Identifikační číslo ČVUT: _____ **Podpis**: _____

Úloha	1	2	3	4	*Celkem*
Maximum	25	25	25	25	100
Počet bodů					

1. (25 bodů) Náhodný vektor (X, Y) má kovarianční matici $\Sigma_{(X,Y)} = \begin{pmatrix} 4 & -3 \\ -3 & 5 \end{pmatrix}$. Určete:

   a) korelační matici $\varrho_{(X,Y)}$ náhodného vektoru (X, Y);

   b) korelační matici $\varrho_{(U,V)}$ náhodného vektoru (U, V), kde $U = 2X + 1$ a $V = -3Y$.

Řešení:

a) Z kovarianční matice dostaneme, že
$\mathrm{D}X = \mathrm{cov}(X, X) = 4,\ \mathrm{D}Y = \mathrm{cov}(Y, Y) = 5 \Rightarrow$

$$\varrho(X, Y) = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{D}X\,\mathrm{D}Y}} = \frac{-3}{\sqrt{4 \cdot 5}} \doteq -0.6708\,.$$

Korelační matice náhodného vektoru (X, Y) je tedy

$$\varrho_{(X,Y)} \doteq \begin{pmatrix} 1 & -0.6708 \\ -0.6708 & 1 \end{pmatrix}.$$

b) Z vlastností rozptylu a kovariance dostaneme

$$\begin{aligned} \mathrm{cov}(U, U) &= \mathrm{D}U = \mathrm{D}(2X + 1) = 4\,\mathrm{D}X = 16\,, \\ \mathrm{cov}(V, V) &= \mathrm{D}V = \mathrm{D}(-3Y) = 9\,\mathrm{D}Y = 45\,, \\ \mathrm{cov}(U, V) &= \mathrm{cov}(2X + 1, -3Y) = 2 \cdot (-3) \cdot \mathrm{cov}(X, Y) = (-6) \cdot (-3) = 18\,, \\ \varrho(U, V) &= -\varrho(X, Y) \doteq 0.6708\,. \end{aligned}$$

Kovarianční a korelační matice náhodného vektoru (U, V) jsou

$$\Sigma_{(U,V)} = \begin{pmatrix} 16 & 18 \\ 18 & 45 \end{pmatrix}, \qquad \varrho_{(U,V)} \doteq \begin{pmatrix} 1 & 0.6708 \\ 0.6708 & 1 \end{pmatrix}.$$

2. (25 bodů) Počty narozených dětí v České republice během let 2006 až 2011 (zaokrouhleny na tisíce) jsou uvedeny v tabulce. Testujte na hladině významnosti 5% hypotézu, že mají data rovnoměrné rozdělení. Posuďte předpoklady a použitelnost metody.

rok	2006	2007	2008	2009	2010	2011
četnost	106 000	115 000	120 000	118 000	117 000	107 000

Řešení:

Použijeme χ^2-test dobré shody. Máme $k = 6$ tříd a celkový počet je $n = 683\,000$. Předpokládáme rovnoměrné rozdělení, tedy pravděpodobnosti příslušnosti k jednotlivým třídám jsou stejné, a to $p_i = 1/6,\ 1 \le i \le 6$, $n\,p_i \doteq 113\,833$. Statistika testu je dána vzorcem

$$T = \sum_{i=1}^{6} \frac{(n_i - np_i)^2}{np_i}$$

a má přibližně rozdělení χ^2 s 5 stupni volnosti. Po dosazení získáme její realizaci

$$t = \frac{1}{113\,833}\left(7\,833^2 + 1\,167^2 + 6\,167^2 + 4\,167^2 + 3\,167^2 + (-6\,833)^2\right) \doteq$$

$$\doteq \frac{10^6}{113\,833}(61.356 + 1.362 + 38.032 + 17.363 + 10.029 + 46.69) =$$

$$= \frac{174.833 \cdot 10^6}{113\,833} \doteq 1\,536\,.$$

Porovnáme s kvantilem $q_{\chi^2(5)}(0.95) \doteq 11.07$ rozdělení $\chi^2(5)$. Protože je $t = 1\,536 \gg 11.07$, hypotézu, že data jsou výběrem z rovnoměrného rozdělení, **zamítáme**, a to na velmi vysoké hladině významnosti. Je to tím, že data neodpovídala přesně rovnoměrnému rozdělení a bylo jich mnoho. Na výsledku nic nemění skutečnost, že se předpokládanému jen asymptoticky blížíme a že data byla zaokrouhlena.

3. (25 bodů) Při trojnásobném souboji v paintballu po sobě současně střílí tři soupeři A, B a C. V každém kole každý nevyřazený hráč jednou vystřelí na nejlepšího střelce z ostatních. Kola se opakují tak dlouho, až zbyde jediný vítěz nebo jsou všichni vyřazeni. Střelec A strefí soupeře s pravděpodobností 1/2, střelec B s pravděpodobností 1/3 a střelec C s pravděpodobností 1/4. Sestrojte matici přechodu a vypočtěte pravděpodobnosti vítězství jednotlivých střelců. (*Návod: Uvažujte řetězec o sedmi stavech ABC, AC, BC, A, B, C a $\emptyset$.*)

Řešení:

Uvažujme Markovův řetězec o sedmi stavech. Stav 1 je absorpční stav (všichni byli zasaženi), stavy 2, 3 a 4 jsou absorpční stavy (po řadě A, B, C vyhrál), stavy 5, 6 a 7 jsou přechodné (po řadě AC, BC a ABC jsou stále ve hře). Kanonický tvar matice přechodu je

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1/8 & 3/8 & 0 & 1/8 & 3/8 & 0 & 0 \\ 1/12 & 0 & 1/4 & 1/6 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix} = \begin{pmatrix} \mathbf{I}_4 & \mathbf{0} \\ \mathbf{R} & \mathbf{Q} \end{pmatrix},$$

kde

$$\mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{R} = \begin{pmatrix} 1/8 & 3/8 & 0 & 1/8 \\ 1/12 & 0 & 1/4 & 1/6 \\ 0 & 0 & 0 & 1/4 \end{pmatrix}, \qquad \mathbf{Q} = \begin{pmatrix} 3/8 & 0 & 0 \\ 0 & 1/2 & 0 \\ 1/4 & 1/4 & 1/4 \end{pmatrix}.$$

Fundamentální matice řetězce je

$$\mathbf{F} = \big(\mathbf{I}_3 - \mathbf{Q}\big)^{-1} = \begin{pmatrix} 5/8 & 0 & 0 \\ 0 & 1/2 & 0 \\ -1/4 & -1/4 & 3/4 \end{pmatrix}^{-1} = \frac{1}{15}\begin{pmatrix} 24 & 0 & 0 \\ 0 & 30 & 0 \\ 8 & 10 & 20 \end{pmatrix}$$

$$\mathbf{F}\,\mathbf{R} = \begin{pmatrix} 1/5 & 3/5 & 0 & 1/5 \\ 1/6 & 0 & 1/2 & 1/3 \\ 11/90 & 1/5 & 1/6 & 23/45 \end{pmatrix}.$$

Pravděpodobnost výhry hráče A je 1/5, hráče B 1/6 a hráče C 23/45. Numericky 0.2, 0.167 a 0.511.

4. (25 bodů) Hodíme dvěma kostkami. Náhodná veličina X je rovna jedné, jestliže součet na kostkách se rovná 6, jinak se rovná nule. Veličina Y se rovná hodnotě na první kostce. Určete $H(Y)$, $H(X|Y)$ a $H(X,Y)$.

Řešení:

Pro veličiny X a Y platí

$$p_X(0) = 31/36, \quad p_X(1) = 5/36, \quad p_Y(i) = 1/6, \quad i = 1, \ldots, 6.$$

Pro náhodný vektor (X, Y) dostaneme tabulku

$p_{XY}(x,y)$	$x=0$	$x=1$
$y=1$	$\frac{5}{36}$	$\frac{1}{36}$
$y=2$	$\frac{5}{36}$	$\frac{1}{36}$
$y=3$	$\frac{5}{36}$	$\frac{1}{36}$
$y=4$	$\frac{5}{35}$	$\frac{1}{36}$
$y=5$	$\frac{5}{36}$	$\frac{1}{36}$
$y=6$	$\frac{1}{6}$	0

a pro podmíněné pravděpodobnosti tabulku

$p_{X\|Y}(x\|y)$	$x=0$	$x=1$
$y=1$	$\frac{5}{6}$	$\frac{1}{6}$
$y=2$	$\frac{5}{6}$	$\frac{1}{6}$
$y=3$	$\frac{5}{6}$	$\frac{1}{6}$
$y=4$	$\frac{5}{6}$	$\frac{1}{6}$
$y=5$	$\frac{5}{6}$	$\frac{1}{6}$
$y=6$	1	0

Dostaneme $H(Y) = \log 6 \doteq 2.585$, $H(X,Y) = \frac{25}{36}\log\frac{36}{5} + \frac{1}{6}\log 6 + \frac{5}{36}\log 36 \doteq 3.1267$ a $H(X|Y) = \frac{5}{6}\left(\frac{5}{6}\log\frac{6}{5} + \frac{1}{6}\log 6\right) \doteq 0.5417$.