STOCKHOLMS UNIVERSITET
MATEMATISKA INSTITUTIONEN
Avd. Matematisk statistik

MT1011, MT102U, MT103U
TENTAMEN
16 februari 2022

Tentamenslösningar – Sannolikhetslära och statistik för lärare

16 februari 2022 kl. 8–13

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

Uppgift 1

a) Om vi istället mäter i cm så blir alla mätvärden 10 gånger mindre. Det samma gäller standardavvikelsen, så den blir 0.54 cm.

b) Enligt **a)** är standardavvikelsen i cm lika med 0.54 cm, så vi får variansen blir $0.54^2 = 0.2916$ cm^2.

c) För 50 oberoende mätningar X_i, $i = 1, 2, \ldots, 50$ som alla har variansen 5.4^2 mm^2 gäller att

$$V(\overline{X}) = V\left(\frac{1}{50}\sum_{i=1}^{50} X_i\right) = \frac{1}{50^2}\sum_{i=1}^{50} V(X_i) = \frac{1}{50^2} \cdot 50 \cdot 5.4^2 = \frac{5.4^2}{50} \text{ mm}\,.$$

Medelvärdets standardavvikelse är därför

$$D(\overline{X}) = \frac{5.4}{\sqrt{50}} = 0.764 \text{ mm}\,.$$

Uppgift 2

a) Populationen antas vara mycket stor, så vi kan anta att vi har två oberoende binomialfördelningar. Andelen grodda frön i järnrik jord är $\hat{p_1} = \frac{26}{50}$ och andelen grodda frön i järnfattig jord är $\hat{p_2} = \frac{33}{50}$. Vi har

$$\begin{aligned} n_1\hat{p_1}(1-\hat{p_1}) &= 50 \cdot \tfrac{26}{50} \cdot \tfrac{24}{50} = \tfrac{624}{50} = 12.48 > 10 \\ n_2\hat{p_2}(1-\hat{p_2}) &= 50 \cdot \tfrac{33}{50} \cdot \tfrac{17}{50} = \tfrac{561}{50} = 11.22 > 10 \end{aligned}$$

Alltså kan formelsamlingens approximativa konfidensintervall användas. I binomialfallet har det formen

$$\begin{aligned} I_{p_1-p_2} &= \hat{p_1} - \hat{p_2} \pm \lambda_{\alpha/2}\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}} \\ &= \tfrac{26}{50} - \tfrac{33}{50} \pm \lambda_{0.975}\sqrt{\frac{\frac{26}{50}(1-\frac{26}{50})}{50} + \frac{\frac{33}{50}(1-\frac{33}{50})}{50}} \\ &= -0.14 \pm 1.96 \cdot \sqrt{\frac{0.52\cdot 0.48 + 0.66 \cdot 0.34}{50}} \\ &= -0.14 \pm 1.96 \cdot \sqrt{0.00948} = -0.14 \pm 0.19\,. \end{aligned}$$

Alternativt kan intervallet skrivas som $(-0.33, 0.05)$.

b) Eftersom intervallet av konfidensgrad $1 - 0.05 = 0.95$ innehåller 0, är skillnaden inte signifikant.

Uppgift 3

a) Formelsamlingen ger $E(X) = \frac{20+30}{2} = 25$.

b) Formelsamlingen ger $V(X) = \frac{(30-20)^2}{12} = \frac{100}{12} = \frac{25}{3} = 8.33$.

c) Kvadratens area är X^2 och X har täthetsfunktionen $f_X(x) = \frac{1}{10}$ för $20 \leq x \leq 30$ och $f_X(x) = 0$ för alla andra x, så väntevärdet av arean ges av

$$E(X^2) = \int_{20}^{30} x^2 \cdot \tfrac{1}{10} \cdot dx = \frac{1}{10}\left[\frac{x^3}{3}\right]_{20}^{30} = \frac{1}{10}\left(\frac{30^3}{3} - \frac{20^3}{3}\right) = \frac{19\,000}{30} = 633.3 \text{ cm}^2\,.$$

En alternativ lösning fås genom att notera att $E(X^2) = V(X) + E(X)^2 = \frac{25}{3} + 25^2 = \frac{1900}{3} = 633.3 \text{ cm}^2$, enligt **a)** och **b)**.

Uppgift 4

a) Låt F beteckna händelsen att du väljer fusktärningen; då blir F^* händelsen att du väljer den vanliga tärningen. Låt S vara händelsen att du slår en sexa. Vad vi vet är att $P(S|F) = \frac{1}{2}$ och att $P(S|F^*) = \frac{1}{6}$, samt att $P(F) = P(F^*) = \frac{1}{2}$. Lagen om total sannolikhet ger att

$$P(S) = P(S|F)P(F) + P(S|F^*)P(F^*) = \tfrac{1}{2}\cdot\tfrac{1}{2} + \tfrac{1}{6}\cdot\tfrac{1}{2} = \tfrac{3+1}{12} = \tfrac{1}{3}\,.$$

b) Med hjälp av Bayes sats beräknar vi

$$\begin{aligned} P(F|S) &= \frac{P(F \cap S)}{P(S)} = \frac{P(S|F)P(F)}{P(S)} = \frac{P(S|F)P(F)}{P(S|F)P(F) + P(S|F^*)P(F^*)} \\ &= \frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{6} \cdot \frac{1}{2}} = \frac{3}{3+1} = \frac{3}{4} = 0.75\,. \end{aligned}$$

Vi kan förstås använda resultatet från **a)** att nämnaren $P(S) = \frac{1}{3}$ och förkorta räkningarna lite.

Uppgift 5

a) Antalet gånger X som spelaren måste spela har en *ffg*-fördelning, vilket enligt formelsamlingen innebär att det förväntade antalet är 1/vinstchansen, det vill säga 37.

b) Sannolikhetsfunktionen för *ffg*-fördelningen är enligt formelsamlingen

$$p_X(k) = p(1-p)^{k-1}, \text{ för } k = 1, 2, \ldots$$

I vårt fall är $p = \frac{1}{37}$ = vinstchansen. Alltså är den sökta sannolikheten

$$P(X \geq 40) = \sum_{k=40}^{\infty} p(1-p)^{k-1}$$

För att kunna använda ledningen behöver vi ordna så att summan startar vid 0 istället för 40. Det gör vi genom att istället använda $j = k - 40$ (så att $k = j + 40$) som summationsindex: när $k = 40$ så är $j = 0$, och vi kan låta a vara lika med första termen i summan, alltså $a = p(1-p)^{39}$. Vi får att

$$\begin{aligned} P(X \geq 40) &= \sum_{k=40}^{\infty} p(1-p)^{k-1} = \sum_{j=0}^{\infty} p(1-p)^{j+39} \\ &= \sum_{j=0}^{\infty} p(1-p)^{39}(1-p)^j = \frac{p(1-p)^{39}}{1-(1-p)} = (1-p)^{39} = \left(\frac{36}{37}\right)^{39} = 0.3435\,. \end{aligned}$$

c) Att spelaren har förlorat fyrtio gånger betyder att $X \geq 41$. Vi söker alltså

$$\begin{aligned} P(X \geq 80|X \geq 41) &= \frac{P(X \geq 80 \cap X \geq 41)}{P(X \geq 41)} = \frac{P(X \geq 80)}{P(X \geq 41)} = \frac{(36/37)^{79}}{(36/37)^{40}} \\ &= (36/37)^{39} = 0.3435\,. \end{aligned}$$

Vi kan också resonera så här: eftersom spelen är oberoende av varandra så har det ingen betydelse om spelaren har vunnit eller förlorat tidigare spel. Det är som om hen började om på nytt vid spel nummer 41, och därför blir sannolikheten densamma som i **b)**.

Uppgift 6

a) Beteckna den sökta standardavvikelsen med s. Om X är en slumpmässigt vald persons intelligenskvot, så är enligt antagande $\frac{X-100}{s}$ standardnormalfördelad. Om man ska tro *Encyclopædia Britannica* så gäller $P(85 < X < 115) = \frac{2}{3}$. Alltså är

$$\tfrac{2}{3} = P(\tfrac{85-100}{s} < \tfrac{X-100}{s} < \tfrac{115-100}{s}) = \Phi(\tfrac{15}{s}) - \Phi(\tfrac{-15}{s}) = 2\Phi(\tfrac{15}{s}) - 1\,.$$

Lös ut $\Phi(\frac{15}{s})$: vi får att $\Phi(\frac{15}{s}) = \frac{5}{6} = 0.8333$. Tabell 3 ger att $\frac{15}{s} = 0.97$, så vi får att $s = \frac{15}{0.97} = 15.46$.

b) Vi får att

$$P(X > 130) = P\left(\tfrac{X-100}{15/0.97} > \tfrac{130-100}{15/0.97}\right) = 1 - \Phi\left(\tfrac{30\cdot 0.97}{15}\right) = 1 - \Phi(1.94) = 1 - 0.9738 = 0.026\,.$$

Om vi istället använder $s = 10$, så får vi

$$P(X > 130) = P\left(\tfrac{X-100}{10} > \tfrac{130-100}{10}\right) = 1 - \Phi\left(\tfrac{30}{10}\right) = 1 - \Phi(3) = 1 - 0.9987 = 0.0013\,.$$