1. (15 bodů) Náhodná veličina X má spojité rovnoměrné rozdělení v intervalu $(-3, 4)$ a náhodná veličina $Y = h(X)$ je definována pomocí funkce

$$h(x) = \begin{cases} -2, & x \le -2, \\ x, & -2 < x < 2, \\ 2, & x \ge 2. \end{cases}$$

a) Vyjádřete náhodnou veličinu Y jako směs spojité a diskrétní náhodné veličiny. *(11 bodů)*

b) Vypočtěte střední hodnotu $\mathrm{E}Y$. *(4 body)*

Řešení:

Náhodná veličina X má rozdělení určené distribuční funkcí F_X a náhodná veličina $Y \in \langle -2, 2\rangle$ s distribuční funkcí F_Y, kde

$$F_X(u) = \begin{cases} 0, & u < -3, \\ \frac{u+3}{7}, & -3 \le u < 4, \\ 1, & u \le 4. \end{cases} \qquad F_Y(u) = \begin{cases} 0, & x < -2, \\ \frac{u+3}{7}, & -2 \le x < 2, \\ 1, & -2 \le x. \end{cases}$$

Odtud je $P(Y = -2) = P(X \le -2) = \frac{1}{7}$ a $P(Y = 2) = P(X \ge 2) = \frac{2}{7}$.

a) Označme $Y = \mathrm{Mix}_\alpha(U, V)$, kde U je náhodná veličina s diskrétním rozdělením a V je náhodná veličina, která má spojité rozdělení. Potom pro distribuční funkce platí:

$$F_Y(u) = \alpha F_U(u) + (1-\alpha) F_V(u), \; u \in \mathbb{R}\,.$$

Je $U \in \{-2, 2\}$ a $P(Y = -2 \vee Y = 2) = \frac{3}{7} \Rightarrow \alpha = \frac{3}{7}$ a $1 - \alpha = \frac{4}{7}$.

Pro pravděpodobnostní funkci náhodné veličiny U a hustotu náhodné veličiny V dostaneme:

$$p_U(u) = \frac{7}{3}\big(F_X(u) - F_X(u-)\big), \quad p_U(-2) = \frac{7}{3} \cdot \frac{1}{7} = \frac{1}{3}, \quad p_U(2) = \frac{7}{3} \cdot \frac{2}{7} = \frac{2}{3},$$

$$f_V(u) = \frac{7}{4} F_X'(u), \quad f_V(u) = \begin{cases} \frac{7}{4} \cdot \frac{1}{7} = \frac{1}{4}, & -2 < u < 2, \\ f_V(u) = 0, & \text{jinde.} \end{cases}$$

Pro distribuční funkce pak dostaneme

$$F_U(u) = \begin{cases} 0, & u < 2, \\ \frac{1}{3}, & -2 \le u < 2, \\ 1, & 2 \le u, \end{cases} \qquad F_V(u) = \begin{cases} 0, & u < 2, \\ \frac{u+2}{4}, & -2 \le u < 2, \\ 1, & 2 \le u. \end{cases}$$

Je tedy $Y = \mathrm{Mix}_{3/7, 4/7}(U, V)$. Potom je distribuční funkce náhodné veličiny Y rovna

$$F_Y(u) = \frac{3}{7} F_U(u) + \frac{4}{7} F_V(u), \; u \in \mathbb{R}.$$

b) Pro střední hodnotu náhodné veličiny Y dostaneme vzorec

$$\mathrm{E}Y = \frac{3}{7}\mathrm{E}U + \frac{4}{7}\mathrm{E}V\,. \tag{$\spadesuit$}$$

$$\mathrm{E}U = \sum_u u\, p_U(u) = -2 \cdot p_U(-2) + 2 \cdot p_U(2) = -2 \cdot \frac{1}{3} + 2 \cdot \frac{2}{3} = -\frac{2}{3} + \frac{4}{3} = \frac{2}{3}\,,$$

$$\mathrm{E}V = \int_{-\infty}^{\infty} u f_V(u)\,\mathrm{d}u = \int_{-2}^{2} \frac{u}{4}\,\mathrm{d}u = \left[\frac{u^2}{8}\right]_{u=-2}^{2} = 0\,.$$

Po dosazení do ($\spadesuit$) dostaneme

$$\mathrm{E}Y = \frac{3}{7} \cdot \frac{2}{3} + \frac{4}{7} \cdot 0 = \boldsymbol{\frac{2}{7}} \doteq \mathbf{0.2857}\,.$$

2. (15 bodů) Náhodný vektor (X, Y) má diskrétní rozdělení, četnosti jeho hodnot jsou uvedeny v tabulce.

	1	2	3
1	20	40	30
2	10	10	10

Testujte na hladině významnosti $\alpha = 0.05$ hypotézu
H_0 : náhodné veličiny X a Y jsou nezávislé
proti alternativě
H_1 : náhodné veličiny X a Y jsou závislé.

Řešení:

Nejprve určíme odhady marginálních pravděpodobnostních funkcí, které zapíšeme do rozšířené tabulky

	1	2	3	q_j
1	20	40	30	$q_1 = \frac{3}{4}$
2	10	10	10	$q_2 = \frac{1}{4}$
p_i	$p_1 = \frac{1}{4}$	$p_2 = \frac{5}{12}$	$p_3 = \frac{1}{3}$	

Použijeme testovací statistiku

$$T = \sum_{i=1}^{3} \sum_{j=1}^{2} \frac{(N_{ij} - np_i q_j)^2}{np_i q_j} \sim \chi^2(2).$$

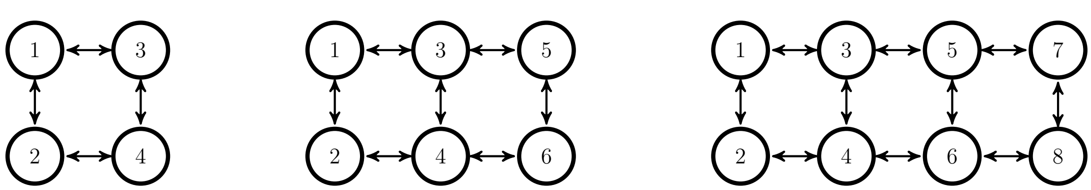
Kritickou hodnotou testu je kvantil $q_{\chi^2(2)}(0.95) \doteq 5.99$.

Po dosazení hodnot z tabulky dostaneme pro realizaci testovací statistiky hodnotu

$$t = \frac{2}{45}\left(20 - \frac{45}{2}\right)^2 + \frac{2}{15}\left(10 - \frac{15}{2}\right)^2 + \frac{2}{75}\left(40 - \frac{75}{2}\right)^2 + \frac{2}{25}\left(10 - \frac{25}{2}\right)^2 + \frac{1}{30} \cdot 0 + \frac{1}{10} \cdot 0 =$$
$$= \frac{2 \cdot 25}{45 \cdot 4} + \frac{2 \cdot 25}{15 \cdot 4} + \frac{2 \cdot 25}{75 \cdot 4} + \frac{2 \cdot 25}{25 \cdot 4} = \frac{5}{18} + \frac{5}{6} + \frac{1}{6} + \frac{1}{2} = \frac{5 + 15 + 3 + 9}{18} = \frac{\mathbf{32}}{\mathbf{18}} \doteq \mathbf{1.7778}\,.$$

Protože je $t \doteq 1.778 < \mathbf{q}_{\chi^2(2)}(\mathbf{0.95}) \doteq \mathbf{5.99}$, hypotézu H_0 **nezamítáme**.

3. (15 bodů) Tři Markovovy řetězce jsou dány přechodovými diagramy na obrázcích. V každém uzlu jsou všechny z něj vycházející hrany stejně pravděpodobné. Ve 3 krocích jsme přešli ze stavu 1 do stavu 3. Který z daných Markovových řetězců nejlépe vyhovuje tomuto pozorování? *(12 bodů)*
Pro tento řetězec oklasifikujte všechny stavy. *(3 body)*



Řešení:

Tabulka ukazuje všechny posloupnosti stavů, které přicházejí v úvahu, a jejich pravděpodobnosti přechodů. První krok má vždy pravděpodobnost 1/2, což na *porovnání* věrohodností nemá vliv; analyzujeme jen zbývající dva kroky mezi stavy x_1, x_2, x_3.

posloupnost stavů	$p_{x_1x_2}$	$p_{x_2x_3}$	$p_{x_1x_2} \cdot p_{x_2x_3}$	$p_{x_1x_2}$	$p_{x_2x_3}$	$p_{x_1x_2} \cdot p_{x_2x_3}$	$p_{x_1x_2}$	$p_{x_2x_3}$	$p_{x_1x_2} \cdot p_{x_2x_3}$
$(1,3,1,3)$	1/2	1/2	1/4	1/3	1/2	1/6	1/3	1/2	1/6
$(1,3,5,3)$	0	0	0	1/3	1/2	1/6	1/3	1/3	1/9
$(1,2,1,3)$	1/2	1/2	1/4	1/2	1/2	1/4	1/2	1/2	1/4
$(1,2,4,3)$	1/2	1/2	1/4	1/2	1/3	1/6	1/2	1/3	1/6
$(1,3,4,3)$	1/2	1/2	1/4	1/3	1/3	1/9	1/3	1/3	1/9
$\sum$			1			31/36			29/36
věrohodnost			**1/2**			**31/72**			**29/72**

Největší věrohodnost dává první řetězec. Všechny jeho stavy jsou trvalé s periodou 2 a tvoří jedinou komponentu.