Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

TMA4240 Statistikk
Høst 2020

Anbefalte oppgaver 6
Løsningsskisse

Oppgave 1

a) Om en passasjer kommer eller ikke kan betraktes som et stokastisk forsøk med to mulige utfall, sukess (kommer) eller fiasko (kommer ikke). For at antall suksesser, X, skal være binomisk fordelt må vi ha at de n forsøkene er uavhengige, dvs at de n passasjerene kommer eller ikke uavhengig av hverandre. Dessuten må alle forsøkene ha samme sannsynlighet for suksess, dvs alle passasjerer må ha samme sannsynlighet p for å komme, og for at man skal ha at 7% av samtlige passasjerer ikke skal komme til bestilt avgang må man ha at $p = 1 - 0.07 = 0.93$. Dermed får vi at

$$\mathrm{E}[X] = np = n \cdot 0.93 = 0.93n$$

og

$$\mathrm{Var}[X] = np(1-p) = n \cdot 0.93 \cdot (1 - 0.93) = 0.0651n.$$

b) Antall passasjerer som møter til avgang er nå $n = 255$ og $p = 0.93$. Sannsynligheten for at alle passasjerene får plass blir

$$P(X \leq 243) = P\left(\frac{X - 0.93n}{\sqrt{0.0651n}} \leq \frac{243 - 0.93 \cdot 255}{\sqrt{0.0651 \cdot 255}}\right) \approx P(Z \leq 1.44),$$

der Z er standard normalfordelt. Ved å bruke tabell over sannsynlighetene i en standard normalfordeling får man dermed

$$P(X \leq 243) \approx P(Z \leq 1.44) = \underline{\underline{0.9251}}.$$

Skal få finne hvor stor n kan være dersom en krever

$$P(X \leq 243) \geq 0.99.$$

Bruker igjen normaltilnærmingen slik at en får kravet

$$\begin{aligned}
P(X \leq 243) &\approx P\left(\frac{X - 0.93n}{\sqrt{0.0651n}} \leq \frac{243 - 0.93n}{\sqrt{0.0651n}}\right) = P\left(Z \leq \frac{243 - 0.93n}{\sqrt{0.0651n}}\right) \\
&= 1 - P\left(Z > \frac{243 - 0.93n}{\sqrt{0.0651n}}\right) \geq 0.99 \\
&\Leftrightarrow P\left(Z > \frac{243 - 0.93n}{\sqrt{0.0651n}}\right) \leq 0.01,
\end{aligned}$$

der Z igjen er standard normalfordelt. Ved å bruke tabellen over kvantiler i standard normalfordelingen får man dermed at kravet blir

$$\begin{aligned}
\frac{243-0.93n}{\sqrt{0.0651n}} &\geq z_{0.01}=2.326\\
243-0.93n &\geq 2.326\cdot\sqrt{0.0651n}=0.5935\sqrt{n}\\
-0.93\left(\sqrt{n}\right)^2-0.5935\sqrt{n}+243 &\geq 0.
\end{aligned}$$

Vi får altså en andregradsulikhet i $\sqrt{n}$. For å løse denne løser vi først den tilhørende andregradsligningen,

$$-0.93\left(\sqrt{n}\right)^2-0.5935\sqrt{n}+243=0.$$

Denne ligningen har løsningene

$$\begin{aligned}
\sqrt{n} &= \frac{-(-0.5935)\pm\sqrt{(-0.5935)^2-4\cdot(-0.93)\cdot 243}}{2\cdot(-0.93)}\\
&= \frac{0.5935\pm 30.072}{-1.86}\Rightarrow\sqrt{n}=-16.487 \text{ eller } \sqrt{n}=15.849.
\end{aligned}$$

Siden det her åpenbart ikke gir mening å ha $\sqrt{n}$ negativ blir løsningen av andregradslingen at

$$\sqrt{n}=15.849.$$

Siden fortegnet til koeffisienten foran $\sqrt{n}$ i ulikheten er negativ blir løsningen til andregradsulikheten

$$\sqrt{n}\leq 15.849\Rightarrow n\leq 15.849^2=251.19.$$

Dessuten, siden antall bestillinger nødvendigvis må være et heltall får vi $n\leq 251$, slik at maksimal overbooking blir $251-243=\underline{\underline{8}}$.

Merk: Man kan også komme frem til dette resultatet ved å prøve seg frem med ulike verdier for n.

c) Tapet dersom man ikke overbooker (dvs $n=243$) blir

$$U=1000\cdot(243-X),$$

og forventet tap blir dermed

$$\mathrm{E}=1000\cdot(243-\mathrm{E}[X])=1000\cdot(243-np)=1000\cdot(243-243\cdot 0.93)=\underline{\underline{17010}}.$$

Dersom man selger $n=255$ bestillinger blir tapet

$$V=1000\cdot\max\{243-X,0\}+4000\cdot\max\{X-243,0\},$$

alternativt

$$V=1000\cdot(243-X)\cdot 1(X<243)+4000\cdot(X-243)\cdot 1(X>243),$$

der $1(\cdot)$ er indikator funksjonen. Forventet tap blir

$$\begin{aligned}
\mathrm{E}[V] &= \mathrm{E}\left[1000\cdot\max\{243-X,0\}+4000\cdot\max\{X-243,0\}\right]\\
&= \underline{\underline{1000\cdot\mathrm{E}\left[\max\{243-X,0\}\right]+4000\cdot\mathrm{E}\left[\max\{X-243,0\}\right]}}.
\end{aligned}$$

Oppgave 2 For eksponentialfordelingen har vi

$$\begin{aligned}
P(X \geq t+s | X > s) &= \frac{P(X \geq t+s \cap X > s)}{P(X > s)} \\
&= \frac{P(X \geq t+s)}{P(X > s)} \\
&= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} \\
&= e^{-\lambda t} = \underline{\underline{P(X \geq t)}}.
\end{aligned}$$

For den geometriske fordelingen har vi, for en vilkårlig $a \in \{0, 1, 2, \ldots\}$,

$$P(X \geq a) = P(\text{Ingen suksesser på de første } a-1 \text{ forsøkene}) = (1-p)^{a-1},$$

og vi får dermed

$$\begin{aligned}
P(X \geq t+s | X > s) &= \frac{P(X \geq t+s \cap X > s)}{P(X > s)} \\
&= \frac{P(X \geq t+s)}{P(X > s)} \\
&= \frac{P(X \geq t+s)}{P(X \geq s+1)} \\
&= \frac{(1-p)^{(t+s)-1}}{(1-p)^{(s+1)-1}} \\
&= \frac{(1-p)^{t+s-1}}{(1-p)^{s}} \\
&= (1-p)^{t-1} = \underline{\underline{P(X \geq t)}}.
\end{aligned}$$

Oppgave 3

a) Vi ønsker å uttrykke A ved hjelp av $A_k,\ k = 1, 2, 3$. Det er klart at hendelsen "spilleren oppnår en 1'er med inntil tre kast" er det samme som at vi enten får en 1'er på første kast, eller at vi feiler på første kast men får en 1'er på andre kast, eller at vi feiler på de to første kastene, men får en 1'er på det tredje kastet. Vi kan derfor uttrykke B som

$$B = A_1 \cup (A_1^C \cap A_2^C) \cup (A_1^C \cap A_2^C \cap A_3).$$

Hendelsen B^C er det samme som at vi verken får en 1'er på kast 1, 2 eller 3. Vi kan derfor uttrykke

$$B^C = A_1^C \cap A_2^C \cap A_3^C.$$

Vi ønsker å regne ut $P(B)$. For å gjøre dette kan vi observere at de tre hendelsene A_1, $(A_1^C \cap A_2^C)$ og $(A_1^C \cap A_2^C \cap A_3)$ er disjunkte, siden A_1 og A_1^C er disjunkte. Vi vet også at de tre mulige kastene er uavhengige av hverandre, slik at A_1, A_2 og A_3 er uavhengige.

Da får vi

$$\begin{aligned}
P(B) &= P(A_1 \cup (A_1^C \cap A_2^C) \cup (A_1^C \cap A_2^C \cap A_3)) \\
&= P(A_1) + P(A_1^C \cap A_2^C) + P(A_1^C \cap A_2^C \cap A_3) \\
&= P(A_1) + P(A_1^C)P(A_2^C) + P(A_1^C)P(A_2^C)P(A_3) \\
&= \frac{1}{6} + \frac{5}{6} \cdot \frac{1}{6} + \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} \\
&= \frac{91}{216} \approx \underline{\underline{0.42}}.
\end{aligned}$$

For B^C får vi

$$P(B^C) = P(A_1^C \cap A_2^C \cap A_3^C) = \left(\frac{5}{6}\right)^3 = \frac{125}{216} \approx \underline{\underline{0.58}}.$$

b) Vi har sett at for é terning er sannsynligheten for å oppnå en 1'er i løpet av tre kast lik $P(B) \approx 0.42$. I løpet av omgang nummer en skal en spiller kaste fem forskjellige terninger opp til tre ganger, for å få så mange 1'ere som mulig. Vi kan derfor anse omgang en som fem uavhengige forsøk, med suksessannsynlighet $P(B)$. Det totale antallet 1'ere, X_1 blir derfor binomisk fordelt med suksessannsynlighet $p = P(B)$, og $n = 5$. Vi får

$$P(X_1 = 5) = \binom{5}{5} p^5 (1-p)^0 = p^5 \approx \underline{\underline{0.013}}.$$

Sannsynligheten $P(X_1 \geqslant 3)$ kan også finnes som

$$\begin{aligned}
P(X_1 \geqslant 3) &= P(X_1 = 3) + P(X_1 = 4) + P(X_1 = 5) \\
&= \binom{5}{3} p^3 (1-p)^2 + \binom{5}{4} p^4 (1-p)^1 + p^5 \\
&\approx \underline{\underline{0.35}}.
\end{aligned}$$

c) Vi kan finne forventningsverdien til Y ved å sette

$$E[Y] = E\left[\sum_{i=1}^{6} iX_i\right] = \sum_{i=1}^{6} iE[X_i].$$

Fra b) vet vi at X_1 er binomisk fordelt med $n = 5$ og $p = P(B)$. Det er klart at sannsynligheten får å få et visst antall 1'ere er lik sannsynligheten for å få det samme antall 2'ere, osv. Vi kan derfor argumentere oss frem til at X_i er binomisk fordelt med $n = 5$ og $p = P(B)$ for alle $i = 1, \ldots, 6$. Da vet vi at $E[X_i] = np$ for alle verdier av i. Derfor får vi

$$E[Y] = \sum_{i=1}^{6} iE[X_i] = \sum_{i=1}^{6} i \cdot 5 \cdot \frac{91}{216} = 21 \cdot 5 \cdot \frac{91}{216} \approx \underline{\underline{44.2}}.$$

På samme måte vet vi at $\mathrm{Var}(X_i) = np(1-p)$ for alle $i = 1, \ldots, 6$. Vi får derfor

$$\begin{aligned}
\mathrm{Var}(Y) &= \mathrm{Var}\left(\sum_{i=1}^{6} iX_i\right) \\
&= \sum_{i=1}^{6} i^2 \mathrm{Var}(X_i) \\
&= \sum_{i=1}^{6} i^2 np(1-p) \\
&= 91 \cdot 5 \cdot \frac{91}{216} \cdot 125216 \\
&\approx 110.93 \approx \underline{\underline{10.53^2}}.
\end{aligned}$$

Vi ønsker nå å tilnærme fordelingen til Y med en normalfordeling. Vi får da

$$\frac{Y - E[Y]}{\sqrt{\mathrm{Var}(Y)}} \sim N(0,1).$$

Den tilnærmede sannsynligheten for å oppnå bonus er den tilnærmede sannsynligheten for at Y er større eller lik 63. Vi finner

$$\begin{aligned}
P(Y \geq 63) &= 1 - P(Y < 63) \\
&= 1 - P\left(\frac{Y - E[Y]}{\sqrt{\mathrm{Var}(Y)}} < \frac{63 - E[Y]}{\sqrt{\mathrm{Var}(Y)}}\right) \\
&\approx 1 - \Phi\left(\frac{63 - 44.2}{10.53}\right) \\
&\approx 1 - 0.96 \\
&= \underline{\underline{0.04}}.
\end{aligned}$$