Løsning eksamen 29. august 2018

Oppgave 1

La J_1=første barn er jente, J_2=andre barn er jente, G_1=første barn er gutt, G_2=andre barn er gutt, E=enegget og T=toegget.

a) For toegget tvillingfødsel er kjønnet til den ene tvillingen uavhengig av kjønnet til den andre slik at:

$$P(J_1 \cap J_2|T) = P(J_1|T) \cdot P(J_2|T) = 0.486 \cdot 0.486 = \underline{\underline{0.236}}$$

Merk at gutt og jente kan skje på to måter, enten er første barn gutt og andre barn jente eller motsatt. Og kombinasjonen gutt og jenter er ikke mulig ved enegget fødsel. Setningen om total sannsynlighet gir da:

$$\begin{aligned} P(G_1 \cap J_2 \cup J_1 \cap G_2) &= P(G_1 \cap J_2 \cup J_1 \cap G_2|E)P(E) + P(G_1 \cap J_2 \cup J_1 \cap G_2|T)P(T) \\ &= 0 + P(G_1 \cap J_2|T)P(T) + P(J_1 \cap G_2|T)P(T) = 2 \cdot 0.486 \cdot (1 - 0.486) \cdot 0.7 = \underline{\underline{0.350}} \end{aligned}$$

b) Setningen om total sannsynlighet og Bayes lov gir:

$$\begin{aligned} P(G_1 \cap G_2) &= P(G_1 \cap G_2|E)P(E) + P(G_1 \cap G_2|T)P(T) = (1 - 0.486) \cdot 0.3 + (1 - 0.486)^2 \cdot 0.7 = \underline{\underline{0.339}} \\ P(E|G_1 \cap G_2) &= \frac{P(G_1 \cap G_2|E)P(E)}{P(G_1 \cap G_2)} = \frac{(1 - 0.486) \cdot 0.3}{0.339} = \underline{\underline{0.455}} \end{aligned}$$

Oppgave 2

Lar X være mengden chips i en tilfeldig pakkke.

a)

$$\begin{aligned} P(X < 225) &= P(\frac{X - 250}{10} < \frac{225 - 250}{10}) = P(Z < -2.50) = \underline{\underline{0.0062}} \\ P(225 < X < 275) &= P(X < 275) - P(X < 225) = P(\frac{X - 250}{10} < \frac{275 - 250}{10}) - 0.0062 \\ &= P(Z < 2.50) - 0.0062 = 0.9938 - 0.0062 = \underline{\underline{0.988}} \end{aligned}$$

Må justere μ slik at $P(X \geq 250) = 0.95$ eller med andre ord $P(X < 250) = 0.05$.

$$\begin{aligned} P(X < 250) &= P(\frac{X - \mu}{10} < \frac{250 - \mu}{10}) = P(Z < \frac{250 - \mu}{10}) = 0.05 \\ &\Rightarrow \frac{250 - \mu}{10} = -1.645 \\ &\quad \mu = 250 + 1.645 \cdot 10 = \underline{\underline{266.5}} \end{aligned}$$

b) Situasjonen her er normalfordeling med ukjent μ og kjent σ. Et $(1 - \alpha)100\%$ konfidensintervall for μ er da gitt ved

$$[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \quad \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}].$$

Med $\alpha = 0.05$ blir $z_{\alpha/2} = z_{0.025} = 1.96$, og med $\bar{x} = 296.1$, $n = 50$ og $\sigma = 10$ blir 95% konfidensintervallet:

$$[296.1 - 1.96\frac{10}{\sqrt{50}}, 296.1 + 1.96\frac{10}{\sqrt{50}}] = \underline{\underline{[293.3\ , 298.9]}}$$

Siden hele intervallet ligger under 300 er det en sterkt indikasjon på at produksjonsutstyret er innstilt på for lav verdi.

Lengden til konfidensintervallet er:

$$\begin{aligned} L &= \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} - (\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \\ \text{Dvs: } L \leq 4 &\Rightarrow 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq 5 \\ & \quad 0.4z_{\alpha/2}\sigma \leq \sqrt{n} \\ & \quad n \geq (0.5z_{\alpha/2}\sigma)^2 = (0.4 \cdot 1.96 \cdot 10)^2 = 61.46 \end{aligned}$$

Dvs det må gjøres minst $\underline{\underline{62}}$ målinger for å få et konfidensintervall med ønsket lengde.

c)

$$H_0: \quad \mu \geq 300 \qquad\qquad \text{mot} \qquad\qquad H_1: \quad \mu < 300$$

Situasjonen her er normalfordeling med ukjent μ og ukjent σ. Dersom H_0 er korrekt er da

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{\bar{X} - 300}{S/\sqrt{n}} \sim t_{n-1}$$

Med signifikansnivå 5%, dvs $\alpha = 0.05$, forkaster vi H_0 dersom $T \leq -t_{0.05,19} = -1.729$.

Observert: $t = \frac{293.9-300}{11.7/\sqrt{20}} = -2.33$

Siden $-2.33 < -1.729$ blir konklusjonen at vi $\underline{\underline{\text{forkaster } H_0}}$. Det betyr i praksis at målingene gir grunnlag for å konkludere at forventet mengde kjøttdeig i pakkene er lavere enn 300 gram.

(Dvs gjennomsnittet på 293.3 er så mye lavere enn 300 at vi her kan konkludere at den reelle forventningsverdien μ (gjennomsnittsvekten av veldig mange pakninger) må være lavere enn 300.)

d)

$$\begin{aligned} P(\mu - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) &= P(-z_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}) = P(-z_{\alpha/2} < Z < z_{\alpha/2}) \\ &= P(Z < z_{\alpha/2}) + P(Z < -z_{\alpha/2}) = 1 - \alpha/2 + \alpha/2 = \underline{\underline{1 - \alpha}} \end{aligned}$$

Merk at sannsynligheten for at en stikkprøver faller *utenfor* intervallet $[\mu - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \mu + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$ dermed blir α.

For Y har vi en situasjon karakterisert ved:

- Gjentatte delforsøk som gir "suksess"/ikke "suksess" - gjentatte stikkprøver hvor resultatet enten faller utenfor kontrollgrensene eller ikke.
- Lik sannsynlighet $p = \alpha$ i alle delforsøk - samme sannsynlighet α for at resultatet av stikkprøven vil falle utenfor kontrollgrensene for hver stikkprøve.
- Uavhengige delforsøk - uavhengig resultat fra stikkprøve til stikkprøve.
- Gjentar delforsøkene inntil første suksess - gjør nye stikkprøver inntil neste alarm.

Dvs, alle betingelsene for geometrisk fordeling er oppfylte og vi har dermed at Y er geometrisk fordelt med parameter $p = \alpha$.

I geometrisk fordeling har vi at $\mathrm{E}(Y) = 1/p$. ARL er forventet tid mellom hver gang en stikkprøve faller utenfor grensene, og siden sannsynligheten for en verdi utenfor grensene er $p = \alpha$ blir ARL $= \mathrm{E}(Y) = 1/p = 1/\alpha$.

Ved å bruke formelen for kumulativ fordelingsfunksjon i geometrisk fordeling og $p = \alpha = 0.01$ får vi:

$$P(Y > 100) = 1 - P(Y \leq 100) = 1 - (1 - (1 - 0.01)^{100}) = 0.99^{100} = \underline{\underline{0.366}}$$

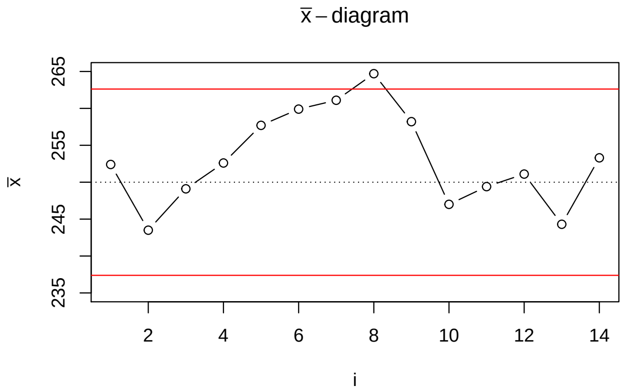
e) Fra ARL $= 1/\alpha$ får vi at vi må bruke $\alpha = 1/\text{ARL} = 1/500 = 0.002$. Fra kvantiltabellen for normalfordeling finner vi at $z_{\alpha/2} = z_{0.001} = 3.09$ Vi bruker grensene og verdiene før punkt d) og får:

Senterlinje: $\mu = 250$.

Øvre kontrollinje: $\mu + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 250 + 3.09\frac{10}{\sqrt{6}} = \underline{\underline{262.6}}$

Nedre kontrollinje: $\mu - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 250 - 3.09\frac{10}{\sqrt{6}} = \underline{\underline{237.4}}$

Plott av $\bar{x}$-diagrammet er gitt under:



Vi ser i en tendens til stadig økende vekt i starten og ved stikkprøve nr 8 er vekten over øver kontrollgrense. Dvs der blir det gitt en alarm om for høy vekt.

Dersom vi ikke kjenner μ og σ men må estimere dem fra data gjør estimeringsfeilen (det faktum at de estimerte verdiene ikke er de helt korrekte verdiene) at vi i praksis kan få en ARL som avviker fra den ønskede. Avviket kan gå i begge retninger avhengige av hvilken vei feilene i estimeringen av μ og σ går. Dvs vi kan enten få en for liten reell ARL eller en for stor reell ARL.

Oppgave 3

a) Første skritt i testen er å regne ut forventet antall i hver celle i tabellen under antagelsen om lik fordeling i gruppene. Vi trenger da også rad- og kolonnesummene. Disse og de forventede verdiene er gitt i tabellen under (de forventede verdiene står i parantesene i hver celle). For eksempel finner vi forventet verdi for kombinasjonen "evaluering 2016" og "for" som: $214 \cdot (110/367) = 64.1$ og for kombinasjonen "evaluering 2018" og "mot" som: $153 \cdot (257/367) = 107.1$.

	for	mot	totalt
evaluering 2016	69 (64.1)	145 (149.9)	214
evaluering 2018	41 (45.9)	112 (107.1)	153
totalt	110	257	367

Testobservatoren blir da:

$$\begin{aligned}\mathcal{Q} &= \sum_{\text{alle celler}} \frac{(\text{observert-forventet})^2}{\text{forventet}} \\ &= \frac{(69-64.1)^2}{64.1} + \frac{(145-149.9)^2}{149.9} + \frac{(41-45.9)^2}{45.9} + \frac{(112-107.1)^2}{107.1} = 1.28\end{aligned}$$

Verdien på testobservatoren skal vi sammenligne med 5% kvantilen i kjikvadratfordelingen med parameter (frihetsgrader) $(r-1)\cdot(k-1) = (2-1)\cdot(2-1) = 1$. Fra kjikvadrat-tabellen finner vi at denne kvantilen har verdi 3.84. Siden $\mathcal{Q} = 1.28 < 3.84$ blir konklusjonen av vi ikke forkaster nullhypotesene om like andeler for/mot de to årene. Dvs, vi kan ikke konkludere at der er noen generell forskjell i andel studenter for/mot videoforelesninger mellom de to årene.

Forutsetningen for testen er at forventet antall observasjoner i hver celle er minst 5. Dette ser vi er oppfylt her.

Videre er det også en forutsetning at dataene vi har representerer et tilfeldig utvalg fra den populasjonen vi ønsker å gjøre en test for. Hvorvidt dette er tilfellet er usikkert. Studentene som svarte på evalueringene i STA100 er neppe representative for alle studenter. De kan være representative for studenter som tar STA100, men selv det er usikkert - det kan være at de som svare på evalueringen har en annen mening enn resten av studentene.

b) $\hat{p}_1 = \frac{x_1}{n_1} = \frac{69}{69+145} = \underline{\underline{0.322}}$ og $\hat{p}_2 = \frac{x_2}{n_2} = \frac{41}{41+112} = \underline{\underline{0.268}}$.

$$\begin{aligned}
\mathrm{E}(\hat{p}_1 - \hat{p}_2) &= \mathrm{E}(\hat{p}_1) - \mathrm{E}(\hat{p}_2) = \frac{\mathrm{E}(X_1)}{n_1} - \frac{\mathrm{E}(X_2)}{n_2} = \frac{n_1 p_1}{n_1} - \frac{n_2 p_2}{n_2} = \underline{\underline{p_1 - p_2}} \\
\mathrm{Var}(\hat{p}_1 - \hat{p}_2) &= \mathrm{Var}(\hat{p}_1) + \mathrm{Var}(\hat{p}_2) = \frac{1}{n_1^2}\mathrm{Var}(X_1) + \frac{1}{n_2^2}\mathrm{Var}(X_2) = \frac{1}{n_1^2} n_1 p_1(1-p_1) + \frac{1}{n_2^2} n_2 p_2 (1-p_2) \\
&= \underline{\underline{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}
\end{aligned}$$

Vi har at $n_1 = 214$ og $n_2 = 153$ og ut fra de estimerte verdiene for p_1 og p_2 er det klart at $n_1 p_1(1-p_1) > 5$ og $n_2 p_2 (1-p_2) > 5$. Dermed følger det at både X_1 og X_2 er tilnærmet normalfordelte. Siden $\hat{p}_1 - \hat{p}_2$ er en lineærkombinasjon av X_1 og X_2 følger det da at $\hat{p}_1 - \hat{p}_2$ er tilnærmet normalfordelte. Forventning og varians er regnet ut over, og vi får altså at: $\hat{p}_1 - \hat{p}_2 \approx \mathrm{N}(p_1 - p_2, \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}})$

c) Fra resultatet i siste del av punkt b) og hintet følger det at

$$\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \approx \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \approx \mathrm{N}(0,1)$$

og vi får dermed:

$$\begin{aligned}
P(-z_{\alpha/2} < \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} < z_{\alpha/2}) &\approx 1-\alpha \\
P(\hat{p}_1 - \hat{p}_2 - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} < p_1 - p_2 < \hat{p}_1 - \hat{p}_2 + z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}) &\approx 1-\alpha
\end{aligned}$$

som gir konfidensintervallet:

$$[\hat{p}_1 - \hat{p}_2 - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}\ ,\ \hat{p}_1 - \hat{p}_2 + z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}]$$

For et 95% intervall setter vi $\alpha = 0.05$ som gir $z_{\alpha/2} = z_{0.025} = 1.96$. Innsatt observerte data får vi:

$$\begin{aligned}
&[0.322 - 0.268 - 1.96 \cdot \sqrt{0.322 \cdot (1-0.322)/214 + 0.268 \cdot (1-0.268)/153}, \\
&\quad 0.322 - 0.268 + 1.96 \cdot \sqrt{0.322 \cdot (1-0.322)/214 + 0.268 \cdot (1-0.268)/153}] \\
= &\ \underline{\underline{[-0.04\ ,\ 0.15]}}
\end{aligned}$$

Siden konfidensintervallet inneholder 0 kan vi ikke konkludere med at det er forskjell i andel mellom de to årene. En test av nullhypotesen $p_1 - p_2 = 0$ mot alternativet $p_1 - p_2 \neq 0$ vil her ikke gi forkastning på 5% nivå siden 0 er inneholdt i 95% konfidensintervallet.

Det er grunn til å være usikker på om det er korrekt at sannsynligheten for å være for video er den samme blant de som besvarte evauleringen som blant alle studenten som tok kurset (det kan være slik, men trenger ikke, kanskje har de som besvarte undersøkelsen andre meninger om videoforelesninger enn resten av studentene, kanskje ikke). Dvs, det er grunn til å være usikker på om de som har besvart er representative for alle studenter. Dersom denne antagelsen ikke er korrekt vil både estimatet og konfidensintervallet gi et feil bilde av hva studentene generelt mente.

Dersom sannsynligheten for å være for er lik blant de som besvarte har vi likvel det forholdet at X_1 og X_2 strengt tatt er hypergeometrisk fordelte, og betingelsen om at vi $N > 10n$ som må være oppfylt for at hypergeometrisk skal kunne tilnærmes med binomisk er ikke helt oppfylt her. Dette medfører i praksis at vi ikke helt har uavhengighet og lik sannsynlighet i hver trekning som binomisk fordeling krever. I praksis medfører dette at bredden på konfidensintervallet ikke blir helt korrekt (estimatet er greit dersom antagelsen om at de som besvarte er representativ holder).