Norges teknisk–naturvitenskapelige universitet
Institutt for matematiske fag

Oppgave 1 Vannverket

a) $P(A_1/A_2)$ er forskjellig fra $P(A_1)$ medører at A_1 og A_2 er avhengige.
$P(A_1) + P(A_2) = 1.6 > 1$ medører at A_1 og A_2 ikke er disjunkte.
$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = 0.8 + 0.8 - 0.9 \cdot 0.8 = 0.88$

b) $0.977 = P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) - P(A_1 \cap A_2) + P(A_3) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$.
Dette gir: $0.977 = P(A_1 \cup A_2) + P(A_3) - P(A_1) \cdot P(A_3) - P(A_2) \cdot P(A_3) + p(A_1 \cap A_2) \cdot P(A_3)$
eller: $P(A_3)(1 - 0.8 - 0.8 + 0.72) = 0.977 - 0.88$ som gir $P(A_3) = \frac{0.117}{0.120} = 0.975$
Vi har uavhengihet mellom dagene, vi registrerer hvorvidt systemet fungerer eller ikke og sannsynligheten for at systemet ikke fungerer er den samme hver dag. Dette medører at T =antall dager (antall forsøk) til 1. svikt er geometrisk fordelt med sannsynlighet $p = 0.003$
$E(T) = \frac{1}{p} = \frac{1}{0.003} = 333$

Oppgave 2 Avviksrapporter

a) For en uke, dvs $t = 1$, er $N \sim Po(1 \cdot 1.5) = Po(1.5)$.
$P(N = 0) = P(N \leq 0) = 0.2231$ (slår opp i tabell for $\mu = 1.5$).

For fire uker, $t = 4$, er $N \sim Po(4 \cdot 1.5) = Po(6.0)$.
$P(N > 2) = 1 - p(N \leq 2) = 1 - 0.062 = 0.9380$ (slår opp i tabell for $\mu = 6.0$)

b) Lar $N(t_1, t_2)$ vere antall meldinger som kommer inn i mellom tid t_1 og t_2. Vi vet at $N(t_1, t_2) \sim Po(\lambda \cdot (t_2 - t_1))$. Skal finne $P(N(0, 1) = 1|N(0, 3) = 1)$:

$$P(N(0, 1) = 1|N(0, 3) = 1) = P(N(0, 1) = 1 \cap N(0, 3) = 1|N(0, 3) = 1)$$

Hendelsen $N(0, 1) = 1 \cap N(0, 3) = 1$, at det kommer inn ei melding første uka og ei melding i løpet av alle tre ukene, han kun bli opppflyt dersom det ikke kommer inn noen meldinger i løpet av dei siste to ukene, dvs

$$P(N(0, 1) = 1 \cap N(0, 3) = 1|N(0, 3) = 1) = P(N(0, 1) = 1 \cap N(1, 3) = 0|N(0, 3) = 1)$$

Da meldingene ankommer uavhengig, er $N(0, 1)$ og $N(1, 3)$ uavhengige.

$$P(N(0,1)=1\cap N(1,3)=0|N(0,3)=1)=\frac{P(N(0,1)=1)P(N(1,3)=0)}{P(N(0,3)=0)}$$
$$=\frac{\frac{(\lambda\cdot 1)^1}{1!}\exp(-\lambda\cdot 1))\cdot\frac{(\lambda\cdot 2)^0}{0!}\exp(-\lambda\cdot 2)}{\frac{(\lambda\cdot 3)^1}{1!}\exp(-\lambda\cdot 3)}=\frac{1}{3}$$

Kan og argumentere for svaret ved at vi har en poissonprosess med konstant intensitet, og dermed er sannsynet for at en hendelse skjer i et tidsintervall proposjonalt med lengden på intervallet.

Ser på kummulativ fordeling for T, tidspunktet når første melding kommer inn:

$$\begin{aligned}
F_T(t) &= P(T<t|N(0,3)=1)\\
&= \frac{P(N(0,t)=1)P(N(t,3)=0)}{P(N(0,3)=0)}\\
&= \frac{\frac{(\lambda\cdot t)^1}{1!}\exp(-\lambda\cdot t))\cdot\frac{(-\lambda\cdot(3-t))^0}{0!}\exp(-\lambda\cdot(3-t))}{\frac{(-\lambda\cdot 3)^1}{1!}\exp(-\lambda\cdot 3)}=\frac{t}{3}
\end{aligned}$$

Deriverer for å finne sannsylighetstettheten, $f_T(t)=\dfrac{dF_T(t)}{dt}=1/3$, altså uniformt fordelt over hele intervallet.

Dette er rimelig når vi har poissonprosess. Kan dele opp de tre ukene i mangen små tidsintervall intervall av lik lengde. Disse vil ha lik sannsynlighet for å inneholde meldingen.

c) Finner likelihoodfunksjonen:

$$L(\lambda;n)=f(n;\lambda t)=\frac{(\lambda t)^n}{n!}\exp(-\lambda t)$$

Finner log-likelihoodfunksjonen:

$$l(\lambda;n)=\log(L(\lambda;n))=\log(t^n/n!)+n\log(\lambda)-\lambda t$$

Finn toppunktet;

$$\frac{dl(\lambda;n)}{d\lambda}=0$$
$$\frac{n}{t}-t=0$$
$$\lambda=\frac{n}{t}$$

$$\hat{\lambda}_{SME} = \frac{N}{t} = \frac{N}{52}$$

$\mathrm{E}(\lambda_{SME}) = \mathrm{E}(\frac{N}{52}) = \mathrm{E}(N)/52 = \lambda \cdot 52/52 = \lambda.$

$\mathrm{Var}(\lambda_{SME}) = \mathrm{Var}(\frac{N}{52}) = \mathrm{Var}(N)/52^2 = \lambda \cdot 52/52^2 = \lambda/52$

Da $\mathrm{E}(\lambda_{SME}) = \lambda$ er λ_{SME} forventningsrett.

Estimat $\lambda^*_{SME} = \frac{n}{52} = 104/52 = 2.0.$

LØSNINGS-SKISSE

Oppgave 3

$K_0 \sim n(k_0; 25, 4)$

a) $P(K_0 \geq 30) = P\left(\frac{K_0-25}{4} \geq \frac{30-25}{4}\right) = P(Z \geq 1.25) = \underline{\underline{0.1056}}$

$P(20 \leq K_0 < 30) = P\left(\frac{20-25}{4} \leq \frac{K_0-25}{4} < \frac{30-25}{4}\right)$

$= P(-1.25 \leq Z < 1.25) = 1 - 2P(Z < -1.25) = \underline{\underline{0.7888}}$

b) $K_{01}, \ldots, K_{010}$ uif $n(k_0; \mu_0, \sigma_0)$

$\hat{\mu}_0 = \frac{1}{10}\sum_{i=1}^{10} K_{0i} \sim n\left(\hat{\mu}_0; \mu_0, \frac{\sigma_0}{\sqrt{10}}\right)$

$\hat{\sigma}_0^2 = \frac{1}{9}\sum_i (K_{0i} - \hat{\mu}_0)^2 = 18.44 = 4.3^2$

$T = \frac{\hat{\mu}_0 - \mu_0}{\hat{\sigma}_0/\sqrt{10}} \sim t(\cdot; 9)$

$\text{Prob}\{-t_{9,0.05} < T < t_{9,0.05}\} = 0.9$

$\Downarrow$

$\text{Prob}\left\{\hat{\mu}_0 - \frac{\hat{\sigma}_0}{\sqrt{10}} t_{9,0.05} < \mu_0 < \hat{\mu}_0 + \frac{\hat{\sigma}_0}{\sqrt{10}} t_{9,0.05}\right\} = 0.9$

90%-konfidensintervall for μ_0:

$\left[\hat{\mu}_0 - \frac{\hat{\sigma}_0}{\sqrt{10}} t_{9,0.05}, \hat{\mu}_0 + \frac{\hat{\sigma}_0}{\sqrt{10}} t_{9,0.05}\right]$

$\left[26 - \frac{4.3}{\sqrt{10}} \cdot 1.83, 26 + \frac{4.3}{\sqrt{10}} 1.83\right]$

$[23.5, 28.5]$

c) $K = K_0 + R$

$K_0 \rightsquigarrow n(k_0; \mu_0, \sigma_0)$

$R \rightsquigarrow n(r; \mu_R, \sigma_R)$

$Corr(K_0, R) = \rho_{0R}$

$K \rightsquigarrow n(k; \mu_K, \sigma_K)$

$$\mu_K = E(K) = E(K_0) + E(R) = \mu_0 + \mu_R$$

$$\sigma_K^2 = Var(K_0) + Var(R) + 2Cov(K_0, R)$$

$$= \sigma_0^2 + \sigma_R^2 + 2\sigma_0\sigma_R\rho_{0R}$$

$$\sigma_K = \left[\sigma_0^2 + \sigma_R^2 + 2\sigma_0\sigma_R\rho_{0R}\right]^{1/2}$$

d) Merk

$$\sigma_K^2 = \left[4^2 + 2^2 + 2\cdot 4\cdot 2\cdot \frac{5}{16}\right] = 25 = 5^2$$

$$K_{A1} = K_0^{A1} + R_{A1}, \quad \ldots \quad K_{A10} = K_0^{A10} + R_{A10} \quad \text{uif} \quad n(k_A; \overbrace{25+\mu_{RA}}^{\mu_{KA}}, \sigma_K)$$

$$K_{B1} = K_0^{B1} + R_{B1}, \quad \ldots \quad K_{B10} = K_0^{B10} + R_{B10} \quad \text{uif} \quad n(k_B; \overbrace{25+\mu_{RB}}^{\mu_{KB}}, \sigma_K)$$

$$\hat{\mu}_{KA} = \frac{1}{10}\sum_i K_{Ai} \Rightarrow n\left(\hat{\mu}_{KA}; \mu_{KA}, \frac{\sigma_K}{\sqrt{10}}\right)$$

$$\hat{\mu}_{KB} = \frac{1}{10}\sum_i K_{Bi} \Rightarrow n\left(\hat{\mu}_{KB}; \mu_{KB}, \frac{\sigma_K}{\sqrt{10}}\right)$$

$$\hat{\Delta}_K = \left[\hat{\mu}_{KA} - \hat{\mu}_{KB}\right] \Rightarrow n\left(\hat{\Delta}_K; \underbrace{\mu_{KA}-\mu_{KB}}_{\mu_{RA}-\mu_{RB}}, \frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}\right)$$

Hypotese:

$H_0: \mu_{RA} = \mu_{RB}$ mot $H_1: \mu_{RA} \neq \mu_{RB}$

$\mu_{RA} - \mu_{RB} = 0$ $\quad\quad$ $\mu_{RA} - \mu_{RB} = 0$

Testobservator:

$$\hat{\Delta}_K \Rightarrow n\left(\hat{\Delta}_K; \mu_{RA}-\mu_{RB}, \frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}\right)$$

Signifikansnivå: 0.1

Kritisk verdi: c

Forkastningskriterium

Forkast $H_0: \mu_{RA} - \mu_{RB} = 0$ hvis $|\hat{\Delta}_K| > c$

Forkastningsstrategi:

$$P(\text{Forkast } H_0 \mid H_0 \text{ sann}) < 0.1$$

$$P(|\hat{\Delta}_K| > c \mid \mu_{RA} - \mu_{RB} = 0) < 0.1$$

$$P(-c < \hat{\Delta}_K < c \mid \mu_{RA} - \mu_{RB} = 0) > 0.9$$

$$P\left(\frac{-c-(\mu_{RA}-\mu_{RB})}{\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}} < \frac{\hat{\Delta}_K-(\mu_{RA}-\mu_{RB})}{\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}} < \frac{c-(\mu_{RA}-\mu_{RB})}{\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}} \,\middle|\, \mu_{RA}-\mu_{RB}=0\right) > 0.9$$

$$P(-z_{0.05} < Z < z_{0.05}$$

herav

$$\frac{c}{\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}} = z_{0.05} \longrightarrow c = z_{0.05}\,\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}$$

Forkastningskriterium:

$$|\hat{\Delta}_K| = |\hat{\mu}_{KA} - \hat{\mu}_{KB}| > z_{0.05}\,\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}$$

$$|35 - 38| > 1.65\,\frac{\sqrt{2}\cdot 5}{\sqrt{10}}$$

$$3 > 3.69$$

Forkaster ikke H_0!

Utled styrke ved $\mu_{RA} - \mu_{RB} = 2$ [eller $= -2$]

$$\beta = P(\text{Forkast } H_0 \mid \mu_{RA} - \mu_{RB} = 2)$$

$$= P\left(|\hat{\Delta}_K| > z_{0.05}\,\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}} \;\middle|\; \Delta_K = 2\right)$$

$$= 1 - P\left(-z_{0.05}\,\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}} < \hat{\Delta}_K < z_{0.05}\,\frac{\sqrt{2}}{\sqrt{10}}\,\sigma_K \;\middle|\; \Delta_K = 2\right)$$

$$= 1 - P\left(-z_{0.05} - \frac{2}{\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}} < \frac{\hat{\Delta}_K - 2}{\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}} < z_{0.05} - \frac{2}{\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}} \;\middle|\; \Delta_K = 2\right)$$

$$= 1 - P\left(-z_{0.05} - \frac{2}{\frac{\sqrt{2}\,\sigma_K}{\sqrt{10}}} < Z < z_{0.05} - \frac{2}{\frac{\sqrt{2}\,\sigma_4}{\sqrt{10}}}\right)$$

$$= 1 - P(-1.65 - 0.89 < Z < 1.65 - 0.89)$$

$$= 1 - P(-2.54 < Z < 0.76)$$

$$= 1 - [.7764 - 0.0055] = \underline{\underline{0.2291}}$$

e) Dette forsøksopplegget er bedre fordi person-til-person variabilitet kan fjernes ved å betrakte $k_A - k_B$.

$$K_{A1} = K_0^1 + R_{A1}, \quad \dots \quad , K_{A10} = K_0^{10} + R_{A10} \quad \text{uif } n(k_A; \mu_{K_A}, \sigma_K)$$

$$K_{B1} = K_0^1 + R_{B1}, \quad \dots \quad , K_{B10} = K_0^{10} + R_{B10} \quad \text{uif } n(k_B; \mu_{K_B}, \sigma_K)$$

$$\Delta_1 = K_{A1} - K_{B1} = R_{A1} - R_{B1} \qquad \Delta_{10} = K_{A10} - K_{B10} = R_{A10} - R_{B10} \quad \text{uif } n(\Delta; \mu_{R_A} - \mu_{R_B}, \sqrt{2}\,\sigma_R)$$

$$\hat{\Delta} = \frac{1}{10}\sum_i \Delta_i \quad \Rightarrow \quad n\left(\hat{\Delta}; \mu_{R_A} - \mu_{R_B}, \frac{\sqrt{2}\,\sigma_R}{\sqrt{10}}\right)$$

Hypotese:

$H_0: \mu_{R_A} = \mu_{R_B}$ mot $H_1: \mu_{R_A} \neq \mu_{R_B}$

$\mu_{R_A} - \mu_{R_B} = 0$ $\qquad$ $\mu_{R_A} - \mu_{R_B} \neq 0$

Testobservator:

$$\hat{\Delta} \quad \Rightarrow \quad n\left(\hat{\Delta}; \mu_{R_A} - \mu_{R_B}, \frac{\sqrt{2}\,\sigma_R}{\sqrt{10}}\right)$$

Som foran....

Forkastningskriterium

$$|\hat{\Delta}| > z_{0.05} \frac{\sqrt{2}\,\sigma_R}{\sqrt{10}}$$

$$|-3| > 1.65 \frac{\sqrt{2}\cdot 2}{\sqrt{10}} = 1.65 \cdot 0.89$$

$$3 > 1.48$$

Forkaster H_0!

Merk at i d) forkaster vi ikke H_0 – dette kan synes rimelig fordi i e) har vi fjernet person-til-person variabiliteten. og har derfor en bedre test.

Utled styrke ved $\mu_{R_A} - \mu_{R_B} = 2$ [eller $= -2$]

$$\beta = P(\text{Forkast } H_0 \mid \mu_{R_A} - \mu_{R_B} = 2)$$
$$= P\left(|\hat{\Delta}| \geq z_{0.05} \frac{\sqrt{2}}{\sqrt{10}} \sigma_R \mid \mu_{R_A} - \mu_{R_B} = 2\right)$$
$$= \text{som over}$$
$$= 1 - P\left(-z_{0.05} - \frac{2}{\frac{\sqrt{2}\sigma_R}{\sqrt{10}}} < Z < z_{0.05} - \frac{2}{\frac{\sqrt{2}\sigma_R}{\sqrt{10}}}\right)$$
$$= 1 - P(-1.65 - 2.25 < Z < 1.65 - 2.25)$$
$$= 1 - P(-3.90 < Z < -0.6)$$
$$= 1 - [0.2743 - 0.0] = \underline{\underline{0.7257}}$$

Vi observerer at styrken for denne testen er mye bedre enn i punkt d) grunnet muligheten for å fjerne person-til-person variabilitet.

For å få samme styrke i d) som i e) må en ha n_d test-personer i d); samt at

$$z_{0.05} - \frac{2}{\frac{\sqrt{2}\sigma_K}{\sqrt{n_d}}} = z_{0.05} - \frac{2}{\frac{\sqrt{2}\sigma_R}{\sqrt{10}}}$$

$$\sqrt{n_d} = \frac{\sigma_K}{\sigma_R}\sqrt{10}$$

$$n_d = \frac{\sigma_K^2}{\sigma_R^2} 10 = \underline{62.5}$$

For å få like god styrke måtte en ha mer en 6.25 ganger så mange personer — altså 63 for hvert middel.