Verjetnost in statistika - zapiski s predavanj prof. Drnovška

Tomaž Poljanšek

študijsko leto 2022/23

Kazalo

1 Verjetnost

1.1 Neformalni uvod v verjetnost

Začetki verjetnosti (kot vede) so v 17. stoletju, motivacija igre na srečo
17. stol: Fermat, Pascal, Bernoulli
18. in 19 stol: Laplace, Poisson, Čebišev, Markov
20. stol: Kolmogorov (okoli 1930), utemeljitelj sodobnega verjetnostnega računa

Definicija 1.1 (Dogodek)**.** Izvajamo poskus, opazujemo nek pojav, ki se lahko zgodi in ga imenujemo dogodek.

Definicija 1.2 (Frekvenca)**.** Poskus ponovimo n-krat. Opazujemo dogodek A.
Naj bo $K_n(A)$ frekvenca dogodka A, t.j. število tistih ponovitev, pri katerih se je dogodek A zgodil.
Relativna frekvenca je $f_n(A) = \frac{K_n(A)}{n} \in [0, 1]$

Dokazati je mogoče, da zaporedje $\{f_n(A)\}$ konvergira, recimo h $p \in [0, 1]$.
Statistična definicija verjetnosti: $P(A) := p$.
Pogosto verjetnost lahko določimo vnaprej:
Klasična definicija verjetnosti: $P(A) = \frac{\text{število ugodnih izidov za dogodek } A}{\text{število vseh izidov}}$ pri pogoju, da imajo vsi izidi enake možnosti

Če je vseh izidov neskončno, si lahko pomagamo z geometrijsko definicijo verjetnosti.

V kvantni mehaniki so kroglice različni delci, posode so energetska stanja.
V primeru (a) imamo Maxwell-Boltzmanovi statistiki, velja za molekule plina.
V primeru (b) imamo Bose-Einsteinovo statistiko, velja za bozone (npr. fotoni).
V primeru (c) imamo Fermi-Diracovo statistiko, velja za fermione (npr elektroni); zanje velja Diracovo izključitveno načelo: v vsakem stanju je največ en delec.

1.2 Aksiomatična definicija verjetnosti

Kolmogorov (okoli 1930)

Definicija 1.3 (Dogodek)**.** Imamo prostor vseh dogodkov Ω (možna oznaka je G). Dogodki so nekatere (ne nujno vse) podmnožice $A \subseteq \Omega$.

Računanje z dogodki

1. Vsota dogodkov oz. unija dogodkov: $A + B$ oz. $A \cup B$: dogodek, da se zgodi vsaj eden od A in B
2. Produkt dogodkov oz. presek dogodkov: $A \cdot B$ oz. $A \cap B$: dogodek, da se zgodita oda dogodka A in B
3. Nasprotni dogodek oz. komplement dogodka: $\overline{A} = A^C$

Pravila za računanje z dogodki:

- idempotentnost:
$$A \cup A = A = A \cap A$$
- komutativnost:
$$A \cap B = B \cap A$$
$$A \cup B = B \cup A$$
- asociativnost:
$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$
- distributivnost:
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$
$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

  oziroma

$$(A \cdot B) + C = (A + C) \cdot (B + C)$$
$$(A + B) \cdot C = (A \cdot C) + (B \cdot C)$$

- deMorganova zakona:
$$(A \cup B)^C = A^C \cap B^C$$
$$(A \cap B)^C = A^C \cup B^C$$

  še več:

$$(\cup_{i \in I} A_i)^C = \cap_{i \in I} A_i^C$$
$$(\cap_{i \in I} A_i)^C = \cup_{i \in I} A_i^C$$

Definicija 1.4 (σ-algebra)**.** Neprazna družina podmnožic dogodkov $\mathcal{F}$ v Ω je σ-algebra, če velja:

1. $A \in \mathcal{F} \implies A^C \in F$ (zaprtost za komplement)
2. $A_1, A_2 \cdots \in \mathcal{F} \implies \cup_{i=1}^{\infty} A_i \in F$ (zaprtost za števne unije)

Če v 2) zahtevamo manj:
$A, B \in \mathcal{F} \implies A \cup B \in F$ (šibkejši pogoj) pravimo, da je $\mathcal{F}$ algebra.

V algebri imamo zaprtost za končne unije, t.j. $A_1 \dots A_n \in \mathcal{F} \implies \cup_{i=1}^{n} A_j \in F$ (zaradi indukcije). Ker je $\cap_i A_i^C = (\cup_i A_i)^C$ (deMorgan), je algebra zaprta za končne preseke, σ-algebra pa za števne preseke.
Ker je $A \setminus B = A \cap B^C$, je algebra zaprta za razlike dogodkov.
Vsaka algebra vsebuje $\{\emptyset, \Omega\}$: ker je neprazna, obstaja dogodek $A \in F$, potem je $A^C \in F$ in zato je

$$\mathcal{F}\Omega = A \cup A^C \in F, \emptyset = A \cap A^C \in \mathcal{F}$$

Najmanjša (σ-)algebra je $\mathcal{F} = \{\emptyset, \Omega\}$, največja ($\sigma$-)algebra je potenčna množica $P(\Omega)$.

Definicija 1.5 (Nezdružljivost dogodkov)**.** Dogodka A in B sta nezdružljiva ali disjunktna, če je $A \cap B = \emptyset$

Definicija 1.6 (Popoln sistem dogodkov)**.** Zaporedje $\{A_i\}_i$ (končno ali števno mnogo) je popoln sistem dogodkov, če $\Omega = \cup_i A_i$ in $A_i \cap A_j = \emptyset \forall i \neq j$

Definicija 1.7 (Verjetnost)**.** Naj bo $\mathcal{F}\sigma$-algebra na Ω. Verjetnost na $(\Omega, \mathcal{F})$ je preslikava $P : \mathcal{F} \to \mathbb{R}$ z lastnostmi:

1. $P(A) \geq 0 \forall A \in F$
2. $P(\Omega) = 1$
3. Za poljubne paroma nezdružljive dogodke $A_1, A_2 \dots$ velja

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

   števna aditivnost (verjetnostne preslikave)

Lastnosti preslikave P:

(a) $P(\emptyset) = 0$

(b) P je končno aditivna, t.j. za poljubne paroma nezdružljive dogodke $A_1 \dots A_n$ velja

$$P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

(c) P je monotona, t.j. iz $A \subseteq B$ $(A, B \in F)$ sledi $P(A) \subseteq P(B)$, še več: $A \subseteq B \implies P(B \backslash A) = P(B) - P(A)$

(d) $P(A^C) = 1 - P(A)$ za $A \in F$

(e) P je zvezna, t.j.

(a) iz $A_1 \subseteq A_2 \subseteq \dots A_i \in F$ sledi $P(\cup_{i=1}^{\infty} A_i) = \lim_{i \to \infty} P(A_i)$

(b) iz $B_1 \supseteq B_2 \supseteq \dots B_i \in F$ sledi $P(\cap_{i=1}^{\infty} B_i) = \lim_{i \to \infty} P(B_i)$

Definiramo

$$C_1 = A_1$$
$$C_i = A_i - A_{i-1} \text{ za } i \geq 2$$

Potem $C_i \cap C_j = \emptyset$ za $i \neq j$, $A_n = C_1 \cup \dots \cup C_n$ in $\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty} C_i$
Torej imamo

$$P(\cup_{i=1}^{\infty} A_i) = P(\cup_{i=1}^{\infty} C_i) =$$
$$\stackrel{3)}{=} \sum_{i=1}^{\infty} P(C_i) = \lim_{n \to \infty} \sum_{i=1}^{n} P(C_i) =$$
$$\stackrel{b)}{=} \lim_{n \to \infty} P(A_i)$$

$(\Omega, \mathcal{F}, P)$ verjetnostni prostor

$(\Omega, \mathcal{F}, P)$

Verjetnost P definiramo na pravokotnikih s $P((a, b) \times (c, d)) = (b - a)(d - c)$
Ni lahko videti, da je to možno razširiti do števno aditivne preslikave na $P(\Omega)$
Verjetnostna preslikava P (na $\mathcal{F}$) se imenuje Lebesgueova mera
To je geometrijska definicija verjetnosti:

$$\square = \cap_{n=1}^{\infty} (a - \frac{1}{n}, b + \frac{1}{n}) \times (c - \frac{1}{n}, d + \frac{1}{n})$$

1.3 Pogojna verjetnost

Definicija 1.8 (Pogojna verjetnost)**.** Fiksirajmo dogodek B s $P(B) > 0$. Pogojna verjetnost dogodka A pri pogoju B je

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Iz definicije sledi $P(A \cap B) = P(B) \cdot P(A \mid B)$
Za poljubne dogodke A, B, C velja

$$\begin{aligned} &P(A \cap (B \cap C)) = P(B \cap C) \cdot P(A \mid B \cap C) = \\ &= P(C) \cdot P(B \mid C) \cdot P(A \mid B \cap C) \end{aligned}$$

oz. "lepše"

$$P(A \cap B \cap C) = P(A) \cdot P(B \mid A) \cdot P(C \mid A \cap B)$$

To posplošimo na n dogodkov $A_1, A_2 \dots A_n$:

$$\begin{aligned} &P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2 \mid A_1) \dots P(A_n \mid A_1 \cap \dots \cap A_{n-1}) = \\ &= P(A_1) \cdot \prod_{i=2}^{n} P(A_i \mid \cap_{j=1}^{i-1} A_j) \end{aligned}$$

Desna stran:

$$P(A_1) \cdot \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \cdots \frac{P(A_1 \cap \dots \cap A_n)}{P(A_1 \cap \dots \cap A_{n-1})}$$

Imejmo poskus v dveh korakih (fazah). V 1. koraku se zgodi natanko en dogodek iz popolnega sistema dogodkov $H_1, H_2 \dots$ (končno/števno mnogo). V drugem koraku nas zanima dogodek A. Izrazimo $P(A)$ z verjetnostmi $P(H_1), P(H_2 \dots)$ in $P(A \mid H_1), P(A \mid H_2) \dots$.
Ker je $A = A \cap \Omega = A \cap (\cup_i H_i) = \cup_i (A \cap H_i)$ in ker so $\{A \cap H_i\}_i$ paroma nezrdužljivi dogodki (zaradi H_i), je

$$P(A) = \sum_i P(A \cap H_i) = \sum_i P(H_i) \cdot P(A \mid H_i)$$

To je formula o popolni verjetnosti

Pri dvofaznem poskusu nas zanima

$$P(H_k \mid A) = \frac{P(H_k \cap A)}{P(A)} = \frac{P(H_k) \cdot P(A \mid H_k)}{\sum_i P(H_i) \cdot P(A \mid H_i)}$$

- Bayesova formula

Matematično ekvivalenten problem je presejalni test, npr. program DORA. (Pogojna) verjetnost, da je oseba bolna, če je test pozitiven, je majhna.
Dogodka A in B sta neodvisna, če je $P(A \cap B) = P(A) \cdot P(B)$
Če je $P(B) > 0$, potem lahko ta pogoj zapišemo kot $P(A) = \frac{P(A \cap B)}{P(B)} = P(A \mid B)$

Definicija 1.9 (Neodvisnost)**.** A in B sta neodvisna, če $P(A \cap B) = P(A) \cdot P(B)$
Dogodki $\{A_i\}_i$ so neodvisni, če za poljuben končen nabor različnih dogodkov $A_{i_1}, A_{i_2} \dots A_{i_n}$ velja

$$P(A_{i_i} \cap \dots \cap A_{i_n}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_n})$$

Če zahtevamo le za $n = 2$, t.j. $P(A_i \cap A_j) = P(A_i) \cdot P(A_j), i \neq j$, tedaj so dogodki paroma neodvisni

Očitno iz neodvisnosti sledi paroma neodvisnost. Obratno ne velja

Trditev 1.10. Naj bosta A in B neodvisna dogodka. Potem sta neodvisna tudi A in B^C. Prav tako tudo A^C in B ter A^C in B^C (komplementiranje ohranja neodvisnost)

1.4 Zaporedja neodvisnih ponovitev poskusa

Definicija 1.11. Imejmo zaporedje n neodvisnih ponovitev poskusa, določenega v verjetnostnem prostoru $(\Omega, \mathcal{F}, P)$, v katerem je možen A s $P(A) = p \in (0, 1)$. Potem je $q := P(A^C) = 1 - p$

Z $A_n(k)$ označimo dogodek, da se v k ponovitvah poskusa A zgodi natanko n-krat, $k = 0, 1 \dots n$
Pokažimo, da je njegova verjetnost $P_n(k) := P(A_n(k)) = \binom{n}{k} p^k q^{n-k}$ - Bernoullijeva formula

$A_n(k)$ je disjunktna unija $\binom{n}{k}$ dogodkov, da se A zgodi na predpisanih k mestih, A^C pa na preostalih $(n-k)$ mestih. Verjetnost le teh je produkt p-jev in q-jev: $p^k q^{n-k}$. Od tod sledi Bernoullijeva formula

Brez računala je to težko izračunati tudi če uporabimo Stirlingovo formulo na $n!$:

$$n! \sim \sqrt{2\pi n}(\frac{n}{e})^n$$

Tukaj $\sim$ pomeni: $a_n \sim b_n$ če $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$
Torej je $\lim_{n\to\infty} \frac{\sqrt{2\pi n}}{n!}(\frac{n}{e})^n = 1$

1.4.1 Aproksimacijski formuli za $P_n(k)$

1.4.1.1 Poissonova formula

Če je n velik in k majhen, je $P_n(k) \approx \frac{\lambda^k}{k!}e^{-\lambda}$, kjer je $\lambda = np$

1.4.1.2 Laplaceova lokalna formula

Če je n velik, potem je $P_n(k) \approx \frac{1}{\sqrt{2\pi npq}} \cdot e^{-\frac{(k-np)^2}{2npq}}$
Kasneje (2. semester) bomo dokazali splošnejši izrek (centralni limitni izrek)
Narišimo zaporedje $\{P_n(k)\}_{k=0}^n$, n fiksen

$$\begin{aligned} P_n(0) &= q^n \\ P_n(1) &= npq^{n-1} \\ P_n(2) &= \frac{n(n-1)}{2}p^2q^{n-2} \end{aligned}$$

Pomaknjena in raztegnjena funkcija $f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

$$\begin{aligned} &P_n(k) \le P_n(k+1)? \\ &\frac{n!}{k!(n-k)!}p^kq^{n-k} \le \frac{n!}{(k+1)!(n-k-1)!}p^{k+1}q^{n-k-1} \\ &\frac{q}{n-k} \le \frac{p}{k+1} \iff kq+q \le np-kp \iff \\ &\iff k(p+q)+q \le np \iff k+q \le np \end{aligned}$$

Neenakost se obrne pri $k \approx np$

1.4.1.3 Laplaceova integralska formula

Zanima nas dogodek $B_n(k_1, k_2)$, da se n n ponovitvah poskusa dogodek A zgodi vsaj k_1-krat in manj kot k_2-krat, $0 \leq k_1 < k_2 \leq n+1$
Ker je

$$B_n(k_1, k_2) = A_n(k_1) \cup A_n(k_1 + 1) \cup \cdots \cup A_n(k_2 - 1)$$

(disjunktna unija), je

$$P_n(k_1, k_2) := P(B_n(k_1, k_2)) = \sum_{k=k_1}^{k_2-1} |A_n(k)| = \sum_{k=k_1}^{k_2-1} P_n(k)$$

Po Laplaceovi lokalni formuli je

$$P_n(k_1, k_2) \approx \frac{1}{\sqrt{2\pi npq}} \sum_{k=k_1}^{k_2-1} e^{-\frac{(k-nq)^2}{2npq}} =$$
$$\doteq \frac{1}{\sqrt{2\pi}} \sum_{k=k_1}^{k_2-1} e^{-\frac{1}{2}x_k^2} \Delta x_k$$

kjer je

$$x_k := \frac{k - np}{\sqrt{npq}}$$
$$\implies \Delta x_k := x_{k-1} - x_k = \frac{k + 1 - np}{\sqrt{npq}} - \frac{k - np}{\sqrt{npq}} = \frac{1}{\sqrt{npq}}$$

To je integralaska (Riemannova) vsota za funkcijo $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
$P_n(k_1, k_2) \approx \sum_{k=k_1}^{k_2-1} f(x_k)\Delta x_k$ na intervalu $a = \frac{k_1 - np}{\sqrt{npq}}, b = \frac{k_2 - np}{\sqrt{npq}}$
Za velik n torej velja:

$$P_n(k_1, k_2) \approx \int_a^b f(x)dx = \int_{\frac{k_1-np}{\sqrt{npq}}}^{\frac{k_2-np}{\sqrt{npq}}} e^{-\frac{x^2}{2}} dx$$

- Laplaceova integralska formula
$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ - verjetnostni integral

Vpeljimo verjetnostni integral

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

Φ je liha funkcija, zvezno odvedljiva in strogo naraščajoča
$\Phi(0) = 0$ in $\Phi(x) = f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
Pokažimo, da je $\lim_{x\to\infty} \Phi(x) = \frac{1}{2}$. S pomočjo Γ funkcije imamo

$$\begin{aligned}
&\lim_{x\to\infty} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{x^2}{2}} dt = \\
&x = \frac{t^2}{2}, dx = tdt, dt = \frac{dx}{t} = \frac{dx}{\sqrt{2x}} \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-x} \frac{dx}{\sqrt{2x}} = \\
&= \frac{1}{2\sqrt{\pi}} \int_0^\infty x^{-\frac{1}{2}} e^{-x} dx = \\
&\overset{\Gamma(\frac{1}{2})=\sqrt{\pi}}{=} \frac{1}{2}
\end{aligned}$$

Laplaceova formula se glasi:

$$P_n(k_1, k_2) = \Phi(\frac{k_2 - np}{\sqrt{npq}}) - \Phi(\frac{k_1 - np}{\sqrt{npq}})$$

1.5 slučajne spremenljivke

Danemu poskusu priredimo določeno številsko količino, katere verjetnost je odvisna od slučajna

Definicija 1.12 (Slučajna spremenljivka)**.** Realna slučajna spremenljivka na verjetnostnem prostoru (Ω, Φ, P) je funkcije $X : \Omega \to \mathbb{R}$ z lastnostjo, da je za $\forall x \in \mathbb{R}$ množica $\{\omega \in \Omega : X(\omega) \leq x\}$ v Φ, se pravi dogodek

Oznaka: $\{\omega \in \Omega : X(\omega) \leq x\} \equiv X^{-1}((-\infty, x]) \equiv (X \leq x)$ (ali $\{X \leq x\}$)

Definicija 1.13 (Porazdelitvena funkcija)**.** Porazdelitvena funkcija $F_X : \mathbb{R} \to \mathbb{R}$ je funkcija, definirana s predpisom $F_X(x) = P(X \leq x) \equiv P((X \leq x))$

Dogovor: $P((X \leq x)) \leftrightarrow P(X \leq x)$
Lastnosti porazdelitvene funkcije $F_X \equiv F$:

1. $0 \leq F(X) \leq 1$ za $\forall x \in \mathbb{R}$ (verjetnost)
2. F je naraščajoča funkcija, t.j. iz $x_1 < x_2$ sledi $F(x_1) \leq F(x_2)$
3. $\lim_{x\to\infty} F(x) = 1, \lim_{x\to-\infty} F(x) = 0$
4. F je zvezna z desne, t.j. $F(X+) = F(X)\ \forall x \in \mathbb{R}$
5. $F(X-) = P(X < x) \neq F(x)$ v splošnem

$$
\begin{aligned}
&P(x_1 < X \leq x_2) = P((X \leq x_2)\backslash(X \leq x_1)) = \\
&= P(X \leq x_2) - P(X \leq x_1) = F(x_2) - F(x_1) \\
&P(x_1 < X < x_2) = P(X < x_2) - P(X \leq x_1) = F(x_2-) - F(x_1) \\
&P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1-) \\
&P(x_1 \leq X < x_2) = F(x_2-) - F(x_1-)
\end{aligned}
$$

Opomba. V nekaterih učbenikih je porazdelitvena funkcija definirana z $F(x) = P(X < x)$ - zvezna z leve

Najpomembnejša razreda slučajnih spremenljivk sta

1.5.1 Diskretna slučajna spremenljivka

Definicija 1.14 (Diskretna slučajna spremenljivka)**.** Slučajna spremenljivka $X : \Omega \to \mathbb{R}$ je diskretno porazdeljena, če je njena zaloga vrednosti končna ali števna množica. Naj bo $\{x_1, x_2 \dots\}$ zaloga vrednosti slučajne spremenljivke X.

Vpeljimo verjetnostno funkcijo $p_n := P(X = x_n)\ n = 1, 2 \dots$. Potem je

$$\sum_n p_n = P(\cup_n (X = x_n)) = P(\Omega) = 1$$

in

$$F_X(x) = P(X \leq x) = P(\cup_{n:x_n \leq x}(X = x_n)) =$$
paroma nezdružljivi dogodki
$$= \sum_{n:x_n \leq x} P(X = x_n) = \sum_{n:x_n \leq x} p_n$$

npr. naj bodo $x_1 < x_2 < x_3$ v zalogi vrednosti slučajne spremenljivke X
F je odsekoma konstantna

$$X : \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$$

Pomembnejše diskretne porazdelitve:

1.5.1.1 Enakomerna diskretna porazdelitev

na n točkah

$$X : \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{pmatrix}$$

1.5.1.2 Binomska porazdelitev

$Bin(n, p), n \in \mathbb{N}, p \in (0, 1), n-$krat ponovimo poskus, gledamo dogodek A z verjetnostjo $P(A) = p$, X je frekvenca dogodka A v n ponovitvah

$$X : \begin{pmatrix} 0 & 1 & \dots & n \\ p_0 & p_1 & \dots & p_n \end{pmatrix}$$
$$p_k = \binom{n}{k} p^k q^{n-k}$$

1.5.1.3 Poissonova porazdelitev

$Poi(\lambda), \lambda > 0$

$$p_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \; k = 0, 1, 2 \dots$$
$$\sum_{k=0}^{\infty} p_k = (\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}) e^{-\lambda} = e^{\lambda} e^{-\lambda} = 1$$

$\implies$ to je res porazdelitev ($p_i \geq 0, \sum p_i = 1$)

1.5.1.4 Geometrijska porazdelitev

$Geo(p),\ p \in (0,1)$
Ponavljamo poskus, v katerem opazujemo dogodek A s $P(A) = p, q = 1 - p$. $(X = k)$ je dogodek, da se A zgodi prvič v k-ti ponovitvi

$$p_k = P(X = k) = p \cdot q^{k-1}\ k = 1, 2 \ldots$$

$$\sum_{k=1}^{\infty} p_k = p \cdot \sum_{k=1}^{\infty} q^{k-1} = p \sum_{k=0}^{\infty} q^k = p \frac{1}{1-q} = \frac{p}{p} = 1$$

1.5.1.5 Pascalova ali negativna binomska porazdelitev

$Pas(m,p),\ m \in \mathbb{N}, p \in (0,1)$
Ponavljamo poskus, v katerem nas zanima dogodek A s $P(A) = p$. $(X = k)$ je dogodek, da se A zgodi m-tič v k-ti ponovitvi poskusa. Torej $Pas(1,p) = Geo(p)$

$$p_k = P(X = k) = \binom{k-1}{m-1} p^m q^{k-m}\ k = m, m+1 \ldots$$

(A se zgodi $(m-1)$-krat, $\overline{A}$ pa $(k-m)$-krat)
DN: Enakost $\sum_{k=m}^{\infty} p_k = 1$ analitično preverimo z $(m-1)$-kratnim odvajanjem geometrijske vrste

$$\sum_{k=0}^{\infty} q^{k-1} = \frac{1}{1-q}$$

oz. z direktno uporabo binomske vrste:

$$(1-q)^{-m} = \sum_{j=0}^{\infty} \binom{-m}{j} q^j$$

1.5.1.6 Hipergeometrijska porazdelitev

$Hip(n; M, N),\ 0 < M < N, n, M, N \in \mathbb{N}, n \leq \min\{M, N - M\}$
V posodi je N kroglic, od tega M belih, ostale črne. Slučajno izberemo n

kroglic (brez vračanja). X je število belih kroglic med izbranimi kroglicami. Torej $(X = k)$ je dogodek, da je med izbranimi n kroglicami k belih

$$p_k = P(X = k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} \; k = 0, 1 \dots n$$

$\binom{m}{k} \dots k$ belih

$\binom{N-m}{n-k} \dots$ ostale črne

$\binom{N}{n} \dots$ izberemo n izmed N

Ker je $\{(X = k)\}^n$ popoln sistem dogodkov, je jasno, da je $\sum_{k=0}^{n} p_k = 1$
Torej velja binomska identiteta

$$\sum_{k=0}^{n} \binom{m}{k}\binom{N-m}{n-k} = \binom{N}{n}$$

- verjetnostni dokaz

Če je $n << \min\{M, N-M\}$, potem je $Hip(n; M, n) \approx Bin(n, \frac{M}{N})$:

$$p_k = \frac{\frac{M(M-1)\dots(M_k+1)}{k!}\frac{(N-m)(N-m+1)\dots(N-m-n+k+1)}{(n-k)!}}{\frac{N(N-1)\dots(N-n+1)}{n!}} \approx$$

$$\overset{\substack{k\leq m \\ n\leqq N}}{\approx} \frac{\frac{M^k}{k!}\frac{(N-m)^{n-k}}{(n-k)!}}{\frac{N^n}{n!}} = \binom{n}{k}(\frac{M}{N})^k(\frac{N-M}{N})^{n-k} = \binom{n}{k}p^k q^{n-k}$$

Intuicija: vzemanje kroglic, $n << \min\{M, N-M\}$
Če je $n << \min\{M, N-M\}$, ne naredimo velike napake, če kroglice vračamo. Tedaj je število belih izvlečenih kroglic binomsko porazdeljeno: $X \sim Bin(n, \frac{M}{N})$

1.5.2 Zvezno porazdeljene slučajne spremenljivke

Definicija 1.15 (Zvezna porazdelitev)**.** Slučajna spremenljivka X je zvezno porazdeljena (zvezna), če obstaja nenegativa integrabilna funkcije p_X, imenovana gostota porazdelitve, da je

$$F_X(x) = \int_{-\infty}^{x} p_X(t)dt \text{ za } \forall x \in \mathbb{R}$$

Analogija z diskretnimi porazdelitvami: $F_X(x) = \sum_{n:X_n \leq x} p_k, X : \begin{pmatrix} x_1 & \dots \\ p_1 & \dots \end{pmatrix}$
Tedaj je F_X zvezna funkcija. V točkah, kjer je p_X zvezna, je F_X zvezno odvedljiva in velja $F_X'(x) = p_X(x)$
Ker je $\lim_{x\to\infty} F_X(x) = 1$, je $\int_{-\infty}^{\infty} p_X(t)dt = 1$
Za $x_1 < x_2$ velja

$$P(x_1 < X < x_2) = F_X(x_2-) - F_X(x_1+) = \int_{-\infty}^{x_2} p_X(t)dx - \int_{-\infty}^{x_1} p_X(t)dt = \int_{x_1}^{x_2} p_X(t)dt$$

Pomembnejše zvezne porazdelitve:

1.5.2.1 Enakomerna zvezna porazdelitev na $[a, b]$

$$p_X(x) = \begin{cases} \frac{1}{b-a} \text{ če } a < x < b \\ 0 \text{ sicer} \end{cases}$$

$$F_X(x) = \begin{cases} 0 \text{ če } x \leq a \\ \frac{x-a}{b-a} \text{ če } a < x < b \\ 1 \text{ če } x \geq b \end{cases}$$

1.5.2.2 Normalna ali Gaussova porazdelitev

$N(\mu, \sigma),\ \mu \in \mathbb{R}, \sigma > 0$

$$p_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

$N(0, 1) : p_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ - standardizirana normalna porazdelitev

σ velik:
σ majhen:
Porazdelitvena funkcija:

$$F(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}(\frac{t-\mu}{\sigma})^2} dt =$$
$$u = \frac{t-\mu}{\sigma}, du = \frac{dt}{\sigma}$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{1}{2}u^2} du =$$
$$= \frac{1}{\sqrt{2\pi}} (\int_{-\infty}^{0} \cdots + \int_{0}^{\frac{x-\mu}{\sigma}} \dots) =$$
$$= \frac{1}{2} + \Phi(\frac{x-\mu}{\sigma})$$

Laplaceova integralaska formula pravi, da je $Bin(n,p) \approx N(np, \sqrt{npq})$ za velik n :

$$P_n(k) = \frac{1}{\sqrt{2\pi npq}} - \frac{1}{2}(\frac{k-np}{\sqrt{npq}})^2$$

1.5.2.3 Eksponentna porazdelitev

$$Exp(\lambda), \quad \lambda > 0$$
$$p(x) = \begin{cases} \lambda e^{-\lambda x} & \lambda \geq 0 \\ 0 & \text{sicer} \end{cases}$$

$$F(x) = \begin{cases} 1 - e^{-\lambda x} \text{ če } x \geq 0 \\ 0 \quad \text{če } x \leq 0 \end{cases}$$

1.5.2.4 Porazdelitev gama

$$\Gamma(b,c), \quad b,c > 0$$
$$p(x) = \begin{cases} \frac{c^b}{\Gamma(b)} x^{b-1} e^{-cx} & x > 0 \\ 0 & \text{sicer} \end{cases}$$

Očitno je $Exp(\lambda) = \Gamma(1,\lambda)$

$$\Gamma(y) = \int_0^\infty x^{y-1} e^{-x} dx$$

$$\int_{-\infty}^{\infty} p(x)dx = \frac{c^b}{\Gamma(b)} \int_0^{\infty} x^{b-1}e^{-cx}dx =$$
$$t = cx, dt = cdx$$
$$= \frac{c^b}{\Gamma(b)} \int_0^{\infty} (cx)^{b-1}e^{-cx}cdx =$$
$$= \frac{1}{\Gamma(b)} \cdot \Gamma(b) = 1$$

- je porazdelitev

1.5.2.5 Porazdelitev $\chi^2(n)$

(hi-kvadrat), $n \in \mathbb{N}$, n je število prostorskih stopenj

$$\chi^2(n) = \Gamma(\frac{n}{2}, \frac{1}{2})$$
$$p(x) = \begin{cases} \frac{1}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})} x^{\frac{n}{2}-1}e^{-\frac{x}{2}} x > 0 \\ 0 \qquad \text{sicer} \end{cases}$$

1.5.2.6 Cauchyjeva porazdelitev

$$p(x) = \frac{1}{\pi(1+x^2)} \quad x \in \mathbb{R}$$

$$F(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{dt}{1+t^2} = \frac{1}{\pi} \arctan t|_{-\infty}^{x} =$$
$$= \frac{1}{\pi} \arctan x - \frac{1}{\pi} \cdot \frac{\pi}{2} = \frac{1}{\pi} \arctan x + \frac{1}{2}$$

1.6 Slučajni vektorji

Definicija 1.16 (Slučajni vektor)**.** Naj bo (Ω, Φ, P) verjetnostni prostor. Slučajni vektor je n-terica slučajnih spremenljivk $x = (x_1 \dots x_n) : \Omega \to \mathbb{R}^n$ z lastnostjo, da je množica

$$(X_1 \leq x_1 \dots X_n \leq x_n) := \{\omega \in \Omega : X_1(\omega) \leq x_1 \dots X_n(\omega) \leq x_n\}$$

dogodek za vse n-terice $x = (x_1 \dots x_n)$, se pravi v Φ za $\forall x = (x_1 \dots x_n) \in \mathbb{R}^n$

Definicija 1.17 (Porazdelitvena funkcija)**.** Porazdelitvena funkcija slučajnega vektorja $X = (X_1 \dots X_n)$ je funkcija, definirana z

$$F_X(x) = F_{(X_1 \dots X_n)}(x_1 \dots x_n) := P(X_1 \leq x_1 \dots X_n \leq x_n)$$

Torej $F_X : \mathbb{R}^n \to \mathbb{R}$

F_X ima podobne lastnosti kot v primeru $n = 1$
Očitno je $0 \leq F_X(x) \leq 1$ za $\forall x \in \mathbb{R}^n$, glede na vsako spremenljivko je F_X naraščajoča in z desne zvezna, velja še:

$$\lim_{\substack{x_1 \to \infty \\ \vdots \\ x_n \to \infty}} F_{(X_1 \dots X_n)}(x_1 \dots x_n) = 1$$

Definicija 1.18 (Robna porazdelitev)**.** Če pošljemo v ∞ samo nekatere spremenljivke, dobimo porazdelitveno funkcijo slučajnega podvektorja, npr.

$$\lim_{\substack{x_2 \to \infty \\ \vdots \\ x_n \to \infty}} F_{(X_1 \dots X_n)}(x_1 \dots x_n) = F_{X_1}(x_1)$$

ali pa

$$\lim_{x_n \to \infty} F_{(X_1 \dots X_n)}(x_1 \dots x_n) = F_{X_1 \dots X_{n-1}}(x_1 \dots x_{n-1})$$

Takim porazdelitvam rečemo robne (marginalne) porazdelitve

Oglejmo si dvorazsežni primer ($n = 2$):

$$(X, Y) : \Omega \to \mathbb{R}^2$$

za $\forall (x, y) \in \mathbb{R}^2$ je

$$(X \leq x, Y \leq y) := \{\omega \in \Omega : X(\omega) \leq x, Y(\omega) \leq y\}$$

dogodek
Porazdelitvena funkcija $F_{(X,Y)} : \mathbb{R}^2 \to \mathbb{R}$ je definirana z

$$F_{(X,Y)}(x,y) := P(X \leq x, Y \leq y)$$
$$\lim_{x\to\infty} F_{(X,Y)}(x,y) = P(Y \leq y) = F_Y(y)$$
$$\lim_{y\to\infty} F_{(X,Y)}(x,y) = P(X \leq x) = F_X(x)$$

Izrazimo $P(a < X \leq b, c < Y \leq d)$ s porazdelitveno fukncijo $F(X,Y) = F$. To bo posplošitev formule

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

ki smo jo imeli v primeru $n = 1$

$(X,Y) : \Omega \to \mathbb{R}^2$ slučajni vektor
$F_{(X,Y)}(x,y) = P(X \leq x, Y \leq y) = P((x,y) \in (-\infty, x] \times (-\infty, y])$

Izrazimo z $F_{(X,Y)} = F$ verjetnost $P(a < X < b, c < Y < d)$. To bo posplošitev formule $P(a < X < b) = F_X(b) - F_X(a)$
Najprej vzemimo posebni primer:

$$P(a < X \leq b, Y \leq d) = P((X \leq b, Y \leq d) \backslash (X \leq a, Y \leq d)) =$$
$$= P(X \leq b, Y \leq d) - P(X \leq a, Y \leq d) = F(b,d) - F(a,b)$$

V splošnem primeru pa imamo

$$P(a < X \leq b, c < Y \leq d) = P((a < X \leq b, Y \leq d) \backslash (a < X \leq b, Y \leq c)) =$$
$$= P(a < X \leq b, Y \leq d) - P(a < X \leq b, Y \leq c) =$$
$$\overset{\text{fiks. y}}{=} (F(b,d) - F(a,d)) - (F(b,c) - F(a,c))$$

Torej je

$$P(a < X \leq b, c < Y \leq d) = F(b,d) - F(a,d) - F(b,c) + F(a,c)$$

Najpomembnejša razreda večrazsežnih porazdelitev sta

1.6.1 Diskretne porazdelitve

Definicija 1.19. Slučajni vektor $X = (X_1 \dots X_n) : \Omega \to \mathbb{R}^n$ je diskretno porazdeljen, če je njegova zaloga vrednosti končna/števna množica točk v $\mathbb{R}^n$. Omejimo se na $n = 2 : \Omega \to \mathbb{R}^2$.
Naj bo $\{x_1, x_2 \dots\}$ zaloga vrednosti slučajne spremenljivke X in $\{y_1, y_2 \dots\}$ zaloga vrednosti slučajne spremenljivke Y. Potem je zaloga vrednosti vektorja (X, Y) vsebovana v $\{(x_i, y_i) : i = 1, 2 \dots j = 1, 2 \dots\}$.
Definiramo verjetnostno funkcijo $p_{ij} := P(X = x_i, Y = y_j) i = 1, 2 \dots j = 1, 2 \dots$
Ker je $\{(X = x_i, Y = y_j)\}_{ij}$ popoln sistem dogodkov, je $\sum_i \sum_j p_{ij} = 1$

$$X : \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$$

$$p_i = P(X = x_i) = P(\cup_j (X = x_i, Y = y_j)) = \sum_j P(X = x_i, Y = y_j) = \sum_j p_{ij} \; i = 1, 2 \dots$$

če je $Y : \begin{pmatrix} y_1 & y_2 & \dots \\ q_1 & q_2 & \dots \end{pmatrix}$, je

$$q_j = P(Y = y_i) = P(\cup_i (X = x_i, Y = y_j)) = \sum_i P(X = x_i, Y = y_j) = \sum_i p_{ij} \; j = 1, 2 \dots$$

1.6.2 Zvezne porazdelitve

Definicija 1.20. Slučajni vektor $X = (X_1 \dots X_n)$ je zvezno porazdeljen, če obstaja integrabilna funkcija $p_X : \mathbb{R}^n \to \mathbb{R}$, imenovana gostota porazdelitve, da je

$$F_X(x) = F_{(X_1 \dots X_n)}(x_1 \dots x_n) =$$
$$= \int_{-\infty}^{x_1} dt_1 \int_{-\infty}^{x_2} dt_2 \cdots \int_{-\infty}^{x_n} p_X(t_1 \dots t_n) dt_n \text{ za } \forall x = (x_1 \dots x_n) \in \mathbb{R}^n$$

Ker je $\lim_{\substack{x_1 \to \infty \\ \vdots \\ x_n \to \infty}} F_X(x_1 \dots x_n) = 1$, je

$$\int \dots_{\mathbb{R}^n} \int p_X(t_1 \dots t_n) dt_1 \dots dt_n = 1$$

Za vsako Borelovo množico $A \subseteq \mathbb{R}^n$ (najmanjša σ-algebra z vsemi odprtimi pravokotniki) je

$$P(X \in A) \equiv P((x_1 \dots x_n) \in A) = \int \dots_A \int p_X(t_1 \dots t_n) dt_1 \dots dt_n$$

Omejimo se na $n = 2 : F_{(X,Y)}(x, y) = \int_{-\infty}^{x} du \int_{-\infty}^{y} p_{(X,Y)}(u, v) dv$
Robni porazdelitvi sta:

$$F_X(x) = \lim_{y \to \infty} F_{(X,Y)}(x, y) = \text{ (brez utemeljevanja)}$$
$$= \int_{-\infty}^{x} du \int_{-\infty}^{\infty} p_{(X,Y)}(u, v) dv$$

ki ima gostoto

$$p_X(x) = \int_{-\infty}^{\infty} p_{(X,Y)}(x, y) dy$$

in

$$F_Y(y) = \lim_{x \to \infty} F_{(X,Y)}(x, y) =$$
$$= \int_{-\infty}^{y} dv \int_{-\infty}^{\infty} p_{(X,Y)}(u, v) du$$

ki ima gostoto

$$p_Y(y) = \int_{-\infty}^{\infty} p_{(X,Y)}(x, y) dx$$

(ekvivalentno vsoti v diskretnem primeru).
Najpomembnejša dvorazsežna zvezna porazdelitev je normalna:

$$N(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho), \mu_x, \mu_y \in \mathbb{R}, \sigma_x, \sigma_y > 0, \rho \in (-1, 1)$$
$$p(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}((\frac{x-\mu_x}{\sigma_x})^2 - 2\rho\frac{x-\mu_x}{\sigma_x}\frac{y-\mu_y}{\sigma_y} + (\frac{y-\mu_y}{\sigma_y})^2)}$$
(μ_x, μ_y) premik, (σ_x, σ_y) razteg
$$N(0, 0, 1, 1, \rho) : p(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)}$$

Nivojnice, izohipse se: $x^2 - 2\rho xy + y^2 = c$

- $\rho = 0$: krožnica
- $\rho \in (-1, 1)$: elipsa

Robni porazdelitvi sta

$$p_X(x) = \int_{-\infty}^{\infty} p(x, y)dx = \cdots = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu_x}{\sigma_x})^2}$$

torej $X \sim N(\mu_x, \sigma_x)$. Podobno $Y \sim N(\mu_y, \sigma_y)$

Dvorazsežna normalna porazdelitev je posebni primer večrazsežne normalne porazdelitve $N(\mu, A)$, kjer je $\mu = (\mu_1 \dots \mu_n)^T$ in A pozitivno definitna matrika.
Gostota v točki $x = (x_1 \dots x_n)^T$ je

$$p(X) = \sqrt{\frac{detA}{(2\pi)^n}} e^{-\frac{1}{2}(x-\mu)^T A(x-\mu)}$$
$$(x-\mu)^T A(x-\mu) = \langle A(x-\mu), x-\mu \rangle$$

Za dokaz enakosti

$$\int \dots_{\mathbb{R}^n} \int p(x)dx_1 \dots dx_n = 1$$

izračunajmo integral

$$\int \dots_{\mathbb{R}^n} \int e^{-\frac{1}{2}(x-\mu)^T A(x-\mu)} dx_1 \dots dx_n = \sqrt{\frac{(2\pi)^n}{detA}}$$

$N(\mu, A), \mu \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}$ pozitivna definitna matrika, t.j. sebi adjungirana matrika, za katero velja

$$x^T Ax = \langle Ax, x \rangle > 0 \; \forall x \in \mathbb{R}^n \backslash \{0^n\}$$

V točki $x = (x_1 \dots x_n)^T$ je

$$p(x) = \sqrt{\frac{detA}{(2\pi)^n}} \cdot e^{-\frac{1}{2}(x-\mu)^T A(x-\mu)}$$

Izračunajmo integral

$$\int \underbrace{\dots}_{\mathbb{R}^n} \int e^{-\frac{1}{2}(x-\mu)^T A(x-\mu)} dx =$$
$$y = x - \mu \implies dy = dx$$
$$= \int \underbrace{\dots}_{\mathbb{R}^n} \int e^{-\frac{1}{2}y^T Ay} dy$$

Ker je A pozitivna definitna matrika, obstaja ortogonalna matrika U in diagonalna matrika $D = diag(\lambda_1 \dots \lambda_n)$, da je $A = U^T DU$

$$= \int \underbrace{\dots}_{\mathbb{R}^n} \int e^{-\frac{1}{2}y^T U^T DUy} dy =$$
$$z = Uy, y = U^T z, dy = |detU^T| dz = dz$$
$$= \int \underbrace{\dots}_{\mathbb{R}^n} \int e^{-\frac{1}{2}z^T Dz} dz =$$
$$= \int \underbrace{\dots}_{\mathbb{R}^n} \int e^{-\frac{1}{2}(\lambda_1 z_1^2 + \dots + \lambda_n z_n^2)} dz_1 \dots dz_n =$$
$$= \int_{\mathbb{R}} e^{-\frac{1}{2}\lambda_1 z_1^2} dz_1 \cdots \int_{\mathbb{R}} e^{-\frac{1}{2}\lambda_1 z_n^2} dz_n =$$

Ker je $\int_{\mathbb{R}} e^{-\frac{1}{2}\lambda z^2} dz = \sqrt{\frac{2\pi}{\lambda}}$ - $z \in \mathbb{R}$ - s pomočjo Γ funkcije, Bronstein, sledi iz

$$\frac{1}{\sqrt{2\pi}\sigma} = \int_{\mathbb{R}} e^{-\frac{1}{2}(\frac{x}{\sigma})^2} dx = 1$$

Gostota za $N(0, \sigma), \lambda := \frac{1}{\sigma^2}, \sigma = \frac{1}{\sqrt{\lambda}}$

$$= \sqrt{\frac{2\pi}{\lambda_1}} \cdot \dots \cdot \sqrt{\frac{2\pi}{\lambda_1}} = \sqrt{\frac{(2\pi)^n}{detA}}$$

Torej je $\int \underbrace{\dots}_{\mathbb{R}^n} \int p(x) dx = 1$

Dvorazšežni primer je posebni primer

$$A = \frac{1}{1-\rho^2}\begin{bmatrix} \frac{1}{\sigma_x^2} & -\frac{\rho}{\sigma_x\sigma_y} \\ -\frac{\rho}{\sigma_x\sigma_y} & \frac{q}{\sigma_y^2} \end{bmatrix}, \mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}$$
$$detA = \frac{1}{1-\rho^2}(\frac{1}{\sigma_x^2\sigma_y^2} - \frac{\rho^2}{\sigma_x^2\sigma_y^2}) \stackrel{?}{=} \frac{1}{\sigma_x^2\sigma_y^2}$$

$K = A^{-1} = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ -\rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$ kovariančna matrika (slučajnemu vektorju X, Y)

1.7 Neodvisnost slučajnih spremenljivk

Definicija 1.21 (Neodvisnost)**.** Slučjane spremenljivke $x_1, x_2 \dots x_n$ v slučjanem vektorju $x = (x_1 \dots x_n)$ so neodvisne, če je

$$F_X(x_1 \dots x_n) = F_{X_1}(x_1) \dots F_{X_n}(x_n) \text{ za } \forall x \in \mathbb{R}^n$$

oziroma

$$P(X_1 \leq x_1, X_2 \leq x_2 \dots X_n \leq x_n) = P(X_1 \leq x_1) \dots P(X_n \leq x_n)$$

oziroma dogodki $(X_1 \leq x_1) \dots (X_n \leq x_n)$ so neodvisni

Oglejmo si dvorazsežni diskretni primer

Trditev 1.22. Naj bo (X, Y) diskretno porazdeljen vektor:

$$p_{ij} = P(X = x_i, Y = y_j), p_i = P(X = x_i), q_j = P(Y = y_j)$$

Potem sta X in Y neodvisni $\iff p_{ij} = p_i \cdot q_j \; \forall i, j$

Torej sta X in Y neodvisni slučajni spremenljivki

Trditev 1.23. Naj bo (X, Y) zvezno porazdeljen slučajni vektore z gostoto $p(x, y)$. Potem sta X in Y neodvisni slučajni spremenljivki $\iff$ $p_{(X,Y)}(x, y) = p_X(x) \cdot p_Y(y)$ za (skoraj) vse $x, y \in \mathbb{R}$

Trditev 1.24. Naj bo (X, Y) zvezno porazdeljen slučajni vektor. Potem sta X in Y neodvisni $\iff p_{(X,Y)}(x, y) = f(x) \cdot g(y)$ za neki integrabilni funkciji f in g

Izrek 1.25. Slučajni spremenljivki X in Y sta neodvisni $\iff$ za vsaki Borelovi množici $A, B \subseteq \mathbb{R}$ velja

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$$

t.j. dogodka $(X \in A)$ in $(Y \in B)$ sta neodvisna
(Borelova σ-algebra: najmanjša σ-algebra z odprtimi množicami)

1.8 Funkcije slučajnih spremenljivk in slučajnih vektorjev

Naj bo $X : \Omega \to \mathbb{R}$ slučajna spremenljivka in $f : \mathbb{R} \to \mathbb{R}$ zvezna. Potem je $Y := f \circ X : \omega \to \mathbb{R}$ tudi slučajna spremenljivka.

$f \circ X = f(X)$
saj je množica

$$\begin{aligned}(Y \leq y) &\equiv \{\omega \in \Omega : f(X(\omega)) \leq y\} = \\ &= \{\omega \in \Omega : f(X(\omega)) \in (-\infty, y]\} = \\ &= \{\omega \in \Omega : X(\omega) \in f^{-1}((-\infty, y])\} = \\ &= \{X \in f^{-1}((-\infty, y])\}\end{aligned}$$

dogodek, ker je $f^{-1}((-\infty, y])$ zaprta množica, torej Borelova.
y je funkcija slučajne spremenljivke X.
Naj bo f strogo naraščajoča funkcija z zalogo vrednosti (a, b), kjer je $-\infty \leq a < b \leq \infty$
Vzemimo $y \in (a, b)$. Potem je

$$\begin{aligned}F_Y(y) &\stackrel{\text{def}}{=} P(Y \leq y) = P(f \circ X \leq y) = \\ &\text{f naraščajoča} \to \text{obrnljiva} \\ &= P(X \leq f^{-1}(y)) = F_X(f^{-1}(y))\end{aligned}$$

kjer je $f^{-1} : (a, b) \to \mathbb{R}$ inverzna funkcija k funkciji f
če je $y \geq b$ je $F_Y(y) = 1$
če je $y \leq a$ je $F_Y(y) = 0$
Če je še f zvezno odvedljiva in X zvezno porazdeljena slučajna spremenljivka, potem je y tudi zvezno porazdeljena z gostoto Φ

$$\Phi_Y(y) = F_Y^{'}(y) = F_X^{'}(f^{-1}(y)) \cdot (f^{-1}(y))^{'}$$

za $y \in (a,b)$, če je $y \notin (a,b)$, je $p_Y(y) = 0$
Podobno ravnamo v primeru, ko je f strogo padajoča ((a,b) zaloga vrednosti)

$$F_Y(y) = P(Y \leq y) = P(f \circ X \leq y) = P(X \geq f^{-1}(y)) =$$
$$= 1 - P(X \leq f^{-1}(y)) = 1 - F_X(f^{-1}(y)-)$$

Trditev 1.26. Če sta X in Y neodvisni slučajni spremenljivki, f in $g : \mathbb{R} \to \mathbb{R}$ zvezni funkciji, potem sta tudi $U = f(X)$ in $V = g(Y)$ neodvisni slučajni spremenljivki

Izrek 1.27. Naj bo $X = (X_1 \dots X_n) : \Omega \to \mathbb{R}^n$ slučajni vektor in $f = (f_1 \dots f_m) : \mathbb{R}^n \to \mathbb{R}^m$ zvezna preslikava. Potem je $Y = f \circ X \equiv f(X) : \Omega \to \mathbb{R}^m$ tudi slučajni vektor (brez dokaza).

Y je funkcija slučajnega vektorja X in ima m komponent $Y = (Y_1 \dots Y_m)$.
Porazdelitvena funkcija za $Y_j, (j = 1, 2 \dots m)$ je

$$F_{Y_j}(y) = P(f_i(x_1 \dots x_n) \leq y) = P((x_1 \dots x_n) \in f_j^{-1}((-\infty, y])) \text{ množica v} \mathbb{R}^n$$

Če je $X = (X_1 \dots X_n)$ zvezno porazdeljena, je torej

$$F_{Y_j}(y) = \int \underbrace{\dots}_{f^{-1}((-\infty,y])} \int p_X(x_1 \dots x_n) dx_1 \dots dx_n$$

Vzemimo posebni primer $X \sim \chi^2(m)$, $Y \sim \chi^2(n)$, torej

$$p_X(x) = \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} x^{\frac{m}{2}-1} e^{-\frac{x}{2}} \text{ za } x > 0 \text{ in 0 sicer}$$

za $p_Y(y)$ podobno.
Po zadnji formuli je $p_Z(z) = \int_{-\infty}^{\infty} p_X(x) \cdot p_Y(z-x) dx = 0$ za $z \leq 0$,
sicer je za $z > 0$

$$p_Z(z) = \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})2^{\frac{m}{2}}\Gamma(\frac{n}{2})}e^{-\frac{z}{2}}\int_0^z x^{\frac{m}{2}-1}(z-x)^{\frac{m}{2}-1+\frac{n}{2}-1+1}e^{-\frac{x}{2}}e^{-\frac{z-x}{2}}dx =$$
$$= \frac{1}{2^{\frac{m+n}{2}}\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}e^{-\frac{z}{2}}\int_0^z x^{\frac{m}{2}-1}(z-x)^{\frac{n}{2}-1}dx =$$

$$B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1}dt$$
$$x = tz\ dx = zdt$$

$$= \frac{1}{2^{\frac{m+n}{2}}\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}e^{-\frac{z}{2}}z^{\frac{m}{2}-1+\frac{n}{2}-1+1}\int_0^1 t^{\frac{m}{2}-1}(1-t)^{\frac{n}{2}-1}dt =$$

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$
$$\to B(\frac{m}{2},\frac{n}{2}) = \frac{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}{\Gamma(\frac{m+n}{2})}$$

$$= \frac{1}{2^{\frac{m+n}{2}}\Gamma(\frac{m+n}{2})}e^{-\frac{z}{2}}z^{\frac{m+n}{2}-1}$$

Torej $X + Y \sim \chi^n(m+n)$
Dokazali smo

Trditev 1.28. Naj bosta neodvisni slučajni spremenljivki $X \sim \chi^2(m), Y \sim \chi^2(n)$. Potem je $X + Y \sim \chi^2(m+n)$

Posledica 1.29. Če so $X_1, X_2 \dots X_n$ neodvisne slučajne spremenljivke, porazdeljene $N(0,1)$, potem je $Y := X_1^2 + \dots + X_n^2$ porazdeljena po $\chi^2(n)$

Oglejmo si transformacijo $f : \mathbb{R}^2 \to \mathbb{R}^2, (x,y) \to (u,v)$, ki preslika zvezno porazdeljen slučajni vektor (x,y) v zvezno porazdeljen slučajni vektor (u,v), torej $U = u(x,y), V = v(x,y)$
Označimo še $A_{u,v} = (-\infty, u] \times (-\infty, v]$
Potem je

$$F_{(U,V)}(u,v) = \iint\limits_{A_{u,v}} p_{(U,V)}(s,t)dsdt$$

Pot drugi strani pa je

$$F_{(U,V)}(u,v) = P((U,V) \in A_{u,v}) = P((X,Y) \in f^{-1}(A_{u,v})) = \iint\limits_{f^{-1}(A_{u,v})} p_{(X,Y)}(x,y)dxdy$$

Privzemimo še, da je f zvezno odvedljiva bijekcija. Potem je $f : \mathbb{R}^2 \to \mathbb{R}^2, (u,v) \to (x,y)$ tudi zvezno odvedljiva. Z zamenjavo spremenljivk $x = X(u,v), y = Y(u,v)$ v zadnjem intergalu dobimo

$$F_{(U,V)}(u,v) = \iint\limits_{A_{u,v}} p_{(X,Y)}(x(s,t), y(s,t)) \cdot |J(s,t)|dsdt$$

kjer je

$$J(u,v) = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} (u,v)$$

Jacobijeva determinanta.

Zaradi 1.8 imamo torej $p_{(U,V)}(u,v) = p_{(X,Y)}(x(u,v), y(u,v))|J(u,v)|$

Oglejmo si poseben primer

1.9 Matematično upanje oz. pričakovana vrednost

V primeru $X : \begin{pmatrix} x_1 & \dots & x_n \\ p_1 & \dots & p_n \end{pmatrix}$ je matematično upanje oz. pricakovana vrednost vsota $E(X) := \sum_{k=1}^{n} x_k \cdot p_k$

V posebnem primeru $p_1 = \dots = p_n = \frac{1}{n}$ je $E(X) = \frac{1}{n}\sum_{k=1}^{n} x_k = \frac{x_1+\dots+x_n}{n}$ - povprečje števil $x_1 \dots x_n$

expected value, expectation, mean value

Naj bo X diskretno porazdeljena slučajna spremenljivka z neskončno zalogo vrednosti:

$$X : \begin{pmatrix} x_1 & x_2 & x_3 & \dots \\ p_1 & p_2 & p_3 & \dots \end{pmatrix}$$

X ima matematično upanje oz. pričakovano vrednost, če je $\sum_{k=1}^{\infty} |x_k| p_k < \infty$
Tedaj je matematično upanje definirano kot $E(X) = \sum_{k=1}^{\infty} x_k \cdot p_k$
Naj bo sedaj X zvezno porazdeljena slučajna spremenljivka z gostoto p_X. Potem ima X matematično upanje, če je $\int_{-\infty}^{\infty} |x| \cdot p_X(x)dx < \infty$. Tedaj je je matematično upanje slučajne spremenljivke X enako $E(X) = \int_{-\infty}^{\infty} x \cdot p_X(x)dx$

Trditev 1.30. Naj bo $f : \mathbb{R} \to \mathbb{R}$ zvezna funkcija

(a) Če je $X : \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$
potem je $E(f \circ X) \equiv E(f(X)) = \sum_{k=1}^{\infty} f(x_k) \cdot p_k$ (če le to matematično upanje obstaja)

(b) Če je X zvezno porazdeljena z gostoto p_X, potem je $E(f \circ X) = \int_{-\infty}^{\infty} f(x) \cdot p_X(x)dx$

Posledica 1.31. Slučajna spremenljivka X ima matematično upanje $\iff$ X ima matematično upanje. Tedaj velja $|E(X)| \leq E(|X|)$

Posledica 1.32. Za $\forall a \in \mathbb{R}$ in vsako slučjano spremenljivko X z matematičnim upanjem velja $E(a \cdot X) = a \cdot E(X)$ (homogenost)

Podobno kot zadnjo trditev se dokaže

Trditev 1.33. Naj bo $f : \mathbb{R}^2 \to \mathbb{R}$ zvezna funkcija in (X, Y) slučajni vektor

(a) Naj bo (X, Y) diskretno porazdeljen $p_{ij} := P(X = x_i, Y = y_j)$. Potem je $E(f(X, Y)) = \sum_i \sum_i f(x_i, y_j) \cdot p_{ij}$ (če le vrsta (oz. končna vsota) absolutno konvergira)

(b) Naj bo (X, Y) zvezno porazdeljen z gostoto $p(X, Y)$. Potem je $E(f(X, Y)) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} f(x, y) p_{(X,Y)}(x, y)dy$ (če le integral absolutno konvergira)

Posledica 1.34. Če slučajni spremenljivki X in Y imata matamatično upanje, potem ga ima tudi $X+Y$ in velja $E(X+Y) = E(X)+E(Y)$ (aditivnost)

Posledica 1.35. Za slučajne spremenljivke $X_1 \dots X_n$, ki imajo matematično upanje, velja $E(a_1X_1+\dots+a_nX_n) = a_1E(X_1)+\dots+a_nE(X_n)$ za $\forall a_1 \dots a_n \in \mathbb{R}$

$$E(X+Y) = \int_{-\infty}^{\infty} x \cdot p_{X+Y}(x)dx \stackrel{?}{=} E(X) + E(Y) \text{ ni očitno iz tega}$$

Posebej to (v 2. zgledu) velja v primeru, ko so $\{X_k\}_{i=1}^n$ neodvisne. To velja tudi za Bernoullijevo zaporedje ponovitev poskusa: opazujemo dogodek A s $P(A) = p$. X je frekvenca dogodka A v n ponovitvah poskusa. Potem je $X \sim Bin(n,p)$ in $X = X_1 + \cdots + X_n$, kjer je $(X_k = 1)$ dogodek, da se A zgodi v k-ti ponovitvi poskusa, sicer je $(X_k = 0)$. Po zgornjem je $E(X) = n \cdot p$. Do tega lahko pridemo tudi direktno:

$$\begin{aligned} E(X) &= \sum_{k=0}^{n} k \cdot p_k = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k q^{n-k} = \\ &= \sum_{k=1}^{n} k \cdot \frac{n}{k} \binom{n-1}{k-1} p^k q^{n-k} = np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-k} \stackrel{j=k-1}{=} \\ &= np(\sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{n-1-j}) = np(p+q)^{n-1} = np \end{aligned}$$

Trditev 1.36 (Cauchy-Schwartzova neenakost)**.** Če obstajata $E(X^2)$ in $E(Y^2)$, potem obstaja tudi $E(X,Y)$ in velja $E(|X \cdot Y|) \leq \sqrt{E(X^2) \cdot E(Y^2)}$. Enačaj velja samo v primeru $|Y| = \sqrt{\frac{E(Y^2)}{E(X^2)}}|X|$ z verjetnostjo 1

Posledica 1.37. Če obstaja $E(X^2)$, potem obstaja $E(X)$ in velja $(E(X))^2 \leq E(X^2)$

Trditev 1.38. Naj bosta X in Y neodvisni slučajni spremenljivki, ki imata matematični upanji. Potem ima matematično upanje tudi $X \cdot Y$ in velja $E(X \cdot Y) = E(X) \cdot E(Y)$

Definicija 1.39 (Nekoreliranost)**.** Slučajni spremenljivki X in Y sta nekorelirani, če velja $E(X \cdot Y) = E(X) \cdot E(Y)$, sicer sta korelirani.

Po trditvi iz neodvisnosti sledi nekoreliranost. Obratno pa ne velja:

Trditev 1.40. $X : \begin{pmatrix} x_1 & x_2 \\ p_1 & p_2 \end{pmatrix}$, $Y : \begin{pmatrix} y_1 & y_2 \\ q_1 & q_2 \end{pmatrix}$
Potem sta X in Y neodvisni $\iff$ nekorelirani
$\iff E(X \cdot Y) = E(X) \cdot E(Y)$

1.10 Disperzija, kovarianco in korelacijski koeficient

Definicija 1.41 (Disperzija)**.** Naj obstaja $E(X^2)$. Disperzija oz. varianca slučajne spremenljivke X je $D(X) \equiv var(X) := E((X - E(X))^2)$

Disperzija meri razpršenost slučajne spremenljivke X okoli $E(X)$
Ker je $E((X-E(X))^2) = E(X^2-2E(X)X+(E(X))^2) = E(X^2)-2E(X)E(X)+(E(X))^2 = E(X^2)-(E(X))^2$, je $D(X) = E(X^2)-(E(X))^2$

Lastnosti disperzije:

- $D(X) \geq 0$ in $D(X) = 0 \iff P(X = E(X)) = 1$, t.j. X je izrojena slučajna spremenljivka
- $D(a \cdot X) = a^2 D(X)$ $a \in \mathbb{R}$
- $\forall a \in \mathbb{R}$ velja: $E((X-a)^2) \geq D(X)$. Enakost velja le v primeru $a = E(X)$

Definicija 1.42 (Standardna deviacija)**.** Standardna deviacija ali standardni odklon slučajne spremenljivke X je $\sigma(X) := \sqrt{D(X)}$

Zanjo velja $\sigma(aX) = |a| \cdot \sigma(X)$ za $\forall a \in \mathbb{R}$
Primeri nekaterih $E(X)$ in $D(X)$

1. enakomerna diskretna porazdelitev: $\begin{pmatrix} x_1 & \cdots & x_n \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix}$

$$E(X) = \frac{x_1 + \cdots + x_n}{n},$$
$$D(X) = E(X^2) - (E(X))^2 = \frac{x_1^2 + \cdots + x_n^2}{2} - (\frac{x_1 + \cdots + x_n}{2})^2$$

2. Binomska porazdelitev $Bin(n,p), n \in \mathbb{N}, p \in (0,1), q = 1-p$

$$E(X) = n \cdot p, D(X) = npq, \sigma(X) = \sqrt{npq}$$

3. Poissonova porazdelitev $Poi(\lambda), \lambda > 0$

$$E(X) = \lambda, D(X) = \lambda$$

4. Geometrijska porazdelitev $geo(p), p \in (0,1), q = 1-p$

$$E(X) = \frac{1}{p}, D(X) = \frac{q}{p^2}$$

5. Pascalova porazdelitev $Pas(m,p), m \in \mathbb{N}, p \in (0,1)$

$$E(X) = \frac{m}{p}, D(X) = \frac{mq}{p^2}$$

6. Enakomerna zvezna porazdelitev Ed na $[a, b]$

$$E(X) = \frac{a+b}{2}, D(X) = \frac{(b-a)^2}{12}$$

7. Normalna porazdelitev $N(\mu, \sigma)$

$$E(X) = \mu, D(X) = \sigma^2, \sigma(X) = \sigma$$

8. Porazdelitev gama $\gamma(b, c)$

$$E(X) = \frac{b}{c}, D(X) = \frac{b}{c^2}$$

9. Porazdelitev $\chi^2(n) = \gamma(\frac{n}{2}, \frac{1}{2})$

$$E(X) = n, D(X) = 2n$$

10. Eksponentna porazdelitev $Exp(\lambda), \lambda > 0 = \gamma(1, \lambda)$

$$E(X) = \frac{1}{\lambda}, D(X) = \frac{1}{\lambda^2}, \sigma(X) = \frac{1}{\lambda}$$

Preverimo, da je $D(X) = \sigma^2$ za $X \sim N(\mu, \sigma)$

$$D(X) = E((X - E(X))^2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 \cdot e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx$$

$$t = \frac{x-\mu}{\sigma} \implies x - \mu = \sigma t, dx = \sigma dt$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-\frac{1}{2}t^2} =$$

$$u = t, dv = t \cdot e^{-\frac{1}{2}t^2}$$
$$du = dt, v = -e^{-\frac{1}{2}t^2}$$

$$\frac{\sigma^2}{\sqrt{2\pi}}(-te^{-\frac{1}{2}t^2}|_{-\infty}^{\infty}) + \int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2} dt =$$
$$= \frac{\sigma^2}{\sqrt{2\pi}}(0 + \sqrt{2\pi}) = \sigma^2$$

Definicija 1.43 (Kovarianca)**.** Kovarianca slučajnih spremenljivk $K(X,Y) \equiv Cov(X,Y) := E((X - E(X))(Y - E(Y)))$

Ker je

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

je $cov(X,Y) = E(XY) - E(X)E(Y)$

Lastnosti:

1. $K(X,X) = D(X)$
2. $K(X,Y) = 0 \iff X$ in Y sta nekorelirani
3. K je simetrična in bilinearna funkcija:
   - $K(X,Y) = K(Y,X)$
   - $K(aX + bY, Z) = aK(X,Z) + bK(Y,Z) \forall a, b \in \mathbb{R}$
4. Če obstajata $D(X)$ in $D(Y)$, potem obstaja tudi $K(X,Y)$. Tedaj velja

   $$|K(X,Y)| \leq \sqrt{D(X) \cdot D(Y)} = \sigma(X) \cdot \sigma(Y)$$

   To sledi iz Cauchy-Schwartzove neenakosti ($|E(U \cdot V)| \leq \sqrt{E(U^2) \cdot E(V^2)}$) za slučajni spremenljivki $X - E(X)$ in $Y - E(Y)$. Enačaj v neenakosti velja $\iff Y - E(Y) \pm \frac{\sigma(Y)}{\sigma(X)}(X - E(X))$ z verjetnostjo 1
5. Če X in Y imata disperziji, potem jo ima tudi $X+Y$ in valja $D(X+Y) = D(X) + D(Y) + 2K(X,Y)$
   če sta X in Y nekorelirani (posebej neodvisni), potem je $D(X+Y) = D(X) + D(Y)$
6. Posplošitev: $D(X_1 + \cdots + X_n) = D(X_1) + \cdots + D(X_n) + 2\sum_{i<j} K(X_i, X_j)$
   Če so $X_1 \ldots X_n$ paroma nekorelirani (posebej neodvisni), potem je $D(X_1 + \cdots + X_n) = D(X_1) + \cdots + D(X_n)$

Definicija 1.44 (Standardizacija slučajne spremenljivke)**.** Standardizacija skučajne spremenljivke X je slučajna spremenljivka $X_s = \frac{X - E(X)}{\sigma(X)}$

Zanjo velja:

- $E(X_s) = 0$

- $D(X_s) = \frac{1}{\sigma(x)^2} \cdot D(X - E(X)) = \frac{1}{\sigma(X)^2} D(X) = 1$

Definicija 1.45 (Korelacijski koeficient)**.** Korelacijski koeficient slučajnih spremenljivk X in Y je

$$r(X,Y) = \frac{K(X,Y)}{\sigma(X)\sigma(Y)} = \frac{E((X - E(X))(Y - E(Y)))}{\sigma(X)\sigma(Y)} = E(X_s \cdot Y_s)$$

Lastnosti:

1. $r(X,Y) = 0 \iff X$ in Y sta nekorelirani
2. $r(X,Y) \in [-1,1]$, kar sledi iz lastnosti (4) za kovarianco
3. - $r(X,Y) = 1 \iff Y = \frac{\sigma(Y)}{\sigma(X)}(X - E(X)) + E(Y)$ z verjetnostjo 1
   - $r(X,Y) = -1 \iff Y = -\frac{\sigma(Y)}{\sigma(X)}(X - E(X)) + E(Y)$ z verjetnostjo 1

   Tedaj imamo linearno zvezo med X in Y

Torej sta X in Y nekorelirani $\overset{\text{v splošnem}}{\iff} \rho = 0 \overset{\text{ta primer}}{\iff} X, Y$ neodvisni
Kakšna je gostota, če je ρ blizu 1? $\rho \uparrow 1 : \rho \downarrow -1 :$
gostota je "skoraj skoncentrirana" na neki premici, torej med X in Y obstaja skoraj linearna zveza

1.11 Pogojna porazdelitev in pogojno matematično upanje

Izberimo si dogodek B s $P(B) > 0$

Definicija 1.46. Pogojna porazdelitvena funkcija slučajne spremnljivke X glede na B je $F_X(X \mid B) := P(X \leq x \mid B) = \frac{P(X \leq x \wedge B)}{P(B)}$

Ima enake lastnosti kot porazdelitvena funkcija

A Diskreten primer
Naj bo (X,Y) diskretno porazdeljen slučajni vektor z verjetnostno funkcijo $p_{ij} = P(X = x_i, Y = y_j)$ $i,j = 1,2\ldots$
Za pogoj B vzemimo $B = (Y = y_j)$ pri nekem j, torej $q_j = P(Y = Y_j)$
Potem je pogojna porazdelitvena funkcija slučajne spremenljivke X glede

$$F_X(X \mid Y = y_j) := \frac{P(X \leq x \mid Y = y_j)}{P(Y = y_j)} = \frac{1}{q_j} \sum_{j:x_j \leq x} p_{ij}$$

Če vpeljemo pogojno verjetnostno funkcijo

$$p_{i|j} = P(X = x_i \mid Y = y_j) = \frac{p_{ij}}{q_j}$$
$$F_X(X \mid Y = y_j) = \sum_{i:x_i \leq X} p_{i|j}$$

Pogojno matematično upanje slučajne spremenljivke X glede na $Y = y_j$ je matematično upanje te porazdelitve:

$$E(X \mid Y = y_j) := \sum_i x_i \cdot p_{i|j} = \frac{1}{q_j} \sum_i x_j \cdot p_{ij}$$

Regresijska funkcija $\ell(y_j) = \sum(X \mid Y = y_j)$, ki je definirana na zalogi vrednoti slučajne spremenljivke Y
Definirajmo novo slučajno spremenljivko $E(X \mid Y) = \ell(y)$, ki ji rečemo pogojno matematično upanje slučajne spremenljivke X glede slučajne spremenljivke Y
Ta ima shemo $E(X \mid Y) = \begin{pmatrix} \ell(y_1) & \ell(y_2) & \dots \\ q_1 & q_2 & \dots \end{pmatrix} = \begin{pmatrix} E(X \mid Y = y_1) & \dots \\ q_1 & \dots \end{pmatrix}$
Zanjo velja

$$E(E(X \mid Y)) = \sum_j \ell(y_j) \cdot q_j = \sum_j \sum_i x_i \cdot p_{ij} = \sum_i x_i (\sum_j p_{ij}) = \sum_i x_i \cdot p_i = E(X)$$

kjer je $p_i = P(X = x_i)$
Kaj dobimo, če sta X in Y neodvisni slučajni spremenljivki?
Tedaj je $p_{i|j} = \frac{p_{ij}}{q_j} = \frac{p_i \cdot q_j}{q_j} = p_i$ in $\ell(y_j) = E(E(X \mid Y = y_j)) = \sum_i x_i \cdot p_{i|j} = \sum_i x_i \cdot p_i = E(X)$, torej je regresijska funkcija kar konstanta $E(X)$ oz. je $E(X \mid Y)$ izrojena slučajna spremenljivka z vrednostjo $E(X)$

B Zvezni primer
Naj bo (X, Y) zvezno porazdeljen slučajni vektor z gostoto $p_{(X,Y)}(x, y)$.
Vzemimo $B = (y < Y \leq y + k)$ za nek $y \in \mathbb{R}, k > 0$.
Potem je $F_X(X \mid y < Y \leq y + k) = P(x \leq x \mid y < Y \leq y + k) = \frac{P(X \leq x, y < Y \leq y+k)}{P(y < Y \leq y+k)} = \frac{F_{(X,Y)}(x,y+k) - F_{(X,Y)}(x,y)}{F_Y(y+k) - F_Y(y)}$

Pogojna porazdelitvena funkcija slučajne spremenljivke X glede na dogodek $(Y = y)$ je limita, če obstaja:

$$F_X(x \mid Y = y) = \lim_{h \downarrow 0} F_X(x \mid y < Y \leq y+h) = \lim_{h \downarrow 0} \frac{F_{(X,Y)}(x, y+h) - F_{(X,Y)}(x, y)}{F_Y(y+h) - F_Y(y)}$$

Denimo sedaj, da sta $p_{X,Y}$ in p_Y zvezni funkciji. Tedaj je $F_X(X \mid Y = y) = \frac{\frac{\partial}{\partial y} F_{(X,Y)}(x,y)}{F'_Y(y)} = \frac{1}{p_Y(y)} \int_{-\infty}^{x} p_{(X,Y)}(x, v)dv$
Če vpeljemo pogojno pogojno gostoto $p_X(x \mid Y = y) := \frac{p_{(X,Y)}(x,y)}{p_Y(y)}$, je torej

$$F_{(X,Y)}(x \mid Y = y) = \int_{-\infty}^{x} p_X(u \mid y)du$$

Pogojno matematično upanje slučajne spremenljivke X glede na dogodek $(Y = y)$ je

$$E(X \mid Y = y) := \int_{-\infty}^{\infty} x \cdot p_X(x|y)dx = \frac{1}{p_Y(y)} \cdot \int_{-\infty}^{\infty} x p_{(X,Y)}(x, y)dx$$

Vpeljimo regresijsko funkcijo $l(y) := E(X \mid Y = y)$, definirano na zalogi vrednosti slučajne spremenljivke Y. Tako dobimo novo slučajno spremenljivko $E(X \mid Y) := l(y)$: pogojno matematično upanje slučajne spremenljivke X glede na slucajno spremenljivko Y.
Kot v diskretnem primeru se pokaže enakost $E(E(X \mid Y)) = E(X)$

1.12 Višji momenti in vrstilne karakteristike

Definicija 1.47 (Momenti)**.** Naj bo $k \in \mathbb{N}$ in $a \in \mathbb{R}$. Moment reda k glede na točko a je $m_k(a) := E((X - a)^k)$ (če obstaja)

Za a običajno vzamemo

1. $a = 0$: $z_k := m_k(0) = E(X^k)$ začetni moment reda k
2. $a = E(X)$: $m_k := m_k(E(X))$ cenralni moment reda k

Ocitno je $z_1 = E(X), m_2 = D(X)$

Trditev 1.48. Če $\exists m_n(a)$, potem obstajaj tudi moment $m_k(a)$ za vse $k < n$

Trditev 1.49. Če obstaja zacetni moment z_n, potem obstaja $m_n(a)$ glede na poljubno točko $a \in \mathbb{R}$

Centralne momente lahko izrazimo z začetnimi:

$$m_n(a) = E((X - a)^n) = \sum_{k=0}^{n} \binom{n}{k} (-a)^{n-k} E(X^k)$$

$$a = E(X) \implies m_k = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} z_1^{n-k} z_k$$

Asimetrija slučajne spremenljivke X je $A(X) := E(X_s^3) = E((\frac{X-E(X)}{\sigma_x})^3) = \frac{m_3}{m_2^{\frac{3}{2}}}$ $m_2 = \sigma^2 = D(X)$
$A(N(\mu, \sigma)) = 0$, ker

$$A(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^3 e^{-\frac{1}{2}x^2} dx = 0$$

Sploščenost (kurtozis) $K(X) := E(X_s^4) = \frac{m_4}{m_2^2}$
$K(N(\mu, \sigma)) = 3$
Če momenti ne obstajajo (npr. že $E(X)$ ne), potem si lahko pomagamo z vrstilnimi karakteristikami

Definicija 1.50 (Mediana)**.** Mediana slučajne spremenljivke X je vsaka vrednost $x \in \mathbb{R}$, za katero velja $P(X \leq x) \leq \frac{1}{2}$ in $P(Y \geq x) \geq \frac{1}{2}$ oz. $(1 - P(X < x) = 1 - F(x-))$

Če je F porazdelitvena funkcija za X, je to ekvivalentno s pogojem $F(x-) \leq \frac{1}{2} \leq F(x)$
Če je X zvezno porazdeljena slučajna spremenljivka, dobimo $F(X) = \frac{1}{2}$ oz. $\int_{-\infty}^{\infty} p(t)dx = \frac{1}{2}$

Te vrednosti (lahko jih je več) označimo z $X_{\frac{1}{2}}$

Definicija 1.51 (Kvantil)**.** Kvantil reda p ($p \in (0, 1)$) je vsaka vrednost x_p, za katero velja $P(X \leq x_p) \geq p$ in $P(X \geq x_p) \geq 1 - p$
Ekvivalentno je $F(x_p-) \leq p \leq F(x_p)$

Če je X zvezno porazdeljena, je pogoj $F(x_p) = p$ t.j. $\int_{-\infty}^{\infty} p(t)dt = p$

- Kvartili: $X_{\frac{1}{4}}, X_{\frac{2}{4}}, X_{\frac{3}{4}}$
- Percentili: $X_{\frac{1}{100}}, X_{\frac{2}{100}}, \ldots X_{\frac{99}{100}}$

Definicija 1.52 ((Semiinter)kvartilni razmik)**.** $s := \frac{1}{2}(x_{\frac{3}{4}} - x_{\frac{1}{4}})$

je nadomestek (analog) za standardno deviacijo

1.13 Rodovne funkcije

Definicija 1.53. Naj bo X slučajna spremenljivka z vrednostmi v $\mathbb{N} \cup \{0\}$: $p_k = P(X = k) k = 0, 1, 2 \ldots$ $p_k \geq 0, \sum_{k=0}^{\infty} = 1$

Rodovna funkcija skučajne spremenljivke X je

$$G_X(s) = p_0 + p_1 s + p_2 s^2 + \cdots = \sum_{k=0}^{\infty} p_k \dots s^k$$

za $\forall s \in \mathbb{R}$, za katere vrsta absolutno konvergira.

Očitno je $G_X(0) = p_0, G_X(1) = \sum_{k=0}^{\infty} p_k = 1$
Ker je $s^X : \begin{pmatrix} s^0 & s^1 & s^2 & \cdots \\ p_0 & p_1 & p_2 & \cdots \end{pmatrix}$, je $G_X(s) = E(s^X)$
Za $s \in [-1, 1]$ velja $|p_k \cdot s^k| \leq P_k$ in $\sum_{k=0}^{\infty} p_k = 1$. Zato je vrsta konvergentna, če je $|s| \leq 1$. Torej je konvergenčni radij vrste vsaj 1

Iz teorije Taylorjevih vrst sledi

Izrek 1.54 (O enoličnosti)**.** Naj imata X in Y rodovni funkciji G_X in G_Y. Potem je $G_X(s) = G_Y(s)$ za $\forall s \in [-1, 1] \leftrightarrow P(X = k) = P(Y = k)$ za vse $k = 0, 1, 2 \dots$
Tedaj velja $P(X = k) = \frac{1}{k!} G_X^k(0)$

$G_X(s) = \sum_{k=0}^{\infty} p_k s^k$, $p_k = P(X = k)$
Naj ima rodovna funkcija G_X slučajne spremenljivke X konvergenčni radij R > 1. Potem za $\forall s \in (-R, R)$ velja $G_X^{'}(s) = \sum_{k=1}^{\infty} k \cdot p_k s^{k-1}$
Če postavimo $s = 1$, dobimo $G^{'}(1) = \sum_{k=1}^{\infty} k \cdot p_k = E(X)$

Izrek 1.55. Naj ima X rodovno funkcijo $G_X(s)$ in naj bo $n \in \mathbb{N}$. Potem je

$$G_X^{(n)}(1-) \equiv \lim_{s \nearrow 1} G_X^{(n)}(s) = E(X(X-1)(X-2)\dots(X-N+1))$$

Posledica 1.56.

$$E(X) = G_X^{'}(1-)$$

$$\begin{aligned} &D(X) = E(X^2) - (E(X))^2 = \\ &E(X(X-1)) + E(X) - (E(X))^2 = \\ &G_X^{(2)}(1-) + G_X^{(1)}(1) - (G_X^{(1)}(1-))^2 \end{aligned}$$

Izrek 1.57. Naj bosta X in Y neodvisni slučajni spremenljivki z rodovnima funkcijama G_X in G_Y. Potem je $G_{X+Y}(s) = G_X(s) \cdot G_Y(s)$ za $s \in [-1, 1]$

Posplošitev 1.58. Če so $X_1, X_2 \dots X_n$ neodvisne slučajne spremenljivke, potem je za vse $s \in [-1, 1] G_{X_1+\cdots+X_n}(s) = G_{X_1}(s) \cdot \dots \cdot G_{X_n}(s)$.
Če so $X_1, X_2 \dots X_n$ enako porazdeljene in neodvisne, potem je

$$G_{X_1+\cdots+X_n}(s) = (G_X(s))^n \tag{1}$$

Izrek 1.59. Naj bodo za $\forall n \in \mathbb{N}$ slučajne spremenljivke $N, X_1, X_2 \dots X_n$ neodvisne. Naj ima N rodovno funkcijo G_N, X_n pa rodovno funkcijo G_X. Potem ima slučajna spemenljivka $S := X_1 + X_2 + \cdots + X_n$ rodovno funkcijo enako $G_S = G_N \circ G_X$ oz. $G_S(s) = G_N(G_X(s))$ za $s \in [-1, 1]$

To je posplošitev formule 1: $P(N = n) = 1, G_N(s) = 1 \cdot s^n = s^n$

Posledica 1.60. Pri predpostavkah iz izreka velja Waldova enakost:

$$E(S) = E(N) \cdot E(X)$$

1.14 Momentno rodovna funkcija

Definicija 1.61 (Momentno rodovna funkcija)**.** Momentno rodovna funkcija je $M_X(t) = E(e^{tX})$ za $t \in \mathbb{R}$, za katere obstaja matematično upanje

V primeru zvezne porazdelitve je $M_X(t) = \int_{-\infty}^{\infty} e^{tx} p_X(x) dx$
To je Laplaceova transformacija funkcije p_X
V diskretnem primeru $X : \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$ je $M_X(t) = \sum_i e^{tx} p_i$

V posebnem primeru, ko ima X nenegative celoštevilske vrednosti, je

$$M_X(t) = \sum_{i=0}^{\infty} e^{it} p_i =$$
$$= \sum_{i=0}^{\infty} p_i (e^t)^i = G_X(e^t)$$

$$M_X(t) = E((e^t)^X) = G_X(e^t), G_X(s) = E(s^X)$$

Očitno je $M_X(0) = E(e^0) = E(1) = 1$

Izrek 1.62. Naj bo $M_X(t) < \infty$ (obstaja, $< \infty$ zato, ker je $e^t > 0$) za $\forall t \in (-\delta, \delta)$ pri nekem $\delta > 0$. Potem je porazdelitev za X natanko določena z M_X, vsi začetni momenti obstajajo, $z_k = E(X^k) = M_X^k(0)$ za $\forall k \in \mathbb{N}$ in velja $M_X(t) = \sum_{k=0}^{\infty} \frac{z_k}{k!} t^k$ za $\forall t \in (-\delta, \delta)$

Trditev 1.63. $M_{aX+b}(t) = e^{bt} M_X(at), a \neq 0, b \in R$

Izrek 1.64. Če sta X in Y neodvisni slučajni spremenljivki, potem je $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$

Trditev 1.65. Naj bosta X in Y neodvisni slučajni spremenljivki in $X \sim N(\mu_x, \sigma_x), Y \sim N(\mu_y, \sigma_y)$. Potem je $X + Y \sim N(\mu_x + \mu_y, \sqrt{\sigma_x^2 + \sigma_y^2})$

Opomba. Če bi vedeli, da je $X + Y$ porazdeljena normalno, bi "samo" izračunali parametra

1.15 Šibki in krepki zakon velikih števil

Definicija 1.66 (Verjetnostna konvergenca)**.** Zaporedje slučajnih spremenljivk $\{X_n\}_{n\in\mathbb{N}}$ verjetnostno konvergira proti slučajni spremenljivki X, če za $\forall\epsilon > 0$ velja

$$\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0$$

oziroma

$$\lim_{n\to\infty} P(|X_n - X| < \epsilon) = 1$$

Definicija 1.67 (Skoraj gotova konvergenca)**.** Zaporedje slučajnih spremenljivk $\{X_n\}_{n\in\mathbb{N}}$ skoraj gotovo konvergira proti slučajni spremenljivki X, če velja

$$P(\lim_{n\to\infty} X_n = X) = 1$$

Tukaj je

$$\begin{aligned}(\lim_{n\to\infty} X_n = X) &= \{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} \\ &= \{\omega \in \Omega : \forall k(\in \mathbb{N})\ \exists m \in \mathbb{N}\ \forall n \geq m : |X_n(\omega) - X(\omega)| < \frac{1}{k}\} \\ &= \{\cap_{k\in\mathbb{N}}\ \cup_{m\in\mathbb{N}}\ \cap_{n\geq m}\ \omega \in \Omega : |X_n(\omega) - X(\omega)| < \frac{1}{k}\}\end{aligned}$$

Opomba. Števne unije in preseki $\implies$ smo v $\sigma-$algebri, torej je to res dogodek

Trditev 1.68. Če $X_n \xrightarrow{n\to\infty} X$ skoraj gotovo, potem za $\forall\epsilon > 0\, \lim_{m\to\infty} P(|X_n - X| < \epsilon \text{ za } n \geq m) = 1$

Posledica 1.69. Če $X_n \xrightarrow{n\to\infty} X$ skoraj gotovo, potem $X_n \xrightarrow{n\to\infty} X$ verjetnostno konvergira.

Opomba. Obratna implikacija ne velja

Definicija 1.70. Naj bo $X_1, X_2, X_3 \ldots$ zaporedje slučajnih spremenljivk, ki imajo matematično upanje. Definirajmo $Y_n = \frac{S_n - E(S_n)}{n} = \frac{X_1 + \cdots + X_n}{n} - \frac{E(X_1) + \cdots + E(X_n)}{n}$
Potem je $E(Y_n) = 0$
Za $\{X_n\}_{n\in\mathbb{N}}$ velja šibki zakon velikih števil (ŠZVŠ), kadar

$$Y_n \overset{n\to\infty}{\to} 0$$

verjetnostno, torej za

$$\forall \epsilon > 0 \lim_{n\to\infty} P(|y| < \epsilon) = 1 = \lim_{n\to\infty} P(|\frac{S_n - E(S_n)}{n}| < \epsilon)$$

Za $\{X_n\}_{n\in\mathbb{N}}$ velja krepki zakon velikih števil (KZVŠ), kadar

$$Y_n \overset{n\to\infty}{\to} 0$$

skoraj gotovo, torej

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

Če velja KVZŠ, potem velja ŠVZŠ

Izrek 1.71.

a Neenakost Markova: če slučajna spremenljivka X ima matematično upanje, potem je $P(|X| \geq a) \leq \frac{E(|X|)}{a}$ za $\forall a > 0$

b Neenakost Čebiševa: če slučajna spremenljivka X ima disperzijo, potem je $P(|X - E(X)| \geq a \cdot \sigma(X)) \leq \frac{1}{a^2}$ za $\forall a > 0$ (pomembno za $a \geq 1$, ker je verjetnost ≤ 1)
oz. če pišemo $\epsilon = a \cdot \sigma(X) \implies P(|X - E(X)| \geq \epsilon) \leq \frac{D(X)}{\epsilon^2}$ za $\forall \epsilon > 0$

Izrek 1.72 (Markov)**.** Če za zaporedje slučajnih spremenljivk $\{X_n\}_{n\in\mathbb{N}}$ velja $\frac{D(S_n)}{n^2} \overset{n\to\infty}{\to} 0$, potem velja ŠZVŠ. Tukaj je $S_n := X_1 + \cdots + X_n$

Posledica 1.73 (Izrek Čebišev)**.** Če so $X_1, X_2, \ldots$ paroma nekorelirane slučajne spremenljivke in $\sup_{n\in\mathbb{N}} D(X_n) = c < \infty$, potem za $\{X_n\}_{n\in\mathbb{N}}$ velja ŠZVŠ

Izrek 1.74 (Kolmogorov)**.** Če za neodvisne slučajne spremenljivke $\{X_n\}_{n\in\mathbb{N}}$ velja $\sum_{n=1}^{\infty} \frac{D_n}{n^2} < \infty$, potem velja KZVŠ, t.j. $P(\lim_{n\to\infty} \frac{S_n - E(S_n)}{n} = 0) = 1$. Posebej je pogoj za vrsto izpolnjen, če je $\sup_n D(X_n) < \infty$

1.16 Centralni limitni izrek

Definicija 1.75. Naj bo $\{X_n\}_{n\in\mathbb{N}}$ zaporedje slučajnih spremenljivk s končnimi disperzijami. Definiramo $S_n := X_1 + \cdots + X_n$ in standardizirajmo:

$$Z_n = \frac{S_n - E(S_n)}{\sigma(S_n)}$$

torej imamo

$$E(Z_n) = 0, D(Z_n) = 1$$

Za $\{X_n\}_{n\in\mathbb{N}}$ velja centralni limitni izrek, če je

$$F_{Z_n}(x) = P(Z_n \leq x) \overset{n\to\infty}{\to} F_{N(0,1)} \; \forall x \in \mathbb{R}$$

to je

$$P(\frac{S_n - E(S_n)}{\sigma(S_n)} \leq x) \overset{n\to\infty}{\to} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \text{ za } \forall x \in \mathbb{R}$$

Pravimo, da $\{Z_n\}_{n\in\mathbb{N}}$ po porazdelitvi konvergira proti standardizirani normalni porazdelitvi.

Izrek 1.76 (Centralni limitni izrek (CLI, osnovna verzija))**.** Naj bodo $X_1, X_2 \ldots$ neodvisne in enako porazdeljene slučajne spremenljivke. Potem zanje velja centralni limitni zakon, t.j

$$P(\frac{S_n - E(S_n)}{\sigma(S_n)} \leq x) \overset{n\to\infty}{\to} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

$\forall x \in \mathbb{R}$

Dokazal je Ljapunov (1900), s tem je posplošil Laplaceov izrek iz leta 1812. V dokazu bomo uporabili

Izrek 1.77 (O zveznosti rodovne funkcije)**.** Naj za zaporedje $\{Z_n\}_{n\in\mathbb{N}}$ slučajnih spremenljivk velja:
$M_{Z_n}(t) \to M_{N(0,1)}(t) = e^{\frac{t^2}{2}}$ za vse $t \in (-\delta, \delta)$ pri nekem $\delta > 0$
Potem $F_{Z_n}(x) \to F_{N(0,1)}(x)$ za $\forall x \in \mathbb{R}$

V splošnem se CLI dokaže s pomočjo karakterističnih funkcij:
naj bo X slučajna spremenljivka,

$$\ell_X(t) := E(e^{itX}) = E(cos(tX)) + iE(sin(tX)) \; t \in \mathbb{R}$$

za razliko od momentno rodovnih funkcij karakteristične funkcije vedno odstajajo
v zveznem primeru je $\int_{-\infty}^{\infty} e^{itx} p(x)dx$ - Fourierova transformacija funkcije $p_X(x)$
$X_1, X_2 \dots X_n$ neodvisne, enako porazdeljene

$$\mu := E(X_n), \sigma := \sigma(X_n)$$
$$E(S_n) \overset{\text{neodvisnost}}{=} E(X_1) + \cdots + E(X_n) = n\mu$$
$$D(S_n) \overset{\text{neodvisnost}}{=} D(X_1) + \cdots + D(X_n) = n\sigma^2$$

$X_1, X_2 \dots X_n$ neodvisne slučajne spremenljivke

$$Z_n = \frac{S_n - E(S_n)}{\sigma(S_n)} = \frac{S_n - n\mu}{\sqrt{n}\sigma} = \frac{\frac{S_n}{n} - \mu}{\frac{\sigma}{\sqrt{n}}}$$
$$\overline{X} := \frac{S_n}{n} = \frac{X_1 + \cdots + X_n}{n} \implies Z_n = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Po CLI za velike n velja $Z_n \approx N(0,1)$, zato je $\overline{X} \approx N(\mu, \frac{\sigma}{\sqrt{n}})$ oz. $S_n \approx N(n\mu, \sigma\sqrt{n})$
Če so $X_1, X_2 \dots$ porazdeljene normalno $N(\mu, \sigma)$, potem je $Z_n \sim N(0,1)$, torej $F_{Z_n}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$

2 Statistika

2.1 Osnovni pojmi

Kot vedo statistiko razdelimo na:

1. opisno statistiko: zbiranje, razvrščanje, prikazovanje podatkov, računanje osnovnih količin
2. analitično statistiko: uporaba podatkov pri sklepanju glede zakonitosti danega področja

Definicija 2.1 (Populacija)**.** Populacija je končna ali neskončna množica elementov, pri katerih merimo ali opazujemo neko količino

Matematični pogled: na verjetnostnem prostoru $(\Omega, \mathcal{F})$ imamo slučajno spremenljivko X.
Praviloma ne moremo izmeriti cele populacije, ampak meritve opravimo na relativno majhnem delu populacije, na vzorcu. Le-ta mora biti reprezentativen, izbran nepristransko in dovolj velik.
Matematični pogled: vzorec velikosti n je slučajni vektor $(x_1 \dots x_n)$, kjer so komponente enako porazdeljene kot slučajna spremenljivka X in med seboj neodvisne.
Vrednost tega slučajnega vektorja pri enem naboru n meritev je realizacija vzorca: $(x_1 \dots x_n)$: to so konkretni podatki, ki jih analiziramo. Pri opisni statistiki predstavimo in obdelamo te podatke.
Iz teh vzorčnih podatkov želimo oceniti nekatere lastnosti populacije, kot sta:

1. sredina populacije μ, t.i. matematično upanje slučajne spremenljivke X
2. povprečni odklon σ od sredine populacije, t.i. Standardna deviacija slučajne spremenljivke X

Ocene za μ so:

- vzorčno povprečje: $\overline{x} = \frac{x_1 + \dots + x_n}{n}$
- vzorčni modus: najpogostejša vrednost v vzorcu
- vzorčna mediana: srednja vrednost v vzorcu, urejenem po velikosti

Ocene za σ so:

- vzorčni razmak: razlika med največjo in najmanjšo vrednostjo v vzorcu
- vzorčna disperzija: $s_0^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$
- popravljena vzorčna disperzija: $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2 = \frac{n}{n-1} s_0^2$

2.2 Vzorčne statistike in cenilke

Definicija 2.2 (Vzorčna statistika)**.** Naj bo $(X_1, X_2 \dots X_n)$ vzorec t.i. slučajni vektor, kjer so $X_1 \dots X_n$ enako porazdeljene kot slucajna spremenljivka X in med seboj neodvisne.
Vzorčna statistika je simetrična funkcija vzorca $y = g(X_1, X_2 \dots X_n)$, kjer je g simetricna funkcije n spremenljivk

Praviloma vzorčna statistika ocenjuje vrednost nekega parametra ξ. Tedaj je y cenilka za parameter.

y je odvisna od n, zato pišemo tudi $y_n = g(X_1 \dots X_n)$.

Definicija 2.3 (Nepristranskost, doslednost)**.** Če je $E(Y) = \xi$, je Y nepristranska cenilka za parameter ξ
Cenilka $Y = Y_n$ je dosledna, če $Y_n \overset{n\to\infty}{\to} \xi$ verjetnostno, t.i. $\forall \epsilon > 0$ je $\lim_{n\to\infty} P(|Y_n - \xi| \geq \epsilon) = 0$ oz. $\lim_{n\to\infty} P(|Y_n - \xi| < \epsilon) = 1$

Definicija 2.4 (Standardna napaka)**.** Standardna napaka vzorčne statistike Y je standardna deviacija slučajne spremenljivke Y: $SE(Y) := \sigma(Y)$

Definicija 2.5 (Vzorčno povprečje)**.** Naj bo X slučajna spremenljivka na populaciji, ki ima matematično upanje $E(X) = \mu$ in standardno deviacijo $\sigma(X) = \sigma$. Naj bo $(X_1 \dots X_n)$ vzorec. Definirajmo vzorčno povprečje

$$\overline{X} = \frac{X_1 + \dots + X_n}{n}$$

ki je vzorčna statistika.

Je cenilka za $\overline{X}$, ki je nepristranska:

$$E(\overline{X}) = \frac{1}{n}(E(X_1) + \dots + E(X_n)) = \frac{1}{n} n \cdot \mu = \mu$$

Po ŠZVŠ (izreku Čebiševa) je to dosledna cenilka za μ.
Ker je

$$D(\overline{X}) \overset{\text{neodv}}{=} \frac{1}{n^2} \sum_{i=1}^{n} D(X_i) = \frac{1}{n^2} n \cdot \sigma^2 = \frac{\sigma^2}{n}$$

je standardna napaka

$$SE(Y) = \frac{\sigma}{\sqrt{n}}$$

- čim večji n, bolje oceni parameter μ
Po CLI je pri velikem n slučajna spremenljivka $Z_n := \frac{S - n\mu}{\sigma\sqrt{n}} = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\overline{X} - \mu}{\sigma}\sqrt{n}$ porazdeljena približno $N(0, 1)$ oz. $\overline{X}$ je porazdeljen približno $N(\mu, \frac{\sigma}{\sqrt{n}})$
Če je X normalno porazdeljena $N(\mu, \sigma)$, potem je $\overline{X}$ porazdeljen $N(\mu, \frac{\sigma}{\sqrt{n}})$ za vsak n

Trditev 2.6. Naj bo Y_n cenilka za ξ. Če je $E(Y_n) \overset{n\to\infty}{\to} \xi$ in $D(Y_n) \overset{n\to\infty}{\to} 0$, potem je $Y = Y_n$ dosledna cenilka za ξ

Definicija 2.7 (Vzorčna disperzija)**.** Naj bo X slučajna spremenljivka na populaciji. Vzorčna disperzija je definirana s

$$s_0^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

popravljena vzorčna disperzija pa je

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

Kako sta porazdeljeni, če je $X \sim N(\mu, \sigma)$?
Raje vzemimo vzorčno statistiko: $\chi^2 := \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{n}{\sigma^2}s_0^2 = \frac{n-1}{\sigma^2}s^2$
Ni lahko izračunati, da je $\chi^2 \sim \chi^2(n-1)$
Ideja izpeljave je $\chi^2 = Z_1^2 + \cdots + Z_{n-1}^2$ za $Z_i \sim N(0,1)$ in med seboj neodvisne.
Potem uporabimo trditev iz verjetnosti: $Z_i^2 \sim \chi^2(1)$, torej $E(\chi^2) = n-1$, $D(\chi^2) = 2(n-1)$. Od tod sledi

$$E(s_0^2) = E(\frac{\sigma^2}{n}\chi^2) = \frac{\sigma^2}{n}E(\chi^2) = \frac{n-1}{n}\sigma^2$$

torej s_0^2 ni nepristranska za σ^2, je pa asimptotično nepristranska, t.i. $E(s_0^2) \overset{n\to\infty}{\to} \sigma^2$
Podobno je $E(s^2) = \frac{\sigma^2}{n-1}E(\chi^2) = \sigma^2$, torej je s^2 nepristranska cenilka za σ^2
Ker je $D(s_0^2) = \frac{\sigma^4}{n^2}D(\chi^2) = \frac{\sigma^4 2(n-1)}{n^4} \overset{n\to\infty}{\to} 0$ in $D(s^2) = \frac{2\sigma^4}{(n-1)^2} \overset{n\to\infty}{\to} 0$, iz trditve sledi, da sta s_0^2 in s^2 dosledni cenilki za σ^2

Studentova t-porazdelitev

$$p(x) = \frac{1}{\sqrt{n}B(\frac{n}{2}, \frac{1}{2})}(1 + \frac{x^2}{n})^{-\frac{n+1}{2}}$$

kjer je $B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ Beta funkcija.

$n = 1: \quad \frac{1}{\pi}(1+x^2)^{-1} = \frac{1}{\pi(1+x^2)}$ Cauchyjeva porazdelitev

ko gre $n \to \infty$, gre $\sqrt{n}B(\frac{n}{2}, \frac{1}{n}) \to \sqrt{2\pi}$ in $(1+\frac{x^2}{n})^{-\frac{n-1}{2}} = ((1+\frac{x^2}{n})^n)^{-\frac{n+1}{2n}} \to e^{-\frac{x^2}{2}}$

torej je pri velikih n gostota približno $N(0,1)$

$n = 2: \quad \frac{1}{\sqrt{2}B(1, \frac{1}{2})}(1+\frac{x^2}{2})^{-\frac{3}{2}}$

za $n \geq 2$ je $E(X) = 0$

$n = 3: \quad c \cdot (1+\frac{x^2}{2})^{-2} \approx \frac{1}{x^4}$ za velike x

za $n \geq 3$ je $D(X) = \frac{n}{n-2} > 1$

Leta 1908 jo je odkril W.S. Gosset, statistik v pivovarni guiness v Dublinu. Student je njegov psevdonim.

Pri normalni porazdelitvi slučajne spremenljivke $X \sim N(\mu, \sigma)$ je vzorčno povprečje $\overline{X}$ porazdeljeno $N(\mu, \frac{\sigma}{\sqrt{n}}), \overline{X} = \frac{X_1+\cdots+X_n}{n}$, torej je $Z := \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\overline{X}-\mu}{\sigma}\sqrt{n}$ porazdeljena N(0,1). Če poznamo σ, potem bomo znali povedati, kako dobra ocena za μ je $\overline{X}$ ($\to$ intervali zaupanja).

Kako ravnati, če σ ne poznamo?
Lahko jo ocenimo s $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2}$, tako da potem vzorčna statistika $T = \frac{\overline{X}-\mu}{s}\sqrt{n}$ ni več porazdeljena po $N(0,1)$, niti približno normalna, razen če je n velik in je s potem skoraj konstanta σ.

Kako je porazdeljena vzorčna statistika T?
Ker je $\chi^2 = \frac{1}{\sigma^2}\sum_{i=1}^n (X_i - \overline{X})^2 = \frac{(n-1)S^2}{\sigma^2}$, je $\frac{Z}{T} = \frac{S}{\sigma} = \sqrt{\frac{\chi^2}{n-1}}$, torej je $T = \frac{Z}{\sqrt{\frac{\chi^2}{n-1}}}$
Izkaže se, da sta $Z \sim N(0,1)$ in $\chi^2 \sim \chi^2(n-1)$ neodvisni slučajni spremenljivki. Od tod lahko izračunamo, da ima T Studentovo porazdelitev z $n-1$ prostorskimi stopnjami:

$$p_T(t) = \frac{1}{(n-1)B(\frac{n-1}{2}, \frac{1}{2})} \cdot \frac{1}{(1+\frac{x^2}{n-1})^{\frac{n}{2}}}$$

2.3 Metode za pridobivanje cenilk

2.3.1 Metoda momentov

Definicija 2.8 (Vzorčni moment)**.** Naj bo $(X_1, X_2 \dots X_n)$ vzorec velikosti n, torej $X_1 \dots X_n$ neodvisne slučajne spremenljivke, porazdeljene kot slučajna spremenljivka X. Začetni moment reda k je $z_k = E(X^k)$. Definiramo k-ti vzorčni moment

$$z_k := \frac{X_1^k + \dots + X_n^k}{n}$$

Le ta je nepristranska cenilka za z_k:

$$E(Z_k) = \frac{1}{n}(E(X_1^k) + \dots + E(X_n^k)) = z_k$$

Z_k je tudi dosledna cenilka za z_k.

Naj bo gostota slučajne spremenljivke X odvisna od parametrov $\xi_1 \dots \xi_n$: $p(X; \xi_1 \dots \xi_m)$.
Naj odstajajo začetni momenti

$$z_k = E(X^k) = \int_{-\infty}^{\infty} p(x; \xi_1 \dots \xi_n)dx, k = 1, 2 \dots m$$

Denimo, da iz teh m enačb lahko izrazimo parametre:

$$\xi_k = \phi_k(z_1, z_2 \dots z_m), k = 1 \dots m$$

Za neko funkcijo ϕ_k. Potem je

$$c_k := \phi_k(z_1 \dots z_m)$$

cenilka za parameter $\xi_k, k = 1 \dots n$

2.3.2 Metoda maksimalne zanesljivosti

oz. največjega verjetja

Definicija 2.9 (Funkcija zanesljivosti)**.** Naj bo gostota slučajne spremenljivke X odvisna od parametra ξ, torej $p(x;\xi)$. Funkcija zanesljivosti (likelihood function) je

$$L(x_1 \dots x_n;\xi) = p(x_1;\xi) \cdot \dots \cdot p(x_n;\xi)$$

Pri danih $x_1 \dots x_n$ izberimo tak ξ_{max}, da ima L tam maksimum. Ta vrednost parametra je odvisna od $x_1 \dots x_n$, torej $\xi_{max} = \phi(x_1, x_2 \dots x_n)$ za neko funkcijo ϕ. Tako dobimo cenilko $c := \phi(x_1 \dots x_n)$ za parameter ξ

2.4 Intervalsko ocenjevanje parametrov

Definicija 2.10 (Interval zaupanja)**.** Naj bo gostota slučajne spremenljivke X odvisna od parametra ξ. Interval $[A, B]$ (odvisen le od $(x_1 \dots x_n)$ in ne od ξ) je interval zaupanja za parameter ξ, pri stopnji tveganja $\alpha \in (0, 1)$, če je

$$P(\xi \in [A, B]) = 1 - \alpha \text{ oz. } P(\xi \notin [A, B]) = \alpha$$

Za α običajno vzamemo vrednost 0.05 (ali 0.01)
A in B sta vzorčni statistiki, $1 - \alpha$ je stopnja zaupanja

2.5 Preizkušanje statističnih hipotez

Definicija 2.11 (Statistična hipoteza)**.** Statistična hipoteza je vsaka domneva o porazdelitvi slučajne spremenljivke X na populaciji

Definicija 2.12 (Enostavnost hipoteze)**.** Hipoteza je enostavna, če natanko določa porazdelitev, sicer je sestavljena

Vedno preizkušamo eno ničelno hipotezo H_0 nasproti alternativni hipotezi H_1

Za H_0 običajno vzamemo enostavno hipotezo, za katero upamo, da jo bomo zavrnili
Hipoteza je lahko pravilna ali nepravilna. Ideal je sprejeti pravilno in zavrniti nepravilno. Odločiti se moramo na osnovi vzorca. Če vzorčni podatki preveč odstopajo od hipoteze, potem niso konsistentni z njo oz. so razlike značilne (signifikantne); tedaj hipotezo zavrnemo
Vnaprej določimo stopnjo značilnosti $\alpha \in [0, 1]$, to je verjetnost, da zavrnemo pravilo hipotezo. Običajno je $\alpha = 0.05$ ali $\alpha = 0.01$. Take teste imenujemo testi značilnosti

Primeri testov znacilnosti

2.5.1 test Z

$X \sim N(\mu, \sigma), \sigma$ znan parameter
Ničelna domneva je $H = (\mu = \mu_0)$, kjer je μ_0 damo realno število.
Pri predpostavki $H_0(\mu = \mu_0)$ je $Z := \frac{\overline{X}-\mu}{\sigma}\sqrt{n}$ porazdeljena $N(0,1)$, saj je $\overline{X} \sim N(\mu_0, \frac{\sigma}{\sqrt{n}})$
Vzemimo $H_1(\mu \neq \mu_0)$. Tedaj H_0 zavrnemo, če vzorčna vrednost za Z leži na kritičnem območju

$$K_\alpha = (-\infty, -z_{\frac{\alpha}{2}}] \cup [z_{\frac{\alpha}{2}}, \infty)$$

kjer je α stopnja značilnosti in $P(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$

$Z \ldots$ testna statistika
Pri stopnji značilnosti α določimo $z_{\frac{\alpha}{2}} > 0$, da je

$$P(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$$

K_α kritično območje
Če je vzorčna vrednost za Z na K_α, hipotezo H_0 zavrnemo

2.5.2 test T

$X \sim N(\mu, \sigma)$
$H_0(\mu = \mu_0) : H_1(\mu \neq \mu_0), \mu_0$ je dano število
Testna statistika

$$T = \frac{\overline{X} - \mu_0}{S}\sqrt{n}$$

S je vzorčna deviacija
Pri predpostavki H_0 je porazdeljena po Student(n-1)
$K_\alpha = (-\infty, -t_{\frac{\alpha}{2}}] \cup [t_{\frac{\alpha}{2}}, \infty)$
Če vzorčna vrednost za T leži na K_α, hipotezo zavrnemo

Definicija 2.13 (P-vrednost)**.** P-vrednost je najmanjša stopnja značilnosti, pri kateri še lahko zavrnemo hipotezo (pri danih vzorčnih podatkih)

V našem primeru je $P = 0.89\% = 0.0089$

2.5.3 Studentov primerjalni test

Imejmo 2 neodvisna vzorca velikosti n in m. Prvi je vzet iz populacije, na kateri ime slučajna spremenljivka $X \sim N(\mu_x, \sigma_x)$, druga pa iz populacije, na kateri imamo $Y \sim N(\mu_x, \sigma_x)$ in privzamemo $\sigma_x = \sigma_y = \sigma$. Predpostavljamo torej enakost disperzij. Če sta s_x^2 in s_y^2 povprečni vzorčni disperziji

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \quad \left(\frac{S_X^2}{\sigma^2} \sim \chi^2(n-1)\right)$$
$$S_Y^2 = \frac{1}{m-1}\sum_{i=1}^{m}(y_i - \overline{y})^2 \quad \left(\frac{S_Y^2}{\sigma^2} \sim \chi^2(m-1)\right)$$

potem definiramo skupno vzorčno varianco

$$S^2 := \frac{(n-1)S_X^2 + (m-1)S_Y^2}{m+n-2} = \frac{1}{m+n-2}(\sum_{i=1}^{n}(x_i - \overline{x})^2 + \sum_{i=1}^{m}(y_i - \overline{y})^2)$$

Testiramo hipotezo $H_0(\mu_x = \mu_y) : H_1(\mu_x \neq \mu_y)$. Testna statistika

$$T = \frac{\overline{X} - \overline{Y}}{S}\sqrt{\frac{mn}{m+n}}$$

Potem je

$$\overline{X} \sim (\mu_x, \frac{\sigma}{\sqrt{n}}),\ \overline{Y} \sim (\mu_y, \frac{\sigma}{\sqrt{m}})$$
$$\overline{X} - \overline{Y} \sim N(\mu_x - \mu_y, \sqrt{(\frac{\sigma}{\sqrt{n}})^2 + (\frac{\sigma}{\sqrt{m}})^2}) =$$
$$= N(0, \sigma\sqrt{\frac{1}{n} + \frac{1}{m}}) = N(0, \sigma\sqrt{\frac{m+n}{mn}})$$

Zato ima spremenljivka

$$Z = \frac{\overline{X} - \overline{Y}}{S}\sqrt{\frac{mn}{m+n}} \sim N(0,1).$$

Ker je spremenljivka

$$U = \frac{(m+n-2)S^2}{\sigma^2} \sim \chi^2(m+n-2)$$

je

$$T = \frac{Z}{\sqrt{\frac{U}{m+n-2}}} = \frac{Z\sigma}{\sqrt{S^2}} \sim Student(m+n-2)$$

Poleg populacijskega povprečja μ lahko testiramo tudi druge količine:

- standardna deviacija $\sigma : H_0(\sigma = \sigma_0), \sigma_0$ je dano število
- tip porazdelitvenega zakona: $H_0(F = F_0)$
- neodvisnost dveh spremenljivk
- korelacijski koeficient

2.6 F-test

Pri prejšnjem testu smo privzeli $\sigma_X = \sigma_Y$- Kako preveriti smiselnost te predpostavke? $H_0(\sigma_X = \sigma_Y) : H_1(\sigma_X \neq \sigma_Y)$

f-statistika:
Če je ($\sigma_X = \sigma_Y$, velja

$$F = \frac{\frac{S_X^2}{\sigma_X^2}}{\frac{S_Y^2}{\sigma_Y^2}} = \frac{S_X^2}{S_Y^2} = \frac{\frac{(n-1)S_X^2}{(n-1)\sigma_X^2}}{\frac{(m-1)S_Y^2}{(m-1)\sigma_Y^2}} = \frac{\chi_{n-1}^2(m-1)}{\chi_{m-1}^2(n-1)} \sim F(n-1, m-1)$$

kjer je $F(\alpha, e)$ **Fischer-Snedecorjeva porazdelitev**.

$$f_{\delta,e}(x) = \frac{1}{B(\frac{\delta}{2}, \frac{e}{2})}(\frac{\delta}{e})^{\frac{\delta}{2}} x^{\frac{\delta}{2}-1}(1 + \frac{\delta}{e}x)^{-\frac{\delta+e}{2}} \quad \text{za} \quad x > 0.$$

2.6.1 Test hi-kvadrat

(Pearson)
Preizkus domneve o tipu porazdelitvenega zakona, torej $H_0(F = F_0) : H_1(F \neq$

F_0), kjer je F_0 dana porazdelitvena funkcija. Zalogo vrednosti slučajne spremenljivke X razdelimo na r razredov (disjunktno) $S_1, S_2 \dots S_r$, da je

$$p_k = P(X \in S_k \mid H_k) > 0 \; \forall k = 1, 2 \dots r$$

Potem je $\sum_{k=1}^{r} p_k = 1, \sum_{k=1}^{r} N_k = n$ ter $N_k \sim Bin(n, p_k)$ in $E(N_k) = p \cdot n_k$ kar je pričakovana vrednost za k-ti razred.
Pri velikem n ima testna statistika $\chi^2 = \sum_{k=1}^{r} \frac{(N_k - np_k)^2}{np_k}$ približno porazdelitev $\chi^2(r-1)$
Če χ^2 zavzame preveliko vrednost, hipotezo H_0 zavrnemo
$K_\alpha = [c_\alpha, \infty), P(\chi^2 > c_\alpha) = \alpha$

Opomba. Če so v testu χ^2 frekvence p_k odvedljivo odvisne od parametra θ, torej $p_k(\theta)$, potem ima statistika $\chi^2 = \sum_{k=1}^{r} \frac{(N_k - np_k(\hat{\theta}))^2}{np_k(\hat{\theta})}$ približno porazdelitev $\chi^2(r-2)$, kjer je $\hat{\theta}$ cenilka za parameter θ po metodi maksimalne zanesljivosti

Opomba. Računi bi bili drugačni če bi imeli $\lambda = 0.73$ podan na začetku ($\chi^2(r-1)$ vs. $\chi^2(r-2)$)

2.7 Linearna regresija

Definicija 2.14 (Linearni regresijski model)**.** Linearni regresijski model: $Y = a + bx + U$
Pri fiksnem $x \in \mathbb{R}$ predpostavljamo, da je $y = a + bx + U$, kjer sta a in b konstanti ter $U \sim N(0, \sigma)$ za nek pozitiven σ oz. $Y \sim N(a + bx, \sigma)$

Za različne vrednosti $x_1, x_2 \dots x_n$ dobimo slučajni vektor $(y_1, y_2 \dots y_n)$, kjer je $y_k \sim N(a + bx_k, \sigma)$
$y = a + bx$ je regresijska premica
y_k je vrednost za $Y_k, k = 1, 2 \dots n$
Radi bi ocenili a in b
Z metodo maksimalne zanesljivosti se dobi cenilki

$$\hat{b} = \frac{\sum_{k=1}^{n}(x_k - \overline{X})(x_k - \overline{Y})}{\sum_{k=0}^{n}(x_k - \overline{X})^2}$$

in

$$\hat{a} = \overline{Y} - \hat{b} \cdot \overline{X}$$

Če vpeljemo naslednje vsote

$$S_x := \sum_{k=1}^{n} x_k$$
$$S_Y := \sum_{k=1}^{n} y_k$$
$$S_{xx} := \sum_{k=1}^{n} x_k^2$$
$$S_{xY} := \sum_{k=1}^{n} x_k y_k$$

je števec

$$\sum_{k=1}^{n}(x_k - \overline{X})(y_k - \overline{Y}) = \sum_{k=1}^{n}(x_k y_k - \overline{X} y_k - x_k \overline{Y} + \overline{XY}) =$$
$$= S_{xx} - \frac{1}{n}\overline{X} S_y - \overline{Y} S_x + n\overline{XY} =$$
$$= S_{xY} - \frac{1}{n} S_x S_Y - \frac{1}{n} S_Y S_x + \frac{1}{n} S_y =$$
$$= S_{xY} - \frac{1}{n} S_x S_Y$$

in imenovalec

$$\sum_{k=1}^{n}(x_k - \overline{X})^2 = \sum_{k=1}^{n}(x_k^2 2 x_n \overline{X})^2 =$$
$$= S_{xx} - 2\overline{X} S_x + n\overline{X}^2 =$$
$$= S_{xx} - 2\frac{1}{n} S_x^2 + \frac{1}{n} S_x^2 =$$
$$= S_{xx} - \frac{1}{n} S_x^2$$

Torej je

$$\hat{b} = \frac{n S_{xY} - S_x S_y}{n S_{xx} - S_x^2}$$

in

$$\hat{a} = \frac{1}{n} S_Y + \hat{b} \frac{1}{n} S_x$$

Do cenilk $\hat{a}$ in $\hat{b}$ lahko pridemo po metodi najmanjših kvadratov: minimiziramo funkcijo

$$f(a, b) := \sum_{k=1}^{n} (y_k - (a - bx_k))^2$$

Porebna pogoja za minimum sta:

$$\begin{aligned} 0 = \frac{\partial f}{\partial a} &= -2 \sum_{k=1}^{n} (y_k - a - bx_k) = (-2)(S_Y - na - bS_x) \\ 0 = \frac{\partial f}{\partial b} &= -2 \sum_{k=1}^{n} x_k (y_k - a - bx_k) = (-2)(S_{xY} - aS_x - bS_{xx}) \end{aligned}$$

Torej

$$\begin{aligned} S_Y &= na + bS_x \quad / \cdot S_x \\ S_{xY} &= aS_x + bS_{xx} \quad \cdot n \end{aligned}$$

Enačbi odštejemo

$$\begin{aligned} S_x S_Y - nS_{xY} &= b(S_x^2 - nS_{xx}) \\ \implies b &= \frac{nS_{xx} - S_x S_y}{nS_{xx} - S_x^2} \\ a &= \frac{1}{n}(S_Y - bS_x) = \frac{1}{n} S_Y - b \frac{1}{n} S_x \end{aligned}$$

$$(X, Y) \sim N(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$$

$$E(Y \mid X = x) = \mu_y + \rho \frac{\sigma_y}{\sigma_x}(x - \mu_x) = \alpha + \beta x$$

kjer je $\beta = \rho\frac{\sigma_y}{\sigma_x}, \alpha = \mu_y - \beta\mu_x$

$$\beta = \rho\frac{\sigma_y}{\sigma_x} = \frac{K}{\sigma_x\sigma_y} \cdot \frac{\sigma_y}{\sigma_x} = \frac{K}{\sigma_x^2}$$

$$\hat{b} = \frac{\frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{X})(y_k - \overline{Y})}{\frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{X})^2}$$

števec: vzorčna kovarianca
imenovalec: vzorčna disperzija za x

$$\hat{a} = \overline{Y} - \beta\overline{X}$$

2.8 Testiranje zanesljivosti

Poseben primer testa χ^2, imenuje se prilagoditveni test (Goodness-of-Fit Test).
Ničelna hipoteza H_0: dogodka A in B sta neodvisna
Če je $p = P(A)$ in $q = P(B)$, imamo 4 razrede (kategorije):

kategorija	$A \cap B$	$A \cap B^C$	$A^C \cap B$	$A^C \cap B^C$
verjetnost	pq	$p(1-q)$	$(1-p)q$	$(1-p)(1-q)$

Če sta p in q znana parametra (običajno nista), uporabimo test χ^2 za $r = 4$:

$$\chi^2 = \frac{(N_{A\cap B} - npq)^2}{npq} + \cdots + \frac{(N_{A^C\cap B^C} - n(1-p)(1-q))^2}{n(1-p)(1-q)}$$

Kjer je $N_{A\cap B}$ opažena frekvenca dogodka $A \cap B$ in n velikost vzorca
Če H_0 velja, ima $\chi^2 \sim \chi^2(3)$ pri velikem n

$$K_\alpha = [c_\alpha, \infty)$$

Če je vzorčna vrednost za χ^2 na K_α, hipotezo H_0 zavrnemo
Običajno p in q nista znana parametra, zato ju ocenimo iz podatkov

Kontingenčna matrika:

	B	B^C
A	X_{11}	X_{12}
A^C	X_{21}	X_{22}

$$X_{11} + X_{12} + X_{21} + X_{22} = n$$

kjer je n velikost vzorca
Cenilki za p in q sta $\hat{p} := \frac{X_{11}+X_{12}}{n}$, $\hat{q} := \frac{X_{11}+X_{21}}{n}$
Statistika χ^2 je potem

$$\chi^2 = \frac{(X_{11} - n\hat{p}\hat{q})^2}{n\hat{p}\hat{q}} + \cdots + \frac{(X_{22} - n(1-\hat{p})(1-\hat{q}))^2}{n(1-\hat{p})(1-\hat{q})}$$

Izkaže se, da je $\chi^2 \sim \chi^2(1)$ za velik n

Opisani test lahko posplošimo na večje kontingenčne tabele:
Denimo, da 1. karakteristika določa r kategorij $A_1, A_2 \dots A_r$, 2. pa s kategorij $B_1, B_2 \dots B_s$.
Naj bo $p_i = P(A_i) i = 1, 2 \dots r$ in $q_j = P(B_j) j = 1, 2 \dots s$ $(\sum_{i=1}^{r} p_i = 1, \sum_{j=1}^{s} q_j = 1)$
Ničelna hipoteza H_0: A_i in B_j sta neodvisna za vsak i in vsak j
Vzorec velikosti n
Opažene frekvence:

	B_1	B_2	$\dots$	B_r	
A_1	X_{11}	X_{12}	$\dots$	X_{1s}	$n \cdot \hat{p_1}$
A_2	X_{21}	X_{22}	$\dots$	X_{2s}	$n \cdot \hat{p_2}$
$\vdots$	$\vdots$				$\vdots$
A_r	X_{r1}	X_{r2}	$\dots$	X_{rs}	$n \cdot \hat{p_r}$
	$n \cdot \hat{q_1}$	$n \cdot \hat{q_2}$	$\dots$	$n \cdot \hat{q_s}$	n

$$\hat{p_i} := \frac{1}{n} \sum_{j=1}^{s} x_{ij}$$
$$\hat{q_j} := \frac{1}{n} \sum_{i=1}^{r} x_{ij}$$

Cenilke za p_i in q_j:
Definiramo

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{(X_{ij} - n\hat{p_i}\hat{q_j})^2}{n\hat{p_i}\hat{q_j}}$$

Izkaže se, da je pri velikih n χ^2 pribliz"no porazdeljena $\chi^2((r-1)(s-1))$

Če je $n\hat{p}_i\hat{q}_j < 5$ za kakšna i in j, potem se priporoča, da se združi nekatere razrede

2.8.1 Teoretične osnove testa χ^2

Oglejmo si primer, ko je $r = 2$

	S_1	S_2
opažene frekvence	N	$n-N$
pričakovane frekvence	np	$n(1-p)$

$p = P(\text{prvi razred})$
$N \ldots$ število vrednosti vzorca, ki padejo v 1. razred S_1
Potem je

$$\chi^2 = \frac{(N-np)^2}{np} + \frac{(n-N-n(1-p))^2}{n(1-p)} =$$
$$= \frac{(N-np)^2}{np} + \frac{(N-np)^2}{n(1-p)} = \frac{(N-np)^2}{np(1-p)}((1-p)+p) =$$
$$= (\frac{N-np}{\sqrt{np(1-p)}})^2$$

Ker je N porazdeljena binomsko $Bin(n,p)$, je pri velikem n slučajna spremenljivka

$$\frac{N-E(N)}{\sigma(N)} = \frac{N-np}{\sqrt{np(1-p)}}$$

porazdeljena po $N(0,1)$ (Laplaceova formula oz. CLI)
Iz verjetnostnega dela vemo, da je kvadrat porazdelitve $N(0,1)$ porazdeljen po $\chi^2(1)$. Torej je χ^2 porazdeljena po $\chi^2(1)$ pri velikem n.
V splošnem primeru (pri poljubnem $r \in \mathbb{N}$) se χ^2 zapiše kot vsota $(r-1)$ kvadratov slučajnih spremenljivk, ki so porazdeljene po $N(0,1)$ in neodvisne. Ker za porazdelitev χ^2 velja $\chi^2(m) + \chi^2(n) \sim \chi^2(m+n)$ (za neodvisne slučajne spremenljivke), je potem χ^2 porazdeljena po $\chi^2(1+1+\cdots+1) = \chi^2(r-1)$

2.9 Test za neznan delež

Naj bo $X : \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$, kjer je p neznan parameter, ki bi ga radi testirali. Testiramo $H_0(p = p_0) : H_1(p \neq p_0)$, kjer je $p \in (0, 1)$ dano število. Vemo, da je $\overline{X}$ nepristranska cenilka za p. Po Laplaceovi formuli je

$$\overline{X} \approx N(p_0, \sqrt{\frac{p_0 q_0}{n}})$$

za velike n, če velja hipoteza H_0. Torej je

$$Z := \frac{\overline{X} - p_0}{\sqrt{p_0 q_0 / n}} \approx N(0, 1)$$

za velike n.
Pri dani stopnji značilnosti $\alpha > 0$ določimo $z_{\frac{\alpha}{2}}$, da je

$$P(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$$

$K_\alpha = (-\infty, -z_{\frac{\alpha}{2}}] \cup [z_{\frac{\alpha}{2}}, \infty)$
Če vzorčna vrednost za Z leži na kritičnem območju K_α, potem hipotezo zavrnemo

Za smiselnost testa je pomembno hipoteze formulirati pred analizo podatkov

2.10 Neparametrični testi

Doslej smo preizkušali hipoteze o neznanih parametrih v danih porazdelitvah (običajno smo privzeli normalno porazdelitev). To so parametrični testi.
Če na porazdelitev slučajne spremenljivke X ne moremo nič privzeti, potem lahko uporabimo neparametrične teste

2.10.1 Test z znaki

To je analog testa T
Na populaciji imamo 2 slučajni spremenljivki: X s porazdelitveno funkcijo F_X in Y s porazdelitveno funkcijo F_Y.
Obravnavamo 2 slučajna vektorja $(X_1, X_2 \dots X_m)$ in $(Y_1, Y_2 \dots Y_n)$, kjer je X_i in Y_i dobimo na istem elementu v populaciji.

Testiramo hipotezo $H_0(F_X = F_Y)$
Definiramo razlike

$$D_i := X_i - Y_i \; i = 1, 2 \dots n$$

Tukaj smo privzeli, da so vrednosti slučajnega vektorja $(D_1, D_2 \dots D_n)$ različne od 0, sicer jih izpustimo in zmanjšamo n. Če velja $H_0(F_X = F_Y)$, potem je $P(D_i > 0) = \frac{1}{2} = P(D_i < 0)$ za vsak $i = 1, 2 \dots n$
Naj bo S^+ število pozitivnih D_i-jev, S^- pa negativnih. Seveda je $S^+ + S^- = n$
Tedaj je $S^+ \sim Bin(n, \frac{1}{2})$, torej je

$$p_k = P(S^+ = k) = \binom{n}{k} 2^{-n} \; k = 0, 1 \dots n$$

Pri dani stopnji značilnosti $\alpha > 0$ je kritično območje

$$H_\alpha = \{k : k \leq k_\alpha \text{ ali } k \geq n - k_\alpha\}$$

kjer je k_α določen z zahtevama

$$\sum_{k=0}^{k_\alpha} p_k = P(S^+ \leq k_\alpha) \leq \frac{\alpha}{2}$$

in

$$\sum_{k=0}^{k_\alpha + 1} p_k = P(S^+ \leq k_\alpha + 1) > \frac{\alpha}{2}$$

Pri velikem n je S^+ približno normalno porazdeljen $N(\frac{n}{2}, \frac{\sqrt{n}}{2})$, $(\sqrt{npq} = \sqrt{n \cdot \frac{1}{2} \cdot \frac{1}{2}})$ torej je slučajna spremenljivka

$$Z := \frac{S^+ - \frac{n}{2}}{\frac{\sqrt{n}}{2}} = \frac{2S^+ - n}{\sqrt{n}}$$

približno normalno $N(0, 1)$

Slabost tega testa je, da gledamo samo predznak razlike in ne velikost

$$Z = \frac{2S^+ - n}{\sqrt{n}},\ n = 9$$

Čeprav n ni velik, izračunajmo velikost za Z

$$Z = \frac{2 \cdot 9 - 8}{\sqrt{9}} = \frac{7}{3} = 2.33$$

2.10.2 Inverzijski test

Wilcoxon-Mann-Whitney, 1945
X, Y naj imata porazdelitveni funkciji F_X in F_Y. Vzorca $(X_1, X_2 \dots X_m)$ in $(Y_1, Y_2 \dots Y_n)$ sta neodvisna in $m \leq n$ (če to ni res zamenjamo vlogi X in Y). Testirajmo hipotezo $H_0(F_X = F_Y)$. Vzorčne vrednosti vzorcev $X_1, X_2 \dots X_m, Y_1, Y_2 \dots Y_n$ razvrstimo po velikosti $z_1 \leq z_2 \leq \dots \leq z_{m+n}$ (zapomnimo si ali je iz X ali Y). Pripišimo mesta (Range)

$$R_i = rang(x_i) = h,\ \text{če je } z_k = x_i$$

Slučajna spremenljivka $r = R_1 + \dots + R_m$ ima vrednosti med $\frac{m(m+1)}{2}$ in $mn + \frac{m(m+1)}{2}$
Vrednost $\frac{m(m+1)}{2}$ dobimo, če so X_i na začetku zaporedja $\{z_m\} : 1 + 2 + \dots + m = \frac{m(m+1)}{2}$
Največjo vrednost pa dobimo, če so X_i na koncu zaporedja: $Y_1, Y_2 \dots Y_n, X_1, X_2 \dots X_m$:

$$(n+1) + (n+2) + \dots + (n+m) = n \cdot m + \frac{m(m+1)}{2}$$

Če velja $H_0(F_X = F_Y)$ in $m + n \geq 0, m, n \geq 4$, potem smemo privzeti, da je V približno normalno porazdeljena

$$V \sim N(\frac{(m+n+1) \cdot m}{2}, \sqrt{\frac{mn(m+n+1)}{12}})$$

oz.

$$Z := \sqrt{\frac{3}{mn(m+n+1)}}(2V - m(m+n+1)) \sim N(0,1)$$

Definicija 2.15 (Inverzija)**.** Inverzija med x_i in y_j se pojavi, če ima y_j manjši rang kot x_i

Zaporedje $x_1, x_2 \dots x_m, y_1 \dots y_n$ nima inverzije
Zaporedje $x_1 \dots x_{m-1}, y_1, x_m, y_2 \dots y_n$ ima eno inverzijo
Naj bo U število vseh inverzij
Vsaka inverzija, ki jo naredimo, poveča število rangov V za 1. Ker je pri

$$U = 0, V = \frac{m(m+1)}{2}$$

je

$$V = U + \frac{m(m+1)}{2}$$

Inverzijski test je neparametrični analog primerjalnega Studentovaga testa
Inverzijski test "gleda" samo urejenost in ne velikost podatkov