Eksamen 1998 - Løsning

1. X = Antall fisk i en teine.
I oppgaven er gitt at X kan betraktes som Poisson-fordelt.
En begrunnelse for at denne antakelsen kan benyttes er følgende:
Vi lar A være hendelsen at en fisk ankommer en bestemt teine.
Gjennomsnittlig antall ankomster er gitt ved: $\lambda = 0.8$ fisk/uke.
Område er her tiden t (med tidsenhet uke).

Poisson-fordel fordi:

1) Forekomster av A i ikke-overlappende områder er uavhengige.
Det er antatt at fiskene går inn i en teine uavhengig av hverandre.
At det er gått en fisk inn i teinen ved tiden t_1, har ingen innvirkning av hvorvidt en annen fisk går inn i teinen ved et annet tidspunkt t_2.
2) Forventet antall forekomster λ av A pr enhet er konstant over hele området.
$\lambda = 0.8$ fisk/uke. Denne er konstant for hver uke.
3) To forekomster av A kan ikke være fullstendig sammenfallende.
To fisk kan ikke ankomme en teine på nøyaktig samme tidspunkt.
Tidspunkt kan vi fin-inndele så mye vi måtte ønske.

a)

X = Antall fisk i en teine

$X \sim \text{Po}(\lambda t) \qquad \lambda = 0.8 \text{ fisk/uke} \qquad t = \text{Antall uker}$

Sannsynligheten for x antall fisk i en teine i løpet av tiden t (antall uker) :

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

Sannsynligheten for 3 fisk i en teine en uke (x = 3, t = 1) :

$$P(X = 3) = \frac{(0.8 \cdot 1)^3}{3!} e^{-0.8 \cdot 1} = \frac{0.8^3}{3!} e^{-0.8} = \underline{\underline{0.0383}}$$

Sannsynligheten for å få minst en fisk (dvs få fisk) i en teine en uke ($x \geq 1$, t = 1) :

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{(0.8 \cdot 1)^0}{0!} e^{-0.8 \cdot 1} = 1 - \frac{0.8^0}{0!} e^{-0.8} = 1 - e^{-0.8} = \underline{\underline{0.55}}$$

b) Y = Antall teiner med fisk.
Et delførsøk er her en ukes utplassering av en teine.
La A være hendelsen at det er fisk i en teine.
I hvert delforsøk registrerer vi hvorvidt hendelsen A har inntruffet eller ikke, dvs for hver teine registrerer vi hvorvidt det er fisk i denne teinen eller ikke.
Siden vi har 4 teiner, har vi n = 4 delforsøk.

Binomisk fordelt fordi:

1) Delforsøkene er uavhengige.
   Ankomst av fisk i en teine er uavhengig av ankomst av fisk i en annen teine.
2) I hvert delforsøk registreres hvorvidt et utfall A inntreffer.
   For hver teine registreres det hvorvidt det finnes fisk i teinen.
3) Sannsynligheten p = P(A) er den samme i alle n delforsøk.
   Samme sannsynlighet (fra a: 0.55) for å få fisk i en vilkårlig av de n = 4 teinene.

n = 4, p = 0.55

Sannsynlig heten for at y av de n = 4 teinene inneholder fisk i løpet av en uke :

$$P(Y = y) = \binom{n}{y} \cdot p^y \cdot (1-p)^{n-y}$$

Sannsynlig heten for at akkurat 3 av teinene inneholder fisk en uke :

$$P(Y = 3) = \binom{4}{3} \cdot 0.55^3 \cdot (1-0.55)^{4-3} = \binom{4}{3} \cdot 0.55^3 \cdot 0.45 = \underline{\underline{0.2995}}$$

Gjennomsni ttlig antall teiner med fisk pr uke =Forventet antall teiner med fisk pr uke :

$$E(Y) = np = 4 \cdot 0.55 = \underline{\underline{2.2}}$$

c) Simulering av antall teiner som det er fisk i pr uke for de 10 ukene:

Beregner den kumulative sannsynlighet P(Y<=y) for y = 0,1,2,3,4.
Ved å ta med 4 desimaler, vil det være naturlig å foreta simuleringen ved å dele inn intervallet [0,9999] etter resultatene fra den kumulative sannsynlighetsberegningen.

y	P(Y<=y)	Intervall
0	0.0410	0000 - 0409
1	0.2415	0410 - 2414
2	0.6090	2415 - 6089
3	0.9085	6090 - 9084
4	1	9085 - 9999

Velger nå 10 randomtall (her kolonne nr 1 fra vedlegget).
Ser hvilket intervall det tilfeldige tallet tilhører.
Antall teiner med fisk er da tilhørende y-verdi.

Uke	Tilfeldige tall	Antall teiner med fisk
1	1306	1
2	0422	1
3	6597	3
4	7965	3
5	7695	3
6	5160	2
7	2961	2
8	1428	1
9	3666	2
10	6543	3

Simuleringen gir gjennomsnittlig antall teiner med fisk pr uke lik:

(1+1+3+3+3+2+2+1+2+3)/10 = 2.1

Dette stemmer rimelig bra med beregnet verdi E(Y) = 2.2 fra oppgave b).

Simulering av antall fisk pr uke de 5 første ukene:

Beregner den kumulative sannsynlighet P(X<=x) inntil vi tilnærmet når kumulativ sannsynlighet 1. Vi må da gjøre beregninger for x = 0,1,2,3,4,5. Ved å ta med 3 desimaler, vil det være naturlig å foreta simuleringen ved å dele inn intervallet [0,999] etter resultatene fra den kumulative sannsynlighetsberegningen.

x	P(X<=x)	Intervall
0	0.449	000 - 448
1	0.809	449 - 808
2	0.953	809 - 952
3	0.991	953 - 990
4	0.999	991 - 998
5	1	999

Velger nå 4 randomtall (siden vi har 4 teiner) for hver av de 5 ukene, dvs i alt 20 randomtall.
Tallene velger vi fra kolonne nr 1 og 2 (og tar med kun de 3 første sifrene).
Ser hvilket intervall det tilfeldige tallet tilhører.
Antall fisk i en teine er da tilhørende x-verdi.

Uke	Tilfeldige tall		Antall fisk pr teine		Antall fisk pr uke
1	118	202	0	0	1
	243	654	0	1	
2	693	055	1	0	2
	785	418	1	0	
3	564	679	1	1	3
	573	064	1	0	
4	616	564	1	1	4
	040	846	0	2	
5	053	431	0	0	2
	453	745	1	1	

Simuleringen gir gjennomsnittlig antall fisk pr teine pr uke lik:

$(1+2+3+4+3)/(4\cdot 5) = 0.6$

Dette stemmer rimelig bra med oppgitt verdi $\lambda = 0.8$ tatt i betraktning få antall simuleringer.

2. a) Sannsynligheten for at den ansatte skal bruke mellom 24 og 28 minutter på et oljeskift:

$$P(24 < X < 28) = G\left(\frac{28-25}{4}\right) - G\left(\frac{24-25}{4}\right) = G(0.75) - G(-0.25)$$
$$= G(0.75) - [1 - G(0.25)] = 0.7734 - [1 - 0.5987] = \underline{\underline{0.3721}}$$

b) 95% konfidensintervall for μ :

X = Oljeskiftt id $\sim N(\mu, \sigma^2) = N(\mu, 4^2)$

$n = 10$

$$\hat{\mu} = \overline{X} = \frac{25+29+28+25+27+26+31+28+27+26}{10} = 27.2$$

$\alpha = 0.05$

$$\hat{\mu} \pm u_{\alpha/2} SD(\hat{\mu}) = \overline{X} \pm u_{0.025} \frac{\sigma}{\sqrt{n}} = 27.2 \pm 1.96 \cdot \frac{4}{\sqrt{10}} = 27.2 \pm 2.5 = \underline{\underline{[24.7, 29.7]}}$$

Et slikt konfidensi ntervall vil i 95% av tilfellen e dekke (inneholde) parametere n μ.

c) Bestemmelse av n = antall oljeskifttidmålinger
for å få et 99% konfidensintervall som ikke er bredere enn
konfidensintervallet vi fant i b):

$\alpha = 0.01 \quad \Rightarrow \quad \alpha/2 = 0.005$

$$u_{0.005} \frac{\sigma}{\sqrt{n}} \le u_{0.025} \frac{\sigma}{\sqrt{10}}$$

$$\sqrt{n} \ge \frac{u_{0.005}}{u_{0.025}} \sqrt{10} = \frac{2.576}{1.96} \sqrt{10} = 4.156$$

$n \ge 17.3$

$\underline{\underline{n = 18}}$

d) Hypotesetest for hvorvidt forventet oljeskifttid er 24 minutter eller mer:

$H_0: \quad \mu = 24$ Nullhypote se

$H_1: \quad \mu > 24$ Alternativ hypotese

$\alpha = 0.05$ Signifikan snivå

$P(H_1 \mid H_0) = 0.05$

$\alpha = 0.05$ er sannsynlig heten for å påstå $\mu > 24$ når $\mu = 24$ er riktig.

$$\mathrm{H_0} \text{ forkastes hvis } \overline{\mathrm{X}} > \mathrm{k} = 24 + u_{0.05} \frac{4}{\sqrt{10}}, \text{ dvs hvis } \mathrm{Z} = \frac{\overline{X} - 24}{\frac{4}{\sqrt{10}}} > u_{0.05} = 1.645$$

$$\mathrm{Z} = \frac{27.2 - 24}{\frac{4}{\sqrt{10}}} = 2.53 > 1.645$$

eller :

$$k = 24 + u_{0.05} \frac{4}{\sqrt{10}} = 24 + 1.645 \cdot \frac{4}{\sqrt{10}} = 26.081$$

$\overline{X} = 27.2$

Vi påstår $\mathrm{H_1}$, dvs $\mathrm{H_0}$ forkastes.

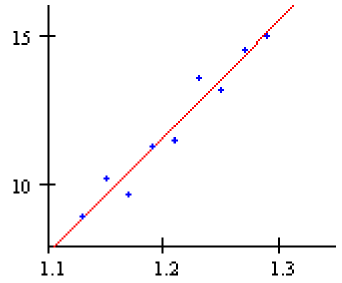
Testen vis er med 95% sikkerhet at eieren tar feil og at $\mu > 24$

3. a) Estimering av α og β vha minste kvadraters metode (MKM):

$$\hat{\beta} = \frac{1}{M}\sum_{i=1}^{n}(x_i - \bar{x})y_i = \frac{0.9414}{0.024} = \underline{\underline{39.225}}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = \frac{1}{9}(107.93 - 39.225 \cdot 10.89) = \underline{\underline{-35.47}}$$

Regresjons linje : $y = 39.225x - 35.47$



b) 90% konfidensintervall for β:

$$S^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - (\hat{\alpha} + \hat{\beta}x_i))^2 = \frac{1}{7} \cdot 2.2175 = 0.3168$$

$S = 0.5628$

$$\hat{\beta} \pm t_{0.05,9-2}\frac{S}{\sqrt{M}} = 39.225 \pm 1.895 \cdot \frac{0.546}{\sqrt{0.024}} = 39.225 \pm 6.679 = [32.54, 45.90]$$

Resultatet tyder på at $\beta \neq 30.0$ siden 30.0 ikke ligger i intervalle t.

4. F = Fargeblind
M = Mann
K = Kvinne

Følgende opplysninger er gitt i oppgaven:

$P(M) = P(K) = 0.5$ Like mange menn og kvinner, dvs sannsynlig heten for at en tilfeld ig uttrukket person er en mann eller en kvinne er begge lik 0.5.

$P(F \mid M) = 0.05$ Sannsynlig heten for at er person er fargeblind gitt at personen er en mann.

$P(F \mid K) = 0.0025$ Sannsynlig heten for at en person er fargeblind gitt at personen er en kvinne.

a) Sannsynligheten for at en tilfeldig uttrukket person er en fargeblind kvinne:

$$P(F \mid K) = \frac{P(F \cap K)}{P(K)} \quad \Rightarrow \quad P(F \cap K) = P(F \mid K)P(K) = 0.0025 \cdot 0.5 = \underline{\underline{0.00125}}$$

b) Sannsynligheten for at en tilfeldig uttrukket person er en mann gitt av personen er fargeblind:

$$P(M \mid F) = \frac{P(M)}{P(F)} P(F \mid M) = \frac{P(M)}{P(F \mid M)P(M) + P(F \mid K)P(K)} P(F \mid M)$$

$$= \frac{0.5}{0.05 \cdot 0.5 + 0.0025 \cdot 0.5} \cdot 0.05 = \underline{\underline{0.9524}}$$

5. Undersøkelse av hvorvidt de to metodene gir jern med forskjellig smeltepunkt:

$n_A = 9 \qquad n_B = 6$

$$\overline{X}_A = \frac{1}{n_A}\sum_{i=1}^{n_a} x_i = \frac{1493 + 1519 + 1512 + 1514 + 1489 + 1507 + 1508 + 1518 + 1495}{9} = 1506.11$$

$$S_A^{\,2} = \frac{1}{n_A - 1}\sum_{i=1}^{n_A}(x_i - \overline{X}_A)^2 = \frac{1}{n_A - 1}[\sum_{I=1}^{n_A} x_i^{\,2} - n_A \overline{X}_A^{\,2}]$$

$$= \frac{1}{9-1}[1493^2 + 1519^2 + 1512^2 + 1514^2 + 1489^2 + 1507^2 + 1508^2 + 1518^2 + 1495^2 - 9 \cdot 1506.11^2] = 133.5$$

$$S_A = \sqrt{S_A^{\,2}} = \sqrt{133.5} = 11.16$$

$$\overline{X}_B = \frac{1}{n_B}\sum_{i=1}^{n_b} x_i = \frac{11509 + 1495 + 1512 + 1483 + 1507 + 1491}{6} = 1499.5$$

$$S_B^{\,2} = \frac{1}{n_B - 1}\sum_{i=1}^{n_B}(x_i - \overline{X}_B)^2 = \frac{1}{n_B - 1}[\sum_{I=1}^{n_B} x_i^{\,2} - n_B \overline{X}_B^{\,2}]$$

$$= \frac{1}{6-1}[1509^2 + 1495^2 + 1512^2 + 1483^2 + 1507^2 + 1491^2 - 6 \cdot 1499.5^2] = 133.403$$

$$S_B = \sqrt{S_B^{\,2}} = 11.55$$

$$S^2 = \frac{1}{n_A + n_B - 2}\left[(n_A - 1)S_A^{\,2} + (n_B - 1)S_B^{\,2}\right]$$

$$= \frac{1}{9+6-2}\left[(9-1)\cdot 11.16^2 + (6-1)\cdot 11.55^2\right] = \frac{1}{13}\left[8 \cdot 11.16^2 + 5 \cdot 11.55^2\right] = 127.952$$

$$S = \sqrt{S^2} = \sqrt{127.952} = 11.312$$

(Det er ikke foretatt dobbeltsjekking på hvorvidt tallene ovenfor er korrekte.)

$\mathrm{H}_0 : \mu_A - \mu_B = 0$ Nullhypotese

$\mathrm{H}_1 : \mu_A - \mu_B \neq 0$ Alternativ hypotese

Påstår H_1 hvis $|\mathrm{T}_0| = \left| \dfrac{\overline{\mathrm{X}}_\mathrm{A} - \overline{\mathrm{X}}_\mathrm{B}}{S\sqrt{\dfrac{1}{n_A} + \dfrac{1}{n_B}}} \right| > t_{0.025,9+6-2} = t_{0.025,13} = 2.160$

$$|\mathrm{T}_0| = \left| \frac{\overline{\mathrm{X}}_\mathrm{A} - \overline{\mathrm{X}}_\mathrm{B}}{S\sqrt{\dfrac{1}{n_A} + \dfrac{1}{n_B}}} \right| = \left| \frac{1506.11 - 1499.5}{11.312\sqrt{\dfrac{1}{9} + \dfrac{1}{6}}} \right| = 1.109$$

Konklusjon : Påstår ikke H_1

Det tyder ikke på at de to metodene gir jern med forskjellig smeltepunkt.