TMA4245 Statistikk
Vår 2016

Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Anbefalte oppgaver 11, blokk II
Løsningsskisse

Oppgave 1

a) En rimelig estimator for forventningsverdien μ er gjennomsnittet

$$\bar{X} = \frac{1}{5}\sum_{i=1}^{5} X_i,$$

som vil være normalfordelt med forventningsverdi $E(\bar{X}) = \mu$ og varians $\text{Var}(\bar{X}) = \sigma^2/5$. En rimelig estimator for variansen er

$$S^2 = \frac{1}{4}\sum_{i=1}^{5}(X_i - \bar{X})^2,$$

som har forventningsverdi $E(S^2) = \sigma^2$. Observasjonene $x_1, \ldots, x_5$ i tabellen gir estimatene

$$\bar{x} = \frac{1}{5}\sum_{i=1}^{5} x_i = \underline{\underline{4.9540}} \quad \text{og} \quad s^2 = \frac{1}{4}\sum_{i=1}^{5}(x_i - \bar{x}) = \underline{\underline{0.3440}}.$$

b) For å utlede et konfidensintervall for μ tar vi utgangspunkt i den tilfeldige variabelen

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/5}} \sim N(0,1)$$

som er standard normalfordelt, siden $\bar{X} \sim N(\mu, \sigma^2/5)$. Når den ukjente variansen σ^2 byttes ut med estimatoren S^2, får vi observatoren

$$T = \frac{\bar{X} - \mu}{\sqrt{S^2/5}} \sim t_4$$

som er t-fordelt med $5 - 1 = 4$ frihetsgrader. Vi ønsker et 95% konfidensintervall for μ, og trenger derfor 0.975-kvantilen i t-fordelingen med 4 frihetsgrader, som er $t_{0.025,4} = 2.7764$. Konfidensintervallet kan nå konstrueres som følger,

$$\begin{aligned}
&P(-t_{0.025,4} \leq T \leq t_{0.025,4}) = 1 - 2 \cdot 0.025 = 0.95 \\
&P\left(-t_{0.025,4} \leq \frac{\bar{X} - \mu}{\sqrt{S^2/5}} \leq t_{0.025,4}\right) = 0.95 \\
&P\left(\bar{X} - t_{0.025,4} \cdot \sqrt{\frac{S^2}{5}} \leq \mu \leq \bar{X} + t_{0.025,4} \cdot \sqrt{\frac{S^2}{5}}\right) = 0.95.
\end{aligned}$$

Når tallverdier settes inn får vi intervallet

$$\bar{x} \pm t_{0.025,4} \cdot \sqrt{\frac{s^2}{5}} = 4.9540 \pm 2.7764 \cdot \sqrt{\frac{0.3440}{5}} = \underline{\underline{[4.2258, 5.6822]}}.$$

Anta at vi skal teste nullhypotesen $H_0: \ \mu = \mu_0$ mot den alternative hypotesen $H_1: \mu \neq \mu_0$. Vi bruker testobservatoren

$$T_0 = \frac{\bar{X} - \mu_0}{\sqrt{S^2/5}},$$

som i likhet med T er t-fordelt med 4 frihetsgrader (for $\mu_0 = \mu$ har vi $T_0 = T$). På signifikansnivå 5% vil vi beholde H_0 hvis vi observerer $-t_{0.025,4} \leq T_0 \leq t_{0.025,4}$. Ellers forkastes H_0. Akseptansekriteriet for H_0 er dermed

$$-t_{0.025,4} \leq \frac{\bar{X} - \mu_0}{\sqrt{S^2/5}} \leq t_{0.025,4}$$

eller, om vi isolerer μ_0 i midten,

$$\bar{X} - t_{0.025,4} \cdot \sqrt{\frac{S^2}{5}} \leq \mu_0 \leq \bar{X} + t_{0.025,4} \cdot \sqrt{\frac{S^2}{5}},$$

som er identisk med 95% konfidensintervallet over. For en gitt verdi av μ_0 kan altså konfidensintervallet brukes til å teste H_0 mot H_1 på signifikansnivå 5%, ved å beholde nullhypotesen kun dersom μ_0 er inneholdt i intervallet.

Oppgave 2

a) μ = populasjonsgjennomsnitt, dvs. eit gjennomsnitt for alle bilane som køyrer på vegstrekningen i ein gitt periode.

$$\widehat{\mu} = \bar{X} = \frac{\sum_{i=1}^{12} X_i}{12}\ , \quad S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{12}(X_i - \bar{X})^2}$$

$$\bar{X} = \frac{\sum_{i=1}^{12} X_i}{12} = \frac{880}{12} = \underline{\underline{73.33}}\ , \quad S = \sqrt{\frac{1}{11}\sum_{i=1}^{12}(X_i - \bar{X})^2} = \sqrt{\frac{1034.7}{11}} = \underline{\underline{9.7}}$$

b) Type 1 feil er å forkaste H_0 når H_0 er rett.

$$H_0: \ \mu \geq 77 \qquad H_1: \ \mu < 77$$

$\alpha = 0.05$, forkast om:

$$\frac{\bar{X} - 77}{\frac{S}{\sqrt{n}}} < -t_{0.05,11} = -1.8$$

$$\frac{73.33 - 77}{\frac{9.7}{\sqrt{12}}} = -1.31 > -1.8$$

dvs. ikkje grunnlag for å påstå at farten er blitt lågare på 5 % nivå.

c) Type 2 feil er å ikkje forkaste når H_0 er gal. La $\beta = P(\text{type 2 feil})$. Då er styrken $1-\beta$.

$$\begin{aligned}
P\left(\frac{\bar{X}-77}{\frac{10}{\sqrt{12}}} < -1.645 \mid \mu = 74\right) &= P\left(\frac{\bar{X}-74}{\frac{10}{\sqrt{12}}} < -1.645 + \frac{3}{\frac{10}{\sqrt{12}}} \mid \mu = 74\right) \\
&= \Phi\left(-1.645 + \frac{3\sqrt{12}}{10}\right) = \Phi(-0.61) \\
&= 1 - 0.729 = \underline{\underline{0.271}}
\end{aligned}$$

$$\begin{aligned}
& P\left(\frac{\bar{X}-77}{\frac{10}{\sqrt{n}}} < -1.645 \mid \mu = 74\right) = 0.9 \\
\Leftrightarrow\ & P\left(\frac{\bar{X}-74}{\frac{10}{\sqrt{n}}} < -1.645 + \frac{3}{\frac{10}{\sqrt{n}}} \mid \mu = 74\right) = 0.9 \\
\Leftrightarrow\ & \Phi\left(-1.645 + \frac{3\sqrt{n}}{10}\right) = 0.9 \\
\Leftrightarrow\ & -1.645 + \frac{3\sqrt{n}}{10} = 1.28 \\
\Leftrightarrow\ & \frac{3\sqrt{n}}{10} = 1.28 + 1.645 = 2.925 \\
\Leftrightarrow\ & n = \frac{(2.925)^2 \cdot 10^2}{3^2} = 95.06
\end{aligned}$$

Dvs. vi må måle farten på 96 bilar eller fleir.

Oppgave 3

a) Sannsynligheten for å få 5 kron er

$$P(5 \text{ kron}) = \frac{1}{2^5} = 1/32 = \underline{\underline{0.031}}.$$

Sannsynligheten for å få 3 kron er lik punktsannsynligheten $P(X=3)$ der X er binomisk fordelt med parametre $n=5$ og $p=0.5$, altså

$$P(X=3) = \binom{5}{3} 0.5^3 \cdot (1-0.5)^{5-3} = 10 \cdot 0.5^3 \cdot 0.5^2 = \underline{\underline{0.3125}}.$$

Fire kron på rad kan inntreffe på 3 forskjellige måter: Kron på alle 5 kastene, kron på de første 4 kastene, og mynt på siste, eller mynt på første kast og kron på de 4 siste. Antall mulige utfall av de fem kastene er $2^5 = 32$, og alle er like sannsynlige, så sannsynligheten for å få fire kron på rad er

$$P(4 \text{ kron på rad}) = \frac{3}{32} = \underline{\underline{0.0938}}.$$

b) Sannsynligheten for at lengste sekvens har lengde 5 eller 6 kan anslås ved å regne ut andelen utfall hvor lengste sekvens var på 5 eller 6 kast, av de 10000 simulasjonene. Fra figuren leser vi av at lengste sekvens hadde lengde 5 i omtrent 2700 tilfeller, og lengde 6 i omtrent 1700 tilfeller, og vi får estimatet

$$P(\widehat{5 \text{ eller}}\ 6) = \frac{2700 + 1700}{10000} = \underline{\underline{0.44}}.$$

I Miriams myntkastsekvens har den lengste uavbrutte sekvensen av kron lengde 2. For en tilfeldig generert myntkastsekvens av lengde 30, vil lengden av lengste uavbrutte sekvens av kron ha en sannsynlighetsfordeling som er svært lik den i figuren. At denne lengden er så lav som 2 er ganske usannsynlig, og Miriams myntkastsekvens er dermed mistenkelig.

Vi vil teste nullhypotesen

$$H_0 : \text{ Sekvensen er tilfeldig generert}$$

mot den alternative hypotesen

$$H_1 : \text{ Sekvensen er ikke tilfeldig generert.}$$

Vi antar at under nullhypotesen er lengden av lengste sammenhengende sekvens av kron fordelt som i figuren. For å avgjøre om nullhypotesen skal forkastes eller ikke, regner vi ut p-verdien, altså sannsynligheten for å observere et like ekstremt eller mer ekstremt utfall. Her er dette lik sannsynligheten for at lengste uavbrutte sekvens av kron er 0, 1 eller 2. Utfra figuren ser det ut som om antall utfall i søylene for 0, 1 og 2 er henholdsvis 0, 0 og 25. Vi får dermed følgende estimat for p-verdien:

$$P(0, \widehat{1 \text{ eller}}\ 2) = \frac{25}{10000} = 0.0025.$$

Dette er en lav p-verdi som tilsier at nullhypotesen forkastes f.eks. på signifikansnivå 0.05. Det er altså grunn til å hevde at Miriam har funnet på tallene.

Oppgave 4

a) Det er mest rimelig med en venstresidig hypotesetest:

$$\begin{aligned} H_0 : &\quad \mu = 16, \\ H_1 : &\quad \mu < 16. \end{aligned}$$

Begrunnelse: forhandleren sier at bilen kan forventes å kjøre **minst** 16 km pr liter. Vi vil avsløre ev. feil i markedsføringen.

NB: Hypotetesetesten skal være uavhengig av målingene. En bør altså ikke velge alternativ hypotese på grunnlag av $\overline{x}$.

$\overline{X}$ er normalfordelt med forventning μ og varians σ^2/n. Variansen er ukjent, derfor kreves T-fordeling med $\nu = n-1 = 19$ frihetsgrader. Gjennomfører testen med $\alpha = 0.05$. Testobservator: $T_{\text{obs}} = \frac{\overline{X}-\mu}{S/\sqrt{n}}$. Observert verdi:

$$t_{\text{obs}} = \frac{\overline{x} - \mu}{s/\sqrt{n}} = \frac{15.56 - 16}{0.94/\sqrt{20}} = -2.093.$$

Fra tabell over kvantiler i T-fordelingen; $-t_{0.05,19} = -1.729$. Altså: $t < -t_{0.05,19}$, dermed skal H_0 forkastes.

Hvis vi hadde valgt å bruke en normalfordelingshypotese, ville kvantilen $-z_{0.05} = -1.645$ gitt samme konklusjon. Imidlertid bør vi da argumentere for at avstanden til denne kvantilen er så stor at høyere varians i T-fordelingen ikke ville påvirket resultatet. Å sammenlikne med denne kvantilen kan ikke regnes som fullgodt svar.

b) P-verdien finnes ved å lete opp verdien på testobservatoren fra a) i tabell. For T-fordeling med $\nu = 19$, finner vi $t_{0.025,19} = 2.093$. Ettersom T-fordelingen er symmetrisk, har vi at $\mathrm{P}(T > t_{\alpha,\nu}) = \mathrm{P}(T < -t_{\alpha,\nu})$. Dermed; $p = \alpha = 0.025 = 2.5\%$.

Testobservatoren er normalfordelt hvis $\sigma = s$. Dette bør være tilnærmet oppfylt for å bruke normalfordeling. Hvis en ikke har ekstra informasjon om σ, er det ikke anbefalt å tilnærme student-fordelingen med en normalfordeling når $n < 30$, da s ikke er et godt nok estimat.

Under normalfordelingen får vi p-verdi

$$\mathrm{P}(Z \leq -2.09) = \Phi(-2.09) = 0.0183.$$

c) Antar $H_1' : \mu = \mu_1 = 15.5$ og $\sigma = s$. Teststyrken er sannsynligheten for å forkaste H_0 under H_1', dvs

$$\mathrm{P}\left(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < -1.645 \mid \mu = \mu_1 = 15.5\right).$$

For å få en normalfordelt variabel, flytter vi alt utenom $\overline{X}$, som er stokastisk, over på høyre side. Deretter trekker vi fra sann forventningsverdi μ_1 og dividerer med standardavviket på begge sider.

$$\begin{aligned}
&\mathrm{P}(\overline{X} < -1.645 \cdot \sigma/\sqrt{n} + \mu_0) \\
&= \mathrm{P}(\frac{\overline{X} - \mu_1}{\sigma/\sqrt{n}} < \frac{-1.645 \cdot \sigma/\sqrt{n} + \mu_0 - \mu_1}{\sigma/\sqrt{n}}) \\
&= \mathrm{P}(Z < 0.7338) \approx 0.767.
\end{aligned}$$

Hvis vi ikke kunne bruke normalfordelingsantakelsen, ville teststyrken blitt svakere. Her er det forutsatt at vi er ganske sikre på variansen, f.eks. på grunnlag av data fra produsent.

Generelt må antall observasjoner økes for å oppnå økt teststyrke. (Dette er fullgodt svar.) Mulig tillegg: Hvis en har mulighet til å gjennomføre forsøket på en måte slik at variansen blir mindre, f.eks. kjøre bilene under mer kontrollerte former i et laboratorium, ville også teststyrken økes. Eventuelt kan en øke signifikansnivået α f.eks. til 0.1, og dermed øke teststyrken, men dette er sjelden aktuelt i praksis.

Oppgave 5

a) For at X skal være binomisk fordelt må sannsynligheten $P(Z_i > Y_i)$ for å gå raskest i siste ytre være lik p for alle løpere $i = 1, \dots, n$, og vi må ha uavhengighet mellom hendelsene $Z_i > Y_i$ for ulike løpere.

Gitt at $n = 20$ og $p = 0.7$ blir $P(X \le 10) = 0.048$ (tabell) og

$$P(X \ge 8 | X \le 10) = \frac{P(8 \le X \le 10)}{P(X \le 10)} = \frac{P(X \le 10) - P(X \le 7)}{P(X \le 10)} = \frac{0.048 - 0.01}{0.048} = 0.98.$$

b) Likelihoodfunksjonen blir

$$L(p) = \binom{n}{x} p^x (1-p)^{n-x},$$

og log-likelihoodfunksjonen

$$l(p) = \ln \binom{n}{x} + x \ln p + (n-x) \ln(1-p).$$

Denne har sitt maksimum der

$$\begin{aligned} \frac{dl}{dp} &= 0 \\ \frac{x}{p} - \frac{n-x}{1-p} &= 0 \\ p &= \frac{x}{n}. \end{aligned}$$

SME for p er dermed $\hat{p} = X/n$. Denne er forventningsrett siden

$$E\left(\frac{X}{n}\right) = \frac{1}{n} EX = \frac{1}{n} np = p.$$

Variansen blir

$$\mathrm{Var}\hat{p} = \mathrm{Var}\left(\frac{X}{n}\right) = \frac{1}{n^2}\mathrm{Var}X = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}.$$

c) Vi skal teste om det er en fordel å gå siste ytre. Dette vil i så fall innebære at parameteren $p > 1/2$. Vi lar dette være vår alternative hypotese H_1. Nullhypotesen H_0 blir at $p = 1/2$. Siden vi ikke har tabell over binomisk fordeling for $n = 39$ bruker vi testobservatoren

$$Z = \frac{\hat{p} - 1/2}{\sqrt{(1/2)(1-1/2)/39}}$$

som er tilnærmet standard normalfordelt under H_0. Vi forkaster H_0 hvis $Z > z_{0.05} = 1.65$. Observert verdi av Z blir

$$Z = \frac{24/39 - 1/2}{\sqrt{(1/2)(1-1/2)/39}} = 1.44.$$

Vi beholder dermed H_0.

Testens p-verdi blir tilnærmet

$$P(Z > 1.44) = 0.075.$$

d) Vi antar at differansene $D_1, D_2, \ldots, D_n$ mellom løpstid med og uten siste ytre til hver enkelt løper er uavhengig $N(\mu, \sigma^2)$. Vi ønsker å undersøke om siste ytre gir en fordel, altså at $EY_i < EZ_i$, som vil innebære at parameteren $\mu = EY_i - EX_i < 0$ (alternativ hypotese H_1). Nullhypotesen H_1 blir $\mu = 0$.

Vi lar $S_D^2 = \frac{1}{n-1}\sum_{i=1}^{37}(D_i - \bar{D})^2$. Ved å bruke at $\bar{D} \sim N(\mu, \sigma_D^2/n)$ og at $S_D^2(n-1)/\sigma_D^2$ er kji-kvadrat med $n-1$ frihetsgrader, følger det at

$$T = \frac{\frac{\bar{D}}{\sigma_D/\sqrt{n}}}{\sqrt{\frac{S^2(n-1}{\sigma_D^2}/(n-1)}} = \frac{\bar{D}}{S/\sqrt{n}}$$

under H_0 er t-fordelt med $n - 1 = 37 - 1 = 36 \approx 35$ frihetsgrader. Vi forkaster dermed H_0 hvis $T < -t_{0.05,37-1} = -1.69$.

Gitt dataene i oppgaven får vi $\bar{d} = -2.654/37 = -0.0717$, $\sum(d_i - \bar{d})^2 = \sum d_i^2 - n\bar{d}^2 = 1.362$, $s_D^2 = \sum(d_i - \bar{d})/(n-1) = 0.0378$, og

$$t = \frac{-0.0717}{\sqrt{0.0378}/\sqrt{37}} = -2.24.$$

Basert på testantakelsene kan vi dermed forkaste H_0 og konkludere med at siste ytre gir en liten fordel (H_1).

e) For test 1 blir teststyrken for $p = 0.64$

$$\begin{aligned}
P(Z > z_\alpha) &= P(\frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} > z_\alpha) \\
&= P(\hat{p} > p_0 + z_\alpha\sqrt{p_0(1-p_0)/n}) \\
&= 1 - \phi(\frac{p_0 - p + z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}) \\
&= \phi(\frac{p - p_0 - z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}) \\
&= \phi(\frac{0.14 - 1.64\sqrt{0.25/39}}{\sqrt{0.64 \cdot 0.36/39}}) \\
&= 0.545.
\end{aligned}$$

Test 1, hvor vi kun benyttet binær informasjon om hvorvidt siste ytre ga beste tid for hver enkelt løper, ga hverken forkastning eller størst teststyrke sammenlignet med Test 2. Dette er forventet ut i fra at Test 2 er basert på all informasjon om de observerte løpstidene i motsetning til Test 1.

På den annen side bygger Test 2 på et skjevt utvalg siden de to løperne som falt er tatt ut av dataene. Dette vil forskyve $\bar{D}$ mot mer positive verdier (negative verdier gir støtte for H_1). At vi da likevel får forkastning tyder da på at det er en reell forskjell.

Den skjeve utvalget med lange løpstider i siste indre tatt ut av dataene vil kunne gjøre at antakelsen om normalfordeling ikke er oppfylt. Men dette vil i enda større grad kunne gjelde også før sensurering. Ut i fra histogrammet av observerte d_i kan det vanskelig konkluderes med at dataene avviker fra antakelsen om normalfordeling siden utvalgsstørrelsen i dette henseende er liten.

Fordelen med Test 1 er at denne ikke forutsetter normalfordeling.

Totalt sett gir dataene grunn for å konkludere med at siste ytre gir en fordel.