

Matematično upanje

Diskretno porazdeljena sl. sprem.

X

∼

⎛

x

1

x

2

⋯

p

1

p

2

⋯

⎞

{\displaystyle X\sim {\begin{pmatrix}x_{1}&x_{2}&\ldots \\p_{1}&p_{2}&\ldots \end{pmatrix}}}

$$E(X)=\sum_{k=1}^{\infty}x_kp_k\quad \text{če}\quad \sum_{k=1}^{\infty}|x_k|p_k<\infty$$

Zvezno porazdeljena sl. sprem. *X* z gostoto *p**X*

$$E(X)=\int_{-\infty}^{\infty}xp_X(x)dx\quad \text{če}\quad \int_{-\infty}^{\infty}|x|p_X(x)dx<\infty$$

Lastnosti

Naj bo *f* : ℝ → ℝ zvezna funkcija. Potem je

E
(
f
(
X
)
)
=

∑

k
=
1

∞

f
(

x

k

)

p

k

{\displaystyle \;E(f(X))=\sum _{k=1}^{\infty }f(x_{k})p_{k}}

če obstaja

E
(
f
(
X
)
)
=

∫

−
∞

∞

f
(
x
)

p

X

(
x
)
d
x

{\displaystyle \;E(f(X))=\int _{-\infty }^{\infty }f(x)p_{X}(x)dx}

če obstaja

Če ima |*X*| mat. up., ga ima tudi *X* in velja

$$|E(X)|\leq E(|X|)$$

Če obstaja *E*(*X*²) in *E*(*Y*²), obstaja tudi *E*(*XY*) in velja:

$$|E(XY)|\leq E(|XY|)\leq {\sqrt{E(X^2)E(Y^2)}}$$

Za poljubne sl. sprem *X*₁, ..., *X*_{*n*} velja:

$$E(a_1X_1+\ldots a_nX_n)=a_1E(X_1)+\cdots +a_nE(X_n)$$

Disperzija (varianca)

$$D(X)=E((X-E(X))^2)=E(X^2)-(E(X))^2$$

Lastnosti:

- D*(*X*) ≥ 0
- D*(*X*) = 0 ⇔ *P*(*X* = *E*(*X*)) = 1
- D*(*aX*) = *a*²*D*(*X*)

Standardna diviacija/odklon:

$$\sigma (X)=\sqrt{D(X)}$$

zanjo velja σ(*aX*) = |*a*|σ(*X*).

Nekoreliranost

Sl. sprem. *X* in *Y* sta nekorelirani, če velja:

$$E(XY)=E(X)E(Y)$$

X,*Y* neodvisni ⟹ *X*,*Y* nekorelirani

Če imata *X* in *Y*, je nekoreliranost ekvivalentna zvezi:

$$D(X+Y)=D(X)+D(Y)$$

Kovarianca

$$\begin{aligned}K(X,Y)&=E((X-E(X))(Y-E(Y)))\\&=E(XY)-E(X)E(Y)\end{aligned}$$

- K*(*X*,*X*) = *D*(*X*)
- K*(*X*,*Y*) = 0 ⇔ *X*,*Y*nekorelirani
- K*(*aX*,*bY*,*Z*) = *aK*(*X*,*Z*) + *bK*(*Y*,*Z*)
- K*(*X*,*Y*) = *K*(*Y*,*X*)
- K*(*aX* + *b*, *cY* + *d*) = *acK*(*X*,*Y*)
- |*K*(*X*,*Y*)| ≤ √*D*(*X*)*D*(*Y*)
- D*(*X* + *Y*) = *D*(*X*) + *D*(*Y*) + 2*K*(*X*,*Y*)
- D*(*X*₁ + ⋯ + *X*_{*n*}) = *D*(*X*₁) + ⋯ + *D*(*X*_{*n*}) + 2∑_{*i*=1}^{*n*−1} ∑_{*j*=*i*+1}^{*n*} *K*(*X*_{*i*},*X*_{*j*})

Standardizacija

$$X_S=\frac{X-E(X)}{\sigma (X)}$$

Korelacijski koeficient

$$r(X,Y)=\frac{K(X,Y)}{\sigma (X)\sigma (Y)}=E(X_S,Y_S)$$

Lastnosti:

- r*(*X*,*Y*) = 0 ⟺ *X*,*Y*nekorelirani

- −1 ≤ *r*(*X*,*Y*) ≤ 1

- r*(*X*,*Y*) = 1 ⟺ *P*(*X**S* = *Y**S*) = 1

- r*(*X*,*Y*) = −1 ⟺ *P*(*X**S* = −*Y**S*) = 1

- r*(*aX* + *b*, *cY* + *d*) = *r*(*X*,*Y*)

Pogojne porazdelitve

Pogojna porazdelitev sl. sprem. *X* glede na dogodek *B*:

$$X|B\sim \left(P(X=a_1|B)\quad P(X=a_2|B)\quad \ldots \right)$$

Pogojna porazdelitvena funkcija

sl. sprem. *X* glede na dogodek *B*:

$$F_{X|B}(x)=F_X(x|B)=P(X\leq x|B)=\frac{P((X\leq x)\cap B)}{P(B)}$$

Če je pogojna porazdelitev zvezna, obstaja tudi **pogojna porazdelitvena gostota**:

$$p_{X|B}(x)=F_{X|B}'(x)$$

Pogojna gostota

$$p_X(x|Y=y)\equiv p_X(x|y)=\frac{p_{(X,Y)}(x,y)}{p_Y(y)}$$

Pogojno matematično upanje

$$E(h(X)|B)=\sum_xh(x)P(X=x|B)$$

$$E(X|Y=y)=\frac{1}{p_Y(y)}\int_{-\infty}^{\infty}xp_{(X,Y)}(x,y)dx$$

Za vsako slučajno spremenljivko *X* in dogodek *B* veleja:

$$E(X|B)=\frac{E(XZ)}{P(B)}=\frac{E(XZ)}{E(Z)}$$

kjer je sl. sprem. *Z* indikator dogodka *B*.

Za vsako sl. sprem. *X* z mat. up. in popoln sistem dogodkov *H*₁,*H*₂, ... velja **izrek o polni pričakovani vrednosti**

$$E(X)=P(H_1)E(X|H_1)+P(H_2)E(X|H_2)+\ldots$$

Regresijska funkcija

$$\varphi (y)=E(X|Y=y)$$

Za vsako sl. sprem. *X* z mat. up. in diskretno sl. sprem. *Y* velja:

$$\begin{aligned}E(Xg(Y)|Y)&=E(X|Y)g(Y)\\E(Xg(Y))&=E(E(X|Y)g(Y))\\E(X)&=E(E(X|Y))\end{aligned}$$

Za vsak dododek *A* in vsako sl. sprem *Y* velja:

$$E(P(A|Y))=P(A)$$

Momenti

Moment reda *k* glede na točko *a* je

$$m_k(a)=E((X-a)^k)\quad \textit{če obstaja}$$

- Začetni moment** *z*_{*k*} := *m*_{*k*}(0) = *E*(*X*^{*k*})
- Centralni moment** *m*_{*k*} := *m*_{*k*}(*E*(*X*)) = *E*((*X* − *E*(*x*))^{*k*})

$$z_1=E(X)\qquad m_2=D(X)$$

Če obstaja *m*_{*n*}(*a*), obstaja tudi *m*_{*k*}(*a*) za ∀*k* < *n*.

Če obstaja *z*_{*n*}, obstaja tudi *m*_{*n*}(*a*) za ∀*a* ∈ ℝ

Centralne momente lahko izračunamo iz začetnih:

$$m_n=\sum_{k=0}^n{\binom{n}{k}}(-1)^{n-k}z_1^{n-k}z_k$$

Asimetrija

$$A(X)=E(X_S^3)=E\left(\left(\frac{X-E(X)}{\sigma (X)}\right)^3\right)=\frac{m_3}{m_2^{\frac{3}{2}}}$$

∀λ > 0 : *A*(λ*X*) = *A*(*X*)

Sploščenost (kurtozis)

$$K(X)=E\left[\left(\frac{X-E(X)}{\sqrt{D(X)}}\right)^4\right]=\frac{m_4}{m_2^2}$$

Presežna sploščenost:

$$K^{*}(X)=K(X)-3$$

Vrstilne karakteristike

Kvantil reda *p*

je vsaka vrednost *x*_{*p*}, za katero velja:

$$P(X\leq x_p)\geq p\; \mathrm{in}\; P(X\geq x_p)=1-p$$

oz. *F*(*x*_{*p*}−) ≤ *p* ≤ *F*(*x*_{*p*})

- Mediana: *x*_{1⁄2}

- Kvartili: *x*1⁄4, *x*2⁄4, *x*3⁄4

- (Per)centili: *x*1⁄100, ..., *x*99⁄100

Semi interkvartilni razmik

$$s=\frac{1}{2}\left(x_{\frac{3}{4}}-x_{\frac{1}{4}}\right)$$

Rodovne funkcije

Naj bo *X* sl. sprem. z zalogo vrednosti ℕ ∪ {0}:

$$p_k=P(X=k)\qquad k=0,1,2,\ldots$$

Rodovna funkcija sl. sprem. *X*:

$$G_X(s)=p_0+p_1s+p_2s^2+\cdots=\sum_{k=0}^{\infty}p_ks^k$$

Obstaja za vse |*s*| ≤ 1.

$$P(X=k)=\frac{G_X^{(k)}(0)}{k!}$$

$$G_X(0)=p_0\quad G_X(1)=1\quad G_X(s)=E(s^X)$$

Izrek o enoličnosti:

$$\begin{aligned}\forall s\in[-1,1]:G_X(s)&=G_Y(s)\\&\iff P(X=k)=P(Y=k)\;\forall k=0,1,2,\ldots\end{aligned}$$

$$\lim_{s\uparrow 1}G_X'(s)=\lim_{s\uparrow 1}\sum_{k=1}^{\infty}kp_ks^{k-1}=\sum_{k=1}^{\infty}\lim_{s\uparrow 1}kp_ks^{k-1}E(X)$$

Naj bo *X* sl. sprem. z rodovno funkcijo *G*_{*X*}, potem je:

$$G_X^{(n)}(1-)=E(X)(X-1)(X-2)\ldots(X-n+1))$$

Naj bosta *X*₁, ..., *X*_{*n*} nedovisne sl. sprem. z rodovnimi funkcijami *G*_{*X*1}, ... *G*_{*X**n*}:

$$G_{X_1+\cdots +X_n}=G(X_1)\ldots G(X_n)\qquad \forall s\in[-1,1]$$

Naj bodo ∀*n* ∈ ℕ sl. sprem *N*,*X*₁, ..., *X*_{*n*} neodvisne. Naj ima *N* rodovno funkcijo *G*_{*N*} in *X*_{*i*} rodovno funkcijo *G*_{*X*} (*X*₁, ..., *X*_{*n*} so enako porazdeljene). Naj bo *S* = *X*₁ + ⋯ + *X*_{*n*}. Potem je:

$$G_S=G_N(G_X(s))\qquad \forall s\in[-1,1]$$

Velja tudi *E*(*S*) = *E*(*N*)*E*(*X*).

$$G_{2X}(s)=G_X(s^2)$$

Znane rodovne funkcije

$$\begin{aligned}\sum_{n=0}^{\infty}q^n&=\frac{1}{1-q}\qquad \sum_{n=0}^bq^n=\frac{1-q^{b+1}}{1-q}\\ \sum_{n=a}^{\infty}q^n&=\frac{q^a}{1-q}\qquad \sum_{n=a}^bq^n=\frac{q^a-q^{b+1}}{1-q}\end{aligned}$$

$$a^n-b^n=(a-b)(a^{n-1}+a^{n-2}b+...+ab^{n-2}+b^{n-1})$$

$$\frac{a_0+...+a_{k-1}x^{k-1}}{1-x^k}=a_0+...+a_{k-1}x^{k-1}+a_0^k+...+a_{k-1}x^{2k-1}+...$$

$$(x+y)^n=\sum_{k=0}^n{\binom{n}{k}}x^{n-k}y^k$$

$$\frac{1}{(1-x)^n}=\sum_{k=0}^n{\binom{n+k-1}{k}}x^k$$

$$B_{\lambda}(x)=\sum_n{\binom{\lambda}{n}}x^n=(1+x)^{\lambda};\qquad {\binom{\lambda}{n}}=\frac{\lambda^n}{n!}$$

Momentno rodovna funkcija

$$\begin{aligned}M_X(t)&=E(e^{tX})\qquad \forall t\in R\quad \textit{če obstaja}\\&=1+z_1t+\frac{z_2}{2!}t^2+\frac{z_3}{3!}t^3+\ldots\end{aligned}$$

V primeru, ko ima *X* zalogo vrednosti ℕ ∪ {0}, je

$$M_X(t)=E(e^{tX})=G_X(e^t)$$

Za zvezno porazdeljeno sl. sprem. *X* velja:

$$M_X(t)=\int_{-\infty}^{\infty}e^{tx}p_X(x)dx$$

Naj pri nekem δ > 0 *M*_{*X*}(*t*) obstaja za vse *t* ∈ (−δ,δ). Potem je porazdelitev za *X* natanko določana z *M*_{*X*} in vsi začetni momenti obstajajo:

$$z_k=E(X^k)=M_X^{(k)}(0)\qquad \forall k\in \mathbb{N}$$

$$M_X(t)=\sum_{k=0}^{\infty}\frac{z_k}{k!}t^k\qquad \forall t\in(-\delta,\delta)$$

Trditev:

$$M_{aX+b}(t)=e^{bt}M_X(at)$$

Če sta *X* in *Y* neodvisni, je:

$$M_{X+Y}(t)=M_X(t)M_Y(t)$$

Izreka o velikih številih

Zaporedje sl. sprem. {*X*_{*n*}}*n*∈ℕ **verjetnostno** konvergira proti sl. sprem. *X*, če

$$\begin{aligned}\forall \varepsilon > 0: \lim_{n\rightarrow \infty}P(|X_n-X|\geq \varepsilon)&=0\\ \forall \varepsilon > 0: \lim_{n\rightarrow \infty}P(|X_n-X|\leq \varepsilon)&=1\end{aligned}$$

Zaporedje sl. sprem. {*X*_{*n*}}*n*∈ℕ **skoraj gotovo** (s.g.) konvergira proti sl. sprem. *X*, če

$$P\big(\lim_{n\rightarrow \infty}=1$$

$$\begin{aligned}\text{Če }X_n\stackrel{s.g.}{n\rightarrow \infty}\rightarrow X, \text{ potem}\\ \forall \varepsilon > 0: \lim_{m\rightarrow \infty}P(|X_n-X|<\varepsilon, \forall n\geq m)=1\end{aligned}$$

$$\text{Če }X_n\stackrel{s.g.}{n\rightarrow \infty}\rightarrow X, \text{ potem }X_n\stackrel{ver.}{n\rightarrow \infty}\rightarrow X$$

Naj bo *X*₁,*X*₂,*X*_{*n*}, ... zaporedje sl. sprem. z mat. up. in naj bo *S* = *X*₁ + ⋯ + *X*_{*n*}, *Y*_{*n*} =

S

n

−
E
(

S

n

)

n

. Potem je *E*(*Y*_{*n*}) = 0.

Za {*X*_{*n*}}*n*∈ℕ velja **šibki zakon o velikih številih**, če zap. {*Y*_{*n*}}*n*∈ℕ konvergira proti 0 verjetnostno, tj.

$$\forall \varepsilon > 0: \lim_{n\rightarrow \infty}P\left(\frac{|S_n-E(S_n)|}{n}<\varepsilon\right)=1$$

Za {*X*_{*n*}}*n*∈ℕ velja **kreпки zakon o velikih številih**, če zap. {*Y*_{*n*}}*n*∈ℕ konvergira proti 0 skoraj gotovo, tj.

$$P\left(\lim_{n\rightarrow \infty}\frac{S_n-E(S_n)}{n}=0\right)=1$$

Neenakost Markova

Če je *X* sl. sprem. z mat. up., potem je

$$P(|X|\geq a)\leq \frac{E(|X|)}{a}\qquad \forall a>0$$

Neenakost Čebiševa

Če ima *X* disperzijo, je

$$P(|X-E(X)|\geq a\sigma (X))\leq \frac{1}{a^2}\qquad \forall a>0$$

oziroma za ε := *aσ*(*X*)

$$P(|X-E(X)|\geq \varepsilon)\leq \frac{D(X)}{\varepsilon ^2}$$

Izrek Markova: Če za sl. sprem. {*X*_{*n*}}*n*∈ℕ velja

D
(

S

n

)

n

2

{\displaystyle {\frac {D(S_n)}{n^2}}}

→
0,

n
→
∞

{\displaystyle \rightarrow 0,n\rightarrow \infty }

, potem velja ŠZVŠ.

Izrek Čibiševa: Če so sl. sprem. *X*₁,*X*₂, ... paroma nekorelirane in je sup*n*∈ℕ *D*(*X*_{*n*}) < ∞, potem velja ŠZVŠ.

Izrek Kolmogorova Naj za neodvisne slučajne spremenljivke *X*₁,*X*₂, ... velja pogoј