

Verjetnost in Statistika - Flashcards

Osnovni pojmi verjetnosti

Q: Kaj je dogodek? A: Izvajamo poskus, opazujemo nek pojav, ki se lahko zgodi in ga imenujemo dogodek.

Q: Kaj je relativna frekvenca? A: Če poskus ponovimo n -krat in opazujemo dogodek A , je relativna frekvenca $f_n(A) = K_n(A)/n \in [0, 1]$, kjer je $K_n(A)$ število ponovitev, pri katerih se je dogodek A zgodil.

Q: Naštej tri definicije verjetnosti A:

1. Statistična: $P(A) := p$ (kjer p je limita relativnih frekvenc)
2. Klasična: $P(A) = \text{število ugodnih izidov} / \text{število vseh izidov}$ (pri enakih možnostih)
3. Geometrijska: za neskončno izidov

Q: Kaj je σ -algebra? A: Neprazna družina podmnožic F v Ω je σ -algebra, če:

1. $A \in F \Rightarrow A^c \in F$ (zaprtost za komplement)
2. $A_1, A_2, \dots \in F \Rightarrow \bigcup_{i=1}^{\infty} A_i \in F$ (zaprtost za števne unije)

Q: Aksiomatična definicija verjetnosti (Kolmogorov) A: Verjetnost $P: F \rightarrow \mathbb{R}$ z lastnostmi:

1. $P(A) \geq 0 \forall A \in F$
2. $P(\Omega) = 1$
3. Za paroma nezdružljive dogodke: $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Pogojna verjetnost in neodvisnost

Q: Formula za pogojno verjetnost A: $P(A|B) = P(A \cap B)/P(B)$, kjer je $P(B) > 0$

Q: Formula o popolni verjetnosti A: $P(A) = \sum_i P(H_i) \cdot P(A|H_i)$, kjer je $\{H_i\}$ popoln sistem dogodkov

Q: Bayesova formula A: $P(H_k|A) = [P(H_k) \cdot P(A|H_k)] / [\sum_i P(H_i) \cdot P(A|H_i)]$

Q: Kdaj sta dogodka neodvisna? A: A in B sta neodvisna, če $P(A \cap B) = P(A) \cdot P(B)$

Q: Bernoullijeva formula A: $P_n(k) = \binom{n}{k} p^k q^{n-k}$, kjer je $q = 1-p$

Aproksimacijske formule

Q: Poissonova aproksimacija A: Če je n velik in k majhen: $P_n(k) \approx \lambda^k e^{-\lambda}/k!$, kjer $\lambda = np$

Q: Laplaceova lokalna formula A: $P_n(k) \approx 1/\sqrt{2\pi npq}) \cdot e^{-(k-np)^2/(2npq)}$

Q: Laplaceova integralska formula A: $P_n(k_1, k_2) \approx \Phi((k_2-np)/\sqrt{npq}) - \Phi((k_1-np)/\sqrt{npq})$

Q: Verjetnostni integral $\Phi(x)$ A: $\Phi(x) = 1/\sqrt{2\pi} \int_0^x e^{-t^2/2} dt$, z lastnostjo $\lim_{x \rightarrow \infty} \Phi(x) = 1/2$

Slučajne spremenljivke

Q: Kaj je slučajna spremenljivka? A: Funkcija $X: \Omega \rightarrow \mathbb{R}$ z lastnostjo, da je $\{\omega \in \Omega : X(\omega) \leq x\}$ dogodek za vse $x \in \mathbb{R}$

Q: Porazdelitvena funkcija A: $F_X(x) = P(X \leq x)$

Q: Lastnosti porazdelitvene funkcije A:

1. $0 \leq F(x) \leq 1$
2. F je naraščajoča
3. $\lim_{x \rightarrow \infty} F(x) = 1$, $\lim_{x \rightarrow -\infty} F(x) = 0$
4. F je zvezna z desne

Q: Diskretna slučajna spremenljivka A: Slučajna spremenljivka z končno ali števno zalogo vrednosti. Določena z verjetnostno funkcijo $p_n = P(X = x_n)$.

Q: Zvezna slučajna spremenljivka A: Obstaja gostota $p_X(x) \geq 0$, da je $F_X(x) = \int_{-\infty}^x p_X(t) dt$

Diskretne porazdelitve

Q: Binomska porazdelitev $\text{Bin}(n, p)$ A: $p_k = \binom{n}{k} p^k q^{n-k}$, $k = 0, 1, \dots, n$

Q: Poissonova porazdelitev $\text{Poi}(\lambda)$ A: $p_k = \lambda^k e^{-\lambda} / k!$, $k = 0, 1, 2, \dots$

Q: Geometrijska porazdelitev $\text{Geo}(p)$ A: $p_k = p \cdot q^{k-1}$, $k = 1, 2, \dots$ (prvi uspeh v k -ti ponovitvi)

Q: Pascalova porazdelitev $\text{Pas}(m, p)$ A: $p_k = \binom{k-1}{m-1} p^m q^{k-m}$, $k = m, m+1, \dots$ (m -ti uspeh v k -ti ponovitvi)

Q: Hipergeometrijska porazdelitev $\text{Hip}(n; M, N)$ A: $p_k = [\binom{M}{k} \binom{N-M}{n-k}] / \binom{N}{n}$

Zvezne porazdelitve

Q: Enakomerna porazdelitev na $[a, b]$ A: $p(x) = 1/(b-a)$ za $a < x < b$, sicer 0

Q: Normalna porazdelitev $N(\mu, \sigma)$ A: $p(x) = 1/(\sigma\sqrt{2\pi}) \cdot e^{-1/2((x-\mu)/\sigma)^2}$

Q: Eksponentna porazdelitev $\text{Exp}(\lambda)$ A: $p(x) = \lambda e^{-\lambda x}$ za $x \geq 0$, sicer 0

Q: Gama porazdelitev $\Gamma(b, c)$ A: $p(x) = (c^b / \Gamma(b)) \cdot x^{b-1} e^{-cx}$ za $x > 0$

Q: Hi-kvadrat porazdelitev $\chi^2(n)$ A: $\chi^2(n) = \Gamma(n/2, 1/2)$

Matematično upanje in disperzija

Q: Matematično upanje diskretne slučajne spremenljivke A: $E(X) = \sum_i x_i \cdot p_i$ (če vrsta absolutno konvergira)

Q: Matematično upanje zvezne slučajne spremenljivke A: $E(X) = \int_{-\infty}^{\infty} x \cdot p_X(x) dx$ (če integral absolutno konvergira)

Q: Linearnost matematičnega upanja A: $E(aX + bY) = aE(X) + bE(Y)$

Q: Disperzija A: $D(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2$

Q: Lastnosti disperzije A:

- $D(X) \geq 0$
- $D(aX) = a^2 D(X)$
- $D(X+Y) = D(X) + D(Y) + 2\text{Cov}(X,Y)$

Q: Kovarianca A: $\text{Cov}(X,Y) = E((X-E(X))(Y-E(Y))) = E(XY) - E(X)E(Y)$

Q: Korelacijski koeficient A: $r(X,Y) = \text{Cov}(X,Y)/(\sigma(X)\sigma(Y)) \in [-1,1]$

Neodvisnost slučajnih spremenljivk

Q: Neodvisnost slučajnih spremenljivk A: X in Y sta neodvisni, če $F_{\{X,Y\}}(x,y) = F_X(x) \cdot F_Y(y)$

Q: Neodvisnost za diskretne spremenljivke A: X,Y neodvisni $\Leftrightarrow p_{\{ij\}} = p_i \cdot q_j$

Q: Neodvisnost za zvezne spremenljivke A: X,Y neodvisni $\Leftrightarrow p_{\{X,Y\}}(x,y) = p_X(x) \cdot p_Y(y)$

Q: Lastnost za neodvisne spremenljivke A: Če sta X,Y neodvisni, potem $E(XY) = E(X)E(Y)$

Momenti in karakteristične funkcije

Q: Začetni moment reda k A: $z_k = E(X^k)$

Q: Centralni moment reda k A: $m_k = E((X - E(X))^k)$

Q: Rodovna funkcija A: $G_X(s) = E(s^X) = \sum_{k=0}^{\infty} p_k s^k$ (za nenegativne celoštevilске vrednosti)

Q: Momentno rodovna funkcija A: $M_X(t) = E(e^{tX})$

Q: Zvezo med rodovnima funkcijama A: $M_X(t) = G_X(e^t)$

Zakoni velikih števil

Q: Šibki zakon velikih števil (ŠZVŠ) A: $(S_n - E(S_n))/n \rightarrow 0$ verjetnostno

Q: Krepi zakon velikih števil (KZVŠ) A: $P(\lim_{n \rightarrow \infty} (S_n - E(S_n))/n = 0) = 1$

Q: Neenakost Čebiševa A: $P(|X - E(X)| \geq \varepsilon) \leq D(X)/\varepsilon^2$

Q: Centralni limitni izrek A: $(S_n - E(S_n))/\sigma(S_n) \rightarrow N(0,1)$ po porazdelitvi

Statistika - osnovni pojmi

Q: Populacija A: Končna ali neskončna množica elementov, pri katerih merimo neko količino

Q: Vzorec A: Slučajni vektor (X_1, \dots, X_n) , kjer so komponente enako porazdeljene kot X in neodvisne

Q: Vzorčna statistika A: Simetrična funkcija vzorca $Y = g(X_1, \dots, X_n)$

Q: Nepristranskost cenilke A: Cenilka Y je nepristranska za parameter ξ , če $E(Y) = \xi$

Q: Doslednost cenilke A: Cenilka Y_n je dosledna, če $Y_n \rightarrow \xi$ verjetnostno

Vzorčne statistike

Q: Vzorčno povprečje A: $\bar{X} = (X_1 + \dots + X_n)/n$; $E(\bar{X}) = \mu$, $D(\bar{X}) = \sigma^2/n$

Q: Vzorčna disperzija A: $s_0^2 = (1/n)\sum_i (X_i - \bar{X})^2$ (pristranska)

Q: Popravljen vzorčna disperzija A: $s^2 = (1/(n-1))\sum_i (X_i - \bar{X})^2$ (nepristranska)

Q: Studentova t-porazdelitev A: $T = (\bar{X} - \mu)/(s/\sqrt{n}) \sim t(n-1)$, če $X \sim N(\mu, \sigma)$

Testiranje hipotez

Q: Ničelna in alternativna hipoteza A: H_0 - hipoteza, ki jo testiramo; H_1 - alternativna hipoteza

Q: Stopnja značilnosti A: α - verjetnost zavrnitve pravilne hipoteze (tip I napaka)

Q: P-vrednost A: Najmanjša stopnja značilnosti, pri kateri še lahko zavrremo hipotezo

Q: Z-test A: Za $X \sim N(\mu, \sigma)$, σ znan: $Z = (\bar{X} - \mu_0)/(\sigma/\sqrt{n}) \sim N(0,1)$

Q: t-test A: Za $X \sim N(\mu, \sigma)$, σ neznan: $T = (\bar{X} - \mu_0)/(s/\sqrt{n}) \sim t(n-1)$

Q: F-test A: Za primerjavo varianc: $F = s_1^2/s_2^2 \sim F(n_1-1, n_2-1)$

Hi-kvadrat test

Q: Hi-kvadrat test prilagojenosti A: $\chi^2 = \sum_i (N_i - np_i)^2/(np_i) \sim \chi^2(r-1)$

Q: Test neodvisnosti A: $\chi^2 = \sum_{ij} (X_{ij} - n \cdot \hat{p}_i \cdot \hat{q}_j)^2/(n \cdot \hat{p}_i \cdot \hat{q}_j) \sim \chi^2((r-1)(s-1))$

Regresijska analiza

Q: Linearni regresijski model A: $Y = a + bx + U$, kjer $U \sim N(0, \sigma)$

Q: Cenilki po metodi najmanjših kvadratov A:

- $\hat{b} = (nS_{xy} - S_x S_y)/(nS_{xx} - S_x^2)$

- $\hat{a} = \bar{Y} - \hat{b}\bar{X}$

Neparametrični testi

Q: Test z znaki A: $S^+ \sim \text{Bin}(n, 1/2)$; testira $H_0: P(D_i > 0) = 1/2$

Q: Wilcoxon-Mann-Whitney test A: Primerja dve porazdelitvi na osnovi rangov; $V = R_1 + \dots + R_m$

Q: Prednosti neparametričnih testov A: Ne zahtevajo predpostavk o porazdelitvi; robustni na osamelce