

Vlastnosti

- 1) $0 \leq P(A) \leq 1$
- 2) $A, B \in \mathcal{A}, A \subset B \Rightarrow P(A) \leq P(B)$
- 3) $P(A^c) = 1 - P(A)$
- 4) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- 5) $P(A \cap B) = P(A) \cdot P(B)$ jestli jsou nezávislé
- 6) $A, B \in \mathcal{A}, A \subset B \Rightarrow P(B-A) = P(B) - P(A)$

Bayesova věta

$$P(A_i | B) = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{j=1}^{\infty} P(A_j) \cdot P(B|A_j)}$$

, $i = 1, 2, \dots$

Nezávislost jevů

Jevy jsou nezávislé, když

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

Jevy A_i jsou vzájemně nezávislé, když

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot \dots \cdot P(A_n)$$

Distribuční funkce

$$F(x) = P(X \leq x)$$

pst jevů = rozdíl skoků (Δ)

Distribuční funkce diskrétní n.v.

$\{x_i\}$ - posloupnost r. čísel

$\{p_i\}$ - posloupnost nezáporných čísel

$$p_i = P(X=x_i), \sum_{i=1}^{\infty} p_i = 1$$

$$F(x) = P(X \leq x) = P(X=x_i) = \sum_{\{i: x_i \leq x\}} p_i = \sum_{\{i: x_i \leq x\}} p_i$$

$$P(a < X < b) = F(b) - F(a) = \sum_{\{i: a < x_i < b\}} p_i = \sum_{\{i: a < x_i < b\}} p_i$$

$$= \sum_{\{i: a < x_i < b\}} p_i$$

A za podmínky jevu B

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

$$P(A \cap B) = P(A|B) \cdot P(B)$$

Kombinační číslo

$\binom{n}{k}$ - počet k-tic z n prvků

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Věta

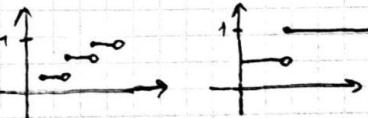
$$P(B) = \sum_{i=1}^n P(A_i) \cdot P(B|A_i)$$

Distribuční funkce spojité

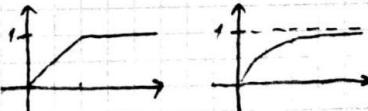
$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$$

Příklady diskrétní a spojité d.

a) Diskrétní



b) Spojitá



Mužtota $f(x)$

$$1) f(x) = \frac{d}{dx} F(x)$$

$$2) \int_{-\infty}^{\infty} f(x) dx = 1$$

$$3) P(a < X < b) = F(b) - F(a) = \int_a^b f(x) dx$$

Smešené rozdělení Diskretní a Spojité n.v.

$$X = M_i x_c(D, S)$$

$$X = M_1 x_c(D, S) = D$$

$$X = M_0 x_c(D, S) = S$$

$$4) F_X(x) = c F_D(x) + (1-c) F_S(x) = c \sum_{i: x_i \leq x} p_i + (1-c) \int_{-\infty}^x f(t) dt$$

$$5) P(X=x) = F_X(x) - \lim_{t \rightarrow x^-} F_X(t)$$

Může nastat

$$M_i x_c(D_1, D_2), \text{ tedy } X \text{ bude diskretní}$$

$$M_i x_c(S_1, S_2), \text{ tedy } X \text{ bude spojité}$$

Kvantilová funkce (LIR)

$$q(\lambda) = \inf_{t \in \mathbb{R}} F_X(t) \leq \lambda \leq \sup_{t \in \mathbb{R}} F_X(t) \geq \lambda$$

inverzní k $F_X(x)$, rotace o 90°

Momenty, rozptyl, kovariace

$$1) EX^n - n-tý moment náh vel X$$

$$2) E(X - EX)^n - n-tý centrální moment$$

$$3) E|X - EX| - absolutní moment$$

$$4) \text{var } X = E(X - EX)^2 - rozptyl$$

$$5) \text{cov}(X, Y) = E(X - EX)(Y - EY) - kovariancia$$

$$EX^2 < \infty, EY < \infty$$

Vlastnosti

$$1) \text{var } X = EX^2 - (EX)^2$$

$$2) \text{var } a = 0$$

$$3) \text{var } (\alpha X) = \alpha^2 \text{var } X$$

$$4) \text{var } (X + a) = \text{var } X$$

$$5) \text{Necht } Z = \frac{X - EX}{\sqrt{\text{var } X}} \Rightarrow EZ = 0, \text{var } Z = 1$$

$$6) \text{var } (X + Y) = \text{var } X + \text{var } Y + 2\text{cov}(X, Y)$$

$$7) \text{cov}(X, Y) = E(XY) - EXEY$$

Střední hodnota

$$1) EX = \int_{-\infty}^{\infty} x f(x) dx$$

pokud integral ex.

a) diskrétní

$$2) EX = \sum_{i=1}^{\infty} x_i \cdot P(X=x_i) \quad E(X^2) = \sum_{i=1}^{\infty} x_i^2 \cdot P(X=x_i)$$

b) spojité

$$3) EX = \int_{-\infty}^{\infty} x f(x) dx \quad E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$$

c) směs

$$EX = cED + (1-c)ES$$

Vlastnosti

$$1) E(a) = a$$

$$2) E(aX + bY) = aEX + bEY$$

$$3) X_1 \leq X \leq X_2 \Rightarrow EX_1 \leq EX \leq EX_2$$

$$4) X \geq 0 \Rightarrow EX \geq 0$$

Střední hodnota funkce náh. vel

$$1) E\phi(X) = \int_{-\infty}^{\infty} \phi(x) dF_X(x)$$

a) diskrétní

$$2) E\phi(X) = \sum_{i=1}^{\infty} \phi(x_i) \cdot P(X=x_i)$$

b) spojité

$$3) E\phi(X) = \int_{-\infty}^{\infty} \phi(x) f(x) dx$$

$$4) \text{směs } E\phi(X) = cE\phi(D) + (1-c)E\phi(S)$$

Čebýševova nerovnost

Pro každé $\varepsilon > 0$ platí:

$$P(|X - \mathbb{E}X| \geq \varepsilon) \leq \frac{\text{Var}X}{\varepsilon^2}$$

Alternativní rozdělení $X \sim \text{alt}(p)$

1) X nabývá hodnot 0 a 1 s pravděpodobnostmi $p, 1-p$

2) p - parametr $P(X=k) = p^k(1-p)^{n-k}$

3) Distribuční funkce

$$F(x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

4) $\mathbb{E}X = p$, $\text{Var}X = p(1-p)$

Binomické rozdělení $X \sim \text{Binom}(n, p)$

1) X nabývá hodnot $k=0, 1, \dots, n$

2) $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

3) Distribuční funkce

$$F(x) = \begin{cases} 0 & x < 0 \\ \sum_{k=0}^{x-1} \binom{n}{k} p^k (1-p)^{n-k} & 0 \leq x < n \\ 1 & x \geq n \end{cases}$$

4) $\mathbb{E}X = np$, $\text{Var}X = np(1-p)$

Poissonovo rozdělení $X \sim \text{Po}(\lambda)$

1) X nabývá hodnot $k=0, 1, 2, \dots$

2) $P(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}$

3) Distribuční funkce

$$F(x) = \sum_{j=0}^x e^{-\lambda} \frac{\lambda^j}{j!} \quad x \in \mathbb{N}_0$$

4) $\mathbb{E}X = \text{Var}X = \lambda$

Rovnoměrné rozdělení na $[a, b]$ $X \sim R(a, b)$

1) Nabývá hodnot z intervalu $[a, b]$

2) $f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{jinak} \end{cases}$

3) $F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$

4) $\mathbb{E}X = \frac{a+b}{2}$, $\text{Var}X = \frac{1}{12}(b-a)^2$

Geometrické rozdělení $X \sim Ge(p)$

1) X nabývá hodnot $k=0, 1, 2, \dots$

2) $P(X=k) = p(1-p)^k$

3) Distribuční funkce

$$F(x) = \begin{cases} 0 & x < 0 \\ \sum_{k=0}^x p(1-p)^k & x \geq 0 \end{cases}$$

4) $\mathbb{E}X = \frac{1-p}{p}$, $\text{Var}X = \frac{1-p}{p^2}$

Hypergeometrické rozdělení $X \sim Hyp$

$X \sim \text{Hyp}(N, K, n)$

1) $N, K, n \in \mathbb{N}$, $N \geq K$, $n \geq n$

2) X nabývá hodnot $k \in \mathbb{N}$:
 $\max\{0, n+K-N\} \leq k \leq \min\{n, K\}$

3) $P(X=k) = \frac{\binom{K}{k} \cdot \binom{N-K}{n-k}}{\binom{N}{n}}$

4) Distribuční funkce

$$F(x) = \begin{cases} 0 & x < 0 \\ \sum_{k=0}^x \frac{\binom{K}{k} \cdot \binom{N-K}{n-k}}{\binom{N}{n}} & 0 \leq x < n \\ 1 & x \geq n \end{cases}$$

a) pro $x < \max\{0, n+K-N\}$

b) pro $x \in \left[\max\{0, n+K-N\}, \min\{n, K\}\right]$

c) pro $x \geq \min\{n, K\}$

5) $\mathbb{E}X = n \frac{K}{N}$

$$\text{Var}X = n \frac{K}{N} \left(1 - \frac{K}{N}\right) \frac{N-n}{N-1}$$

Exponenciální rozdělení $X \sim Exp(\lambda)$

1) Nabývá hodnot $(0, \infty)$

2) $f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$

3) $F(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-\lambda x} & x > 0 \end{cases}$

4) $\mathbb{E}X = \int_0^\infty x \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$

$$\text{Var}X = \frac{1}{\lambda^2} \quad P(X \leq a) = \int_0^a \lambda e^{-\lambda x} dx$$

Použití rozdělení

1) Alt(p) - modelování úspěchu a neúspěchu.

p - pravděpodobnost úspěchu

2) Binom(n, p) - počet úspěchů v nzávislostech pokusech

n - počet vš. pokusů, p - pravděpodobnost úspěchu

3) Po(λ) - počet nezav. událostí, které nastoupí v několika intervalech
 λ - počet událostí v jednotce času (v intervalu) Př 23

4) Ge(p) - počet neúspěchů před prvním úspěchem

p - pravděpodobnost úspěchu

5) Hyp(N, K, n) - ~~N objektů s vlastností K, z nichž k má nejákon vlastnost. Náhodně vybereme n objektů.~~

N - počet objektů s vlastností K (z n objektů)

6) Ro(a, b) - doba čekání na pravidelně přicházející události
 (a, b) - interval

7) Exp(λ) - doba čekání na událost, které chodi vzájemně nezávisle

Obecné Normální rozdělení $X \sim N(\mu, \sigma^2)$

1) Nabývá hodnot \mathbb{R}

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

3) Distribuční funkce

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

4) $\mathbb{E}X = \mu$, $\text{Var}X = \sigma^2$

Normované norm. rozd $X \sim N(0, 1)$

1) Nabývá hodnot \mathbb{R}

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

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

4) $\mathbb{E}X = 0$, $\text{Var}X = 1$

5) $\Phi(x)$ - tabulková hodnota

$$\Phi(x) = 1 - \Phi(-x)$$

Transformace $X \sim N(\mu, \sigma^2) \rightarrow Z \sim N(0, 1)$

$$X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X - \mu}{\sigma}$$

$$\tilde{P}: P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P(Z \leq \text{const}) = \Phi(\text{const})$$

Součet normálních rozdělení

$$Z = X - Y \quad X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$$

$$\mathbb{E}Z = \mu_1 - \mu_2$$

$$\text{Var}Z = \sigma_1^2 + \sigma_2^2 \quad (\text{vždy plus})$$

Funkce náh. vel

X je náh. vel ; $\Omega: \mathbb{R} \rightarrow \mathbb{R}$; $Y = \varphi(X)$

- diskrétní $G(y) = \sum_{\{x_i : \varphi(x_i) \leq y\}} p_i \quad \forall y \in \mathbb{R}$
- spojité $G(y) = \int_{\{x : \varphi(x) \leq y\}} f(x) dx \quad \forall y \in \mathbb{R}$

Konvoluce $H = F \cdot G$; $Z = X + Y$

- X, Y diskrétní (F & G distrib. & ce)
 $F(x) = \sum_{n=0}^{\infty} p_n \quad G(g) = \sum_{n=0}^{\infty} q_n$
 $H(z) = \sum_{n=0}^{\infty} h_n, \quad h_n = \sum_{k=0}^z p_k \cdot q_{n-k}$
- X, Y spojité
 $h(z) = \int_{-\infty}^z (z-y) g(y) dy$

Náhodný vektor $X = (X_1, \dots, X_m)^T$

- $\mathbb{E}X = (\mathbb{E}X_1, \dots, \mathbb{E}X_m)^T$ vektor s. i. hodn.
- Variacní matice $\text{var}X$ s prvky
 $\text{cov}(X_i, X_j) = \mathbb{E}(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j) \quad i \neq j$
- $\text{corr}X$ s prvky (korelační matici)
 $\text{corr}(X_i, X_j) = \frac{\text{cov}(X_i, X_j)}{\sqrt{\text{var}X_i} \sqrt{\text{var}X_j}} \quad i \neq j$
 $= 0 \text{ iff } X_i \text{ a } X_j \text{ jsou nezávislé}$

Pozn: $-1 \leq \text{corr}(X, Y) \leq 1$ $\boxed{\text{corr}(X_i, X_i) = 1}$

Nezávislost náh. vel v náh. vektorech

- diskrétní $P(X_1=x_1^{(i)}, \dots, X_n=x_n^{(i)}) = \prod_{j=1}^n P(X_j=x_j^{(i)})$
- spojité
 $f_X(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$

Nezávislost & kovariace

Nechť X a Y jsou nezáv.

- $\mathbb{E}XY = (\mathbb{E}X)(\mathbb{E}Y)$
- $\mathbb{E}X^2 < \infty$ a $\mathbb{E}Y^2 < \infty$: $\text{cov}(X, Y) = 0$

Základní konvoluce $Z = X + Y$

- $X \sim \text{Alt}(p)$, $Y \sim \text{Alt}(p)$
 $Z \sim \text{Binom}(2, p)$
- $X \sim \text{Binom}(n_1, p)$, $Y \sim \text{Binom}(n_2, p)$
 $Z \sim \text{Binom}(n_1 + n_2, p)$
- $X \sim \text{Po}(\lambda_1)$, $Y \sim \text{Po}(\lambda_2)$
 $Z \sim \text{Po}(\lambda_1 + \lambda_2)$
- $X \sim \text{Ro}(a, b)$, $Y \sim \text{Ro}(c, d)$
 $f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{jinak} \end{cases}$
 $g(y) = \begin{cases} \frac{1}{d-c} & c \leq y \leq d \\ 0 & \text{jinak} \end{cases}$

Pro $d-c \geq b-a$ platí

$$h(z) = \begin{cases} 0 & z < a+c/b+d \leq z \\ \frac{z-(a+c)}{(b-a)(d-c)} & a+c \leq z \leq b+c \\ \frac{1}{d-c} & b+c \leq z \leq a+d \\ \frac{(b+d)-z}{(b-a)(d-c)} & a+d \leq z \leq b+d \end{cases}$$

- $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$
 $Z \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
- $X \sim \text{Exp}(\lambda)$, $Y \sim \text{Exp}(\lambda)$
 $h(z) = \begin{cases} \lambda^2 z \exp(-\lambda z), & z > 0 \\ 0, & z \leq 0 \end{cases}$

Konvergence náh. vel.

$\hookrightarrow X_n$ konverguje s. j. k X , jestliže
 $P\{w: \lim_{n \rightarrow \infty} X_n(w) = X(w)\} = 1$

$\hookrightarrow X_n$ konverguje k X s prst. jestliže pro vš. $\epsilon > 0$ platí
 $\lim_{n \rightarrow \infty} P\{w: |X_n(w) - X(w)| > \epsilon\} = 0$

Pozn: Konverg. s. j. \Rightarrow Konverg. v prst.

Centralní limitní věta

X_1, \dots, X_n jsou nez. stejně rozdělené náh. vel $\Rightarrow \mathbb{E}X_i = \mu$ a $\text{var}X_i = \sigma^2$

$Z_n = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma^2}, \quad n=1, 2, \dots$

Pak $\lim_{n \rightarrow \infty} F_n(x) = \Phi(x)$

Rozdělení používaná ve Statistiky

- ~~$Y \sim \chi^2_m$~~ $\Rightarrow X_1, \dots, X_m \sim N(0, 1)$
 $Y = \sum_{i=1}^m X_i^2$ (chi-kvadrat rozdělení)
- $Z \sim t_n$, $X \sim N(0, 1)$, $Y \sim \chi^2_n$
 $Z = \frac{X}{\sqrt{Y/n}}$ (Studentovo t- rozdělení)
- Fisherovo - Snedecorovo rozd. s param. n, m
 $W \sim F_{n,m}$, $U, V \sim \chi^2_n, \chi^2_m$
 $W = \frac{U/n}{V/m}$

Kvantily

Nechť F je spojitá a monotonní distrib. & ce a $0 < \beta \leq 1$. Pak $\text{kvantil } F$

Z_β je β -kvantil, pokud $F(Z_\beta) = \beta$

$$\bullet P(Z_{\alpha/2} < X < Z_{1-\alpha/2}) = F(Z_{1-\alpha/2}) - F(Z_{\alpha/2}) = 1 - \alpha$$

Emperická distribuční funkce

$(x_1, \dots, x_n)^T$ je realizace náh. výb. $(X_1, \dots, X_n)^T$

$F_{\text{emp}}(x) = \frac{\#\{x_i : x_i \leq x\}}{n}$

kde $\#$ počet prvků

Kvantily, mediany

- $Z_{1/4} = 1.$ kvartil
- $Z_{2/4} = 2.$ kvartil = median
- $Z_{3/4} = 3.$ kvartil

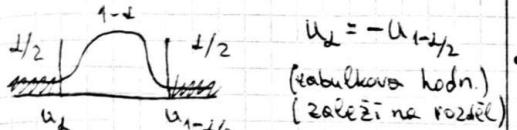
Bodový odhad

- nestraný bodový odhad
 $\mathbb{E} \hat{\theta}(X_1, \dots, X_n) = \theta$
- Konzistentní bodový odhad

- $\lim_{n \rightarrow \infty} \mathbb{E} \hat{\theta}(X_1, \dots, X_n) = \theta$
- $\lim_{n \rightarrow \infty} \text{var} \hat{\theta}(X_1, \dots, X_n) = 0$
- Eficienční bodový odhad - nestraný bodový odhad \Rightarrow nejmenší rozptyl v rozložení $\hat{\theta}(X_1, \dots, X_n)$

Intervalový odhad (CI)

Interval spolehlivosti 95% $\Rightarrow \alpha = 5\%$



$$(\bar{X}_n + U_2 \cdot \frac{s_n}{\sqrt{n}}, \bar{X}_n + U_{1-\alpha/2} \cdot \frac{s_n}{\sqrt{n}})$$

Jednovýberový t-test (hodina 2%)

$$T = \frac{\bar{X}_n - \mu_0}{\sqrt{s_n^2}} \sim t_{n-1}$$

$$T_0 = \frac{\bar{X}_n - \mu_0}{\sqrt{s_n^2}} \cdot \sqrt{n} \quad \text{testovací statistika}$$

Testování: $H_0: \mu = \mu_0$

$H_A: \mu \neq \mu_0$

$|T_0| \geq t_{1-\alpha/2, n-1}$, zamítáme H_0
ve prospěch H_A .

Testování: $H_0: \mu = \mu_0$
 $H_A: \mu > \mu_0$

$T_0 \geq t_{1-\alpha, n-1}$, zamítáme H_0
ve prospěch H_A

$$\bar{X}_{us} = 8,4 \quad n=48$$

$$s_{us}^2 = 2,56 \quad \text{hodina} = 5\%$$

$$H_0: \mu = 8; \quad H_A: \mu \neq 8$$

$$t_{48, 0,975} = 2,01$$

$$T_0 = \frac{8,4 - 8}{\sqrt{2,56}} \cdot \sqrt{48} = 1,75$$

$$|T_0| \not\geq t_{48, 0,975} \Rightarrow$$

NEzamítáme H_0 ve prospěch H_A

Pr. Máme 2 výbery $X \rightarrow m=1, \bar{X}_1=1, S_1=0,2$
 $\alpha = 5\%$
 $\rightarrow n=21, \bar{X}_{21}=1/2, S_{21}=0,3$

1) $H_0: \text{var}X = \text{var}Y; \quad H_A: \text{var}X \neq \text{var}Y$

$$F_0 = \frac{S_{11}^2}{S_{21}^2} = \frac{0,04}{0,09} = 0,4$$

$$F_0 < F_{10, 20, 0,975} \checkmark \quad F_0 > F_{20, 10, 0,025} \checkmark$$

Intervalový odhad $A\ell(p) \rightarrow P$

$$(\bar{X}_n - U_{1-\alpha/2} \cdot \frac{\sqrt{X_{n+1}-\bar{X}_n}}{\sqrt{n}}, \bar{X}_n + U_{1-\alpha/2} \cdot \frac{\sqrt{X_{n+1}-\bar{X}_n}}{\sqrt{n}})$$

Intervalový odhad $P_C(\lambda) \rightarrow \lambda$

$$(\bar{X}_n - U_{1-\alpha/2} \cdot \sqrt{\frac{X_n}{n}}, \bar{X}_n + U_{1-\alpha/2} \cdot \sqrt{\frac{X_n}{n}})$$

Párový t-test

Máme $(y_1, z_1), (y_2, z_2) \dots (y_n, z_n)$

Položime $X_i = y_i - z_i; X_n = y_n - z_n$

Testujeme $H_0: EY_i = EZ_i = \mu_0$

Použijeme jednovýberový t-test
pro $X_1, X_2 \dots X_n$.

Testování shody rozptylu dvoch výb
 $(x_1, \dots, x_n) \sim N(\mu_1, \sigma_1^2) \quad (y_1, \dots, y_n) \sim N(\mu_2, \sigma_2^2)$

$$1) F \sim F_{m-1, n-1}$$

$$2) F_0 = \frac{S_x^2}{S_y^2}$$

$$3) H_0: \sigma_1^2 = \sigma_2^2; \quad H_A: \sigma_1^2 \neq \sigma_2^2$$

$$4) F_0 \leq F_{\alpha/2, m-1, n-1} \text{ or } F_0 \geq F_{1-\alpha/2, m-1, n-1}$$

Zamítáme H_0 ve prospěch H_A

Dvouvýberový t-test

Máme $(x_1, \dots, x_n) \sim N(\mu_1, \sigma^2)$
 $(y_1, \dots, y_n) \sim N(\mu_2, \sigma^2)$

1) Testujeme shodu rozptylu

2) $H_0: \mu_1 - \mu_2 = \mu_0; \quad H_A: \mu_1 - \mu_2 \neq \mu_0$

$$3) T_0 = \frac{\bar{X} - \bar{Y} - \mu_0}{\sqrt{(m-1)s_x^2 + (n-1)s_y^2}} \cdot \sqrt{\frac{mn(m+n-2)}{mn}}$$

4) $|T_0| \geq t_{1-\alpha/2, m+n-2}$, zamítáme
 H_0 ve prospěch H_A .

2) $H_0: EY = EY \Leftrightarrow EX = EY = 0$

$H_A: EX \neq EY$

$$3) T_0 = -1,984 \quad \text{zobrazit kalkulačku s rozložením na } 30,0,975 = 2,9 \quad 2,9 \cdot 0,055 = 2,04$$

4) H_0 nezamítáme ve prospěch H_A

Pearsonov χ^2 test dobré shody

1) H_0 : marginalní pravděpodobnosti jsou rovny hodnotám p_1, p_2, \dots, p_r

H_A : alespoň jedno p_i je jiné

$$2) \chi^2_0 = \sum_{i=1}^r \frac{(O_i - E_i)^2}{E_i}$$

$$3) \chi^2_0 > \chi^2_{1-\alpha, r-1} \Rightarrow \text{zamítáme}$$

H_0 ve prospěch H_A

Test nezávislosti v kontingenční tabulce

Máme náh výb $(y_1, z_1), \dots, (y_n, z_n)$

$y_k = 1, \dots, r; \quad z_k = 1, \dots, c \quad k = 1, \dots, n$

1) H_0 : " $y_k \perp z_k$ jsou v.r. nez."

H_A : " $y_k \perp z_k$ nejsou nez"

$$2) n_{ij} = (y_k=i, z_k=j)$$

matice n ($r \times c$) je kontingenční tabulka

$$3) n_{i.} = \sum_j n_{ij} \quad n_{.j} = \sum_i n_{ij}$$

$$4) \chi^2_0 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - n_{i.} \cdot n_{.j})^2}{n_{i.} \cdot n_{.j}}$$

$$5) \chi^2_0 \geq \chi^2_{1-\alpha, (r-1)(c-1)}$$

zamítáme H_0 ve pr. H_A

Pr. přechodů

1) Pr. přechodů řádu 1

$$P(X_{t+1} = 1, X_1 = 1, X_2 = 3, X_3 = 2) = P(X_0 = 1) \cdot P_{11} \cdot P_{13} \cdot P_{32}$$

2) Pr. přechodů vyšších řádu

$$M=2: \quad P(X_{n+2} = j | X_n = i) = \sum_{k=1}^M P_{ik} P_{kj}$$

$$P^{(1)} = P \cdot P = P^2$$

$$M > 2: \quad P_{ij}^{(m)} = \sum_{k=1}^{M-1} P_{ik} P_{kj} P_{jm} = P^{(m-1)} \cdot P \cdot P^m$$

Pr. 120x hárku kosočka a pozorujeme

1 2 3 4 5 6
15 21 24 18 18 21
Testuje se zda kostka OK

1) $H_0: p_1 = p_2 = \dots = p_6 = \frac{1}{6}$. p_i - pravd. i pravd. k.

H_A : alespoň jedno $p_i \neq \frac{1}{6}$

$$2) \chi^2_0 = \frac{(15-20)^2}{20} + \frac{(21-20)^2}{20} \dots = 4,05$$

$\chi^2_{5, 0, 95} = 11,07$

3) H_0 nezamítáme ve prospěch H_A

Nahodné procesy $\{X_t, t \in T\}$

$\mu_t = E X_t$ - sřední hodnota procesu $\{X_t\}$

Autokovarianční funkce

$$R(s, t) = E(X_s - \mu_s)(X_t - \mu_t)$$

Rozptyl v čase t

$$R(t, t)$$

$R(s, t) = R(s+h, t+h) \Rightarrow$ proces je slabě stacionární

$$P(X_t = x_1, \dots, X_{t+h} = x_h) = P(X_{t+h} = x_1, \dots, X_{t+h} = x_h)$$

\Rightarrow silně stacionární (diskretní)

Markovský řetězec s diskretním časem

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i)$$

$\uparrow \quad \uparrow \quad \uparrow$
prišti krok minulest (nezaleží na minulosti)

Podmíněná pravděpodobnost

$$P(X_{n+m} = j | X_n = i) = p_{ij}(n, n+m) \quad \text{předch řádku } m$$

pr. přechodu ze stavu i do st. j v čase m

Homogenní Markovský řetězec

nezaleží na n , jen na m .

Strukturálně stacionární Mark. řetězec

$$p_j(n) = P(X_n = j) = \pi_j, \quad j \in S \quad \text{je siac}$$

Stacionární rozdělení

$$\begin{aligned} \rightarrow \pi_j &= \sum_{k \in S} \pi_k p_{kj}, j \in S \\ \pi_j &\geq 0, j \in S, \sum_{j \in S} \pi_j = 1 \end{aligned}$$

$\tilde{\pi}_r$

$$\begin{aligned} \pi_1 &= \pi_1 p_{11} + \pi_2 p_{21} + \pi_3 p_{31} \\ \pi_2 &= \pi_1 p_{12} + \pi_2 p_{22} + \pi_3 p_{32} \\ \pi_3 &= \pi_1 p_{13} + \pi_2 p_{23} + \pi_3 p_{33} \\ \pi &= (\pi_1, \pi_2, \pi_3) \\ \pi_1 + \pi_2 + \pi_3 &= 1 \end{aligned}$$

$\tilde{\pi}_r$

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 2/3 \\ 0 & 1/4 & 3/4 \end{pmatrix}$$

$$\begin{array}{c} \text{Diagram of state transitions:} \\ \text{State 1: } 1 \xrightarrow{1/3} 2, 1 \xrightarrow{1/3} 3, 2 \xrightarrow{1/3} 1, 2 \xrightarrow{1/3} 3, 3 \xrightarrow{1/4} 1, 3 \xrightarrow{3/4} 2 \\ \text{State 2: } 1 \xrightarrow{2/3} 3, 2 \xrightarrow{1/3} 1, 2 \xrightarrow{1/3} 3, 3 \xrightarrow{1/4} 2, 3 \xrightarrow{3/4} 1 \\ \text{State 3: } 1 \xrightarrow{1/3} 2, 2 \xrightarrow{1/3} 3, 2 \xrightarrow{1/3} 1, 3 \xrightarrow{1/4} 2, 3 \xrightarrow{3/4} 1 \end{array}$$

$$\begin{aligned} P(\tau_1^+ = \infty) &= \frac{2}{3} \Rightarrow 1. \text{ stav je trvalý} \\ \text{P.s., že narrativem se} \\ P(\tau_2^+ = \infty) &= \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{3}{4} = \frac{2}{3} \cdot \left(\frac{3}{4}\right)^2 = \\ &= \frac{2}{3} \cdot 0 = 0 \Rightarrow 2. \text{ trvalý stav} \\ P(\tau_3^+ = \infty) &= \frac{1}{4} \cdot \left(\frac{1}{3}\right)^{\infty} = 0 \Rightarrow 3. \text{ stav je trvalý} \end{aligned}$$

Absorpní stavy, Nerozložitelné řetězce

\hookrightarrow Nerozložitelný - když stav je dosažitelný z jiného stavu

\hookrightarrow Absorpní - jednobodová mn. stavů $\{j\}$ je uzavřená. $(p_{ij}=1)$ (konečný stav)

\hookrightarrow Absorpní řetězec - vše trvalé stavy jsou absorpní (konečný řet.)

• Perioda Markovského řetězce = počet kroků, který máme udělat, abychom dostali počáteční matice přechodů ($P = P^n$)

• $\tilde{\pi}_r$ 2. ner. komponenty

$$\begin{array}{ll} \text{Diagram:} & \pi_{\text{naře}} = \left(\frac{2}{5}, \frac{3}{5}, 0, 0, 0 \right) \\ \text{Matrix:} & \pi_{\text{dolu}} = \left(0, 0, \frac{1}{7}, \frac{4}{7}, \frac{2}{7} \right) \end{array}$$

$$\text{S nulou: } \pi = c \cdot \pi_{\text{naře}} + (1-c) \cdot \pi_{\text{dolu}}$$

• Klasifikace stavů

τ_j - čas prvního navratu do stavu j

μ_j - střední hodnota doby prvního navratu = $E[\tau_j | X_0=j]$

d_j - největší společný dělitel $\bar{c} \quad n \geq 1$, pro které $p_{ij} > 0$

\hookrightarrow Přechodný stav - doba navratu jde k ∞

$$p_j(\tau_j = \infty) > 0 \quad \lim_{n \rightarrow \infty} p_j^{(n)} = 0$$

\hookrightarrow Trvalý stav

$$p_j(\tau_j < \infty) = 1$$

\hookrightarrow Neronulový stav $\forall j < \infty$

\hookrightarrow Nulový stav: $\mu_j = \infty$

\hookrightarrow Periodický stav: $d_j > 1$, d_j - perioda

\hookrightarrow Neperiodický stav: $d_j = 1$

\hookrightarrow Trvalý + Nulový: $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$

\hookrightarrow Trvalý + Neronulový + Neperiodický: $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \frac{1}{d_j}$

\hookrightarrow Trvalý + Neronulový + Periodický: $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \frac{d_j}{d_j}$

\hookrightarrow Trvalý je Nulový iif $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$

• Dosažitelnost stavů

\hookrightarrow Dosažitelný: existuje cesta $i \rightarrow j$

\hookrightarrow Nedosažitelný $\exists i \text{ do } j \nexists$ neexistuje cesta $i \rightarrow j$

\hookrightarrow Mn. stavů C je uzavřená, když neexistuje cesta do C

\hookrightarrow Mn. stavů C je komponenta, když neexistuje cesta z C do lib. mln. mln. C

• Limithní chování řetězce s absorpními stavy

\hookrightarrow Llkvž. rozložitelný M.R. se stavů

$\hookrightarrow j \in S$, \exists z nich m jsou absorpní

\hookrightarrow Ex. permutace

$$P = \begin{pmatrix} I_m & 0 \\ R & Q \end{pmatrix} \quad I_m \text{ jednotková}$$

\hookrightarrow Pak

$$P^\infty = \lim_{n \rightarrow \infty} P^n = \begin{pmatrix} I_m & 0 \\ M & 0 \end{pmatrix} \quad \text{hledáme}$$

$$\hookrightarrow M = FR, F = (I_{1 \times m} - Q)^{-1}$$

$\uparrow p_{ij}, i = m+1, \dots, M \quad j = 1, \dots, m$ (indexy původní P)
prvky vyjadřujou pšt, že řetězec, který vysel ze stavu i , skončí v absorpním stavu j .

• II metoda

$$1) \text{Dfn. } A_i = P(X_\infty = 3 | X_0 = i)$$

$$A_1 = aA_1 + bA_2 + cA_3$$

$$A_2 = dA_2 + eA_3$$

$$A_3 = 1$$

$$A_4 = 0 \quad (\text{jelikož cyklus, t.j. mln. neustanovné do } 3)$$

$$A_5 = fA_4 + gA_1$$

$$a = P_{11}, b = P_{12}, c = P_{13} \dots$$

$$\text{P.s., že začneme v 1 a skončíme v 3} = A_1$$

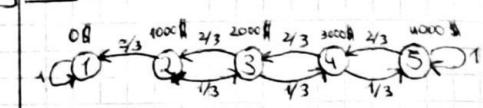
$\tilde{\pi}_r$ faktračování kasino

$$1) \text{Dfn. } A_i = P(X_\infty = 5 | X_0 = i)$$

2)

$$\begin{cases} A_1 = 0 \\ A_2 = \frac{1}{3} A_3 + \frac{2}{3} A_1 \\ A_3 = \frac{2}{3} A_2 + \frac{1}{3} A_4 \dots \\ A_4 = \frac{2}{3} A_3 + \frac{1}{3} A_5 \\ A_5 = 1 \end{cases} \quad \begin{cases} A_1 = 0 \\ A_5 = 1 \\ A_3 = \frac{1}{3} \\ A_4 = \frac{1}{3} \\ A_2 = \dots \end{cases} \quad \text{odpověď}$$

$\tilde{\pi}_r$ Kasino



① a ⑤ jsou konečné stavy

• Cíl: pšt, že vyhráme 4000\$, když přijdeme s 3000\$

$$1) P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2/3 & 0 & 1/3 & 0 & 0 \\ 0 & 2/3 & 0 & 1/3 & 0 \\ 0 & 0 & 2/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$2) P^n = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 2/3 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 2/3 & 0 & 1/3 \\ 0 & 1/3 & 0 & 2/3 & 0 \end{pmatrix} \quad \begin{matrix} R \\ Q \end{matrix}$$

$$3) P^\infty = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ M & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$4) F = \frac{1}{5} \cdot \begin{pmatrix} 7 & 3 & 1 \\ 6 & 9 & 3 \\ 4 & 6 & 7 \end{pmatrix}$$

$$5) M = \frac{1}{5} \begin{pmatrix} 7 & 3 & 1 \\ 6 & 9 & 3 \\ 4 & 6 & 7 \end{pmatrix} \cdot \frac{1}{3} \begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{15} \begin{pmatrix} 14 & 0 & 0 \\ 12 & 3 & 0 \\ 8 & 7 & 1 \end{pmatrix}$$

$$6) P^\infty = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 14/15 & 1/15 & 0 & 0 & 0 \\ 12/15 & 3/15 & 0 & 0 & 0 \\ 8/15 & 7/15 & 0 & 0 & 0 \end{pmatrix} \quad \text{odpověď}$$

Příklad Exp rozdělení

	M	N	J	
Oprava	45	25	5	75
Prodaj	5	5	15	25
	50	30	20	100

Do obchodu chodí 2 člověka za hodinu.

Cíl: Pst, že na dalšího zakazníka počítajícího opravu budeme čekat alespoň 1,5 hodiny.

Řešení:

1) Pst, že člověk potřebuje opravu.
 $\text{Pop} = 0,75$

2) 1 h. - 2 ⋅ 0,75
 $\text{Č. potřebující opravu } 1,5 \text{ h.} - X = 2,25$

2,25 lidí potřebujou opravu $\geq 1,5 \text{ h.}$

3) $X \sim \text{Exp}(1,5)$.

$$P(X \geq 1,5 \text{ h.}) = \int_{1,5}^{\infty} f(x) dx$$

$$= 1 - P(X \leq 1,5 \text{ h.}) = 1 - F(1,5 \text{ h.}) = 1 - (1 - e^{-1,5 \cdot 1,5}) = e^{-2,25}$$

Příklad χ^2 test nezávislosti

	0	1	2	3	4	5
0	L	S _n	B	S _a		
děti	15	5	5	15	40	
dospělí	15	15	25	5	60	
Σ	30	20	30	20	100	

$$\chi^2_0 = \sum_{ij} \frac{(K_{ij} - \frac{K_{i,j} \cdot K_{2,j}}{K_{2,5}})^2}{K_{i,j} \cdot K_{2,j}}$$

$$\chi^2_0 = \frac{(15 - \frac{40 \cdot 30}{100})^2}{\frac{40 \cdot 30}{100}} + \frac{(5 - \frac{40 \cdot 20}{100})^2}{\frac{40 \cdot 20}{100}} + \frac{(5 - \frac{40 \cdot 30}{100})^2}{\frac{40 \cdot 30}{100}} + \frac{(15 - \frac{60 \cdot 20}{100})^2}{\frac{60 \cdot 20}{100}} + \frac{(15 - \frac{60 \cdot 30}{100})^2}{\frac{60 \cdot 30}{100}} + \frac{(15 - \frac{60 \cdot 20}{100})^2}{\frac{60 \cdot 20}{100}} + \frac{(15 - \frac{60 \cdot 30}{100})^2}{\frac{60 \cdot 30}{100}} = 20,14$$

$$\chi^2_0 \geq \chi^2_{1, d.f.} \quad \checkmark$$

Záhláme Ho, jsou závislé

Příklad Poissonovo rozdělení

Stejný příklad

- 1) $\text{Pop} = 0,75$
- 2) 1 h. - 1,5 ē
- 3) $X \sim \text{Po}(2,25)$

1,5 čekáme na ē, který bude potřebovat opravu $\Rightarrow 0$ (adcesti) $\Rightarrow P(X=0) = e^{-2,25} \frac{2,25^0}{0!} = e^{-2,25}$

Příklad CLV

Dáno: děti přichází k automatu s bonbony 1-krát za 2 minuty. V automatu 100 bonbonů, vydává automat alespoň 3 hodiny bez nabídky?

1) 2 min → 1 bonbon

1 h. → 30 bonbonů

2) $X \sim \text{Exp}(30)$ $E[X] = \frac{1}{\lambda} = \frac{1}{30}$ $\text{var}[X] = \frac{1}{\lambda^2} = \frac{1}{30^2}$

3) $P(\sum_i X_i \geq 3 \text{ hodiny}) =$

suma dob čekání (když přijde dítě, vynuluje se, nakonec máme počet minut, když je automat prázdný)

$$= P\left(\frac{\sum X_i - n \cdot E[X]}{\sqrt{n \cdot \text{var}[X]}} \geq \frac{3 - n \cdot E[X]}{\sqrt{n \cdot \text{var}[X]}}\right) = Z \sim N(0,1)$$

$$= P(Z \geq \frac{3 - 100 \cdot \frac{1}{30}}{\sqrt{\frac{1}{300} \cdot 100}}) = P(Z \geq -1) =$$

~~1 - P(Z < -1) = 1 - Φ(-1) = 1 - (1 - Φ(1)) =~~

$$= 1 - Φ(-1) = 1 - Φ(-1) = 1 - (1 - Φ(1)) = \\ = Φ(1) \approx 0,8431$$

$$\chi^2_0 = \frac{(15 - \frac{40 \cdot 30}{100})^2}{\frac{40 \cdot 30}{100}} + \frac{(5 - \frac{40 \cdot 20}{100})^2}{\frac{40 \cdot 20}{100}} + \frac{(5 - \frac{40 \cdot 30}{100})^2}{\frac{40 \cdot 30}{100}} + \frac{(15 - \frac{60 \cdot 20}{100})^2}{\frac{60 \cdot 20}{100}} + \frac{(15 - \frac{60 \cdot 30}{100})^2}{\frac{60 \cdot 30}{100}} + \frac{(15 - \frac{60 \cdot 20}{100})^2}{\frac{60 \cdot 20}{100}} + \frac{(15 - \frac{60 \cdot 30}{100})^2}{\frac{60 \cdot 30}{100}} = 20,14$$

$$\chi^2_0 \geq \chi^2_{1, d.f.} \quad \checkmark$$

Záhláme Ho, jsou závislé

Příklad Mix rozdělení: $X, y \rightarrow \text{Mix}$

Dáno:

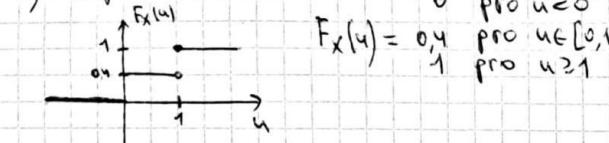
X - diskrétní rozdělení

$$P(X=0) = 0,4, \quad P(X=1) = 0,6$$

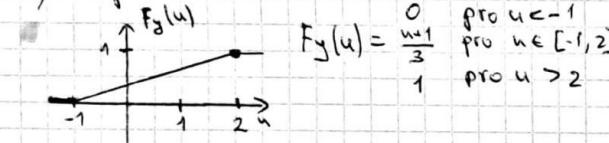
Y - spojité rovnoramné rozdělení na $[-1, 2]$

Z - $\text{Mix}_1(X, y)$

1) Rozepíšeme X



2) Rozepíšeme Y



3) Najdeme $F_Z(u)$

$$F_Z(u) = \frac{1}{4} F_X + \frac{3}{4} F_Y$$

$$\begin{aligned} F_Z(u) &= \frac{1}{4} \cdot 0 + \frac{3}{4} \cdot 0 = 0 & u < -1 \\ &= \frac{1}{4} \cdot 0 + \frac{3}{4} \cdot \frac{u+1}{3} = \frac{u+1}{4} & u \in [-1, 0] \\ &= \frac{1}{4} \cdot \frac{4}{10} + \frac{3}{4} \cdot \frac{u+1}{3} = \frac{10u+14}{40} & u \in [0, 1] \\ &= \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot \frac{u+1}{3} = \frac{u+2}{4} & u \in [1, 2] \\ &= \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot 1 = 1 & u \in [2, \infty) \end{aligned}$$

4) Nacházení $P(Z = \dots)$

$$P(0 \leq Z \leq 1) = F_Z(1) - \lim_{u \rightarrow 0} F_Z(u) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$$

Jiný způsob

$$P(0 \leq Z \leq 1) = (P(0 \leq X \leq 1) + 1 - c)P(0 \leq Y \leq 1)$$

Příklad Mix rozdělení: $\text{Mix} \rightarrow D, S$

Dáno:

$$F_X(u) = \begin{cases} 0 & x \in (-\infty, -2) \\ 0,1x + 0,5 & x \in [-2, -1] \\ 0,4 & x \in [-1, 0] \\ 0,3x + 0,6 & x \in [0, 1] \\ 1 & x \in [1, \infty) \end{cases}$$

1) Najdeme C

$$C = P(X=-2) + P(X=-1) + P(X=0) + P(X=1)$$

$$P(X=-2) = F_X(-2) - \lim_{u \rightarrow -2} F_X(u) = 0,3$$

$$P(X=-1) = F_X(-1) - \lim_{u \rightarrow -1} F_X(u) = 0$$

$$P(X=0) = F_X(0) - \lim_{u \rightarrow 0} F_X(u) = 0,2$$

$$P(X=1) = F_X(1) - \lim_{u \rightarrow 1} F_X(u) = 0,1$$

$$C = 0,3 + 0,2 + 0,1 = 0,6$$

2) Najdeme F_D

$$F_D(u) = \begin{cases} 0 & x \in (-\infty, -2) \\ 0,3 & x \in [-2, 0] \\ 0,3 + 0,2 & x \in [0, 1] \\ 0,3 + 0,2 + 0,1 & x \in [1, \infty) \end{cases}$$

$$F_D(u) = \begin{cases} 0 & x \in (-\infty, -2) \\ 0,5 & x \in [-2, 0] \\ 0,83 & x \in [0, 1] \\ 1 & x \in [1, \infty) \end{cases}$$

3) Najdeme F_S

$$(1-c)F_S(u) = \begin{cases} 0 & x \in (-\infty, -2) \\ 0,1x + 0,2 & x \in [-2, 0] \\ 0,1 & x \in [0, 1] \\ 0,3x + 0,1 & x \in [0, 1] \\ 0,4 & x \in [1, \infty) \end{cases}$$

$$F_S(u) = \begin{cases} 0 & x \in (-\infty, -2) \\ \frac{x+2}{4} & x \in [-2, 0] \\ 0,25 & x \in [0, 1] \\ \frac{3x+1}{4} & x \in [0, 1] \\ 1 & x \in [1, \infty) \end{cases}$$