Diskrétní pravděpodobnostní prostor (Ω, P): Ω – libovolná konečná/ spočetná množina, $\omega \in \Omega$ – elementární jev
$\Omega = \cup\omega$ – jistý jev, $\emptyset$ – nemožný jev; P: $\Omega \to [0, 1]$, $P(\Omega)=1$, $P(\emptyset)=0$; $A \subseteq \Omega$ - náhodné jevy
Podmíněná pravděpodobnost (Ω, P): $P(A|B)=P(A \cap B)/P(B)$, $P(B)>0$ – $P(A)$, pokud nastane i B
Jev A nezávisí na B, pokud $P(A|B)=P(A)$ => jevy A, B, nezávislé, pokud $P(A \cap B)=P(A)P(B)$
Jevy $A_i \in \Omega$ jsou vzájemně nezávislé, pokud $\forall k$ $1<k\leq n$ $P(A_{i\_1} \cap ... \cap A_{i\_n})=\prod P(A_{i\_j})$
Jevy $A_i \in \Omega$ jsou vzájemně nezávislé po k=2, 3, …, n, jestliže $\forall A_{i\_j}$ $P(A_{i1} \cap ... \cap A_{i\_n})=\prod P(A_{i\_j})$
Věta o úplné pravděpodobnosti: jestliže jevy H_i tvoří Ω a navzájem mají $\emptyset$ průnik, pak $P(A)=\sum P(A|H_i)P(H_i)$
Bayersova věta: $H_i, A \in \Omega$, $\cup H_i=\Omega$, $H_i \cap H_j=\emptyset$. $P(H_i|A)=P(A \cap H_i)/P(A)=P(A|H_i)P(H_i)$ / $\sum P(A|H_j)P(H_j)$
Náhodná veličina: $X:(\Omega, \mathcal{A}, P) \to (R_1, \mathcal{B}_1)$; X nabývá x_1; $[X=x_i]=\{\omega|X(\omega)=x_i\}$; $p_i=P(X=x_i)=P\{\omega|X(\omega)=x_i\}$; $\sum p_i=1$
Rozdělení pravděpobnosti v X: $\{x_i$ nad $p_i\}$ kde x_i jsou hodnoty, jichž lze nabýt a p_i příslušné pravděpobnosti
Distribuční fce: $F(x)=P(X\leq x)$ – neklesající, $\lim_{x\to-\infty}F(x)=0$, $\lim_{x\to\infty}F(x)=1$; zprava spojitá; jednoznačně určuje pst
Střední hodnota dnv: $EX=\sum X(\omega)P(\omega)=\sum x_i p_i$; $Ea=a$; $E(a+X)=a+EX$; $E(a+bX+cY)=a+bEX+cEY$; $Eg(X)=\sum p_i g(x_i)$
Složky vektoru X_1, …, X_n jsou nezávislé, pokud sdružené rozdělení $P(X_1\leq x_1, \dots, X_n\leq x_n)=\prod P(X_i\leq x_i)$
Věta o konvoluci: $X\sim(x_i$ nad $p_i)$, $Y\sim(y_j$ nad $q_j)$ a X a Y nezávislé, pak $Z=X+Y\sim(z_k$ nad $r_k)$, kde r_k je konvolucí z p_i, q_j
Vytvořující funkcí k n. v. X nazveme $A(x)=p_i x^i$; $A'(1)=EX$, $A'(0)=p_1$
Vytvořující funkce jednoznačně definuje distribuci (rozdělení) n. v. X – $p_n=A^{(n)}(0)$ / n!
Nechť X a Y jsou n. v. ne nutně nezávislé, potom $E(X+Y)=EX+EY$
Nechť X je n. v., potom rozptylem nazveme $varX=E(X-EX)^2$; $var(c)=0$, $var(c+X)=varX$, $var(cX)=c^2 varX$ (blbé)
$var(X+Y)=varX+varY+2E(X-EX)(Y-EY)$; $E(X-EX)(Y-EY)=cov(X,Y)=E(XY)-EX\cdot EY$, pro X, Y nezávislé =0
Čebyševova nerovnost: n. v. X, $varX$ existuje, $\varepsilon>0$. Pak platí $P(|X-EX|\geq\varepsilon) \leq varX / \varepsilon^2$ Dk: $varX=\sum(x_i-\mu)^2p_i=...$
X_1, ..., X_n nezávislé, $EX_i=\mu$, $varX_i=\sigma^2$, $\overline{X}_n=(X_1+\dots+X_n)/n$ (výběrový průměr), $E\,\overline{X}_n=\mu$, $var\,\overline{X}_n=\sigma^2/n$
X_1, …, X_n konverguje k n. v. X v pravdě $X_n\to X$, pokud $\forall\varepsilon>0$ platí $limP(|X_n-X|\geq\varepsilon)=0$
Slabý zákon velkých čísel: X_1, ..., X_n nezávislé, $EX_i=\mu$, $varX_i=\sigma^2$, $\overline{X}_n=(X_1+\dots+X_n)/n$ platí $\overline{X}_n\to\mu$ Dk: Čebyš.
n.v. X má spojité rozdělení s hustotou f když $\forall a\leq b\in R$: $P(X\in(a, b))=\int_a^b f(x)$; $F(x)=\int_{-\infty}^x f(t)$; $f(x)=F'(x)$; $EX=\int_{-\infty}^{\infty} xf(x)$
Centrální limitní věta 1: X_i stejně rozdělené nezávislé n.v. $Ex_i=\mu$ $varX=\sigma^2$: $P((\Sigma x_i-n\mu)/\sqrt{(n\sigma^2)}<x)\approx F(x)$ z $N(0, 1)$
X … $f(x)$, $F(x)$, potom kvantilovou funkcí n. v. X rozumíme $F^{-1}(t)$, $t\in(0, 1)$, pokud existuje
Zobecněná kvantilová funkce – $\sup\{x|F(x)<t\}$ $\inf\{x|F(x)>t\}$
Testování hypotéz - P(zamítnu H | H platí) – chyba prvního druhu; P(H přijmu | H neplatí) – chyba druhého dr.
X_1, …, X_n n. v. a $TR_n\to R_p$, $p\geq1$, potom $T(X_1, \dots, X_n)$ nazveme statistikou, je zřejmé, že je také náhodná veličina
Je-li T_n odhad parametru Θ a $ET_n=\Theta$, říkáme, že odhad je nestranný/ nevychýlený
Odhad s nejmenším rozptylem v dané třídě odhadů se nazývá nejlepší
konzistence: $P(|T_n-\Theta|>\varepsilon)\to0$; robustnost: necitlovost na změnu předpokladů
Metoda maximální věrohodnosti: hledám pro jaké Θ to je nejpravděpodobnější arg max $\prod p(X_i=x_i)$
n-tice X_1, …, X_n náhodný výběr; $\underline{X}_n=1/n\ \Sigma x_i$ výběrový průměr; $S_n^2=1/(n-1)\ \Sigma(X_i-\underline{X}_i)^2$ výběrový rozptyl
Intervalový odhad Θ o spolehlivosti 1-α $P(L, \Theta, P)=1-\alpha$; X_β takové, že $F(X_\beta)=\beta$ je β-kvantil, F spojitá...
Oboustrané odhady $(\underline{X}_n - \mu_{1-\alpha/2}\,\sigma/\sqrt{n},\ \underline{X}_n + \mu_{1-\alpha/2}\,\sigma/\sqrt{n})$; $(\underline{X}_n \mp t_{1-\alpha/2,\,n-1}\,S_n/\sqrt{n})$; $(S_n^2(n-1)/\chi_{1-\alpha/2,\,n-1}^2;\ S_n^2(n-1)/\chi_{\alpha/2,\,n-1}^2)$
Kritický obor – interval, kde hypotézu zamítnu; hladina testu – pravděpodobnost chyby prvního druhu
Síla testu – doplněk k pravděpodobnosti chyby druhého druhu
Dvojvýběrový test $|\underline{X}_n-\underline{Y}_m| > \mu_{1-\alpha/2}\,\sigma\,\sqrt{(n+m)/nm}$; $|\underline{X}_n-\underline{Y}_m| > \mu_{1-\alpha/2,\,n+m-2}\,S^*\,\sqrt{(n+m)/nm}$; $F_{1-\alpha/2,\,n-1,\,m-1}<S_n^2/S_m^2<F_{\alpha/2,\,n-1,\,m-1}$
$S^{*2}=1/(n+m-2)\ [(n-1)S_{1,\,n}^2+(m-1)S_{1,\,m}^2]$
Párový test (Y_i, Z_i), z nich X_i a tu zkoumám...
Regrese zkoumá závislost veličin na jiných. O chybě předpokládáme $N(0, \sigma^2)$; arg $\min\Sigma p(Y_i-a-bX_i)$ (…) rezidua
Náhodná veličina se řídí Alternativním rozdělením $Alt(p)$, pokud nabývá 0, 1 s pstí $1-p=q$ a p; $EX=p$ $varX=pq$
N. v. se řídí Binomický rozdělením $Bi(n, p)$, nabývá-li 1, 2, …, n s pstí (n nad k)$p^k(1-p)^{n-k}$; $EX=np$ $varX=npq$
N. v. se řídí Geometrickým rozdělením $Ge(p)$: pq^{n-1} (čekání na 1. zdar); $EX=q/p$ $varX=q/p^2$
N. v. se řídí Negativním binomickým rozdělením $NBi(r, p)$: (n-1 nad r-1)$p^rq^{n-r}=\sum Ge_i(p)$; $EX=rq/p$ $varX=rq/p^2$
N. v. se řídí Hypergeometrickým rozdělením $HGe(A, B, n)$: (A nad k)(B nad $n-k$) / ($A+B$ nad n) (# tažených)
$EX=nA/(A+B)$ $varX=n\ A/(A+B)\ B/(A+B)\ (A+B-n)/(A+B-1)$
N. v. se řídí Rovnoměrným rozdělením $R(n)$, nabývá-li hodnot z $\{1, \dots, n\}$ s pravděpobnosti $1/n$
N. v. se řídí Poissonovým rozdělením $Po(\lambda)$, nabývá-li hodnot $\{0, 1, 2, \dots\}$ s psti $e^{-\lambda}\lambda^i/i!$; $EX=\lambda$ $varX=\lambda$
N. v. se řídí Multinomickým rozdělením $M(k, n_i, p_i)$ $n!$ / $n_0!n_1!...n_k!$ $p_0^{n\_0}p_1^{n\_1}\dots p_k^{n\_k}$ (pro k=1 ~ Bi)
$X\sim U(a, b)$ Rovnoměrné na (a, b) – tam stejné, jinde 0; $EX=(a+b)/2$; $varX=(b-a)^2/12$; $\alpha X+\beta\sim U(\alpha a+\beta, \alpha b+\beta)$
$X\sim N(\mu, \sigma^2)$ Normální rozdělení – $f(x)=1/\sqrt{2\pi\sigma^2}\ e^{-(x-\mu)\wedge2\ /\ 2\sigma\wedge2}$; $EX=\mu$; $varX=\sigma^2$; $N(0, 1)$ standartní/ normalizované
$\alpha X+\beta\sim N(\alpha\mu+\beta, \alpha^2\sigma^2)$; $X_1+X_2\sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$ (z věty o konvoluci)
$X\sim Exp(\lambda)$ Exponenciální – $f(x)=\lambda e^{-\lambda x}$, $x>0$; $f(x)=0$, $x<0$; $F(x)=1-e^{-\lambda x}$; $EX=1/\lambda$; $varX=1/\lambda^2$; $\alpha X\sim Exp(\lambda/\alpha)$
„rozdělení bez paměti“ – $P(X<t+\Delta t|X>t)=P(X<\Delta t)$; $|.X.|\sim Ge(1-e^{\lambda})$; nezávislá X_1, X_2 – $\min\{X_1, X_2\}\sim Exp(\lambda_1+\lambda_2)$