Facit till tenta 14-03-2022

U1 (a) Med $S=\{(1,1),(1,2),\dots,(1,6),(2,1),\dots,(6,6)\}$ och $X((i,j))=\min\{i,j\}$ följer att möjliga värden för X är $\{1,2,3,4,5,6\}$.

Motsvarande slh. funktion ges av

k	1	2	3	4	5	6
$p_X(k)$	$11/36$	$9/36$	$7/36$	$5/36$	$3/36$	$1/36$

exempelvis:
$p_X(3) = \mathbb{P}(X=3) =$
$\mathbb{P}(\{(3,3),(3,4),(3,5),(3,6),(4,3),(5,3),(6,3)\}) = \frac{7}{36}$.

(b) $\mathbb{E}X = \sum_{k=1}^{6} k\cdot p_X(k) = 1\cdot\frac{11}{36}+\dots+6\cdot\frac{1}{36} = \frac{91}{36}$

$\mathbb{E}X^2 = \sum_{k=1}^{6} k^2\, p_X(k) = 1\cdot\frac{11}{36}+4\cdot\frac{9}{36}+\dots+36\,\frac{1}{36} = \frac{301}{36} \quad\Rightarrow \operatorname{var}(X) = \mathbb{E}X^2-(\mathbb{E}X)^2 = \frac{2555}{1296} \approx 1.97$

(c) Med $B=\{X=4\}$ och $A=\{(4,4)\}$ följer $A\cap B = A$ och $\mathbb{P}(A\mid B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \frac{1/36}{5/36} = \frac{1}{5}$.

U2 (a) Då $n=65$ kan fördelningen av $\overline{X}_n$ approximeras med normalfördelningen $\mathcal{N}(\mu,\frac{\sigma^2}{n})$ enligt Centrala gränsvärdessatsen (CGS). Här $\mu=24$, $\sigma^2/n = \frac{100}{65} = \frac{20}{13}$

Med $\overline{X}_n \sim \mathcal{N}(24,\frac{20}{13})$ approx. följer:

$$\mathbb{P}(22\le \overline{X}_n \le 25) = \mathbb{P}\left(\frac{-2}{10/\sqrt{65}} \le \underbrace{\frac{\overline{X}_n-24}{10/\sqrt{65}}}_{=Z\sim\mathcal{N}(0,1)\text{ approx.}} \le \frac{1}{10/\sqrt{65}}\right) \approx \Phi\left(\frac{\sqrt{65}}{10}\right)-\Phi\left(-\frac{\sqrt{65}}{5}\right)$$

$$= \Phi(0.806)-\Phi(-1.612) \overset{A.3}{\approx} 0.79 - 0.0535 = 0.735$$

(b) $Y\sim \mathrm{Bin}(65, 0.04)$

Då $np = 2.6 < 10$ lämpar sig inte normalfördelningen utan Poisson-fördelningen (med samma väntevärde) här, dvs. $\mathrm{Bin}(65,0.04)\approx \mathrm{Poiss}(\underbrace{2.6}_{=\lambda})$.

$\underbrace{\mathbb{P}(Y=3)}_{=\binom{65}{3}0.04^3\cdot 0.96^{62}\approx 0.2225 \text{ exakt.}} \approx e^{-\lambda}\frac{\lambda^3}{3!} \approx 0.2176$ enligt Poisson approx.

$= \binom{65}{3}\,0.04^3\cdot 0.96^{62} \approx 0.2225$ exakt.

U3 $f(x,y) = 2x\cdot e^{-y}$ för $x\in[0,1]$, $y\ge 0$ är den gemensamma täthetsfkt.

(a) $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^{\infty} 2x\,e^{-y}\,dy = 2x\,[-e^{-y}]_0^{\infty} = 2x$ för $x\in[0,1]$, 0 annars.

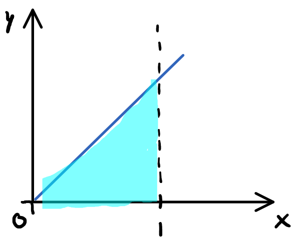
$f_Y(y) = \int_0^1 2x\,e^{-y}\,dx = e^{-y}$ för $y\in[0,\infty)$.

$\mathbb{E}X = \int_0^1 x\cdot 2x\,dx = \left[\frac{2x^3}{3}\right]_0^1 = \frac{2}{3}$

(b) $f(x,y) = f_X(x)\cdot f_Y(y)$ för alla $(x,y)\in\mathbb{R}^2 \Rightarrow X,Y$ är oberoende

(c) $\mathbb{P}(X\ge Y) = \iint f(x,y)\,dy\,dx = \int_0^1\int_0^x 2x\,e^{-y}\,dy\,dx$

$= \int_0^1 2x(1-e^{-x})\,dx = \frac{4}{e}-1 \approx 0.47.$



(U4) (a) Första stickprovs moment: $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$

Första moment av $X \sim \text{unif}(\{1,2,\dots,\theta\})$ är $\mathbb{E}X = \frac{1}{\theta}\sum_{k=1}^{\theta} k = \frac{\theta+1}{2}$

Likställer man dessa fås $\hat{\theta}_{MM} = \frac{2}{n}\sum_{k=1}^{n} X_k - 1$

(b) En skattare $\hat{\theta}$ för parametern θ är VVR om $\mathbb{E}\hat{\theta} = \theta$ och konsistent om $\operatorname{var}(\hat{\theta}) \to 0$ när $n \to \infty$.

Här $\mathbb{E}\hat{\theta}_{MM} = \frac{2}{n}\sum_{k=1}^{n} \underbrace{\mathbb{E}X_k}_{=\mathbb{E}X = \frac{\theta+1}{2}} - 1 = \frac{2}{n}\cdot n \cdot \left(\frac{\theta+1}{2}\right) - 1 = \theta \to \hat{\theta}_{MM}$ är VVR.

$\operatorname{var}(\hat{\theta}_{MM}) = \operatorname{var}\left(\frac{2}{n}\sum_{k=1}^{n} \mathbb{E}X_k\right) = \frac{4}{n^2}\sum_{k=1}^{n}\operatorname{var}(X_k) = \frac{4}{n}\cdot \operatorname{var}(X) \to 0$ när $n \to \infty$ (konstant m.a.p. n)

$\to \hat{\theta}_{MM}$ är även konsistent.

(c) $\text{unif}(\{1,\dots,\theta\})$ har slh. fkt. $p_X(k) = \frac{1}{\theta}$ för $k \in \{1,\dots,\theta\}$

$\Rightarrow L(x_1,\dots,x_n;\theta) = \left(\frac{1}{\theta}\right)^n$ om $x_i \in \{1,\dots,\theta\}$ för alla $i = 1,\dots,n$; 0 annars

Maximeras genom att välja $\theta = \max\{x_1,\dots,x_n\}$, vilket är det minsta möjliga värde på θ utan att göra L lika med noll. Alltså

$\hat{\theta}_{ML} = \max\{X_1,\dots,X_n\}$

Då $X_1,\dots,X_n$ är ett stickprov ifrån $X \sim \text{unif}(\{1,\dots,\theta\})$ är $\mathbb{P}(X \le \theta) = 1$ och och därmed $\mathbb{P}(\hat{\theta}_{ML} \le \theta) = \mathbb{P}(X_1 \le \theta, \dots, X_n \le \theta) = \mathbb{P}(X \le \theta)^n = 1$

Dock är $\mathbb{P}(\max\{X_1,\dots,X_n\} = \theta) < 1 \Rightarrow \mathbb{E}\hat{\theta}_{ML} < \theta$ (underskattas parametern systematiskt) $\Rightarrow \hat{\theta}_{ML}$ ej VVR.

(U5) (a) För det givna stickprovet är

$\bar{x} = \frac{150}{10} = 15$ och $s^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2 = \frac{710.92}{9} \approx 79.0$

(b) Med hjälpvariabeln $T = \frac{\overline{X}_n - \mu}{S/\sqrt{n}} \sim t_{n-1}$ och motsvarande kvantil från t_9 (tabell A.5): $t_{0.025,9} = 2.262$ ($\alpha = 0.05$, tvåsidigt) fås

$I_\mu = \left(\bar{x} - t_{0.025,9}\frac{s}{\sqrt{n}}, \bar{x} + t_{0.025,9}\frac{s}{\sqrt{n}}\right) = (8.64, 21.36)$

(c) Med hjälpvariabeln $X = \frac{n-1}{\sigma^2}S^2 \sim \chi^2_{n-1}$ och motsvarande kvantiler från χ^2_9 (tabell A.7): $\chi^2_{0.975,9} = 2.7$ och $\chi^2_{0.025,9} = 19.022$ fås

$I_{\sigma^2} = \left(\frac{n-1}{\chi^2_{0.025,9}}\cdot s^2, \frac{n-1}{\chi^2_{0.975,9}} s^2\right) = (37.38, 263.33)$

(d) $H_0: \mu \le 11 \; (=\mu_0)$, $H_1: \mu > 11$ (ensidig hypotesprövning, $\alpha = 0.05$)

Testvariabeln T (se ovan) får utfallet $t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{4}{\sqrt{7.9}} = 1.423$

Med $t_{0.05,9} = 1.833$ (A.5) kan vi dra slutsatsen att $t_0 < 1.833$, alltså inte ligger i det kritiska området och behåller därför H_0.

$\to$ Påståendet är inte orimligt m.a.p. det givna stickprovet (och signifikans $\alpha = 0.05$).

(46) Stickprov från två populationer, båda normalfördelade
$(X_1,\dots,X_{n_1})$ med $n_1 = 25,\ \bar{x} = 5.8,\ \sigma_1^2 = 3$
$(Y_1,\dots,Y_{n_2})$ med $n_2 = 22,\ \bar{y} = 4.9,\ \sigma_2^2 = 2.5$

(a) $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$
$\alpha = 0.1$, tvåsidigt test

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$ är $N(0,1)$ och givet nollhypotesen $(\mu_1 = \mu_2)$

får testvariabeln värdet $z_0 = \frac{\bar{x} - \bar{\mu}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{0.9}{\sqrt{\frac{3}{25} + \frac{2.5}{22}}} \approx 1.862$

$z_{\alpha/2} = z_{0.05} = 1.645$ (A.3) vilket betyder att z landar utanför $(-z_{\alpha/2}, z_{\alpha/2})$
och vi förkastar H_0 ($\rightarrow$ finns anledning att anta $\mu_1 \neq \mu_2$ istället).

(b) $\Phi(1.862) \approx \Phi(1.86) = 0.9682$ (A.3)
$\Rightarrow P = 2 \cdot \mathbb{P}(Z \geq z_0) = 2(1 - 0.9682) = 0.0628$

(c) Om $\mu_1 - \mu_2 = \frac{1}{2}$, då är $Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ inte längre $N(0,1)$-fördelad
utan har väntevärde

$$\mathbb{E}Z = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = 1.0344$$ och varians 1.

Slh. att förkasta H_0 när den är (så mycket) fel blir då
$\mathbb{P}(Z \notin (-1.645, 1.645)) = 1 - \mathbb{P}(-1.645 \leq Z \leq 1.645) = 1 - \beta$
där slh. för typ II-fel är

$$\beta = \mathbb{P}(-2.68 \leq \underbrace{Z - \mathbb{E}Z}_{\sim N(0,1)} \leq 0.61) = \Phi(0.61) - \Phi(-2.68) = 0.7291 - 0.0037 = 0.7254$$

$\Rightarrow$ Testets effektivitet är då $1 - \beta = 0.2746$ (inte så strålande faktiskt ...).

(47) (a) För $A_0 = \{X = 0\}, A_1 = \{X = 1\}, A_2 = \{X = 2\}, A_3 = \{X = 3\}, A_4 = \{X = 4\}$
och $A_5 = \{X \geq 5\}$ där $X \sim \text{Poiss}(2)$
fås $p_0 = \mathbb{P}(A_0) = e^{-2} \cdot \frac{2^0}{0!} = \frac{1}{e^2} \approx 0.135$, $p_1 = \mathbb{P}(A_1) = e^{-2}\frac{2^1}{1!} = \frac{2}{e^2} \approx 0.271$
$p_2 = e^{-2}\frac{2^2}{2!} = p_1$, $p_3 = \frac{4}{3e^2} \approx 0.18$, $p_4 = \frac{2}{3e^2} \approx 0.09$ och
$p_5 = 1 - \sum_{k=0}^{4} p_k = 1 - \frac{7}{e^2} \approx 0.053$.

(b) Med $n = 100$ har alla 6 kategorier ett förväntat antal $np_i \geq 5$,
och antalet observationer är $(N_0, \dots, N_5) = (16, 20, 26, 24, 5, 9)$.

$H_0: (p_0, p_1, \dots, p_5) = \left(\frac{1}{e^2}, \frac{2}{e^2}, \frac{2}{e^2}, \frac{4}{3e^2}, \frac{2}{3e^2}, 1 - \frac{7}{e^2}\right)$, H_1: minst ett p_i skiljer sig från värdet ovan.

$$X = \sum_{k=0}^{5} \frac{(N_i - np_i)^2}{np_i} \sim \chi^2_5$$ och har utfall $X_0 = 8.73$.

Tröskelvärdet $\chi^2_{0.05,5} = 11.07$ ($\alpha = 0.05$, 6 kategorier) är större, därför förkastar vi ej H_0. Poisson-fördelningsantagandet kan stämma.