quad \mathrm{E}[X/y] = \mathrm{E}[X], \tag{92 a}$$

$$q_x(y) = f_2(y), \quad \mathrm{E}[Y/x] = \mathrm{E}[Y], \tag{92 b}$$

iz čega se vidi da uvjetne razdiobe ne ovise o odabranoj fiksiranoj vrijednosti i jednake su odgovarajućim marginalnim razdiobama. Uvjetna očekivanja također ne ovise o odabranoj fiksiranoj vrijednosti, tako da su u tom slučaju funkcije regresije konstante. Krivulje regresije su pravci usporedni s koordinatnim osima koji se sijeku u središtu razdiobe (u točki (m_1, m_2)).

Stavi li se $\mathrm{E}[XY] = \sum_i \sum_j x_i y_j p_{ij}$ za diskretni slučajni vektor, odnosno $\mathrm{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x y f(x,y)\,\mathrm{d}x\,\mathrm{d}y$ za kontinuirani slučajni vektor (X, Y), može se reći da je $\mathrm{E}[XY]$ očekivanje umnoška slučajnih varijabli X i Y. Ako su X i Y nezavisne slučajne varijable, onda je $\mathrm{E}[XY] = \mathrm{E}[X] \cdot \mathrm{E}[Y]$. Općenito se veličina μ_{11}, odnosno $\mathrm{Cov}(X, Y)$, zove *korelacijski moment*, odnosno *kovarijanca* slučajnih varijabli X i Y:

$$\mu_{11} = \mathrm{Cov}(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X] \cdot \mathrm{E}[Y]. \tag{93}$$

Ako je $\mu_{11} = 0$, onda su X i Y *nekorelirane* slučajne varijable. Nezavisne slučajne varijable su, dakako, i nekorelirane, ali obrat općenito ne vrijedi.

Ako je $\mathrm{D}[X] > 0$ i $\mathrm{D}[Y] > 0$, onda je

$$r = \frac{\mu_{11}}{s_1 s_2} \tag{94}$$

koeficijent korelacije slučajnih varijabli X i Y. Vrijedi da je $r^2 \leqq 1$ i $r^2 = 1$ onda i samo onda ako između varijabli X i Y postoji funkcijska ovisnost oblika $AX + BY + C = 0$ ($A \neq 0$, $B \neq 0$). Ako je $|r| < 1/2$, kaže se da su X i Y *slabo korelirane* slučajne varijable.

Pravac $y = ax + b$, koji u smislu metode najmanjih kvadrata najbolje aproksimira krivulju regresije $x \mapsto \mathrm{E}[Y/x]$, i pravac $x = cy + d$, koji u istom smislu najbolje aproksimira krivulju regresije $y \mapsto \mathrm{E}[X/y]$ (sl. 10), zovu se *pravci regresije*.

Vrijedi da je

$$a = \frac{\mu_{11}}{s_1^2}, \quad b = m_2 - \frac{\mu_{11}}{s_1^2}, \quad c = \frac{\mu_{11}}{s_2^2}, \quad d = m_1 - \frac{\mu_{11}}{s_2^2} m_2, \tag{95}$$

tako da su

$$y - m_2 = \frac{\mu_{11}}{s_1^2}(x - m_1), \quad x - m_1 = \frac{\mu_{11}}{s_2^2}(y - m_2) \tag{96}$$

jednadžbe pravaca regresije. Pravci se sijeku u točki (m_1, m_2) i za kut φ između njih vrijedi

$$\tan\varphi = \frac{1 - r^2}{r} \cdot \frac{s_1 s_2}{s_1^2 + s_2^2}. \tag{97}$$

Ako su X i Y nekorelirane slučajne varijable, onda je $r = 0$, $a = 0$, $c = 0$ i $\varphi = \pi/2$, tj. pravci regresije usporedni su s koordinatnim osima. Za $r = 1$ pravci regresije međusobno se poklapaju ($\varphi = 0$) i tada je dvodimenzijska razdioba »degenerirala« u razdiobu na pravcu $y - m_2 = \frac{s_2}{s_1}(x - m_1)$.

Za trinomnu razdiobu $\mathrm{B}(m, p_1, p_2)$ funkcije su regresije $i \mapsto \mathrm{E}[Y/i] = (m - i)\frac{p_2}{1 - p_1}$ i $j \mapsto \mathrm{E}[X/j] = (m - j)\frac{p_1}{1 - p_2}$. Budući da se radi o diskretnoj dvodimenzijskoj razdiobi, krivulje se regresije sastoje od diskretnog skupa točaka koje leže na istom pravcu. Korelacijski je moment $\mu_{11} = -m p_1 p_2$ i jednadžbe su pravaca regresije $y = (m - x)\frac{p_2}{1 - p_1}$ i $x = (m - y)\frac{p_1}{1 - p_2}$, iz čega se vidi da ti pravci sadrže točke odgovarajućih krivulja regresije. Koeficijent je korelacije $r = -\sqrt{\frac{p_1 p_2}{(1 - p_1)(1 - p_2)}}$, pa za $p_1 + p_2 = 1$ iznosi $r = -1$ i tada je dvodimenzijska razdioba degenerirala u razdiobu uzduž pravca $y = m - x$, odnosno varijable X i Y funkcijski su zavisne, što je izraženo jednadžbom $X + Y = m$.

Za dvodimenzijsku normalnu razdiobu $\mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ je $\mu_{11} = \rho\sigma_1\sigma_2$, tako da su pravci regresije $y - \mu_2 = \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1)$, $x - \mu_1 = \rho\frac{\sigma_1}{\sigma_2}(y - \mu_2)$ ujedno i krivulje regresije. Oni se sijeku u točki (μ_1, μ_2) pod kutom φ, za koji vrijedi $\tan\varphi = \frac{1 - \rho^2}{\rho} \cdot \frac{\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2}$. Ako je $\mu_{11} = 0$, onda je i $\rho = 0$, što znači da nekoreliranost slučajnih varijabli X i Y implicira njihovu nezavisnost, pa se može reći da su za dvodimenzijsku normalnu razdiobu nezavisnost i nekoreliranost ekvivalentni pojmovi.

Višedimenzijska razdioba vjerojatnosti. Ako je ishod slučajnog eksperimenta uređena n-torka ($n = 2, 3, \ldots$) brojeva, odnosno ako se u slučajnom pokusu mjeri n varijabilnih veličina $X_1, \ldots, X_n$, onda je matematički model za tu pojavu n-dimenzijski slučajni vektor $(X_1, \ldots, X_n)$. Slučajni se vektor $(X_1, \ldots, X_n)$ definira kao određena funkcija koja ishodima slučajnog pokusa pridružuje uređene n-torke realnih brojeva $(x_1, \ldots, x_n) \in \mathbf{R}^n$.

Funkcija razdiobe slučajnog vektora $(X_1, \ldots, X_n)$ realna je funkcija n realnih varijabli:

$$F(x_1, \ldots, x_n) = \mathrm{P}(X_1 \leqq x_1, \ldots, X_n \leqq x_n), \tag{98 a}$$

gdje $\mathrm{P}(X_1 \leqq x_1, \ldots, X_n \leqq x_n)$ označuje vjerojatnost da X_1 bude manje ili jednako x_1, X_2 manje ili jednako x_2 itd.

Ako bar jedna od varijabli $x_1, \ldots, x_n$ ima vrijednost $-\infty$, onda funkcija razdiobe poprima vrijednost nula, a ako sve varijable poprime vrijednost ∞, funkcija razdiobe poprima vrijednost jedan. Funkcija

$$x_i \mapsto F_i(x_i) = F(\infty, \ldots, \infty, x_i, \infty, \ldots, \infty) = \mathrm{P}(X_i \leqq x_i) \tag{98 b}$$

marginalna je *funkcija razdiobe komponente* X_i ($i = 1, \ldots, n$). Ona opisuje statističko ponašanje slučajne varijable X_i same za sebe.

Ako vrijedi $F(x_1, \ldots, x_n) = F_1(x_1) \cdot \ldots \cdot F_n(x_n)$, onda su $X_1, \ldots, X_n$ *nezavisne slučajne varijable.*

Uređena n-torka $(m_1, \ldots, m_n)$, gdje je $m_i = \mathrm{E}[X_i]$ ($i = 1, \ldots, n$), jest *vektor očekivanja* slučajnog vektora $(X_1, \ldots, X_n)$ ili *središte* n-dimenzijske razdiobe.

Ako je $n > 2$, onda se mogu promatrati tzv. *dvodimenzijske marginalne razdiobe* zadanoga slučajnog vektora. Tako je $F_{12}(x_1, x_2) = F(x_1, x_2, \infty, \ldots, \infty) = \mathrm{P}(X_1 \leqq x_1, X_2 \leqq x_2)$ marginalna funkcija razdiobe slučajnog vektora (X_1, X_2), dok je $F_{13}(x_1, x_3) = F(x_1, \infty, x_3, \infty, \ldots, \infty) = \mathrm{P}(X_1 \leqq x_1, X_3 \leqq x_3)$ marginalna funkcija razdiobe slučajnog vektora (X_1, X_3). U vezi s n-dimenzijskom razdiobom može se promatrati $\binom{n}{2}$ dvodimenzijskih marginalnih razdioba. Općenito se govori o marginalnoj funkciji razdiobe $F_{ij}(x_i, x_j) = F(\infty, \ldots, \infty, x_i, \infty, \ldots, \infty, x_j, \ldots, \infty, \ldots, \infty) = \mathrm{P}(X_i \leqq x_i, X_j \leqq x_j)$ slučajnog vektora (X_i, X_j), pri čemu su X_i i X_j ($i < j$; $i, j = 1, \ldots, n$) komponente slučajnog vektora $(X_1, \ldots, X_n)$. To omogućuje da se definira tzv. *disperzijska* ili *kovarijancna matrica* S slučajnog vektora $(X_1, ..., X_n)$, tako da se za elemente uzmu brojevi:

$$s_{ij} = \mathrm{E}[X_i X_j] - \mathrm{E}[X_i]\mathrm{E}[X_j], \quad i, j = 1, \dots, n. \tag{99}$$

Disperzijska matrica S simetrična je matrica kojoj dijagonalni element $s_{ii} = s_i^2$ znači disperziju slučajne varijable X_i $(i = 1, ..., n)$, dok izvandijagonalni element s_{ij} $(i \neq j)$ znači kovarijancu slučajnih varijabli X_i i X_j.

Ako je rang disperzijske matrice manji od dimenzije slučajnog vektora razdioba je *degenerirana*.

Ako su $X_1, ..., X_n$ nezavisne slučajne varijable, onda je S dijagonalna matrica. Ako je pak S dijagonalna matrica, onda su $X_1, ..., X_n$ nekorelirane slučajne varijable. Nezavisne slučajne varijable su, dakle, i nekorelirane, dok obrat ne vrijedi.

Poopćenjem trinomne razdiobe $\mathrm{B}(m, p_1, p_2)$ dobiva se tzv. *polinomijalna razdioba* $\mathrm{B}(m, p_1, ..., p_n)$ s parametrima m $(m \in \mathbf{N})$ i $p_1, ..., p_n$ $\left(p_i > 0,\ i = 1, \dots, n,\ \sum_{i=1}^{n} p_i < 1\right)$. Ako se m puta ponavlja slučajni pokus u kojem su uočeni disjunktni događaji $A_1, ..., A_n$ $(\mathrm{P}(A_i) = p_i,\ i = 1, ..., n)$ i pri tome se registriraju brojevi $X_1, ..., X_n$, gdje X_i označuje broj nastupa događaja A_i, onda je $(X_1, ..., X_n)$ diskretni slučajni vektor sa skupom vrijednosti:

$$\mathscr{R}(X_1, ..., X_n) = \{(x_1, ..., x_n) \in \mathbf{R}^n : x_i \in \{0, 1, ..., m\},\ x_1 + ... + x_n \leq m\} \tag{100a}$$

i pripadnim vjerojatnostima:

$$\mathrm{P}(X_1 = x_1, ..., X_n = x_n) =$$

$$= \frac{m!}{x_1! \dots x_n! (m - x_1 - \dots - x_n)!} \cdot p_1^{x_1} \dots p_n^{x_n} (1 - p_1 -, \dots, - p_n)^{m - x_1 - \dots - x_n}. \tag{100b}$$

Piše se $(X_1, ..., X_n) \sim \mathrm{B}(m, p_1, ..., p_n)$ i govori da slučajni vektor $(X_1, ..., X_n)$ ima polinomijalnu razdiobu. Marginalne razdiobe komponenata u polinomijalnoj razdiobi odgovarajuće su binomne razdiobe, a dvodimenzijske marginalne razdiobe odgovarajuće su trinomne razdiobe, tako da je $(m p_1, ..., m p_n)$ vektor očekivanja, dok su $s_{ij} = -m p_i p_j$ $(i \neq j)$, $s_{ii} = s_i^2 = m p_i (1 - p_i)$ elementi disperzijske matrice.

Višedimenzijska normalna razdioba najvažniji je primjer kontinuirane n-dimenzijske razdiobe. Slučajni vektor $(X_1, ..., X_n)$ ima normalnu razdiobu s vektorom očekivanja $\mu = (\mu_1, ..., \mu_n)$ i disperzijskom matricom $\Sigma = \{\sigma_{ij} : \sigma_{ij} = \sigma_{ji},\ \sigma_{ii} = \sigma_i^2\}$ (Σ je regularna i pozitivno definitna) ako se pripadna funkcija razdiobe vjerojatnosti može prikazati u obliku

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(t_1, \dots, t_n) \mathrm{d}t_1, \dots, \mathrm{d}t_n, \tag{101}$$

gdje je

$$F(t_1, \dots, t_n) = (2\pi \det S)^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{ij} (t_i - \mu_i)(t_j - \mu_j)\right] \tag{102}$$

pripadna funkcija gustoće vjerojatnosti, pri čemu su λ_{ij} elementi matrice $\Lambda = \Sigma^{-1}$. Piše se $(X_1, ..., X_n) \sim \mathcal{N}(\mu, \Sigma)$. Komponenti X_i pripada kao marginalna razdioba $\mathcal{N}(\mu_1, \sigma_i^2)$ $(i = 1, ..., n)$, iz čega slijedi da će $X_1, ..., X_n$ biti nezavisne slučajne varijable onda i samo onda ako je Σ, dakle i Λ, dijagonalna matrica, tj. ako su $X_1, ..., X_n$ nekorelirane slučajne varijable.

Funkcije slučajnih varijabli. Ako je $(x_1, ..., x_n) \mapsto h(x_1, ..., x_n)$ određena realna funkcija n $(n \in \mathbf{N})$ realnih varijabli i ako je $(X_1, ..., X_n)$ zadani slučajni vektor, onda je $Y = h(X_1, ..., X_n)$ slučajna varijabla za koju se kaže da je *funkcija slučajnih varijabli* $X_1, ..., X_n$. Najvažnije su *linearne funkcije slučajnih varijabli* za koje je $Y = a_1 X_1 + ... + a_n X_n$, gdje su a_i $(i = 1, ..., n)$ zadani realni brojevi, tzv. koeficijenti. Tada je

$$\mathrm{E}[Y] = \mathrm{E}[a_1 X_1 + \ldots + a_n X_n] = a_1 \mathrm{E}[X_1] + \ldots + a_n \mathrm{E}[X_n], \tag{103}$$

$$\mathrm{D}[Y] = \mathrm{D}[a_1 X_1 + \ldots + a_n X_n] = \sum_{i=j}^{n} \sum_{j=1}^{n} a_i a_j s_{ij}, \tag{104}$$

gdje su s_{ij} elementi disperzijske matrice.

Jednadžba (103) pokazuje da je očekivanje linearne kombinacije slučajnih varijabli jednako linearnoj kombinaciji očekivanja komponenata, dok se iz (104) vidi da je disperzija linearne kombinacije slučajnih varijabli jednaka linearnoj kombinaciji elemenata disperzijske matrice zadanoga slučajnog vektora. Ako su $X_1, ..., X_n$ nekorelirane slučajne varijable, onda (104) postaje

$$\mathrm{D}[a_1 X_1 + \ldots + a_n X_n] = a_1^2 \mathrm{D}[X_1] + \ldots + a_n^2 \mathrm{D}[X_n]. \tag{105}$$

Glavni se problemi u vezi s funkcijama slučajnih varijabli odnose na određivanje razdiobe vjerojatnosti slučajne varijable $Y = h(X_1, ..., X_n)$, uz pretpostavku da je zadana funkcija h i razdioba vjerojatnosti slučajnog vektora $(X_1, ..., X_n)$. Za pojedine posebne slučajeve postoje jednostavna rješenja tog problema:

1. Ako su $X_1, ..., X_n$ nezavisne slučajne varijable i $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ $(i = 1, ..., n)$, onda slučajna varijabla $Y = a_1 X_1 + \ldots + a_n X_n$ ima $\mathcal{N}(\mu, \sigma^2)$, gdje je $\mu = a_1 \mu_1 + \ldots + a_n \mu_n$ i $\sigma^2 = a_1^2 \sigma_1^2 + \ldots + a_n^2 \sigma_n^2$, tj. linearna kombinacija nezavisnih normalnih slučajnih varijabli također je normalna slučajna varijabla.

2. Ako su $X_1, ..., X_n$ nezavisne slučajne varijable i $X_i \sim \mathrm{B}(m_i, p_i)$, $(i = 1, ..., n,\ m_i \in \mathbf{N})$, onda je $Y = X_1 + \ldots + X_n \sim \mathrm{B}(m, p)$, gdje je $m = m_1 + \ldots + m_n$, tj. zbroj nezavisnih binomnih slučajnih varijabla parametra p također je binomna slučajna varijabla.

3. Ako su $X_1, ..., X_n$ nezavisne slučajne varijable i $X_i \sim \mathrm{Po}(a_i)$ $(i = 1, ..., n)$, tada je $Y = X_1 + \ldots + X_n \sim \mathrm{Po}(a)$, gdje je $a = a_1 + \ldots + a_n$, tj. zbroj nezavisnih Poissonovih slučajnih varijabli također je Poissonova slučajna varijabla.

4. Ako su $X_1, ..., X_n$ nezavisne slučajne varijable i $X_i \sim \mathrm{G}(\alpha, \beta_i)$, $(i = 1, ..., n)$, onda je $Y = X_1 + \ldots + X_n \sim \mathrm{G}(\alpha, \beta)$, gdje je $\beta = \beta_1 + \ldots + \beta_n$. Posebno, ako $X_i \sim \mathrm{Ex}(\alpha) = \mathrm{G}(\alpha, 1)$, onda je $Y \sim \mathrm{G}(\alpha, n)$, tj. zbroj n nezavisnih eksponencijalnih slučajnih varijabli parametra α slučajna je varijabla gama-razdiobe $\mathrm{G}(\alpha, n)$.

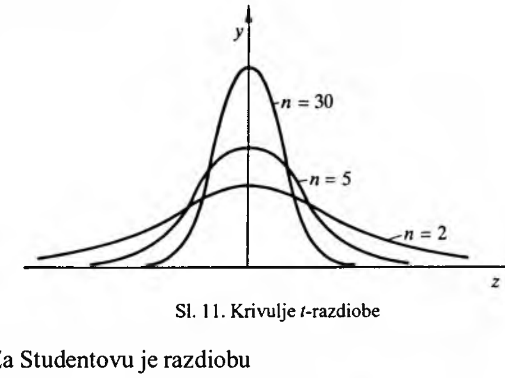
Ako je $X_i \sim \chi^2(n_i) = \mathrm{G}\left(\frac{1}{2}, \frac{n_i}{2}\right)$, onda je $Y \sim \mathrm{G}\left(\frac{1}{2}, \frac{n_1 + \ldots + n_n}{2}\right) = \chi^2(n_1 + \ldots + n_n)$, tj. zbroj nezavisnih slučajnih varijabli hikvadratne razdiobe također je slučajna varijabla hikvadratne razdiobe.

5. Ako su $X_1, ..., X_n$ nezavisne slučajne varijable i $X_i \sim \mathcal{N}(0, 1)$, $(i = 1, ..., n)$, onda je $Y = X_1^2 + \ldots + X_n^2 \sim \chi^2(n)$. To pokazuje da zbroj kvadrata nezavisnih standardnih normalnih slučajnih varijabli ima hikvadratnu razdiobu.

6. Ako su X i Y nezavisne slučajne varijable i $X \sim \mathcal{N}(0, 1)$, a $Y \sim \chi^2(n)$, onda slučajna varijabla $Z = X\sqrt{\frac{n}{Y}}$ $(n \in \mathbf{N})$ ima tzv. *Studentovu* ili *t-razdiobu* s *n stupnjeva slobode*. Piše se $Z \sim t(n)$. Studentovoj razdiobi (sl. 11) pripada funkcija gustoće vjerojatnosti:

$$f(z) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\, \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{z^2}{n}\right)^{-\frac{n+1}{2}}, \quad z \in \mathbf{R}. \tag{106}$$

Za $n = 1$ dobiva se tzv. *Cauchyjeva razdioba*, koja je zanimljiva po tome što nema konačno očekivanje, pa ni bilo koji moment višeg reda.



Sl. 11. Krivulje t-razdiobe

Za Studentovu je razdiobu

$$\mathrm{E}[Z] = 0 \ (n > 1), \quad \mathrm{D}[Z] = \frac{n}{n-2} \ (n > 2). \tag{107}$$

Radi se o simetričnoj razdiobi oko ishodišta.

7. Ako su X i Y nezavisne slučajne varijable i $X \sim \chi^2(r)$, $Y \sim \chi^2(s)$ $(r, s \in \mathbf{N})$, onda slučajna varijabla $Z = \frac{sX}{rY}$ ima tzv. *F-razdiobu* s (r, s) *stupnjeva slobode*. Piše se $Z \sim \mathrm{F}(r, s)$, a pripadna funkcija gustoće vjerojatnosti glasi

$$f(z)=\frac{\Gamma\left(\frac{r+s}{2}\right)r^{\frac{r}{2}}s^{\frac{s}{2}}z^{\frac{r}{2}-1}}{\Gamma\left(\frac{r}{2}\right)\Gamma\left(\frac{s}{2}\right)(rz+s)^{\frac{r+s}{2}}},\quad z>0. \tag{108}$$

Očekivanje postoji za $s>2$, a disperzija za $s>4$ i vrijedi

$$\mathrm{E}[Z]=\frac{s}{s-2},\quad \mathrm{D}[Z]=\frac{2s^2(r+s-2)}{r(s-2)^2(s-4)}. \tag{109}$$

8. Ako su X i Y nezavisne slučajne varijable i $X\sim\mathrm{G}(1,a)$, $Y\sim\mathrm{G}(1,b)$, onda slučajna varijabla $Z=\dfrac{Y}{X+Y}$ ima tzv. *beta-razdiobu* s parametrima a i b (gdje je $a>0$, $b>0$). Piše se $Z\sim\beta(a,b)$, a pripadna funkcija gustoće vjerojatnosti glasi

$$f(z)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}z^{a-1}(1-z)^{b-1},\quad 0<z<1. \tag{110}$$

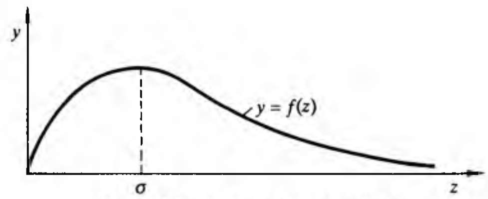
Dalje je

$$\mathrm{E}[Z]=\frac{a}{a+b},\quad \mathrm{D}[Z]=\frac{ab}{(a+b)^2(a+b+1)}. \tag{111}$$

9. Ako su X i Y nezavisne slučajne varijable i $X\sim\mathcal{N}(\mu_1,\sigma^2)$, $Y\sim\mathcal{N}(\mu_2,\sigma^2)$, onda slučajna varijabla $Z=\sqrt{(X-\mu_1)^2+(Y-\mu_2)^2}$, koja se može interpretirati kao udaljenost »slučajne točke« (X,Y) od fiksirane točke (μ_1,μ_2) dane ravnine, ima tzv. *Rayleighovu razdiobu parametra* $\sigma>0$. Njezina funkcija gustoće vjerojatnosti (sl. 12) glasi

$$f(z)=\frac{z}{\sigma^2}\exp\left(-\frac{z^2}{2\sigma^2}\right),\quad z>0, \tag{112}$$

dok je $\mathrm{E}[Z]=\sigma\sqrt{\pi/2}$, a $\mathrm{D}[Z]=\sigma^2(2-\pi/2)$.

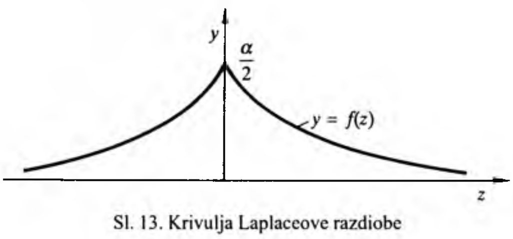


Sl. 12. Krivulja Rayleighove razdiobe

10. Ako su X i Y nezavisne slučajne varijable i obje imaju eksponencijalnu razdiobu $Ex(\alpha)$, onda slučajna varijabla $Z=X-Y$ ima tzv. *Laplaceovu razdiobu* (sl. 13) parametra $\alpha>0$. Pripadna funkcija gustoće vjerojatnosti glasi

$$f(z)=\frac{\alpha}{2}\exp(-\alpha|z|),\quad z\in\mathbf{R}. \tag{113}$$

Zatim je $\mathrm{E}[Z]=0$ i $\mathrm{D}[Z]=2/\alpha^2$.



Sl. 13. Krivulja Laplaceove razdiobe

11. Ako su $X_1,\ldots,X_n$ nezavisne slučajne varijable sa zajedničkom funkcijom distribucije F, onda slučajnoj varijabli $Y=\max(X_1,\ldots,X_n)$ pripada funkcija razdiobe:

$$\mathrm{G}(y)=\mathrm{P}(Y\leqq y)=[\mathrm{F}(y)]^n,\quad y\in\mathbf{R}, \tag{114}$$

a slučajnoj varijabli $Z=\min(X_1,\ldots,X_n)$ pripada funkcija razdiobe:

$$H(z)=\mathrm{P}(Z\leqq z)=1-[1-\mathrm{F}(z)]^n,\quad z\in\mathbf{R}. \tag{115}$$

NIZOVI SLUČAJNIH VARIJABLI

Kao što se u matematičkoj analizi proučavaju nizovi brojeva, vektora, funkcija i sl., tako se u teoriji vjerojatnosti proučavaju nizovi slučajnih varijabli. Glavni je problem u vezi s beskonačnim nizovima ispitivanje konvergencije niza, pa se proučavaju i različiti tipovi konvergencije niza slučajnih varijabla. Zbog posebnosti te problematike dobiveni se rezultati obično iskazuju u obliku teorema koji se nazivaju *zakoni velikih brojeva* ili *granični teoremi.* Najpoznatiji su *Bernoullijev zakon velikih brojeva* i *centralni granični teorem.* U zakonima velikih brojeva obično se navode uvjeti uz koje određeni niz slučajnih varijabli konvergira prema nekoj konstanti (degeneriranoj slučajnoj varijabli), dok se u graničnim teoremima navode uvjeti uz koje zadani niz slučajnih varijabli konvergira prema određenoj slučajnoj varijabli s pripadnom funkcijom razdiobe.

Zakoni velikih brojeva. Ako je A određeni događaj vjerojatnosti $\mathrm{P}(A)=p(0<p<1)$, onda se može promatrati slučajna varijabla $X_i\sim\mathrm{B}(1,p)$ $(i=1,2,\ldots)$, pri čemu $\{X_i=0\}$ označuje događaj da u i-tom ponavljanju slučajnog pokusa A nije nastupio, dok $\{X_i=1\}$ označuje događaj da je A nastupio. Slučajna varijabla $S_n=X_1+\ldots+X_n$ označuje stoga frekvenciju događaja A u nizu od n ponavljanja slučajnog pokusa i $S_n\sim\mathrm{B}(n,p)$, a $\dfrac{S_n}{n}$ može se interpretirati kao relativna frekvencija. Primijeni li se Čebiševljeva nejednakost (68) na slučajnu varijablu $\dfrac{S_n}{n}$, za koju je $\mathrm{E}\left[\dfrac{S_n}{n}\right]=p$ i $\mathrm{D}\left[\dfrac{S_n}{n}\right]=\dfrac{p(1-p)}{n}$, dobiva se

$$\mathrm{P}\left(\left|\frac{S_n}{n}-p\right|\geqq\varepsilon\right)\leqq\frac{p(1-p)}{n\varepsilon^2}\leqq\frac{1}{4n\varepsilon^2}, \tag{116}$$

gdje je $\varepsilon>0$ proizvoljan realan broj. Iz toga se vidi da se vjerojatnost događaja »da apsolutna vrijednost razlike između relativne frekvencije i vjerojatnosti događaja A bude veća ili jednaka od po volji malenog ε« može učiniti po volji malenom, samo treba uzeti dovoljno veliko n. To se piše i ovako:

$$\lim_{n\to\infty}\mathrm{P}\left(\left|\frac{S_n}{n}-p\right|\geqq\varepsilon\right)=0. \tag{117}$$

Relacijom (117) iskazan je *Bernoullijev zakon velikih brojeva.* Pojednostavljeno, njime se izražava činjenica da se vjerojatnost p događaja A može aproksimirati relativnom frekvencijom događaja A uz veliki broj n ponavljanja slučajnog pokusa.

Postoje različita poopćenja Bernoullijeva zakona velikih brojeva. Tako, npr., ako je X_i $(i=1,2,\ldots)$ niz nezavisnih slučajnih varijabli sa zajedničkim očekivanjem $\mathrm{E}[X_i]=\sigma^2$ i zajedničkom konačnom disperzijom $\mathrm{D}[X_i]=\sigma$, onda za niz $Y_n=(1/n)(X_1+\ldots+X_n)$ $(n=1,2,\ldots)$ vrijedi

$$\mathrm{P}(|Y_n-\mu|\geqq\varepsilon)\leqq\frac{\sigma^2}{n\varepsilon^2}, \tag{118}$$

odnosno

$$\lim_{n\to\infty}\mathrm{P}(|Y_n-\mu|\geqq\varepsilon)=0. \tag{119}$$

Kaže se da niz aritmetičkih sredina nezavisnih slučajnih varijabli sa zajedničkim očekivanjem μ i disperzijom σ^2 *stohastički konvergira* prema zajedničkom očekivanju μ.

Centralni granični teorem. Ako je zadan niz Y_n $(n=1,2,\ldots)$ slučajnih varijabli, onda se može promatrati i niz pripadnih funkcija razdiobe $F_n(y)=\mathrm{P}(Y_n\leqq y)$, $y\in\mathbf{R}$, pa se nameće pitanje o uvjetima postojanja funkcije F koja bi bila limes niza funkcija razdiobe i koja bi i sama bila funkcija razdiobe neke slučajne varijable Y. Tada se kaže da niz Y_n konvergira po razdiobi prema slučajnoj varijabli Y i piše $Y_n\xrightarrow{\mathrm{D}}Y$.

Prvi rezultat tipa centralnog graničnog teorema dokazali su početkom XVIII. st. Moivre i Laplace i on se sastoji u tome da niz

$Y_n = \dfrac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{D} Y \sim \mathcal{N}(0,1)$, tj. niz centriranih ($\mathrm{E}[Y_n]=0$) i normiranih ($\mathrm{D}[Y_n]=1$) slučajnih varijabli, koje se mogu interpretirati kao normirane razlike frekvencije S_n događaja A u nizu od n ponavljanja slučajnog pokusa i broja np ($np=\mathrm{E}[S_n]$), konvergira po razdiobi prema slučajnoj varijabli Y standardne normalne razdiobe. To znači da se za velike n može uzeti da slučajna varijabla S_n, kojoj inače pripada binomna razdioba $\mathrm{B}(n,p)$, približno ima normalnu razdiobu $\mathcal{N}(np, np(1-p))$, odnosno da se binomna razdioba $\mathrm{B}(n,p)$ može aproksimirati normalnom razdiobom $\mathcal{N}(np, np(1-p))$. Posebno, za veliko n i $a<b$ vrijedi

$$\mathrm{P}(a < S_n < b) \approx \Phi\left(\frac{b-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right). \quad (120)$$

Nešto općenitija verzija centralnog graničnog teorema potječe od P. Levyja. Ako je, naime, X_i ($i=1,2,\ldots$) niz nezavisnih slučajnih varijabli iste razdiobe vjerojatnosti s očekivanjem μ i disperzijom σ^2, gdje je $0<\sigma^2<\infty$, i $S_n = \sum_{i=1}^{n} X_i$, tada niz $Y_n = \dfrac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} Y \sim \mathcal{N}(0,1)$. Može se, dakle, reći da za veliko n slučajna varijabla S_n, koja u ovom slučaju označuje zbroj od n nezavisnih slučajnih varijabli iste razdiobe vjerojatnosti, približno ima normalnu razdiobu $\mathcal{N}(n\mu, n\sigma^2)$. Ako je $\bar{X}_n = S_n/n$, dobiva se da je $Y_n = \dfrac{\bar{X}_n - \mu}{\sigma}\sqrt{n}$, pa se vidi da će za veliko n aritmetička sredina n nezavisnih i jednako raspoređenih slučajnih varijabli približno imati normalnu razdiobu $\mathcal{N}\left(\mu, \dfrac{\sigma^2}{n}\right)$, tj. za $a<b$ vrijedi

$$\mathrm{P}(a \leqq \bar{X}_n \leqq b) \approx \Phi\left(\frac{b-\mu}{\sigma}\sqrt{n}\right) - \Phi\left(\frac{a-\mu}{\sigma}\sqrt{n}\right). \quad (121)$$

Granične razdiobe ekstrema slučajnih varijabli. Ako je X_i ($i=1,2,\ldots$) niz nezavisnih slučajnih varijabli sa zajedničkom funkcijom razdiobe F, onda se mogu promatrati nizovi slučajnih varijabli $Y_n = \max(X_1,\ldots,X_n)$ i $Z_n = \min(X_1,\ldots,X_n)$ ($n=1,2,\ldots$), kao i pripadni nizovi funkcija razdiobe $\mathrm{G}_n(y) = \mathrm{P}(Y_n \leqq y) = [\mathrm{F}(y)]^n$ i $\mathrm{H}_n(z) = \mathrm{P}(Z_n \leqq z) = 1 - [1-\mathrm{F}(z)]^n$. Promotri li se $\lim_{n\to\infty} \mathrm{G}_n(y)$, odnosno $\lim_{n\to\infty} \mathrm{H}_n(z)$, uočit će se da će granična razdioba redovito biti degenerirana razdioba, pa je stoga potrebno istražiti uvjete uz koje postoje nizovi konstanti a_n, b_n ($n=1,2,\ldots$) takvi da niz slučajnih varijabli $\hat{Y}_n = \dfrac{Y_n}{a_n} - b_n$, odnosno $\hat{Z}_n = \dfrac{Z_n}{a_n} - b_n$, konvergira po razdiobi prema nekoj nedegeneriranoj slučajnoj varijabli. Budući da vrijedi $\min(X_1,\ldots,X_n) = -\max(-X_1,\ldots,-X_n)$, dovoljno je proučiti niz Y_n, odnosno ponašanje niza pripadnih funkcija razdiobe:

$$\hat{\mathrm{G}}_n(y) = \mathrm{P}(\hat{Y}_n \leqq y) = \mathrm{G}_n\left(\frac{y+b_n}{a_n}\right) = \left[\mathrm{F}\left(\frac{y+b_n}{a_n}\right)\right]^n, y \in \mathbf{R}. \quad (122)$$

Problemi u vezi s maksimumom i minimumom niza nezavisnih slučajnih varijabli uočeni su početkom dvadesetog stoljeća prilikom proučavanja različitih stvarnih pojava, kao što su maksimalna opterećenja tehničkih naprava, maksimalne vrijednosti meteoroloških pojava i sl. U četrdesetim se godinama došlo do prilično cjelovita rješenja (B. V. Gnedenko). Utvrđeno je, naime, da vrijedi sljedeće:

1. Ako je zajednička funkcija razdiobe F tzv. *eksponencijalnog tipa*, onda niz slučajnih varijabli $\hat{Y}_n = Y_n - \ln\beta n$ ($n=1,2,\ldots$) konvergira po razdiobi prema slučajnoj varijabli Y kojoj pripada tzv. Gumbellova razdioba (sl. 14) s funkcijom razdiobe:

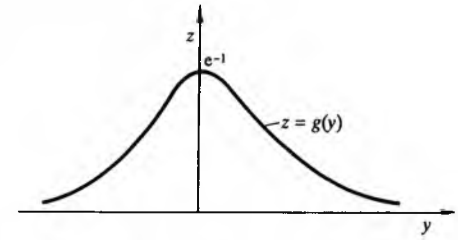
$$\mathrm{G}(y) = \exp[-\exp(-y)], \quad y \in \mathbf{R}, \quad (123)$$

odnosno s pripadnom funkcijom gustoće vjerojatnosti:

$$g(y) = \exp\{-[y + \exp(-y)]\}, \quad y \in \mathbf{R}, \quad (124)$$

te očekivanjem $\mathrm{E}[Y] = \gamma \approx 0{,}57722\ldots$ (*Eulerova konstanta*) i disperzijom $\mathrm{D}[Y] = \pi^2/6$.

U razdiobe eksponencijalnog tipa idu *normalne razdiobe, gama-razdiobe* i još neke druge razdiobe.



Sl. 14. Krivulja Gumbellove razdiobe

2. Ako je zajednička funkcija razdiobe F tzv. *Cauchyjeva tipa*, onda niz slučajnih varijabli $\hat{Y}_n = (n\beta)^{-1/\alpha} Y_n$ ($n=1,2,\ldots$) konvergira po razdiobi prema slučajnoj varijabli Y kojoj pripada tzv. *Frechetova razdioba* parametra α s funkcijom razdiobe:

$$\mathrm{G}(y) = \begin{cases} 0, & \text{za } y \leqq 0 \\ \exp(-y^{-\alpha}), & \text{za } y > 0 \end{cases} \quad (125)$$

i pripadnom funkcijom gustoće vjerojatnosti:

$$g(y) = \begin{cases} 0, & \text{za } y \leqq 0 \\ \alpha x^{-\alpha-1}\exp(-y^{-\alpha}), & \text{za } y > 0, \end{cases} \quad (126)$$

te očekivanjem $\mathrm{E}[Y] = \Gamma\left(1 - \dfrac{1}{\alpha}\right)$ za $\alpha > 1$ i disperzijom $\mathrm{D}[Y] = \Gamma\left(1-\dfrac{2}{\alpha}\right) - \Gamma^2\left(1 - \dfrac{1}{\alpha}\right)$ za $\alpha > 2$.

3. Ako je zajednička razdioba vjerojatnosti f odozgo ograničena brojem a, onda niz slučajnih varijabli $\hat{Y}_n = (n\beta)^{1/\alpha}(Y_n - a)$ konvergira po razdiobi prema slučajnoj varijabli Y kojoj pripada tzv. *Weibullova razdioba* parametra α. Pripadna funkcija gustoće vjerojatnosti glasi

$$g(y) = \begin{cases} \alpha(-y)^{\alpha-1}\exp[-(-y)^\alpha], & \text{za } y < 0 \\ 0, & \text{za } y \geqq 0 \end{cases} \quad (127)$$

Očekivanje je $\mathrm{E}[Y] = -\Gamma\left(\dfrac{1}{\alpha}+1\right)$ i disperzija $\mathrm{D}[Y] = \Gamma\left(\dfrac{2}{\alpha}+1\right) - \Gamma^2\left(\dfrac{1}{\alpha}+1\right)$.

Budući da razdioba vjerojatnosti definirana sa (127) ima pozitivne gustoće vjerojatnosti na negativnom dijelu apscisne osi, dok na pozitivnom dijelu apscisne osi ima gustoću nula, katkada se Weibullovom razdiobom naziva simetrijom preslikana razdioba s obzirom na ishodište i tada pripadna funkcija gustoće vjerojatnosti glasi

$$\bar{g}(y) = \begin{cases} 0, & \text{za } y \leqq 0 \\ \alpha y^{\alpha-1}\exp(-y^\alpha), & \text{za } y > 0 \end{cases} \quad (128)$$

Očekivanje je tada $\Gamma\left(\dfrac{1}{\alpha}+1\right)$.

U zaključku se može reći da se vjerojatnosne (probabilističke) metode i probabilistički pristup određenim tehničkim problemima (sigurnost, pouzdanost, učinkovitost i sl.) sve više primjenjuju u različitim područjima tehničkih znanosti. Poznavanje temeljnih načela, te osnovnih pojmova i postupaka teorije vjerojatnosti postaje nužno sredstvo za uspješan istraživački rad u svim područjima tehnike, a također i komponenta stručnog znanja u inženjerskoj praksi.

LIT.: *L. Breiman*, Probability and Stochastic Processes with a View Toward Applications. Houghton Mifflin Company, Boston 1969. – *Б. В. Гнеденко*, Курс теории вероятностей. Наука, Москва 1969. – *A. J. Thomasian*, The Structure of Probability Theory with Applications. McGraw-Hill, New York 1969. – *A. Renyi*, Probability Theory. North-Holland Co., Amsterdam–London 1970. – *W. Feller*, An Introduction to Probability Theory and its Applications, Vol 1, 2. J. Wiley, New York 1968, 1971. – *V. Vranić*, Vjerojatnost i statistika. Tehnička knjiga, Zagreb [3]1971. – *А. Н. Колмогоров*, Основные понятия теории вероятностей. Наука, Москва 1974. – *N. Sarapa*, Teorija vjerojatnosti. Školska knjiga, Zagreb 1987. – *Ž. Pauše*, Vjerojatnost, informacija, stohastički procesi. Školska knjiga, Zagreb [4]1988. – *Ž. Pauše*, Uvod u matematičku statistiku. Školska knjiga, Zagreb 1993.

Ž. Pauše