# Statistikk og Økonomi

Løsningsforslag eksamen 2013 - del 1

## Oppgave 1

La $X$ være antall seksere på $n$ terningkast. Det er fornuftig å anta at $X$ er binomisk fordelt. Altså at:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

der $p$ er sannsynligheten for å få sekser på et terningkast. Siden $p = \frac{1}{6}$ blir da sannsynligheten for å få minst én sekser på $n = 6$ terningkast:

$$\begin{aligned} P(X \geq 1) &= 1 - P(X < 1) = 1 - P(X = 0) \\ &= 1 - \binom{6}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^6 = 1 - 0.335 = \underline{\underline{0.665}} \end{aligned}$$

Sannsynligheten for å få minst to seksere på $n = 12$ terningkast:

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - P(X \leq 1) \\ &= 1 - [P(X = 0) + P(X = 1)] \\ &= 1 - \left[\binom{12}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{12} + \binom{12}{1} \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^{11}\right] \\ &= 1 - 0.381 = \underline{\underline{0.619}} \end{aligned}$$

Det er altså størst sannsynlighet for å få minst èn sekser på seks terningskast.

## Oppgave 2

Vi definerer følgende hendelser:

$$\begin{aligned} A &: \text{Person er syk.} \\ B &: \text{Test indikerer sykdom.} \end{aligned}$$

Fra oppgaveteksten får vi følgende opplysninger:

$$\begin{aligned} P(B|\overline{A}) &= 0.05 \\ P(\overline{B}|A) &= 0.02 \\ P(A) &= 0.03 \end{aligned}$$

Vi skal finne sannsynligheten for at en person ikke er syk, gitt at testen viser smitte, altså $P(\overline{A}|B)$. Vi bruker Bayes lov og får:

$$P(\overline{A}|B) = \frac{P(\overline{A})P(B|\overline{A})}{P(B)}$$

Vi bruker loven om total sannsynlighet for å finne $P(B)$ og får da:

$$P(\overline{A}|B) = \frac{P(\overline{A})P(B|\overline{A})}{P(A)P(B|A) + P(\overline{A})P(B|\overline{A})}$$

De resterende sannsynlighetene finner vi ved komplementsetningen:

$$\begin{aligned} P(\overline{A}|B) &= \frac{[1 - P(A)]P(B|\overline{A})}{P(A)[1 - P(\overline{B}|A)] + [1 - P(A)]P(B|\overline{A})} \\ &= \frac{0.97 \cdot 0.05}{0.03 \cdot 0.98 + 0.97 \cdot 0.05} \approx \underline{\underline{0.623}} \end{aligned}$$

## Oppgave 3

Fra oppgaveteksten får vi oppgitt at vi har en intensitet, $\alpha$, på 2.5 utrykninger per dag. Hvis vi lar $X$ være antall utrykninger i en periode på $t$ dager, får vi da:

$$P(X = x) = \frac{(\alpha t)^x}{x!} e^{-\alpha t}$$

a) Sannsynligheten for å få mer enn tre utrykninger på én dag:

$$\begin{aligned} P(X > 3) = 1 - P(X \leq 3) &= 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)] \\ &= 1 - \left[\frac{2.5^0}{0!}e^{-2.5} + \frac{2.5^1}{1!}e^{-2.5} + \frac{2.5^2}{2!}e^{-2.5} + \frac{2.5^3}{3!}e^{-2.5}\right] = \underline{\underline{0.242}} \end{aligned}$$

b) Over en periode på fire dager, får vi $\alpha t = 2.5 \cdot 4 = 10$. Sannsynligheten for å få minst 10 utrykninger, blir da:

$$\begin{aligned} P(X \geq 10) &= 1 - P(X < 10) = 1 - P(X \leq 9) \\ &= 1 - \sum_{x=0}^{9} \frac{10^x}{x!} e^{-10} \overset{calc.}{=} 1 - 0.458 = \underline{\underline{0.542}} \end{aligned}$$

c) Hvis vi lar $X_i$ være antall utrykninger på dag $i$, får vi fra sentralgrensesetningen at:

$$\overline{X} = \frac{1}{n}(X_1 + X_2 + ...X_n) \approx N\left(\mu, \frac{\sigma^2}{n}\right)$$

der $\mu = E(X_i) = \lambda = \alpha t = 2.5$, $\sigma^2 = Var(X_i) = \lambda = \alpha t = 2.5$ og $n = 365$. Vi kan dermed regne ut sannsynligheten:

$$P(\overline{X} \leq 2.4) \approx G\left(\frac{2.4 - 2.5}{\sqrt{2.5/365}}\right) \approx G(-1.21) = 1 - G(1.21) \overset{\text{tabell}}{=} 1 - 0.8869 \approx \underline{\underline{0.113}}$$

## Oppgave 4

Fra oppgaveteksten får vi opplyst at analysetiden for en tilfeldig valgt jordprøve, $X$, har fordelingen $N(\mu = 150, \sigma^2 = 36^2)$.

a) Sannsynligheten for at X er minst 200:

$$\begin{aligned} P(X \geq 200) &= 1 - P(X < 200) \overset{\text{kont.ford.}}{=} 1 - P(X \leq 200) = 1 - F(200) \\ &= 1 - G\left(\frac{200 - 150}{36}\right) \approx 1 - G(1.39) \overset{\text{tabell}}{=} 1 - 0.9177 = \underline{\underline{0.0823}} \end{aligned}$$

b) Vi skal finne $P(X \leq 250 | X > 200)$. Fra formelhefte ser vi at (betinget sannsynlighet):

$$\begin{aligned} P(X \leq 250 | X > 200) &= \frac{P(X \leq 250 \cap X > 200)}{P(X > 200)} = \frac{P(200 < X \leq 250)}{P(X > 200)} \\ &= \frac{P(X \leq 250) - P(X \leq 200)}{P(X > 200)} = \frac{P(X \leq 250) - P(X \leq 200)}{1 - P(X \leq 200)} \\ &= \frac{G\left(\frac{250-150}{36}\right) - G\left(\frac{200-150}{36}\right)}{1 - G\left(\frac{200-150}{36}\right)} \\ &\approx \frac{G(2.78) - G(1.39)}{1 - G(1.39)} \overset{\text{tabell}}{=} \frac{0.9973 - 0.9177}{1 - 0.9177} \approx \underline{\underline{0.967}} \end{aligned}$$

c) Et 95% konfidensintervall for $\mu$, er gitt ved:

$$\overline{X} \pm U_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

der $\alpha = 0.05$. Fra tabellene finner vi $U_{\alpha/2} = U_{0.025} = 1.96$. Vi får da:

$$\overline{X} \pm U_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 130 \pm 1.96 \frac{36}{\sqrt{15}} = 130 \pm 18.219$$

som gir oss konfidensintervallet: $\underline{\underline{[111.8,\ 148.2]}}$.

d) Vi har $n = 15$, $\sigma = 36$ og vi målte et gjennomsnitt $\overline{X} = 130$. Vi setter opp følgende hypotesetest:

$$\mathrm{H}_0\text{: } \mu = 150 \quad \text{mot} \quad \mathrm{H}_1\text{: } \mu \neq 150$$

Vi bruker $\alpha/2$ kvantilen i $N(0,1)$ fordelingen, og påstår $\mathrm{H}_1$ (evt. forkaster $\mathrm{H}_0$) dersom $\overline{X}$ ligger utenfor intervallet:

$$\begin{aligned} \mu_0 \pm U_{\alpha/2} \frac{\sigma}{\sqrt{n}} &\overset{\text{tabell}}{=} 150 \pm 1.96 \frac{36}{\sqrt{15}} \\ &= 150 \pm 18.219 \approx [131.8, 168.2] \end{aligned}$$

Siden $\overline{X} = 130$ ligger utenfor dette intervallet, påstår vi $\mathrm{H}_1$, altså vi konkluderer med at maskinen <u>ikke</u> har en forventet analysetid på 150 sekunder etter kalibreringen.

**Kommentar:** Siden vi har en tosidig test, kunne vi også brukt konfidensintervallet i oppgave c) for å avgjøre hypotesetesten. Vi ser at $\mu_0 = 150$ ligger utenfor konfidensintervallet og vi har dermed grunnlag til å konkludere at $\mu \neq 150$.

e) Styrkefunksjonen er definert som $P(\text{“Påstå H}_1\text{”}|\mu)$.

i. Hvis analysetiden ikke har forandret seg, altså $\mu = 150$, vil rett konklusjon være å ikke påstå $\mathrm{H}_1$. Fra grafen ser vi at sannsynligheten for å påstå $\mathrm{H}_1$ når $\mu = 150$, er 0.05. Sannsynligheten for å få korrekt konklusjon, vil da bli 1-0.05 = $\underline{\underline{0.95}}$.

   **Kommentar:** Dette kunne vi også funnet ut uten å se på grafen, for vi har definert $\alpha$ som sannsynligheten for at man påstår $\mathrm{H}_1$ gitt at $\mathrm{H}_0$ gjelder (dvs. $\mu$ er uforandret). Sannsynligheten for å ikke påstå $\mathrm{H}_1$ gitt at $\mathrm{H}_0$ gjelder, blir da 1-$\alpha$, som gir samme resultat.

ii. Når $\mu = 130$, vil feil konklusjon være å ikke påstå $\mathrm{H}_1$ ($\mu$ er jo ulik 150). Sannsynligheten for å påstå $\mathrm{H}_1$ når $\mu = 130$ ser vi fra grafen er ca. lik 0.57. Sannsynligheten for feil konklusjon blir da ca 1-0.57 = $\underline{\underline{0.43}}$.