STOCKHOLMS UNIVERSITET
MATEMATISKA INSTITUTIONEN
Avd. Matematisk statistik

MT1011, MT102U, MT103U
TENTAMEN
20 februari 2023

Tentamenslösningar – Sannolikhetslära och statistik för lärare

20 februari 2023 kl. 14–19

Examinator: Gudrun Brattström, gudrun@math.su.se

Uppgift 1

a) Eftersom $5-(-4)=9$ har X tätheten $f_X(x)=\frac{1}{9}$ för $-4 \le X \le 5$. Utanför intervallet är $f_X(x)=0$. Alltså är

$$P(X>0)=\int_0^5 \frac{1}{9}dx = \left[\frac{1}{9}\right]_0^5 = \frac{5}{9} = 0.556\,.$$

Alternativt kan sannolikheten beräknas som arean av en rektangel med sidorna $\frac{1}{9}$ och 5.

b) Enligt formelsamlingen är

$$\begin{aligned} E(X) &= \tfrac{-4+5}{2} = \tfrac{1}{2} = 0.5 \\ V(X) &= \tfrac{(5-(-4))^2}{12} = \tfrac{81}{12} = \tfrac{27}{4} = 6.75 \end{aligned}$$

c) Eftersom en likformig fördelning är symmetrisk och saknar extrema värden räcker 50 oberoende X_i mer än väl för att vi ska kunna använda centrala gränsvärdessatsen. Vi har

$$\begin{aligned} E(\overline{X}) &= 0.5 \\ D(\overline{X}) &= \sqrt{6.75}/\sqrt{50} = \sqrt{0.135} = 0.367 \end{aligned}$$

Enligt centrala gränsvärdessatsen gäller

$$\frac{\overline{X}-0.5}{\sqrt{6.75}/\sqrt{50}} = \frac{\overline{X}-0.5}{\sqrt{0.135}} \sim_{\text{approx.}} N(0,1)\,,$$

så att

$$\begin{aligned} P(\overline{X}>0) &= P\left(\frac{\overline{X}-0.5}{\sqrt{0.135}} > \frac{0-0.5}{\sqrt{0.135}}\right) \approx 1-\Phi\left(-\frac{0.5}{\sqrt{0.135}}\right) = 1-\Phi(-1.36) \\ &= 1-(1-\Phi(1.36)) = \Phi(1.36) = 0.9131\,. \end{aligned}$$

Uppgift 2

a) För en exponentialfördelning gäller att parametern β i tätheten $f_X(x) = \frac{1}{\beta}e^{-\frac{x}{\beta}}$ är lika med väntevärdet. (Se formelsamlingen.) Om vi kallar livslängden för X så blir tätheten

$$f_X(x) = \tfrac{1}{25}e^{-\frac{x}{25}} = 0.04e^{-0.04x}\,.$$

b) Vi får sannolikheten

$$P(X>30) = \int_{30}^{\infty} 0.04e^{-0.04x}dx = \left[-e^{-0.04x}\right]_{30}^{\infty} = 0 - \left(-e^{-0.04\cdot 30}\right) = e^{-0.04\cdot 30} = e^{-1.2} = 0.301.$$

c) På grund av exponentialfördelningens minneslöshet får vi att

$$P(X > 30+10|X > 30) = P(X > 10) = e^{-0.04\cdot 10} = e^{-0.4} = 0.670\,.$$

Det går även att räkna ut den betingade sannolikheten direkt, utan att hänvisa till minneslöshet:

$$\begin{aligned} P(X > 30+10|X > 30) &= \frac{P((X>40)\cap(X>30))}{P(X>30)} = \frac{P(X>40)}{P(X>30)} \\ &= \frac{e^{-0.04\cdot 40}}{e^{-0.04\cdot 30}} = e^{-0.04\cdot 10} = 0.670\,. \end{aligned}$$

Uppgift 3

a) Vi har urvalsstorlekarna $n_1 = n_2 = 200$ och populationsstorlekarna $N_1 = 1132$ och $N_2 = 2510$. Andelen oroliga i A-stad skattas med $\hat{p}_1 = \frac{124}{200} = 0.62$, och i B-stad med $\hat{p}_2 = \frac{101}{200} = 0.505$.

$$\begin{aligned} n_1\hat{p}_1(1-\hat{p}_1)\tfrac{N_1-n_1}{N_1-1} &= 200\cdot\tfrac{124}{200}(1-\tfrac{124}{200})\cdot\tfrac{1132-200}{1132-1} = 47.1 > 10 \\ n_2\hat{p}_2(1-\hat{p}_2)\tfrac{N_2-n_2}{N_2-1} &= 200\cdot\tfrac{101}{200}(1-\tfrac{101}{200})\cdot\tfrac{2510-200}{2510-1} = 50.0 > 10 \end{aligned}$$

Alltså kan formelsamlingens approximativa konfidensintervall för två proportioner användas.

$$\begin{aligned} I_{p_1-p_2} &= \hat{p}_1 - \hat{p}_2 \pm \lambda_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}\cdot\frac{N_1-n_1}{N_1-1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}\cdot\frac{N_2-n_2}{N_2-1}} \\ &= \frac{124}{200} - \frac{101}{200} \pm \lambda_{0.975}\sqrt{\frac{\frac{124}{200}\left(1-\frac{124}{200}\right)}{200}\cdot\frac{1132-200}{1132-1} + \frac{\frac{101}{200}\left(1-\frac{101}{200}\right)}{200}\cdot\frac{2510-200}{2510-1}} \\ &= 0.115 \pm 1.96\cdot\sqrt{0.0021215} = 0.115 \pm 1.96\cdot 0.04606 = 0.115 \pm 0.090\,. \end{aligned}$$

Alternativt kan intervallet skrivas $(0.025, 0.205)$.

b) Eftersom intervallet av konfidensgrad $1-0.05 = 0.95$ inte innehåller 0, är skillnaden signifikant. Vi drar slutsatsen att invånarna i A-stad oroar sig mer än invånarna i B-stad.

Uppgift 4

a) Låt X beteckna vikten av en slumpmässigt vald spik. Standardisera:

$$\begin{aligned} P(X < 6.5) &= P\left(\frac{X - 6.57}{0.081} < \frac{6.5 - 6.57}{0.081}\right) = \Phi\left(\frac{-0.07}{0.081}\right) = \Phi(-0.86) \\ &= 1 - \Phi(0.86) = 1 - 0.8051 = 0.1949\,. \end{aligned}$$

b) Summan $W = \sum_{i=1}^{10} X_i$, där X_i kan antas vara oberoende och normalfördelade med $E(X_i) = 6.57$ och $D(X_i) = 0.081$, är normalfördelad med $E(W) = 10 \cdot 6.57 = 65.7$ och $D(W) = 0.081\sqrt{10} = 0.2561$. Alltså gäller

$$\begin{aligned} P(W < 65) &= P\left(\frac{W - 65.7}{0.081\sqrt{10}} < \frac{65 - 65.7}{0.081\sqrt{10}}\right) = \Phi\left(-\frac{0.7}{0.2561}\right) = \Phi(-2.73) \\ &= 1 - \Phi(2.73) = 1 - 0.9968 = 0.0032\,. \end{aligned}$$

c) Antalet spikar Y som vi väger är *ffg*-fördelat med parameter $p = P(X < 6.5) = 0.1949$. Enligt formelsamlingen gäller att $E(Y) = 1/p = 1/0.1949 = 5.13$. Så vi förväntar oss att behöva väga lite drygt fem spikar.

Uppgift 5

a) $P(X \text{ är udda}) = P(X = 1) + P(X = 3) = p_X(1) + p_X(3) = 0.2 + 0.2 = 0.4$.

b) Enligt formelsamlingen har vi

$$E(X) = \sum_{k=0}^{4} k p_X(k) = 0 \cdot 0.3 + 1 \cdot 0.2 + 2 \cdot 0.2 + 3 \cdot 0.2 + 4 \cdot 0.1 \ = 1.6\,.$$

c) Vi har att $P(X \text{ är udda}) = 0.4$ och $P(X < 2) = p_X(0) + p_X(1) = 0.3 + 0.2 = 0.5$, så att

$$P(X \text{ är udda}) \cdot P(X < 2) = 0.4 \cdot 0.5 = 0.2\,.$$

Å andra sidan är

$$P((X \text{ är udda}) \cap (X < 2)) = P(X = 1) = p_X(1) = 0.2\,.$$

Så $P(X \text{ är udda}) \cdot P(X < 2) = P((X \text{ är udda}) \cap (X < 2))$, vilket enligt definitionen betyder att de båda händelserna är oberoende.

Uppgift 6

a) Låt X vara antalet dragna röda kulor. Då har X hypergeometrisk fördelning, med parametrarna $N = 6 =$ totalt antal kulor i påsen, $n = 2 =$ antalet kulor som dras, och $p = \frac{1}{6} =$ andelen röda kulor i påsen. Vi söker (se formelsamlingen)

$$P(X = 1) = \frac{\binom{6\cdot\frac{1}{6}}{1}\binom{6\cdot\left(1-\frac{1}{6}\right)}{2-1}}{\binom{6}{2}} = \frac{\binom{1}{1}\binom{5}{1}}{\binom{6}{2}} = \frac{1\cdot 5}{\frac{6\cdot 5}{2}} = \frac{1}{3} = 0.333\,.$$

Not: Det går förstås lika bra att räkna på $W =$ antalet dragna vita kulor, som också är hypergeometriskt fördelat, med $N = 6$, $n = 2$, och $p = \frac{5}{6} =$ andelen vita kulor i påsen. Vi får

$$P(W = 1) = \frac{\binom{6\cdot\frac{5}{6}}{1}\binom{6\cdot\left(1-\frac{5}{6}\right)}{2-1}}{\binom{6}{2}} = \frac{\binom{5}{1}\binom{1}{1}}{\binom{6}{2}} = \cdots = 0.333\,.$$

b) Vi börjar med att göra motsvarande kalkyl för sannolikheten att dra en röd och en vit kula ur den nya påsen. Parametrarna är nu $N = 6$ och $n = 2$ som förut, medan $p = \frac{3}{6} = \frac{1}{2}$. Kalla antalet röda kulor som dras ur den nya påsen för Y. Då gäller

$$P(Y = 1) = \frac{\binom{6\cdot\frac{1}{2}}{1}\binom{6\cdot\left(1-\frac{1}{2}\right)}{2-1}}{\binom{6}{2}} = \frac{\binom{3}{1}\binom{3}{1}}{\binom{6}{2}} = \frac{3\cdot 3}{15} = \frac{3}{5} = 0.6\,.$$

Kalla händelsen att den gamla påsen väljs för A. Då blir händelsen A^* att den nya påsen väljs. Lagen om total sannolikhet ger att

$$\begin{aligned} P(1 \text{ röd, } 1 \text{ vit}) &= P(X{=}1|A)\cdot P(A) + P(Y{=}1|A^*)\cdot P(A^*) \\ &= \frac{1}{3}\cdot\frac{1}{2} + \frac{3}{5}\cdot\frac{1}{2} = \frac{7}{15} = 0.467\,. \end{aligned}$$

Man kan nå samma resultat genom att använda träddiagram.