Pravděpodobnost a **st**atistika
Výpisky z cvičení Ondřeje Chocholy

Jan Štětina

29. listopadu 2009

Cvičení 30.9.2009

Úloha: Máme 4 kostky. $\Omega = \{a, b, c, d\}$, $|\Omega| = 6^4$
$|A| = 6 \cdot 5 \cdot 4 \cdot 3 \qquad P(A) = \frac{6 \cdot 5 \cdot 4 \cdot 3}{6^4}$
Nanejvýš l kostek:
m... maximum
$A = [m \leq l], \qquad [m = l] = [m \leq l] \setminus [m \leq l - 1], \qquad P(A) = (\frac{l}{6})^4$

Úloha: Jaká je pravděpodobnost, že ve skupině n lidí existuje dvojice s narozeninami ve stejný den? (rok = 365 dní)

Řešení: Jev $A \in \Omega \qquad \Omega \setminus A = A^C$ doplňkový jev A
Zde je snazší zjistit $P(A^C) = 1 - P(A)$.
$|\Omega| = 365^n, \qquad |A| = \frac{365!}{(365-n)!}, \qquad P(A^C) = \frac{\frac{365!}{(365-n)!}}{365^n}$

Úloha: Mějme N jedinců $\{1, ..., N\}$. Z nich vybíráme n: $\{s_1, ..., s_n\}$. Jaká je pravděpodobnost, že bude vybrán i-tý jedinec?

Řešení: $|\Omega| = \binom{N}{n} = \frac{N!}{(N-n)! \cdot n!}, \qquad |A| = \binom{N-1}{n-1}, \qquad P(A) = \frac{\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N}$

Jaká je pravděpodobnost, že budou vybráni oba i-tý a j-tý jedinec?

Řešení: $|\Omega| = \binom{N}{n}, \qquad |A_{ij}| = \binom{N-2}{n-2}$

Jaká je pravděpodobnost, že bude vybrán i-tý nebo j-tý jedinec?

Řešení: $A_{i \vee j} = A_i \cup A_j, \qquad P(A_{i \vee j}) = P(A_i \cup A_j) = P(A_i) + P(A_j) - P(A_{ij})$

Cvičení 7.10.2009

Pojem: *Maxwellovo schéma.*
Máme r rozlišitelných prvků, n přihrádek.
$r/n \xrightarrow{n \to \infty} \lambda; \qquad p_k = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$

Úloha: Mějme obec o $r = 500$ lidech, $n = 365$ dní v roce. $\lambda = \frac{500}{365}$. Jaká je pravděpodobnost, že dva lidé mají narozeniny ve stejný den?

Řešení: $p_0 \doteq 0.25$

Úloha: B_i = i-tá přihrádka je prázdná.

Řešení:

$$P(\bigcup_{i=1}^{n} B_i) = \sum_{i=1}^{n} P(B_i) - \sum_{1\leq i\leq j\leq n} P(B_i \cap B_j) + \sum_{1\leq i\leq j\leq k\leq n} P(B_i \cap B_j \cap B_k) - ... + (-1)^{n+1} \cdot P(\bigcap_{i=1}^{n} B_i)$$

$P(B_i) = \frac{(n-1)^r}{n^r} \ldots \frac{\text{(všechny ostatní možnosti kromě } B_i\text{)}}{\text{(všechny možnosti)}}$; $\qquad P(B_i \cap B_j) = \frac{(n-2)^r}{n^r}\ \frac{\text{(ostatní kromě } B_i, B_j\text{)}}{\text{(všechny možnosti)}}$

$P(\bigcup_i B_i) = n \cdot \frac{(n-1)^r}{n^r} - \binom{n}{2} \cdot \frac{(n-2)^r}{n^r} + \binom{n}{3} \cdot \frac{(n-3)^r}{n^r} - ...$

Pojem: *Bose-Einstein*

Máme r nerozlišitelných prvků, n přihrádek.
$(s_1, ..., s_r)$, $s_i \in \{1, ..., n\}$ jsou prvky, $(t_1, ..., t_n)$, $\sum_{i=1}^{n} t_i = r$ jsou přihrádky.

Úloha: $\underbrace{\cdot\ \cdot\ \cdot | \cdot\ \cdot | \cdot | \cdot\ \cdot\ \cdot\ \cdot\ \cdot | \cdot\ \cdot}_{n \text{ přihrádek} \Rightarrow n-1 \text{ příček}}$ Kolika způsoby lze rozdělit prvky a příčky?

Řešení: $\underbrace{\binom{r+n-1}{r} = \binom{r+n-1}{n-1}}_{\text{symetrie kombinačních čísel}} = |\Omega|$

Pravděpodobnost, že v přihrádce je k prvků.

Řešení: $p_k = \frac{\binom{r-k+n-2}{r-k}}{\binom{r+n-1}{r}}\ \frac{\forall \text{ ostatní než } k \text{ prvků}}{\forall \text{ možnosti}}$

Pravděpodobnost, že neexistuje prázdná přihrádka = $1 - p_0$ (za předpokladu $r \geq n$).

Řešení: $p_0 = \frac{\binom{(r-n)+n-1}{n-1}}{\binom{r+n-1}{n-1}}\ \frac{\text{v každé přihrádce alespoň jeden prvek} \Rightarrow \text{zbývá } r-n \text{ prvků}}{\text{všechny možnosti}}$

Pojem: *Geometrická pravděpodobnost.*

Obrázek
$P(A) = \frac{\mathcal{U}(A)}{\mathcal{Y}(\Omega)}$

Úloha: Mějme přístav, každý den připlují dvě lodi v náhodný čas, jedna setrvá hodinu, druhá dvě hodiny. Jaká je pravděpodobnost, že se nebudou blokovat?

Řešení: Obrázek

$p_{\text{blok}} = \frac{\frac{1}{2}\cdot(22^2+23^2)}{24^2}\ \frac{\text{neblokování}}{\text{všechny časy}}$
$p_A = 1 - p_{\text{blok}}$

Úloha: $x, y \in [0, 1]$, chceme pravděpodobnost $\{^{x+y<1}_{x\cdot y<0.09}$

Řešení: Obrázky

$S_0 + S_1 + S_3 = 0.1$
$y = 1 - x; \quad y = \frac{9}{100}x \quad \rightsquigarrow \quad \frac{9}{100} \cdot \int_0 .1^0 .9 \frac{1}{x} \cdot dx = \frac{9}{100} \cdot [\ln x]_0 .1^0 .9 \Rightarrow S_2 = \frac{9}{100} \cdot \ln 9$
$P(A) = S = 0.1 + 0.O9 \cdot \ln 9$

Pojem: *Podmíněná pravděpodobnost.*

$A \quad \rightsquigarrow P(A); \qquad B \quad \rightsquigarrow P(A|B) = \frac{P(A\cap B)}{P(B)}$

Příklad: $P(A) = 0.5, P(B) = 0.3$. Pak $P(B|A^C) = \frac{0.3}{0.5} = 0.6$

Pojem: *Úplný systém jevů.*
Nechť $(A_i)_{i=1}^{\infty}$, $\forall i \neq j : A_i \cap A_j = \emptyset$, $\bigcup_{i=1}^{\infty} A_i = \Omega$, $P(A_i) > 0$, pak $P(B) = \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)$

Úloha: Semifinále: $P(K > F) = 0.7, P(> R) = 0.55$, finále: $P(> K) = 0.6, P(> F) = 0.75$.

Cvičení 14.10.2009

Opakování: Jevy A, B jsou nezávislé, pokud $P(A \cap B) = P(A) \cdot P(B)$
$P(B|A) = \frac{P(A \cap B)}{P(A)} \underbrace{=}_{\text{nez.}} = P(B)$
$(A_i)_{i=1}^{\infty}$ nezávislé $\forall n \forall (j_1, ..., j_n) : P(\bigcap_{k=1}^{n} A_{j_k}) = \prod_{k=1}^{n} P(A_{j_k})$

Úloha: Mějme nezávislé jevy A, B, C tž. $P(A) = P(B) = P(C) = 0.1$. Chceme $P(A \cup B \cup C)$

Řešení: $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) = 3 \cdot 0.1 - 3 \cdot 00.1 + 0.001$

Úloha: Mějme lovce A, B, C, kteří zároveň vystřelí na kance k. Pravděpodobnosti zásahu jsou $P(A) = 0.2, P(B) = 0.4, P(C) = 0.6$. Jaká je pravděpodobnost zásahu k jednotlivými lovci právě jednou kulkou?

Řešení: $P(\text{lovec}|[1]) \rightsquigarrow$ nejdříve potřebujeme $P([1]) = P(A \cap B^C \cap C^C) + P(A^C \cap B \cap C^C) + P(A^C + B^C + C) = P(A) \cdot P(B^C) \cdot P(C^C) + P(A^C) \cdot P(B) \cdot P(C^C) + P(A^C) \cdot P(B^C) \cdot P(C)$
$P(A|[1]) = \frac{P(A \cap [1])}{P([1])} = P(A \cap B^C \cap C^C) = P(A) \cdot P(B^C) \cdot P(C^C)$, etc.

Opakování: Bayesova věta.
Mějme $(A_i)_{i=1}^{\infty}, P(A_i) > 0, B, P(B) > 0$. Pak $P(A_j|B) = \frac{P(B|A_j) \cdot P(A_j)}{\sum_{i=1}^{\infty} P(B|A_i) \cdot /(A_i)}$.

Úloha: Mějme třídu, kde je $P(C^C) = 30\%$ dívek a $P(C) = 70\%$ chlapců. Víme, že $P(D|C^C) = 80\%$ dívek a $P(D|C) = 10\%$ chlapců má dlouhé vlasy. Když náhodně vybereme žáka s dlouhými vlasy, jaká je pravděpodobnost $P(C^C|D)$, že půjde o dívku?

Řešení: $P(C^C|D) = \frac{P(D|C^C) \cdot P(C^C)}{P(D|C^C) \cdot P(C^C) + P(D|C) \cdot P(C)}$

Úloha: V populaci se vyskytuje nemoc D. Pravděpodobnost pozitivního výsledku u nemocného člověka (senzitivita) je $P(+|D) = 0.999$, u zdravého člověka je $P(+|D^C) = 0.01$ (a tedy $P(-|D^C) = 0.99$). Nemocí trpí setina populace, $P(D) = 0.1$. Chceme znát pravděpodobnost nakažení při pozitivním testu $P(D|+)$.

Řešení: $P(D|+) = \frac{P(+|D) \cdot P(D)}{P(+|D) \cdot P(D) + P(+|D^C) \cdot P(D^C) = 0.502}$

Pojem: Náhodná veličina.
Binomické rozdělení $Bi(n, p)$. Máme n pokusů a úspěch nastane s pravděpodobností p. X počet úspěchů v n pokusech $\sim B_i(n, p)$
$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$
$x ... 0 ... n \rightsquigarrow \quad \sum_{k=0}^{n} p_k = \sum_{k=0}^{n} \binom{n}{k} p^k \cdot (1 - p)^{n-k} = (p + 1 - p)^n = 1$

Úloha: $X \sim Bi(4, \frac{2}{3}), Y = (X - 2)^2$. Určit rozdělení veličiny Y a spočítat její střední hodnotu.

Řešení: Obecně $E[Z] = \sum_{k=0}^{\infty} k \cdot P(Z = k)$.

$X : 01234$

$Y : 41014$

$q_0 = P(Y = 0) = p_2; \qquad q_1 = p_1 + p_3; \qquad q_4 = p_0 + p_4$ ČÁST CHYBÍ!

$p_0 = 1 \cdot 1 \cdot (1-p)^4 = (1-p)^4; \qquad p_1 = 4 \cdot p(1-p)^3; \qquad p_3 = 6 \cdot p^2(1-p)^2$

Úloha: $X \sim Bi(n, p)$, chceme $E[X] = \sum_{k=0}^{n} k \cdot P(X = k) = \sum k = 0^n k \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = n \cdot p$

$l = k - 1$

$$\sum_{l=0}^{n-1} (l+1) \frac{n!}{(n-(l+))! \cdot (l)!} \cdot p^{l+1} \cdot (1-p)^{n-l-1} = n \cdot p \cdot \sum_{l=0}^{n-1} \overbrace{\frac{(n-1)!}{(n-1-l)! \cdot l!}}^{\binom{n-1}{l}} \cdot p^l \cdot (1-p)^{n-1-l} = (p + (1-p))^{n-1}$$

anebo:

$Bi(n, p) \simeq \sum^n Alt(p)$ ČÁST CHYBÍ!

Úloha: VIZ. FOTO Z TABULE - DOPLNIT!

CVIČENÍ 21.10.2009

Pojem: *Geometrické rozdělení.* Udává počet neúspěchů před prvním úspěchem.

$X \sim Ge(p); \quad q = 1 - p; \quad p_k = P(X = k) = q^k \cdot p$

Úloha: $a, x \in \mathbb{N}$, máme a neúspěchů a chceme znát pravděpodobnost, že bude x dalších.

Řešení: $P(X \geq a + x | x \geq a) = \frac{P(X \geq a+x | X \geq a)}{P(X \geq a)} = \frac{P(X \geq a+x)}{P(X \geq a)} = \frac{q^{a+x}}{q^a} = q^x = P(X \geq x)$

Vlastnosti geometrického rozdělení.

Vytvořující funkce $f(s) = \sum_{k=0}^{\infty} p_k \cdot s^k \quad \rightsquigarrow \quad f'(s) = \sum_{k=0}^{\infty} k \cdot p_k \cdot s^{k-1} = E[X]$.

$f''(s) = \sum_{k=1}^{\infty} k \cdot (k-1) \cdot p_k \cdot s^{k-2}; \quad f''(1) = \sum_{k=1}^{\infty} k \cdot (k-1) \cdot p_k = E[X(X-1)]$

$E[X^2] = E[X(X-1)] + E[X] = f''(1) + f'(1) \quad \rightsquigarrow \quad var[X]$

Úloha: Vytvořující funkce pro $Ge(p)$

Řešení: $f(s) = \sum_{k=0}^{\infty} q^k \cdot p \cdot s^k = p \cdot \sum_{k=0}^{\infty} (q \cdot s)^k$

$f(s) = p \cdot \frac{1}{1 - q \cdot s} = p \cdot (1 - q \cdot s)^{-1} \quad \rightsquigarrow \quad f'(s) = \frac{p \cdot q}{(1 - q \cdot s)^2}$

$E[X] = f'(1) = \frac{p \cdot q}{(1-q)^2} = \frac{p}{q}$

Postup: Obecně,

$E[X] = \{^{\int x \cdot dF_x(x)}_{\sum_{i=1}^{n} x_i \cdot P(X = x_i)} \cdots$ $^{\text{spojitý případ}}_{\text{diskrétní případ}}$

$E[g(X)] = \{^{\int g(x) \cdot f(x) \cdot dx}_{\sum_{i=1}^{\infty} g(x_i) \cdot P(X = x_i)}$

Úloha: Viz předchozí, rozptyl.

Řešení: $f''(s) = \frac{2 \cdot p \cdot q^2}{(1 - q \cdot s)^3}$

$f''(1) = \frac{2 \cdot p \cdot q^2}{p^3} = \frac{2 \cdot q^2}{p^2} = E[X(X-1)]$

$E[X^2] = \frac{2 \cdot q^2}{p^2} + \frac{q}{p}$

$var[X] = f''(1) + f'(1) - (f'(1))^2 = E[X^2] - E[X]^2$

$\frac{2 \cdot q^2}{p^2} + \frac{q}{p} - \frac{q^2}{p^2} = \frac{q^2 + q \cdot p}{p^2} = \frac{q \cdot (q+p)}{p^2} = \frac{q}{p^2}$

Pojem: *Poissonovo rozdělení.*
$X \sim Po(\lambda)$; $\quad p_k = P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \quad k = 0, 1, ...$

Úloha: $E[X]$ Poissonova rozdělení.

Řešení: $f(s) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot s^k = \sum_{k=0}^{\infty} \frac{(\lambda \cdot s)^k}{k! \cdot e^{\lambda}} = e^{\lambda \cdot s - \lambda} = e^{\lambda \cdot (s-1)}$
$f'(s) = \lambda \cdot e^{\lambda \cdot (s-1)}$
$f'(1) = E[X] = \lambda$

Úloha: Máme $N = 14$ koulí, z toho $M = 8$ bílých, vytáhneme $n = 5$ koulí a chceme znát rozdělení: $X = \#$ bílých.
Koule vždy vracíme, pokusy jsou nezávislé.

Řešení: Počet úspěchů v n nezávislých pokusech $\rightsquigarrow$ binomické rozdělení $X \sim Bi(n, p)$, přičemž $p = P(\text{bílá}) = \frac{M}{N}$.

Pokud koule vracíme, pokusy na sobě závisí.

Řešení: $P(Y = k) = \dfrac{\overbrace{\binom{M}{k}}^{k \text{ zbylých vybíráme z } M} \cdot \overbrace{\binom{N-M}{n-k}}^{\text{zbylé nevybrané}}}{\underbrace{\binom{N}{n}}_{\text{v možnosti}}}$, přičemž předpokládáme $k \leq n \wedge M$.
Hypergeometrické rozdělení.

Úloha: Máme dva čtyřstěny se stěnami $O, 1, 2, 3$, $X = \#$ bodů, které padnou. Distribuční funkce.

Řešení: Obrázek.

Pojem: $R(0, 1)$ budiž rovnoměrné rozdělení na intervalu $[0, 1]$.

Poznámky: $F(x) = \int_{-\infty}^{x} f(z) \cdot dz; \quad F'(x) = f(x)$
$F(\infty) = P(X \leq \infty) = 1; \quad \int_{\mathbb{R}} f(x) \cdot d(x) = 1$

$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x) \cdot dx$
$E[X] = \int x \cdot f(x) \cdot dx \quad \rightsquigarrow \quad f(x) = I[0, 1] \quad ... \int_{-\infty}^{\infty} I[0, 1] \cdot dz = \int_0^1 1 \cdot dz$
$E[X^2] = \int x^2 \cdot f(x) \cdot dx$
$F(X) = \int_0^x 1 \cdot dz = x; \quad x \in (0, 1)$
$E[X^2] = \int_0^1 x^2 \cdot dx = \frac{1}{3}$
$var[X] = E[X^2] - E[X]^2 = \frac{1}{3} - (\frac{1}{2})^2 = \frac{1}{12}$

Úloha: Mějme $R(a, b)$. Chceme, aby $c = \frac{1}{b-a}$ byla hustota.

Řešení: $f(x) = \{ {}^{\frac{1}{b-a} \text{pro} x \in (a,b)}_{0 \text{jinak}}$
$F(x) = \int_{-\infty}^{x} f(z) \cdot dz = \int_a^x \frac{1}{b-a} \cdot dz = \frac{x-a}{b-a}$
$E[X] = \frac{a+b}{2}, \quad var[X] = \text{fuj}$

Cvičení 4.11.2009

Pojem: *Spojitá rozdělení.*

(1): $X \sim F: \quad F(x) = P(X < x) = \int_{-\infty}^{x} f(t) \cdot dt$
$F' = f; \quad f \geq 0$ je hustota, vždy se integruje na 1.
$E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx$
$E[g(x)] = \int_{-\infty}^{\infty} g(x) \cdot f(x) \cdot dx$
$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x) \cdot dx$
$var[X] = E[X - E[X]]^2 = E[X^2 - 2 \cdot X \cdot E[X] + E[X]^2] = E[X^2] - 2 \cdot E[X] \cdot E[X] + E[X]^2 = E[X^2] - E[X]^2$

(2): $X \sim Exp(\lambda)$
$f(x) = \lambda \cdot e^{-\lambda \cdot x}; \quad x \geq 0$
$F(x) = \int_0^X \underbrace{\lambda \cdot e^{-\lambda \cdot x}}_{y = \lambda \cdot t;\ dy = \lambda \cdot dt} = \int e^{-y} \cdot dy = -e^{-y} \rightsquigarrow [-e^{-\lambda \cdot x}]_0^x = 1 - e^{-\lambda \cdot x}$
$E[X] = \int_0^\infty x \cdot \lambda \cdot e^{-\lambda \cdot x} \cdot dx = \underbrace{[-X \cdot e^{-\lambda \cdot x}]_0^\infty}_{0} - \int_0^\infty -e^{-\lambda \cdot x} \cdot \frac{\lambda}{\lambda} \cdot dx = \frac{1}{\lambda} \cdot [-e^{-t}]_0^\infty = \frac{1}{\lambda}$
$var[X] = \frac{1}{\lambda^2}$

Úloha: $P(x \in [1, 2)) = F(2) - F(1) = 1 - e^{-2 \cdot \lambda} - 1 + e^{-\lambda} = e^{-\lambda} - e^{-\lambda \cdot 2}$

Pojem: *Funkce přežití* je $H = 1 - F(x) = P(X \geq x) = e^{-\lambda \cdot x}$

Poznámka: $P(X \geq x + y | x \geq y) = P(X \geq x)$

Úloha: Mějme X životnost v letech, $f(x) = c \cdot e^{-3 \cdot x}, \quad x > 0, \quad 1 = \int_0^\infty c \cdot e^{-3 \cdot x} \cdot dx = \frac{c}{3}$.
Určete c.

Řešení: $1 = \frac{c}{3} \quad \rightsquigarrow c = 3$

$P(X \geq 2)$

Řešení: $P(X \geq 2) = 1 - \underbrace{P(X < 2)}_{F(2)} = 1 - (1 - e^{-3 \cdot 2}) = e^{-6}$

Pojem: *Normální rozdělení.*
$X \sim N(0, 1)$
$f(x) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}$
$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{t^2}{2}} \cdot dt$
$Y = a + b \cdot X \quad \rightsquigarrow Y \sim N(a, b^2)$
$N(\mu, \sigma^2) \quad \rightsquigarrow f(x) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}}$
OBRÁZEK.

Příklad: $N(\mu, 1)$
$E[X] = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-\mu)^2}{2}} \cdot dx = \int_{-\infty}^{\infty} (x - \mu) \cdot \frac{e^{-\frac{(x-\mu)^2}{2}}}{\sqrt{2 \cdot \pi}} \cdot dx +$
$+ \underbrace{\int_{-\infty}^{\infty} \mu \cdot \underbrace{\frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-\mu)^2}{2}} \cdot dx}_{=1}}_{=\mu} \quad \underbrace{= \ldots}_{\substack{x+0=(x-\mu)+\mu;\ x-\mu=y \\ z=y^2;\ dz=2 \cdot y \cdot dy}} \quad \int_{-\infty}^{\infty} y \cdot e^{-\frac{y^2}{2}} \cdot dy$

Příklad: $X \sim N(0, 1); \quad P(|X| < 1) = P(-1 < X < 1) = \Phi(1) - \Phi(-1)$

Příklad: $Z \sim N(1,4) \quad \rightsquigarrow \frac{(Z-\mu)}{\sigma} \sim N(0,1)$
$P(Z \in (0,2)) = P(Z < 2) = P(\frac{Z-1}{2} < \frac{2-1}{2}) = P(N < \frac{1}{2}) = \Phi(\frac{1}{2})$

Příklad: $0 < Z < 2 \quad \rightsquigarrow P(-\frac{1}{2} < \frac{Z-1}{2} < \frac{2-1}{2}) = \Phi(\frac{1}{2}) - \Phi(-\frac{1}{2})$

Poznámka: $\Phi(-x) = 1 - \Phi(x); \quad \Phi(-x) + \Phi(x) = 1$

Příklad: $X \sim N(\mu, \sigma); \quad \rightsquigarrow E[|X-\mu|] = \sqrt{\frac{2}{\pi}} \cdot \sigma$

Úloha: $X \sim R(0,10), \quad f = \frac{\frac{1}{10}:\ x \in [0,10]}{0:\ x \notin [1,10]}$. Chceme objem $E[X^3]$.

Řešení: $E[g(x)] = \int g(x) \cdot f(x) \cdot dx$
$E[X^3] = \int_0^1 0x^3 \cdot \frac{1}{10} \cdot dx = \frac{1}{10} \cdot \frac{1}{4} \cdot [x^4]_0^1 0 = \frac{1}{40} \cdot 10000 = 250$

Cvičení 11.11.2009

Definice: Náhodné veličiny $X_1, ..., X_n$ jsou navzájem *nezávislé*, když jejich sdružená distribuční funkce $F_{\underline{X}}(\underline{x}) = P(X_1 \leq x_1, ..., X_n \leq x_n)$ =Nekonzistentní zápis.

Poznámka: „Učebnicová“ definice je ekvivalentní.

Úloha: Mějme $X_1, ..., X_n$ se stejnou distribuční funkcí F.
$Y = min\{X_1, ..., X_n\}, \quad Z = max\{X_1, ..., X_n\}$.
Chceme znát F_Y, F_Z.

Řešení: $F_Z(z) = P(Z \leq z) = P(X_1 \leq z, ..., X_n \leq z) = F(z)^n$.
$F_Y(y) = P(Y \leq y) = 1 - P(Y > y) = 1 - (1 - F(y))^n$.

Úloha: $T \sim Exp(\lambda),\ \lambda > 0,\ f(t) = \lambda \cdot e^{-\lambda \cdot t},\ f \geq 0$.
Chceme znát pravděpodobnost, že nejkratší doba zpracování jednoho z pěti dotazů bude větší než $\frac{1}{30}$ sekund. Zpracování jsou navzájem nezávislá. Spočítat střední hodnotu a rozptyl.

Řešení: $T_i \sim Exp(\lambda),\ i = 1, ..., 5$.
$P(min(T_1, ..., T_5) > \frac{1}{30}) = (P(T_1 > \frac{1}{30}))^5 = (1 - F(\frac{1}{30}))^5 = e^{-\frac{5 \cdot \lambda}{30}}$.
Víme $F(t) = 1 - e^{-\lambda \cdot t}$, tedy $E[T_i] = \frac{1}{\lambda},\ var[T_i] = \frac{1}{\lambda^2}$.
Chceme $E[min(T_1, ..., T_5)]$.
$F_{min}(y) = 1 - e^{-5 \cdot t} \quad \rightsquigarrow f_{min}(y) = F' \quad \rightsquigarrow E[min(T_1, ..., T_5)] = \int_0^\infty y \cdot f_{min}(y) \cdot dy$,
to je ovšem velmi pracné! Lepší je všimnout si, že $F_{min}(y)$ je distribuční funkce $Exp(5 \cdot \lambda)$.

Příklad: $P(T \leq a) = \frac{1}{2}$, určit a (= medián) $\Rightarrow \ \alpha \in (0,1)$, α-kvantil.
$F(a) = \alpha$, je-li rostoucí, monotónní, prostá $\Rightarrow F^{-1}(\alpha)$.
Pro nespojitou $F(x)$ definujeme $F^{-1}(\alpha) := \inf\{x : F(x) \geq \alpha$.
$1 - e^{-\lambda \cdot a} = \alpha \quad \Rightarrow \quad a = -\frac{\log(1-\alpha)}{\alpha}$.
$P(T \leq -\frac{log(a-\lambda)}{\lambda}) = \alpha$.

Poznámka: Medián je mnohem lepší charakteristika než průměr, neovlivní jej tolik extrémní hodnoty (typickým příkladem budiž tzv. průměrný plat).

Téma: Rozdělení náhodných vektorů.
$\underline{X} = (X_1, ..., X_n)$, položky vektoru mohou být navzájem závislé.
$E[\underline{X}] = (\begin{smallmatrix} E[X_1] \\ \dots \\ E[X_n] \end{smallmatrix})$.

$E[\underline{X} - E[\underline{X}]][\underline{X} - E[\underline{X}]]^T \quad \rightsquigarrow \boldsymbol{\Sigma}$ matice $n \times n$,
kde na pozici $\boldsymbol{\Sigma}_{i,j}$ je $E[X_i - E[X_i]][X_j - E[X_j]] = cov(X_i, X_j)$ (a na diagonále máme tedy rozptyl).
$cor(X_i, X_j) = \frac{cov(X_i,X_j)}{\sqrt{var[X_i]}\cdot\sqrt{var[X_j]}}$.
Pro nezávislé $X, Y (X \perp\!\!\!\perp Y)$ platí $E[XY] = E[X] \cdot E[Y] \quad \Rightarrow \quad cov(X, Y) = 0$,
tedy také $X \perp\!\!\!\perp Y \quad \Rightarrow cor(X, Y) = 0$ (ovšem neplatí $\Leftarrow$).
$cor(a \cdot X + b, X) = {}^{\;1}_{-1}, \quad cor(X, Y) \in [-1, 1]$.
Normální rozdělení: $X \sim N(0, 1) \quad \rightsquigarrow f(x) = \frac{1}{\sqrt{2\cdot\pi}} \cdot e^{-\frac{x^2}{2}},\ x \in \mathbb{R}, \quad Y = X^2$
$E[XY] = E[X] \cdot E[Y] = E[X^3] = 0$ (integrál liché funkce)
$cov(X, Y) = E[XY] - E[X] \cdot E[Y] = 0 \quad \Rightarrow cor(X, Y) = 0$, ale závislé jevy!

Příklad: $P(\underline{X} = \underline{x_i}) = p_i$ diskrétní rozdělení.
$P(X_1 = x_{i_1}, ..., X_n = x_{i_n}) = p_i, \quad \sum_i p_i = 1$.
$\underline{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \quad \rightsquigarrow f(x_1, ..., x_n) \quad \rightsquigarrow P(\underline{X} \in \underbrace{B}_{\text{interval}}) = \int \cdots \int_B f(x_1, ..., x_n) \cdot dx_1 \cdot ... \cdot dx_n$.
$f(x_1, x_2) = \frac{\partial^2 F}{\partial x_1\ \partial x_2}$.
$X_1 \perp\!\!\!\perp X_2 \quad \Rightarrow f(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$.
$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(a, b) \cdot da \cdot db$.

Příklad: Diskrétní $\binom{X}{Y}$.

$_Y\backslash^X$	0	1	2	3	P_X
0	0	$\frac{1}{12}$	0	0	$\frac{1}{12}$
1	$\frac{1}{12}$	$\mathbf{0}$	$\mathbf{\frac{1}{3}}$	$\mathbf{0}$	$\mathbf{\frac{5}{12}}$
2	0	$\mathbf{\frac{1}{3}}$	$\mathbf{0}$	$\mathbf{\frac{1}{12}}$	$\frac{5}{12}$
3	0	$\mathbf{0}$	$\mathbf{\frac{1}{12}}$	$\mathbf{0}$	$\frac{1}{12}$
P_Y	$\frac{1}{12}$	$\frac{5}{12}$	$\frac{5}{12}$	$\frac{1}{12}$	

1. $X \perp\!\!\!\perp Y$? Ověřit definici, chceme $P(X = a, Y = b) = P(X = a) \cdot P(Y = b)$, ale $P(X = 0, Y = 0) \neq P(X = 0) \cdot P(Y = O)$, $SPOR$.
2. $cov(X, Y)$, potřebujeme tučně vyznačené psti:
$E[XY] = \cdots = \frac{7}{3}$, $E[X] = E[Y] = 0 \cdot \frac{1}{12} + 1 \cdot \frac{5}{12} + \cdots = \frac{3}{2}$
$cov(X, Y) = \frac{7}{3} - (\frac{3}{2})^2 = \ldots$

Úloha: Jsou e^{X+Y} a e^{X-Y} nezávislé?

Řešení: e nemá vliv, stačí ověřit $X + Y \perp\!\!\!\perp X - Y \quad \rightsquigarrow$ tabulka pravděpodobností...

CVIČENÍ 18.11.2009

Úloha: $X \sim N(1, 4)$, chceme $P(|X| < 3)$.

Řešení: $-3 < X < 3 \quad \rightsquigarrow \quad \frac{-3-1}{2} < \underbrace{\frac{X - 1}{2}}_{Z \sim N(0,1)} < \frac{3-1}{2}$.
$P(-2 < Z < 1) = \Phi(1) - \Phi(-2)$.

Poznámka: platí $\Phi(-a) = 1 - \Phi(a)$.

Věta: Mějme $(X_i)_i^\infty$ nezávislé, stejné rozdělení, $E[X_1] = \mu,\ var[X_1] = \sigma^2$. Pak $P(\frac{\sum_{k=1}^n X_k - n\cdot\mu}{\sqrt{n\cdot\sigma^2}} < x) \underset{n \to \infty}{\longrightarrow} \Phi(x)$

Poznámky: $\frac{n}{\sqrt{n}} \cdot \frac{(\overline{X}_n-\mu)}{\sqrt{\sigma^2}}$, $\overline{X}_n = \frac{1}{n} \cdot \sum_{k=1}^{n} X_k \quad \rightsquigarrow P(\sqrt{n} \cdot \frac{\overline{X}_n-\mu}{\sqrt{\sigma^2}} < x) \to \Phi(x)$

Úloha: Jeden zákazník zabere $T \sim Exp(2)$ sekund, máme $n = 36$ zákazníků $T_1, ..., T_36$. Jaká je pravděpodobnost, že zaberou dohromady méně než 20?

Řešení: $X = \sum_{i=1}^{3} 6T_i \quad \rightsquigarrow \quad P(X < 20)?$
$f(x) = \{^{\lambda \cdot e^{-\lambda \cdot x}: \ x>0}_{0: \quad x \le 0}.$
$E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx = \int_0^{\infty} \underbrace{x}_{G} \cdot \underbrace{\lambda \cdot e^{-\lambda \cdot x}}_{f} \cdot dx = [-x \cdot e^{-\lambda \cdot x}]_0^{\infty} + \int_0^{\infty} e^{-\lambda \cdot x} \cdot dx = [-x \cdot e^{-\lambda \cdot x}]_0^{\infty} -$
$[-\frac{1}{\lambda} \cdot e^{-\lambda \cdot x}]_0^{\infty} = 0 + \frac{1}{\lambda} = \frac{1}{\lambda}.$
$\mu = E[T_1] = \frac{1}{2}, \quad \sigma^2 = var[T_1] = \frac{1}{4}.$
$P(X < 20) = P(\sim_{i=1}^{3} 6T_i < 20) = P(\underbrace{\frac{\sum_i T_i - \frac{1}{2} \cdot 36}{3}}_{X} < \underbrace{\frac{20 - \frac{1}{2} \cdot 36}{3}}_{x}) \approx \Phi(x) = \Phi(\frac{2}{3}) \approx 0.75$

Úloha: $n = 100$krát házíme kostkou, chceme pravděpodobnost, že součet hodů $X = \sum_{i=1}^{1} 00K_i \in (320, 380)$.

Řešení: $\mu = E[K_i] = 3.5 \quad \Rightarrow \quad E[X] = 350.$
$\sigma^2 = var[K_i] = E[K_i^2] - E[K_i]^2 = \frac{91}{6} - 3.5^2 = \frac{35}{12}.$
$P(320 < X < 380) = P((X - 350) < 30) = P(\frac{X-350-350}{10 \cdot \sqrt{\frac{35}{12}}} < \frac{30-350}{10 \cdot \sqrt{\frac{35}{12}}}) \approx 1 - \Phi(\frac{320}{10 \cdot \sqrt{3512}})$

Cvičení 25.11.2009

Úloha: $n = 1000$ lidí, nějaký z nich zemře s pravděpodobností $p = 0.01$. Pojistné činí $a = 150$kč a pojistka při úmrtí $A = 10000$kč. S jakou pravděpodobností bude mít pojišťovna ztrátu?

Řešení: Příjmy $P = a \cdot n$, výdaje $V = X \cdot A$, $X = \sum_{i=1}^{n} X_i$.
$X_i = 1$, pokud i-tý člověk zemřel, jinak 0. $X_i \sim Alt(p)$.
Chceme znát $P(V > P) = P(X \cdot A > a \cdot n) = P(X > \underbrace{\frac{a \cdot n}{A}}_{y=15}) = *$
$\mu = E[X_i] = p, \ E[X^2] = p$
$\sigma^2 = E[X^2] - E[X]^2 = p - p^2 = p \cdot (1 - p) \quad \rightsquigarrow X \sim Bi(n, p) \sim \sum_{i=1}^{n} Alt(p)_i.$
$E[\sum \dots] = \sum E[\dots] = n \cdot p$ a pokud jsou jevy $\perp\!\!\!\perp$, pak $var[\sum \dots] = \sum var[\dots] = n \cdot p \cdot (1 - p)$.
$* = P(\frac{\sum_{i=1}^{n} X_i - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} > \frac{y - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}}) \xrightarrow{n \to \infty} 1 - \Phi(\frac{y - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}}) \approx 0.56$

Úloha: $n = 1000$ hodů symetrickou mincí ($p = 0.5$). Chceme znát pravděpodobnost, že počet líců bude větší než 4100.

Řešení: $P(X > 4100)$, $x = \sum_{i=1}^{n} X_i$, $X_i \sim Alt(0.5)$, $X \sim Bi(10000, \frac{1}{2})$.
$P(\frac{x - n \cdot \mu}{\sqrt{n \cdot \sigma^2}} > \frac{4100 - 10000 \cdot 0.5}{\sqrt{2500}}) \approx 1 - \Phi(-2) = \Phi(2).$

Úloha: Každé ráno a odpoledne cesta metrem, čekání 0 až 3 minuty. Jaká je pravděpodobnost, že během 22 dnů stráví cestující čekáním déle než hodinu?

Řešení: $X_i \sim U(0, 3)$, $X = \sum_{i=1}^{4} 4X_i$, chceme $P(X > 60)$.
$E[X_i] = \frac{3}{2}, \quad E[X_i^2] = \int_0^3 x^2 \cdot \frac{1}{3} \cdot dx = [\frac{x^3}{9}]_0^3 = 3.$
$var[X_i] = E[X_i^2] - E[X_i]^2 = 3 - (\frac{3}{2})^2 = \frac{3}{4}.$
$P(X > 60) = P(\frac{X-66}{\sqrt{33}} > \frac{60-66}{\sqrt{33}}) = 1 - \Phi(-\frac{6}{\sqrt{33}}) = \Phi(\frac{6}{\sqrt{33}})$

Úloha: Hody mincí, V_n četnost líců, chceme $P(|V_n - \frac{1}{2}| \leq 0.05) \geq 0.95$

Řešení 1: Čebyševova nerovnost: Pro X, ϵ platí $P(|X - E[X]| > \epsilon) \leq \frac{var[X]}{\epsilon^2}$ (omezuje pravděpodobnost, že se X odchyluje od své střední hodnoty o více než ϵ).
Poznámka, $P = E[I[|X - E[X]| > \epsilon]] \leq E[\frac{|X-E[X]|^2}{\epsilon^2} \cdot I[\dots]] \leq \frac{|X-E[X]|^2}{\epsilon^2} = \frac{var[X]}{\epsilon^2}$.
$P(|V_n - \frac{1}{2}| > 0.05) \leq \frac{1}{4 \cdot n \cdot (0.05)^2} = 0.05$.

Řešení 2: Centrální limitní věta.
$V_n = \frac{\sum_{i=1}^n X_i}{n}$, $E[V_n] = \frac{1}{2}$, $var[V_n] = \frac{1}{n^2} \cdot \sum^n \frac{1}{4} = \frac{1}{4 \cdot n}$.
$1 - P(|V_n - \frac{1}{2} > 0.05) \geq 0.95 \quad \rightsquigarrow \quad 0.05 \geq P(\dots)\dots$
$P(\frac{|V_n - \frac{1}{2}|}{\sqrt{\frac{1}{4 \cdot n}}} \leq \frac{0.05}{\sqrt{\frac{1}{4 \cdot n}}}) = P(2 \cdot \sqrt{n} \cdot |V_n - \frac{1}{2}| \leq 0.1 \cdot \sqrt{n}) = 0.95$

Poznámky: Pro $N \sim N(0, 1)$:
$P(|N| < U) = 0.95 \quad \rightsquigarrow P(N < U) = 0.95) \quad \dots 95\%$-kvantil
$\Phi(U) \quad \dots U$-kvantil $(\Phi(U_{0.95}) = 0.95,\ \Phi(U_{0.975}) = 0.975,\ P(N > U_{0.975}) = 0.025 \quad \dots$

Poznámka: Shrnutí pro písemku. $E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx$.
$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x) \cdot dx$.
$var[X] = E[X - E[X]]^2 = E[X^2] - E[X]^2$.
Rovnoměrné rozdělení (R, U).
Centrální limitní věta.
$X_1, \dots, X_n$ (iid) $\sim F,\ Y = min(X_1, \dots, X_n), Z = max(\dots),\ G(y) = P(Y < y),\ H(z) = P(Z < z) = P(max(\dots) < z) = [F(z)]^n$.
Vlastnosti f, F, etc.