

Løsning eksamen desember 2017

Oppgave 1

Innfører hendelsene

D: enheten er defekt

K: enheten blir kassert

a)

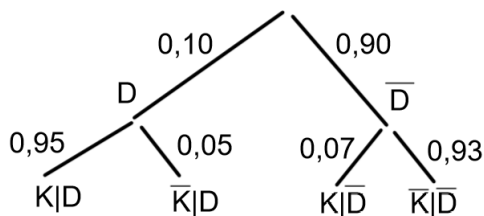
i) Disse sannsynlighetene kan leses ut av oppgaveteksten:

$$P(D) = 0,10$$

$$P(K|\bar{D}) = 0,07$$

$$P(K|D) = 0,95$$

Sannsynlighetstre:



ii) Finner $P(K)$, dvs. sannsynligheten for at en enhet blir kassert. Loven om total sannsynlighet:

$$\begin{aligned} P(K) &= P(K|D) \cdot P(D) + P(K|\bar{D}) \cdot P(\bar{D}) \\ &= 0,95 \cdot 0,10 + 0,07 \cdot 0,90 \\ &= \underline{\underline{0,158}} \end{aligned}$$

b)

i) Finner $P(D|K)$ fra Bayes' teorem:

$$\begin{aligned} P(D|K) &= \frac{P(D) P(K|D)}{P(K)} \\ &= \frac{0,10 \cdot 0,95}{0,158} \\ &= \underline{\underline{0,601}} \end{aligned}$$

Finner $P(D|\bar{K})$ fra Bayes' teorem:

$$\begin{aligned} P(D|\bar{K}) &= \frac{P(D) P(\bar{K}|D)}{P(\bar{K})} \\ &= \frac{0,10 \cdot 0,05}{1 - 0,158} \\ &= \underline{\underline{0,00594}} \end{aligned}$$

ii) Her representerer $P(D|K)$ sannsynligheten for at en kassert enhet faktisk var defekt, mens $P(D|\bar{K})$ er sannsynligheten for at en enhet som ikke ble kassert, var defekt (dvs. "feilvare" som ikke oppdages i kvalitetskontrollen).

c)

i) Ettersom en enhet enten er defekt eller ikke (to utfall); enhetene antas uavhengige og sannsynligheten for at en enhet er defekt, er konstant, er X binomisk fordelt.

ii) Sannsynlighet for to eller flere enheter i et parti på 100:

$$P(X \geq 2) = 1 - P(X \leq 1)$$

Kan enten bruke kumulativ binomisk fordeling på kalkulator; tabell eller punktsannsynlighetene (fordi $X \leq 1$ tilsvarer at $X = 0$ eller $X = 1$). Velger kalkulator:

```
Binomial C.D
Data      :Variable
X         :1
Numtrial :100
P         :6E-03
Save Res:None
Execute
None 0.122
```

Får at

$$\begin{aligned} P(X \geq 2) &= 1 - P(X \leq 1) \\ &= 1 - 0,878 \\ &= \underline{\underline{0,122}} \end{aligned}$$

iii) La Y angi antall feilfrie enheter. Sannsynligheten for at en enhet er feilfri, er $p = 1 - 0,006 = 0,994$. Finner antall enheter som må kjøpes for at sannsynligheten for at varepartiet inneholder minst 100 feilfrie enheter, blir minst 99 % - vi skal altså bestemme n slik at

$$\begin{aligned} P(Y \geq 100) &\geq 0,99 \\ 1 - P(Y \leq 99) &\geq 0,99 \\ P(Y \leq 99) &\leq 0,01 \end{aligned}$$

Prøver på kalkulator til man finner en n som oppfyller kravet (binomisk kumulativ fordeling der man justerer n helt til $P(Y \leq 99) \leq 0,01$):

n	$P(Y \leq 99)$
100	0,452
101	0,123
102	0,0239
103	0,00358

Ser altså at $n = \underline{\underline{103}}$ oppfyller kravet til at det skal være minst 99 % sannsynlig at minst 100 enheter er feilfrie.

Oppgave 2

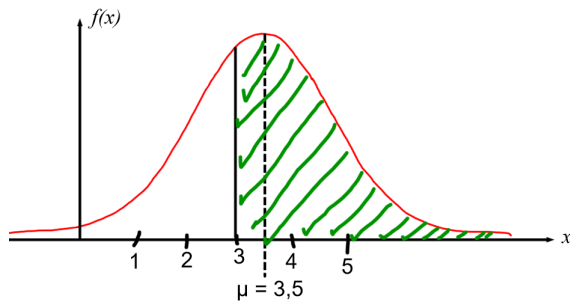
a) Gitt at $X \sim N(3,5, 4,0)$, dvs. X er normalfordelt med $\mu = 3,5$ og $\sigma^2 = 4,0$.

i) Finner $P(X > 3,0)$ på kalkulator:

```
Normal C.D
Lower : 3
Upper : 200
σ : 2
μ : 3,5
Save Res: None
Execute
None List
```

$$P(X > 3,0) = \underline{\underline{0,599}}$$

Skisse av arealet:



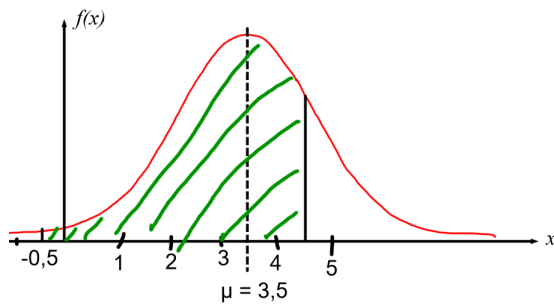
ii) Finner $P(-0,5 < X < 4,5)$ på kalkulator:

```
Normal C.D
Lower : -0,5
Upper : 4,5
σ : 2
μ : 3,5
Save Res: None
Execute
None List
```

Får at

$$P(-0,5 < X < 4,5) = \underline{\underline{0,669}}$$

Skisse:



b) Skal bestemme en k som er slik at

$$P(3,5 - k \leq X \leq 3,5 + k) = 0,99$$

Konverterer til en standard normalfordelt variabel Z ved å trekke fra μ og dele med σ :

$$\begin{aligned} P\left(\frac{3,5 - k - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{3,5 + k - \mu}{\sigma}\right) &= 0,99 \\ P\left(\frac{3,5 - k - 3,5}{2} \leq Z \leq \frac{3,5 + k - 3,5}{2}\right) &= 0,99 \\ P\left(-\frac{k}{2} \leq Z \leq \frac{k}{2}\right) &= 0,99 \\ P\left(Z \leq \frac{k}{2}\right) - P\left(Z \leq -\frac{k}{2}\right) &= 0,99 \end{aligned}$$

Bruker at fordelingsfunksjonen $G(-x) = 1 - G(x)$ (fra formelarket), slik at

$$\begin{aligned} P\left(Z \leq \frac{k}{2}\right) - \left(1 - P\left(Z \leq \frac{k}{2}\right)\right) &= 0,99 \\ 2P\left(Z \leq \frac{k}{2}\right) &= 1,99 \\ P\left(Z \leq \frac{k}{2}\right) &= 0,995 \end{aligned}$$

Leser ut fra normalfordelingstabellen at (verdien ligger midt mellom 2,57 og 2,58, dvs. 2,575):

$$\frac{k}{2} = 2,575 \Rightarrow k = \underline{\underline{5,15}}$$

c) Gitt lineærkombinasjonen

$$Y = 2X - 1,$$

og vi skal bestemme

$$P(Y < 0) = P(2X - 1 < 0)$$

Vi må bestemme hva slags fordeling Y får. Vi vet at lineærkombinasjoner av normalfordelte variable er selv normalfordelt. Finner $\mu_Y = E(Y)$ og $\sigma_Y^2 = Var(Y)$ ved hjelp av regnereglene for forventningsverdi og varians

$$\begin{aligned} E(Y) &= E(2X - 1) \\ &= E(2X) - E(1) \\ &= 2E(X) - 1 \\ &= 2\mu - 1 \\ &= 2 \cdot 3,5 - 1 \\ &= \underline{6} \end{aligned}$$

$$\begin{aligned} Var(Y) &= Var(2X - 1) \\ &= Var(2X + (-1)) \\ &= Var(2X) + Var(-1) \\ &= 2^2 Var(X) + 0 \\ &= 4 \cdot 4 \\ &= \underline{16} \end{aligned}$$

Dette gjør at vi kan finne $P(Y < 0)$ på kalkulator:

```
Normal C.D
Lower : -200
Upper : 0
σ : 4
μ : 16
Save Res: None
Execute
None List
```

Får at

$$P(Y < 0) = \underline{\underline{0,0668}}$$

Oppgave 3

a)

i) Bestemmer sannsynligheten for at en vilkårlig husstand har minst én bil, dvs.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - 0,15 \\ &= \underline{\underline{0,85}} \end{aligned}$$

ii) Forventningsverdi:

$$\begin{aligned} \mu = E(X) &= \sum_{x=0}^3 x \cdot P(X = x) \\ &= 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2 \cdot P(X = 2) + 3 \cdot P(X = 3) \\ &= 1 \cdot 0,55 + 2 \cdot 0,25 + 3 \cdot 0,05 \\ &= \underline{\underline{1,2}} \end{aligned}$$

Varsians:

$$\begin{aligned} \sigma^2 = Var(X) &= \sum_{x=0}^3 x^2 \cdot P(X = x) - \mu^2 \\ &= 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2 \cdot P(X = 2) + 3 \cdot P(X = 3) - 1,2^2 \\ &= 1^2 \cdot 0,55 + 2^2 \cdot 0,25 + 3^2 \cdot 0,05 - 1,2^2 \\ &= \underline{\underline{0,56}} \end{aligned}$$

b) Summen X_{sum} av antall biler til 400 (uavhengige) husstander er via sentralgrenseteoremet normalfordelt som

$$X_{sum} = X_1 + X_2 + \dots + X_{400} \sim N(400\mu, 400\sigma^2) \sim N(480, 224).$$

Summen har altså forventningsverdi lik 480 og varians lik 224, slik at standardavviket er $\sqrt{224} = 14,97$. Vi finner at

$$P(X_{sum} > 480) = 1 - P(X_{sum} \leq 480)$$

Vi kan enten finne $P(X_{sum} > 480)$ direkte på kalkulatoren, eller gå via den kumulative fordelingen på kalkulator eller tabell (må da normalisere først). Det raskeste er å innse at vi skal finne sannsynligheten for at variabelen er større enn forventningsverdien μ , og ettersom normalfordelingen er symmetrisk om μ , er sannsynligheten lik $\frac{1}{2}$.

Bildet under viser hvordan sannsynligheten kan bestemmes på kalkulator:

```
Normal C.D
Lower :480
Upper :10000
σ :14,9666295
μ :480
Save Res:None
Execute
None List
```

Får uansett metode at

$$P(X_{sum} > 480) = \underline{\underline{\frac{1}{2}}}$$

Oppgave 4

a) Gitt en Weibull-fordelt variabel $X \sim W(\beta = 2, \eta = 7)$.

i) Forventningsverdi:

$$\begin{aligned}\mu = E(X) &= \eta \cdot \Gamma\left(1 + \frac{1}{\beta}\right) \\ &= 7 \cdot \Gamma\left(1 + \frac{1}{2}\right) \\ &= 7 \cdot \Gamma(1,5) \\ &= 7 \cdot 0,88623 \\ &= \underline{\underline{6,20}}\end{aligned}$$

ii) Varians og standardavvik:

$$\begin{aligned}\sigma^2 = Var(X) &= \eta^2 \cdot \left(\Gamma\left(1 + \frac{2}{\beta}\right) - \left(\Gamma\left(1 + \frac{1}{\beta}\right)\right)^2\right) \\ &= 49 \cdot \left(\Gamma\left(1 + \frac{2}{2}\right) - \left(\Gamma\left(1 + \frac{1}{2}\right)\right)^2\right) \\ &= 49 \cdot \left(\Gamma(2) - (\Gamma(1,5))^2\right) \\ &= 49 \cdot (1 - 0,88623^2) \\ &= \underline{\underline{10,5}}\end{aligned}$$

b) Det at det **ikke** produseres strøm, tilsvarer altså at $X < 3$ eller $X > 25$. Vi kan enten summere sannsynlighetene for dette, eller ta 1 og trekke fra sannsynligheten for at X ligger i området der det produseres strøm:

$$\begin{aligned}P(X < 3) + P(X > 25) &= 1 - P(3 \leq X \leq 25) \\ &= 1 - (P(X \leq 25) - P(X \leq 3)) \\ &= 1 - \left(\left(1 - e^{-\left(\frac{25}{7}\right)^2}\right) - \left(1 - e^{-\left(\frac{3}{7}\right)^2}\right)\right) \\ &= \underline{\underline{0,168}}\end{aligned}$$

c)

i) La Y angi antall vindmøller som utvikler feil og må repareres i en tilfeldig uke. Ettersom antall vindmøller som utvikler feil antas jevnt fordelt over tidsintervallet (konstant "rate"); de svikter

uavhengige av hverandre og ikke to svikter samtidig, oppfyller Y kravene til en Poisson-fordeling, med $\lambda = 3$.

ii) Finner sannsynligheten for flere enn 3 vindmøller med feil i løpet av en uke, ved hjelp av kumulativ Poisson-fordeling på kalkulator:

$$\begin{aligned} P(Y > 3) &= 1 - P(Y \leq 3) \\ &= 1 - 0,647 \\ &= \underline{\underline{0,353}} \end{aligned}$$

d)

i) Det er i snitt 3 vindmøller som svikter per uke. Intensiteten av hendelser er altså

$$\alpha = \frac{3}{uke}$$

Tiden T mellom hver svikt er eksponentialfordelt med parameter $\alpha = \frac{3}{uke}$. Forventet tid mellom hver svikt er da

$$E(T) = \frac{1}{\alpha} = \frac{1}{\frac{3}{uke}} = \underline{\underline{\frac{1}{3} uke}}$$

ii) Sannsynligheten for at det går mer enn en uke mellom hver svikt er da (måler tiden i uker):

$$\begin{aligned} P(T > 1) &= 1 - P(T \leq 1) \\ &= 1 - (1 - e^{-\alpha \cdot 1}) \\ &= e^{-3 \cdot 1} \\ &= \underline{\underline{0,0498}} \end{aligned}$$

Oppgave 5

a)

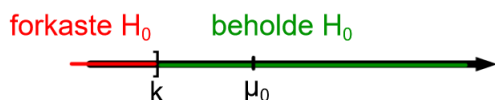
i) Innfører

$$H_0: \mu = 70$$

$$H_1: \mu < 70$$

Med kritisk verdi

Finner kritisk verdi for 1 % signifikansnivå. Starter med en skisse over situasjonen:



Vi kommer altså til å forkaste H_0 dersom snittet i stikkprøven er **mindre** enn kritisk verdi. Bruker definisjonen av signifikansnivå til å bestemme kritisk verdi:

$$\begin{aligned} P(\bar{X} \leq k | H_0) &= \alpha = 0,01 \\ P\left(\frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \leq \frac{k - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right) &= 0,01 \\ P\left(Z \leq \frac{k - 70}{\frac{0,7}{\sqrt{8}}}\right) &= 0,01 \end{aligned}$$

z -verdien vil her tilsvare 1 %-kvantilen $u_{0,01} = 2,326$, men med negativt fortegn:

$$\frac{k - 70}{\frac{0,7}{\sqrt{8}}} = -2,326 \Rightarrow k = 70 - 2,326 \cdot \frac{0,7}{\sqrt{8}} = \underline{\underline{69,42}}$$

Ettersom snittet i stikkprøven er $\bar{X} = 69,36$ er **mindre** enn kritisk verdi, vil vi derfor **forkaste** H_0 , og hevde at det er **for lite** virkestoff i tabletten.

Med testobservator

Testobservatoren er

$$\begin{aligned}U_0 &= \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \\&= \frac{69,36 - 70}{\frac{0,7}{\sqrt{8}}} \\&= \underline{\underline{-2,59}}\end{aligned}$$

Ettersom $|U_0| > u_{0,01} = 2,326$, vil vi derfor **forkaste** H_0 , og hevde at det er **for lite** virkestoff i tabletten.

ii) Signifikansnivået α er sannsynligheten for å feilaktig forkaste nullhypotesen - dvs. å hevde at det er for lite virkestoff ut i fra stikkprøven, selv om det er realiteten er tilfelle at tablettene inneholder 70 mikrogram.

b)

i) Signifikanssannsynligheten p er sannsynligheten for å få en minst like “ekstrem” stikkprøve $\bar{X}$ som den faktiske stikkprøven $\bar{x} = 69,36$, dersom H_0 er riktig:

$$\begin{aligned}p &= P(\bar{X} \leq \bar{x} | H_0) \\&= P(\bar{X} \leq 69,36 | H_0)\end{aligned}$$

Ettersom H_0 antas sann, er her $\mu = 70$ og standardavviket for snittet $\bar{X}$ er $\frac{\sigma}{\sqrt{n}} = \frac{0,7}{\sqrt{8}} = 0,247$, slik at vi kan finne denne sannsynligheten direkte på kalkulator:

```
Normal C.D
Lower  : -200
Upper  : 69,36
σ      : 0,24748737
μ      : 70
Save Res: None
Execute
Calc
```

Får at

$$p = \underline{\underline{0,00486}}$$

ii) Ettersom $p < \alpha = 0,01$, vil vi også her forkaste H_0 - på samme måte som i a).

c) Finner styrkefunksjonen $\beta(\mu = 69)$, dvs. teststyrken når $\mu = 69$, ved å bruke den kritiske verdien fra oppgave a:

$$\begin{aligned}\beta(\mu = 69) &= P(\text{akseptere } H_1 | \mu = 69) \\&= P(\bar{X} \leq k | \mu = 69) \\&= P(\bar{X} \leq 69,42 | \mu = 69)\end{aligned}$$

Denne kan vi finne direkte på kalkulator (standardavviket for $\bar{X}$ er fortsatt $\frac{0,7}{\sqrt{8}} = 0,247$):

```
Normal C.D
Lower  : -200
Upper  : 69,42
σ      : 0,24748737
μ      : 69
Save Res: None
Execute
None List
```

Får at

$$\beta(\mu = 69) = \underline{\underline{0,955}}$$

Det er altså 95,5 % sannsynlighet for at hypotesetesten vår “fanger opp” det faktum at det er for lite virkestoff dersom det i realiteten var slik at tablettene inneholdt 69 mikrogram, mens produsenten hevder at det er 70 mikrogram.

d) Dersom standardavviket ikke hadde vært kjent, måtte vi ha brukt det empiriske standardavviket, som vi finner fra den oppgitte empiriske variansen:

$$S = \sqrt{S^2} = \sqrt{0,5798} = 0,7614$$

Utfører t -test med $n - 1 = 7$ frihetsgrader:

$$\begin{aligned} T_0 &= \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \\ &= \frac{69,36 - 70}{\frac{0,7614}{\sqrt{8}}} \\ &= \underline{\underline{-2,377}} \end{aligned}$$

Sammenlikner denne med kvantilen $t_{\alpha, n-1} = t_{0,01,7} = 2,998$. Ettersom $|T_0| = 2,377 < t_{0,01,7}$, kan vi **ikke** forkaste nullhypotesen - dvs. hvis standardavviket er ukjent, får vi motsatt konklusjon av tilfellet i oppgave a).