1. (18 bodů) Náhodná veličina $U$ má spojité rozdělení určené hustotou $f_U$, kde

$$f_U(u) = \begin{cases} \frac{1}{8}\,(u+1), & -1 < u < 3, \\ 0, & \text{jinde}, \end{cases}$$

a náhodná veličina $V$ má diskrétní rozdělení určené pravděpodobnostní funkcí $p_V$, kde $p_V(0) = \frac{1}{3}$, $p_V(1) = \frac{2}{3}$. Náhodná veličina $X$ je jejich směsí, $X = \mathrm{Mix}_{(\frac{3}{4},\frac{1}{4})}(U, V)$.

a) Vypočtěte distribuční funkci $F_X$ náhodné veličiny $X$. *(6 bodů)*

b) Znázorněte $F_X$. *(2 body)*

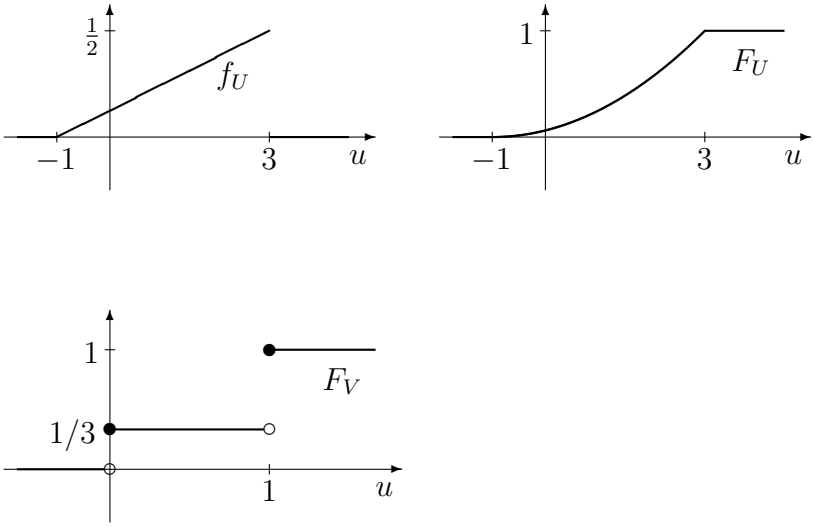
c) Vypočtěte střední hodnotu $\mathrm{E}X$. *(5 bodů)*

d) Uveďte definici kvantilové funkce $q_X$. *(2 body)*

e) Vypočtěte medián $q_X(0.5)$. *(3 body)*

**Řešení:**

a)




Distribuční funkci $F_X$ směsi vypočteme podle vzorce

$$F_X(u) = \frac{3}{4}\,F_U(u) + \frac{1}{4}\,F_V(u)\,, \qquad u \in \mathbb{R}\,. \tag{♣}$$

Distribuční funkci $F_U$ vypočteme pomocí vzorce $F_U(u) = \int_{-\infty}^{u} f_U(t)\,\mathrm{d}t$. Dostaneme:

$F_U(u) = 0$ pro $u \le -1$; $\quad F_U(u) = 1$ pro $u \ge 3$;

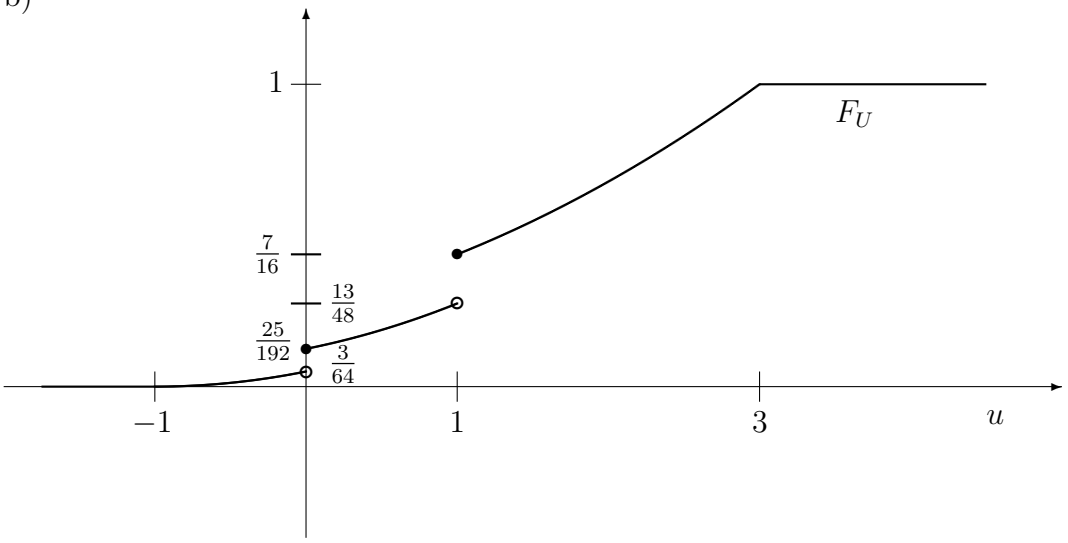
$$F_U(u) = F_U(-1) + \int_{-1}^{u} \frac{1}{8}\,(t+1)\,\mathrm{d}t = 0 + \frac{1}{16}\left[(t+1)^2\right]_{t=-1}^{u} = \frac{1}{16}\,(u+1)^2 \text{ pro } -1 < u < 3.$$

$$F_U(u) = \begin{cases} 0, & u < -1, \\ \frac{1}{16}\,(u+1)^2, & -1 \le u < 3, \\ 1, & u \ge 3, \end{cases} \qquad F_V(u) = \begin{cases} 0, & u < 0, \\ \frac{1}{3}, & 0 \le u < 1, \\ 1, & u \ge 1. \end{cases}$$

Po dosazení do vzorce (♣) dostaneme pro dostribuční funkci $F_X$:

$$\begin{aligned}
u < -1: &\qquad F_X(u) = 0\,;\\
-1 \le u < 0: &\qquad F_X(u) = \frac{\mathbf{3}}{\mathbf{64}}\mathbf{(u+1)^2}\,;\\
0 \le u < 1: &\qquad F_X(u) = \frac{\mathbf{3}}{\mathbf{64}}\mathbf{(u+1)^2} + \frac{\mathbf{1}}{\mathbf{12}}\,;\\
1 \le u < 3: &\qquad F_X(u) = \frac{\mathbf{3}}{\mathbf{64}}\mathbf{(u+1)^2} + \frac{\mathbf{1}}{\mathbf{4}}\,.\\
u \ge 3: &\qquad F_X(u) = 1\,.
\end{aligned}$$

b)




c) Pro střední hodnotu směsi platí vzorec

$$\mathrm{E}X = \frac{3}{4}\,\mathrm{E}U + \frac{1}{4}\,\mathrm{E}V\,. \qquad (\spadesuit)$$

Potom

$$\mathrm{E}U = \int_{-\infty}^{\infty} u f_U(u)\,\mathrm{d}u = \int_{-1}^{3} \frac{1}{8}\,u(u+1)\,\mathrm{d}u = \frac{1}{8}\left[\frac{u^3}{3} + \frac{u^2}{2}\right]_{u=-1}^{3} = \frac{1}{8}\left(9 + \frac{9}{2} + \frac{1}{3} - \frac{1}{2}\right) = \frac{5}{3}\,,$$

$$\mathrm{E}V = \sum_u u\,p_V(u) = 0 \cdot p_V(0) + 1 \cdot p_V(1) = 0 + 1 \cdot \frac{2}{3} = \frac{2}{3}\,.$$

Po dosazení do vzorce (♠) dostaneme pro střední hodnotu

$$\mathrm{E}X = \frac{3}{4} \cdot \frac{5}{3} + \frac{1}{4} \cdot \frac{2}{3} = \frac{5}{4} + \frac{1}{6} = \frac{\mathbf{17}}{\mathbf{12}} \doteq \mathbf{1.417}.$$

d) Pro náhodnou veličinu $X$ a $0 < \alpha < 1$ je hodnotou $q_X(\alpha)$ kvantilové funkce číslo $u_\alpha$, které splňuje podmínky

$$F_X(u_\alpha-) \le \alpha\,, \quad F_X(u_\alpha+) \ge \alpha.$$

Je-li takových čísel víc, tvoří interval, ze kterého vybereme jeho střed. Je-li distribuční funkce $F_X$ spojitá a rostoucí, pak pro kvantilovou funkci platí jednodušší vzorec

$$q_X(\alpha) = F_X^{-1}(\alpha)\,, \qquad 0 < \alpha < 1\,.$$

e) Protože má náhodná veličina $X$ smíšené rozdělení, určíme medián $q_X(0.5) = \tilde{u}$ z podmínek

$$F_X(\tilde{u}-) \le 0.5 \le F_X(\tilde{u}+).$$

Z distribuční funkce v bodech nespojitosti dostaneme:

$$F_X(0-) = \frac{3}{64} < 0.5\,, \qquad F_X(0+) = \frac{3}{64} + \frac{1}{12} = \frac{25}{192} \doteq 0.13 < 0.5\,,$$
$$F_X(1-) = 4 \cdot \frac{3}{64} + \frac{1}{12} = \frac{13}{48} \doteq 0.27 < 0.5\,, \quad F_X(1+) = 4 \cdot \frac{3}{64} + \frac{1}{4} = \frac{7}{16} < 0.5\,.$$

Tedy medián $\tilde{u} = q_X(0.5)$ leží v intervalu $(1,\,3)$ a dostaneme pro něj rovnici

$$F_X(\tilde{u}) = 0.5 \Rightarrow \frac{3}{64}(\tilde{u}^2+1)^2 + \frac{1}{4} = \frac{1}{2} \Rightarrow (\tilde{u}^2+1)^2 = \frac{16}{3} \Rightarrow \tilde{u} = \frac{\mathbf{4}}{\sqrt{\mathbf{3}}} - \mathbf{1} \doteq \mathbf{1.309}\,.$$

2. (18 bodů) Předpokládejme, že náhodná veličina $X$ má normální rozdělení $\mathrm{N}(\mu, \sigma^2)$ s rozptylem $\sigma^2 = 0.22$. Při měření $n = 16$ vzorků jsme zjistili, že výběrový průměr má hodnotu $\bar{\boldsymbol{x}} = 10.2$.

   a) Na hladině významnosti $\alpha = 5$ % otestujte hypotézu, že střední hodnota $\mu$ je **nejvýše rovna** $\mu_0 = 10$. *(9 bodů)*

   b) Mějme nyní hypotézu tvrdící, že na hladině významnosti $\alpha = 5$ % je střední hodnota $\mu$ **právě rovna** $\mu_1 = 9.98$. Můžeme ihned říci, zda ji zamítáme nebo ne, pokud budeme vědět pouze to, jak dopadlo testování hypotézy z části a)? Zdůvodněte svou odpověď konkrétním výpočtem. *(9 bodů)*

**Řešení:**

a) Máme otestovat hypotézu o střední hodnotě při známém rozptylu

$$\mathbf{H}_0 : \mu \le \mu_0\ (= 10)$$

proti alternativní hypotéze:

$$\mathbf{H}_1 : \mu > \mu_0\ (= 10)\ .$$

Můžeme použít testovací statistiku pro hodnotu $\sigma = \sqrt{0.22} \doteq 0.469$:

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

s realizací $t = \frac{\bar{\boldsymbol{x}} - \mu_0}{\sigma}\sqrt{n} \doteq \frac{10.2-10}{0.469} \cdot 4 \doteq 1.706$. Tu porovnáme s kvantilem rozdělení $\mathrm{N}(0,1)$,

$$\Phi^{-1}(1-\alpha) = \Phi^{-1}(0.95) \doteq 1.645\,,$$

hypotézu $\mathbf{H}_0$ **ZAMÍTÁME**.

b) Na první pohled by to svádělo říci, že když jsme zamítli $\mu \leq \mu_0$ $(= 10)$ na hladině $\alpha$, těžko můžeme očekávat, že by mohlo platit $\mu = \mu_1$ $(= 9.98)$ na stejné hladině $\alpha$. Ale kvantily tu teď budeme mít jiné (protože chyba se rozdělí na obě strany a nebude soustředěna jen na jedné straně). Nejdřív tedy zkusíme, jaký vlastně bude výsledek navržené hypotézy.

Máme otestovat hypotézu o střední hodnotě při známém rozptylu

$$\mathbf{H}_0' : \mu = \mu_1 \ (= 9.98)$$

proti alternativní hypotéze:

$$\mathbf{H}_1' : \mu \neq \mu_1 \ (= 9.98) \ .$$

Opět použijeme analogickou testovací statistiku pro hodnotu $\sigma = \sqrt{0.22} \doteq 0.469$ jako v a):

$$T' = \frac{\overline{\boldsymbol{X}} - \mu_1}{\sigma}\sqrt{n}$$

s realizací $t' = \frac{\bar{x}-\mu_1}{\sigma}\sqrt{n} \doteq \frac{10.2-9.98}{0.469} \cdot 4 \doteq 1.876$. Tu porovnáme s kvantilem rozdělení $\mathrm{N}(0,1)$,

$$\Phi^{-1}\left(1 - \tfrac{\alpha}{2}\right) = \Phi^{-1}\left(0.975\right) \doteq 1.96\,,$$

hypotézu $\mathbf{H}_0'$ **NEZAMÍTÁME**.

Není těžké si představit případ (s jinými naměřenými hodnotami), pro který bychom $\mathbf{H}_0'$ naopak zamítli. Obecně tedy nemůžeme říci nic, pokud známe jen odpověď typu zamítnuto/nezamítnuto.

3. (14 bodů) Markovův řetězec má matici přechodu

$$\begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

A. Oklasifikujte všechny stavy. *(2 body)*

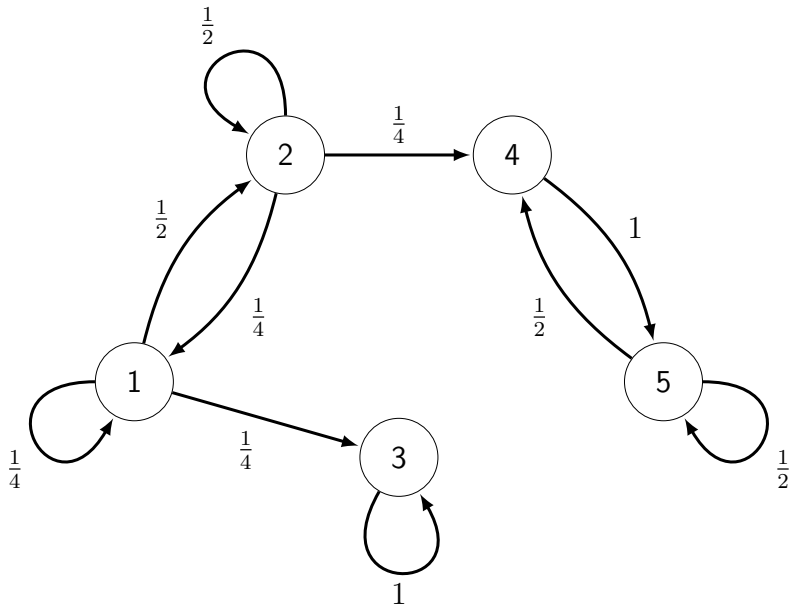
B. Stanovte všechny uzavřené množiny stavů. *(2 body)*

C. Stanovte rozdělení pravděpodobností stavů po třech krocích, jestliže začínáme ve stavu 1. *(3 body)*

D. Odhadněte rozdělení pravděpodobností stavů po 10 000 krocích, jestliže začínáme ve stavu 1. *(7 bodů)*

**Řešení:**

A. Nakreslíme si přechodový graf:




Stavy $1, 2$ jsou přechodné, stav 3 absorpční, $4, 5$ trvalé neperiodické.

B. $\{\emptyset, \{3\}, \{4,5\}, \{3,4,5\}\}$.

C. $\left(\frac{9}{64}, \frac{9}{32}, \frac{23}{64}, \frac{3}{32}, \frac{1}{8}\right)$.

D. Nejprve budeme posuzovat zjednodušený řetězec, v němž sloučíme stavy $4, 5$ do jednoho (označme ho $\{4,5\}$), který bude absorpční. Permutací stavů $(3, \{4,5\}, 1, 2)$ dostaneme matici přechodu ve tvaru

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_2 & \boldsymbol{0} \\ \boldsymbol{R} & \boldsymbol{Q} \end{pmatrix},$$

$$\boldsymbol{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{R} = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}, \qquad \boldsymbol{Q} = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix}.$$

Fundamentální matice je

$$\boldsymbol{F} = (\boldsymbol{I}_2 - \boldsymbol{Q})^{-1} = \begin{pmatrix} \frac{3}{4} & -\frac{1}{2} \\ -\frac{1}{4} & \frac{1}{2} \end{pmatrix}^{-1} = 4 \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}.$$

Matice

$$\boldsymbol{F}\,\boldsymbol{R} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

udává asymptotické pravděpodobnosti dosažení absorpčních stavů z přechodných. Její první řádek odpovídá zadání, že začínáme ve stavu 1. Pravděpodobnosti po 10 000 krocích budou blízké asymptotickým. Asymptotické pravděpodobnosti absorpčních stavů zjednodušeného řetězce, tj. komponent $\{3\}$, $\{4,5\}$ původního řetězce, jsou $\frac{1}{2}$, asymptotické pravděpodobnosti přechodných stavů jsou nulové.

Zbývá dořešit rozdělení pravděpodobností uvnitř komponenty $\{4,5\}$. Ta má matici přechodu

$$\begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

je ergodická, rozdělení pravděpodobností konverguje ke stacionárnímu, $\left(\frac{1}{3}, \frac{2}{3}\right)$. Je třeba vynásobit pravděpodobností $\frac{1}{2}$, že v této komponentě skončíme. Asymptotické rozdělení pravděpodobností všech původních stavů je $\left(0, 0, \frac{1}{2}, \frac{1}{6}, \frac{1}{3}\right)$.