Parcijalni koeficijent korelacije

- Ako je koeficijent korelacije blizak 1, ne mora značiti da su X i Y međusobno zavisne, već da postoji treća promenljiva Z od koje zavise X i Y.
- Definiše se parcijalni koeficijent korelacije pomoću kojeg se nalazi zavisnost X od Y **bez uticaja** Z.
- Neka su ρ_{12}, ρ_{13} i ρ_{23}, koeficijenti korelacija između X i Y, X i Z i između Y i Z. Parcijalni koeficijent korelacije između X i Y **bez uticaja** Z:

$$\frac{\rho_{12} - \rho_{13} \cdot \rho_{23}}{\sqrt{(1-\rho_{13}^2)(1-\rho_{23}^2)}}$$

Koeficijent asimetrije i spljoštenosti

- Pomoću matematičkog očekivanja definišu se i koeficijent asimetrije i spljoštenosti.
- **Definicija. Neka su $E(X)=m$ i $D(X)=\sigma^2$ matematičko očekivanje i disperzija slučajne promenljive X. Ako postoje veličine**

$$f_1 = \frac{E(X-m)^3}{\sigma^3} \qquad f_2 = \frac{E(X-m)^4}{\sigma^4} - 3$$

Koeficijent asimetrije

Koeficijent spljoštenosti

- Koriste se i termini: prvi i drugi Fišerov koeficijent, i prvi i drugi Pirsonov koeficijent.

Koeficijent varijacije, indeks disperzije

- Ako je $E(X) \neq 0$, tada je **koeficijent varijacije**:

$$c_V = \frac{\sigma}{E(X)}$$

- Ako je $E(X) \neq 0$, tada je **indeks disperzije**:

$$I = \frac{D(X)}{E(X)}$$

Karakteristična funkcija

- Pomoću matematičkog očekivanja definiše se funkcionalna karakteristika slučajnih promenljivih – ***karakteristična funkcija.***
- Neka je data sp X. Funkcija φ definisana jednakošću:

$$\varphi(t)=E(e^{itX}),\ t\in R,\ i^2=-1$$

  Naziva se karakteristična funkcija slučajne promenljive X.
- Karakteristična funkcija je jedinstvena za svaku sp.
- Koristi se u teoriji verovatnoće pri dokazivanju mnogih teorema.

Mod raspodele

- **Definicija. Neka je data diskretna sp X (sa konačno ili prebrojivo mnogo vrednosti) svojim zakonom raspodele p_j=P[X=x_j], $j \in J \subseteq N$. Svaka vrednost ili vrednosti sp X čije su odgovarajuće verovatnoće veće od susednih je *mod raspodele*.**
- Neka je data neprekidna sp X svojom gustinom raspodele $g(x)$, $x \in R$. Apscisa svake tačke lokalnog maksimuma funkcije $g(x)$ je mod raspodele.
- Raspodela može imati jedan mod – **unimodalna**, ili dva moda – **bimodalna**, ili više – **polimodalna**.

Kvantili raspodele

- **Definicija. Neka je data sp X i neka je $F(x)$ njena funkcija raspodele. Neka je q realan broj iz intervala (0, 1). *Kvantil reda* q je svaki broj $x_0 \in R$ za koji važe nejednakosti**:

$$F(x_0) \leq q \ \text{ i } \lim_{x \to x_0} F(x) \geq q$$

- Ako je $F(X)$ neprekidna i strogo monotona, tada će postojati jedinstveni kvantili svakog reda.
- Ako $F(X)$ ima intervale konstantnosti, može postojati više brojeva x_0 koji su kvantil nekog reda za posmatranu raspodelu.

Decili i kvartili

- Ako je q=0,1, tada je odgovarajući broj x_0 **prvi decil raspodele**;
- Ako je q=0,2, tada je odgovarajući broj x_0 **drugi decil raspodele**, itd.
- Ako je q=0,25, tada je odgovarajući broj x_0 **prvi kvartil raspodele**;
- Ako je q=0,5, tada je odgovarajući broj x_0 **medijana raspodele**,
- Ako je q=0,75, tada je odgovarajući broj x_0 **treći kvartil raspodele.**
- Ako je q=0,95, tada je odgovarajući broj x_0 **95-ti percentil raspodele.**

Matematičko očekivanje 2Dsp

- Ako X i Y imaju matematička očekivanja, tada je uređeni par ($E(X)$, $E(Y)$) matematičko očekivanje sp (X, Y).
- Neka je (X, Y) dvodimenzionalna sp. Uslovno matematičko očekivanje sp Y pri fiksiranoj vrednosti $X= x$ je

$$E(Y/X=x)=\int_{-\infty}^{\infty} y\cdot g(y/x)\mathrm{d}y$$

gde je $g(y/x)$ uslovna gustina raspodele sp Y pri $X= x$.

- Ako su X i Y diskretne sp, tada je

$$E(Y/X=x_i)=\sum_j y_j P[Y=y_j/X=x_i]$$

Neke važnije raspodele verovatnoće

- Poznavanje raspodela i njihovih svojstava značajno kada na osnovu konačnog broja podataka treba utvrditi kojoj klasi sp pripada i koje su njene osnovne karakteristike.
- Diskretne: ***binomna***, ***negativna binomna*** i ***Puasonova raspodela***.
- ***Normalna raspodela***
- Neprekidne: ***uniformna***, ***eksponencijalna***, ***hi-kvadrat***, ***studentova***, ***Fišerova***, ***Gama raspodela***, itd.

Binomna raspodela

- Neka je verovatnoća realizacije dog. A u n eksperimenata jednaka p. Ako je sp X jednaka broju realizacija dog. A pri istim uslovima, tada X ima **binomnu raspodelu** sa parametrima n i p:

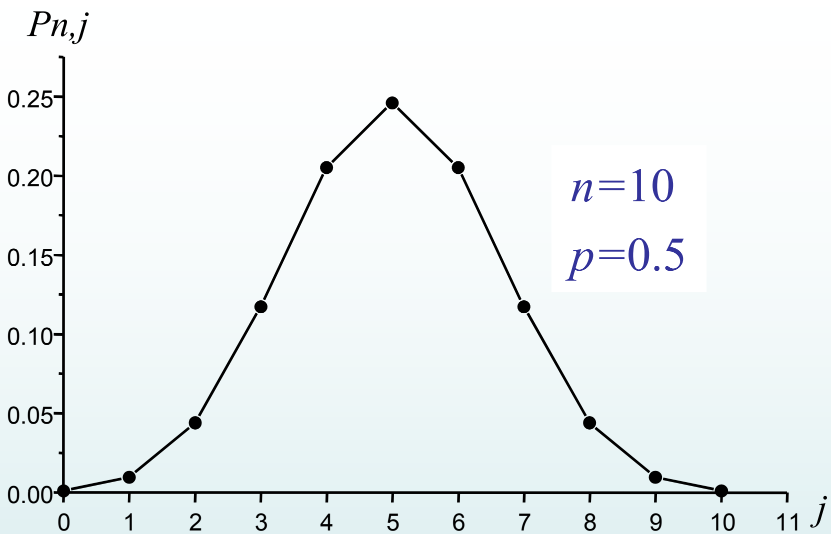
$$X : \mathcal{B}(n, p)$$

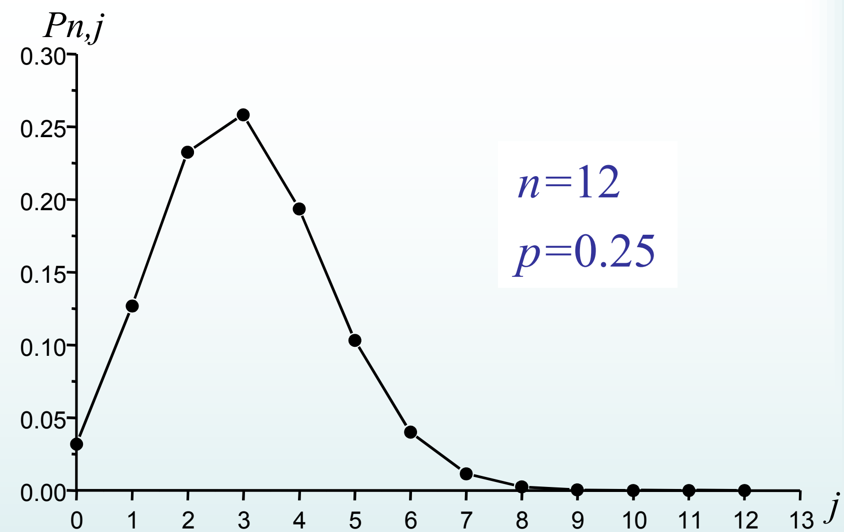
Bernulijeva sl. pr S_n

- Za sp sa binomnom raspodelom je:

$$X : \begin{pmatrix} 0 & 1 & \dots & n \\ p_{n,0} & p_{n,1} & \dots & p_{n,n} \end{pmatrix} \quad \boxed{p_{n,j} = \binom{n}{j} p^j (1-p)^{n-j}} \quad j = 0,1,\dots,n$$

Binomne verovatnoće $p_{n,j}$





Binomna raspodela, nastavak

- Ako sp X ima $\mathcal{B}(n_1, p)$ raspodelu, a sp Y ima $\mathcal{B}(n_2, p)$ raspodelu, i ako su X i Y nezavisne, tada sp $Z= X+Y$ ima $\mathcal{B}(n_1+n_2, p)$ raspodelu.
- Raspodela za Y uslovno u odnosu na $Z=X+Y=k$ u slučaju nezavisnih veličina X i Y sa $\mathcal{B}(n_1, p)$ i $\mathcal{B}(n_2, p)$ raspodelama je

$$P(Y=j/Z=k)=\frac{P(Y=j,Z=k)}{P(Z=k)}=\frac{P(X=k-j)P(Y=j)}{P(Z=k)}=\frac{\binom{n_1}{k-j}\binom{n_2}{j}}{\binom{n_1+n_2}{k}}$$

$$\max\{0,k-n_1\}\le j\le \min\{k,n_2\}$$

Binomna raspodela, nastavak

- Ako sp X ima $\mathcal{B}(n, p)$ raspodelu, tada sp $Y=n-X$ ima $\mathcal{B}(n, 1-p)$ raspodelu.
- Izračunavanje vrednosti izraza $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$

  može biti složeno za velike vrednosti n ili male vrednosti p, pa se može koristiti relacija

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

- Ako je $n\geq 30$ i $np<10$, binomna raspodela se aproksimira ***Puasonovom*** raspodelom
- Ako je $n\geq 30$ i $np>10$, binomna raspodela se aproksimira ***normalnom*** raspodelom

Binomna raspodela, disperzija

- Neka je verovatnoća dog. A u svakom eksperimentu p i neka je izvedeno n eksperimenata pod istim uslovima, nezavisno jedan od drugog. Neka je I_j indikator dog. A u j-tom eksperimentu ($j=1,\dots,n$). Za sp X sa $\mathcal{B}(n, p)$ raspodelom važi

$$X = I_1 + \ldots + I_n \qquad \mathrm{I}_{\mathrm{A}} : \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

$$E(X) = E(I_1) + \ldots + E(I_n) = np$$

$$D(X) = D(I_1 + \ldots + I_n) = np(1-p)$$

Binomna raspodela, koeficijenti

- Neka je data sp X sa $\mathcal{B}(n, p)$ raspodelom. Koeficijent varijacije je

$$\boxed{C_V = \sqrt{\frac{1-p}{np}}}$$

- Indeks disperzije je $\boxed{I = 1 - p < 1}$

- Koeficijent asimetrije je

$$f_1 = \frac{1-2p}{\sqrt{np(1-p)}}$$

- Koeficijent spljoštenosti je

$$f_2 = \frac{1-6p(1-p)}{np(1-p)}$$

Bernulijev zakon velikih brojeva

- Iz nejednakosti Čebiševa sledi da za svako $\varepsilon>0$, za sp X sa $\mathcal{B}(n, p)$ raspodelom važi:

$$P\left[\left|\frac{X}{n} - p\right| \geq \varepsilon\right] \to 0 \qquad n \to \infty$$

- ***Bernulijev zakon velikih brojeva*** ukazuje na to da se relativna frekvencija dog. A čija je verovatnoća p, grupiše oko p.

Puasonova raspodela

- SP X ima **Puasonovu raspodelu** sa parametrom λ, $\mathcal{P}(\lambda)$ ako je $\lambda>0$ i

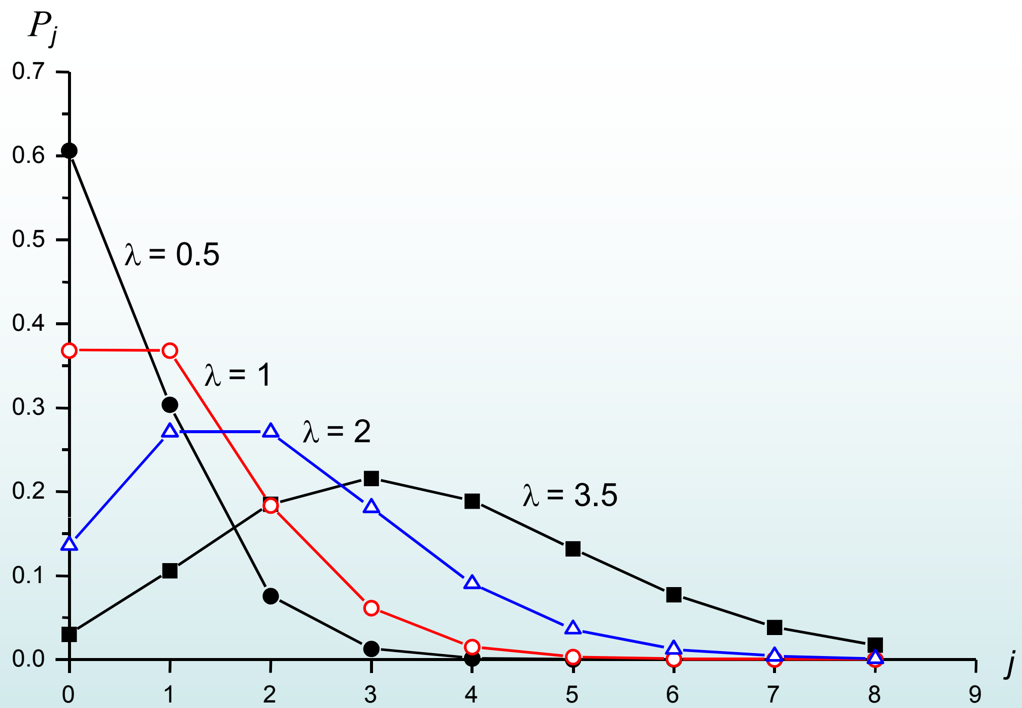
$$X : \begin{pmatrix} 0 & 1 & \ldots & n & \ldots \\ p_0 & p_1 & \ldots & p_n & \ldots \end{pmatrix} \qquad \boxed{p_j = P[X = j] = \frac{\lambda^j}{j!}\mathrm{e}^{-\lambda}} \quad j = 0,1,2,\ldots$$

- Za $n\geq 30$ i $np<10$, razlika između binomne i Puasonove raspodele je vrlo mala.
- Ako verovatnoća p_n realizacije dog. A u binomnom zakonu zavisi od n, tj. S_n:$\mathcal{B}(n, p_n)$ tada ako $np_n \to \lambda$, $n\to\infty$

$$P[S_n = j] \to \frac{\lambda^j}{j!}\mathrm{e}^{-\lambda} \qquad n \to \infty \qquad j = 0,1,2,\ldots$$

za $\lambda<10$

Puasonova raspodela za različite vrednosti parametra λ



- Ako parametar λ Puasonove raspodele nije prirodan broj, tada raspodela ima jedan mod jednak celom delu broja λ, a ako je λ prirodan broj, tada Puasonova raspodela ima dva moda (λ-1) i λ.

Puasonova raspodela, disperzija

- Matematičko očekivanje Puasonove raspodele je

$$E(X)=\sum_{j=0}^{\infty} j\frac{\lambda^j}{j!}\mathrm{e}^{-\lambda}=\lambda \qquad \sum_{j=0}^{\infty}\frac{\lambda^j}{j!}=\mathrm{e}^{\lambda}$$

- Disperzija Puasonove raspodele je

$$D(X)=E(X^2)-(E(X))^2=\sum_{j=0}^{\infty} j(j-1)\frac{\lambda^j}{j!}\mathrm{e}^{-\lambda}+\lambda-\lambda^2=\lambda$$

Puasonova raspodela, koeficijenti

- Neka je data sp X sa $\mathcal{P}(\lambda)$ raspodelom. Koeficijent varijacije je

$$\boxed{C_V = \frac{1}{\sqrt{\lambda}}}$$

- Indeks disperzije je

$$\boxed{I = 1}$$

- Koeficijent asimetrije je

$$f_1 = \frac{1}{\sqrt{\lambda}}$$

- Koeficijent spljoštenosti je

$$f_2 = \frac{1}{\lambda}$$

Puasonova raspodela, nastavak

- Neka su sp X:$\mathcal{P}(\lambda)$ i Y:$\mathcal{P}(\mu)$ nezavisne. Tada je $Z=X+Y$ sp sa Puasonom raspodelom Z:$\mathcal{P}(\lambda+\mu)$.
- Uslovna raspodela za X pri datom $Z=k$ biće $\mathcal{B}(k,p)$, gde je $p= \lambda/(\lambda+\mu)$.
- Ako je X:$\mathcal{P}(\lambda)$ i $\lambda>0$, tada se X može aproksimirati sp koja ima normalnu raspodelu sa parametrima $m = \lambda$ i $\sigma^2 = \lambda$.