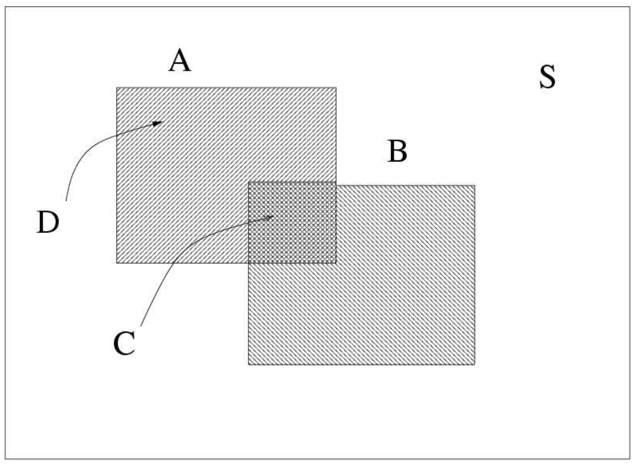
Oppgave 1

$C = A \cap B, D = A \cap B'$



Figur 1: Venndiagram

C og D er disjunkte siden

$$C \cap D \iff (A \cap B) \cap (A \cap B') \iff A \cap B \cap B' = \emptyset$$

siden $B \cap B' = \emptyset$. Dette resultatet kan vi også se ut fra venndiagrammet da C og D ikke overlapper.
C og D er ikke uavhengige, siden

$$P(C \cap D) = P(\emptyset) = 0 \neq P(A)P(B)$$

Den siste biten gjelder fordi $P(A \cap B) > 0$.

Oppgave 2

$$P(X \leq 2) = P\left(\frac{X-1}{\sqrt{1}} \leq \frac{2-1}{\sqrt{1}}\right) = \Phi(1) = 0,8413$$

$$P(X \leq 2 \cap Y \leq 1) \overset{\text{uavh.}}{=} P(X \leq 2) \cdot P(Y \leq 1)$$
$$= 0,8413 \cdot \Phi\left(\frac{1-1}{\sqrt{4}}\right) = 0,413 \cdot 0.5 = 0,4207$$

$X+2Y$ er en lineærkombinasjon av uavhengige normalfordelte variable og derfor normalfordelt med forventning og varians

$$E(X+2Y) = E(X) + 2E(Y) = 1 + 2 \cdot 1 = 3$$
$$\mathrm{Var}(X+2Y) = \mathrm{Var}(X) + 2^2\mathrm{Var}(Y) = 1 + 4 \cdot 4 = 17$$

Dermed blir regninga som over.

$$P(X+2Y > 2) = 1 - P(X+2Y \leq 2) = 1 - \Phi\left(\frac{2-3}{\sqrt{17}}\right)$$
$$= 1 - \Phi(-0,24) = 1 - 0,4052 = 0,5948$$

Oppgave 3

a)

X er hypergeometrisk fordelt fordi:

- man trekker n personer uten tilbakelegging
- alle N personer har samme sannsynlighet for å bli trukket ut
- for hver person som blir trukket ut har man to muligheter: stemmer på P (suksess) eller stemmer ikke på P (fiasko)
- X er antall suksesser

Dermed har vi en nøyaktig situasjon som definerer den hypergeometriske fordelingen.

Siden antall stemmeberettigede, N, vil være mye større enn antall personer som blir spurt, n, vil det i praksis ha svært liten betydning om de n personene trekkes med eller uten tilbakelegging. Vi kan dermed med god tilnærming regne som om de n var valgt med tilbakelegging. Da har vi en situasjon som karakteriserer en binomisk fordeling.

b)

Vi vet følgende:

$$\hat{p} = \frac{X}{n},$$
$$X \sim \mathrm{b}(x; n, p),$$
$$E(X) = np,$$
$$\mathrm{Var}(X) = np(1-p).$$

Dermed er forventning og varians for $\hat{p}$ gitt ved

$$E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{1}{n}\cdot np = p$$
$$\mathrm{Var}(\hat{p}) = \mathrm{Var}\left(\frac{X}{n}\right) = \left(\frac{1}{n}\right)^2 \cdot \mathrm{Var}(X) = \frac{1}{n^2}\cdot np(1-p) = \frac{1}{n}p(1-p)$$

Ønsker n slik at når $p = 0,079$, så er

$$\mathrm{SD}(\hat{p}) = \sqrt{\mathrm{Var}(\hat{p})} \leq 0,01.$$

Da har vi at

$$\sqrt{\frac{1}{n}p(1-p)} \leq 0,01,$$
$$\frac{p(1-p)}{0,01^2} \leq n,$$
$$n \geq \frac{p(1-p)}{0,01^2} = \frac{0,079(1-0,079)}{0,01^2},$$
$$n \geq 727,59$$

Det vil si at man må minst spørre 728 personer.

c)

Skal teste hypotesen

$$H_0 : p = p_0 \quad \text{mot} \quad H_1 : p < p_0$$

Vi bruker testobservatoren

$$Z = \frac{X - np_0}{\sqrt{np_0(1-p_0)}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0+(1-p_0)}{n}}}.$$

Vi vet at Z er standard normalfordelt under H_0. Forkaster H_0 dersom testobservatoren Z er mindre enn k, der k bestemmes ut fra kravet

$$P(Z < k | p = p_0) = \alpha = 0,05$$

Velger k til å være 5% kvantilen i standard normalfordelingen.

$$k = -z_{0,05} = -1,645$$

D.v.s. vi forkaster H_0 dersom $Z < -1,645$.

Insatt observasjonene

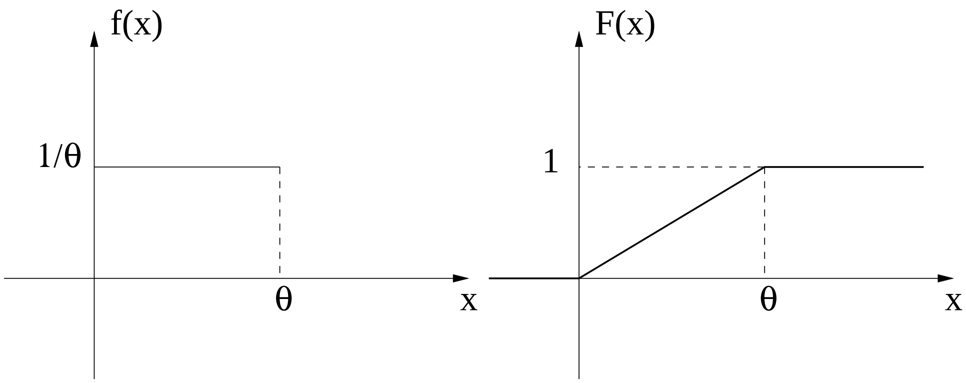
$$\left.\begin{array}{rc} n = & 1000 \\ x = & 52 \\ p_0 = & 0,079 \\ \hat{p} = & \frac{52}{1000} = 0,052 \end{array}\right\} Z = \frac{0,052 - 0,079}{\sqrt{\frac{0,079(1-0,079)}{1000}}} = -3,165$$

D.v.s vi forkaster H_0.

Oppgave 4

a)

$$F(x) = \int_{-\infty}^{x} f(x)\,dx \;=\; \begin{cases} 0 & \text{hvis } x < 0, \\ \frac{x}{\theta} & \text{hvis } 0 \le x \le \theta, \\ 1 & \text{hvis } x > \theta. \end{cases}$$

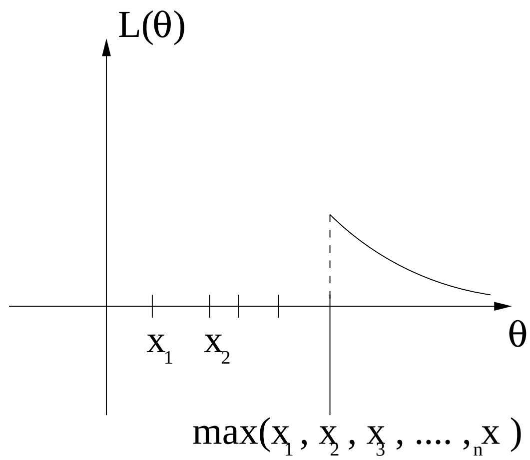


$$P(X \le 0,4 | \theta = 2) = F(0,4) = \frac{0,4}{2,0} = 0,2$$

b)

$$L(\theta) = \prod_{i=1}^{n} f(x_i) \;=\; \begin{cases} \frac{1}{\theta^n} & \text{hvis } 0 \le x_i \le \theta \;\forall\, i = 1, 2, \ldots, n, \\ 0 & \text{ellers} \end{cases}$$

Vi ser fra skissen at $L(\theta)$ maksimeres for den minste verdien av θ der $\theta \ge \max(x_1, x_2, \ldots x_n)$. D.v.s $\widehat{\theta} = \max(X_1, X_2, \ldots X_n)$.



Figur 2: Skisse av $L(\theta)$:

c)

$$
\begin{aligned}
G(y) &= P(Y \leq y) = P(\max(x_1, x_2, \ldots, x_n) \leq y) \\
&= P(x_1 \leq y \cap x_2 \leq y \cap \ldots \cap x_n \leq y) \\
&= P(x_1 \leq y) \cdot P(x_2 \leq y) \cdot \ldots \cdot P(x_n \leq y) \\
&= F(y) \cdot F(y) \cdot \ldots \cdot F(y) = (F(y))^n \\
&= \begin{cases} 0 & \text{hvis } y < 0, \\ \left(\frac{y}{\theta}\right)^n & \text{hvis } 0 \leq y \leq \theta, \\ 1 & \text{hvis } y > \theta. \end{cases}
\end{aligned}
$$

Vi finner sannsynlighetstettheten til Y ved å derivere den kumulative fordelingsfunksjonen:

$$
\begin{aligned}
g(y) = G'(y) &= \begin{cases} 0 & \text{hvis } y < 0, \\ n\left(\frac{y}{\theta}\right)^{n-1} \cdot \theta^{-1} & \text{hvis } 0 \leq y \leq \theta, \\ 0 & \text{hvis } y > \theta. \end{cases} \\
&= \begin{cases} \frac{y^{n-1}}{n \cdot \theta^n} & \text{hvis } 0 \leq y \leq \theta, \\ 0 & \text{ellers.} \end{cases}
\end{aligned}
$$

d)

Skal vise at estimatoren vår ikke er forventningsrett.

$$
\begin{aligned}
E(\widehat{\theta}) &= E(Y) = \int_{-\infty}^{\infty} y \cdot g(y)\ dy \\
&= \int_0^{\theta} y \cdot n\frac{y^{n-1}}{\theta^n}\ dy = \int_0^{\theta} n\frac{y^n}{\theta^n}\ dy \\
&= \left[\frac{n}{n+1} \cdot \frac{y^{n+1}}{\theta^n}\right]_0^{\theta} = \frac{n}{n+1} \cdot \frac{\theta^{n+1}}{\theta^n} \\
&= \frac{n}{n+1} \cdot \theta \neq \theta
\end{aligned}
$$

Dvs. $\widehat{\theta}$ er ikke forventingsrett. Skal bestemme k slik at $\widetilde{\theta} = kY$ er forventningsrett.

$$
\begin{gathered}
E(\widetilde{\theta}) = \theta \\
\Downarrow \\
E(kY) = k \cdot E(Y) = k \cdot \frac{n}{n+1} \cdot \theta = \theta \\
\Downarrow \\
k = \frac{n+1}{n}
\end{gathered}
$$

e)

Finner først fordelinga til $Z = \widetilde{\theta}/(k\theta)$ ved hjelp av transformasjonsformelen

$$f_Z(z) = f_Y(y(z)) \cdot |y'(z)|$$

der

$$z(y) = \frac{\widetilde{\theta}}{k\theta} = \frac{y}{\theta} \qquad \Leftrightarrow \qquad \begin{cases} y(z) = z\theta \\ y'(z) = \theta \end{cases}$$

Fordelinga til Z blir da

$$f_Z(z) = \begin{cases} \frac{n(z\theta)^{n-1}}{\theta^n} \cdot \theta = nz^{n-1} & \text{for } 0 \leq z \leq 1 \\ 0 & \text{ellers} \end{cases}$$

Da har vi at

$$P(L_z \leq Z \leq U_z) = 0,95$$

der L_z og U_z henholdsvis er 2,5% og 97,5% kvantilene til Z-fordelinga. Dermed gjenstår det bare å bestemme kvantilene.

$$\int_0^{L_z} f_Z(z)\,dz = [z^n]_0^{L_z} = L_z^n = 0,025$$
$$\Downarrow$$
$$L_z = \sqrt[n]{0,025}$$

$$\int_0^{U_z} f_Z(z)\,dz = [z^n]_0^{U_z} = U_z^n = 1 - 0,025$$
$$\Downarrow$$
$$U_z = \sqrt[n]{0.975}$$

Dermed kan vi sette

$$P\left(\sqrt[n]{0,025} \leq \frac{\widetilde{\theta}}{k\theta} \leq \sqrt[n]{0.975}\right) = 0,95$$
$$\Downarrow$$
$$P\left(\frac{\widetilde{\theta}}{k\sqrt[n]{0.975}} \leq \theta \leq \frac{\widetilde{\theta}}{k\sqrt[n]{0.025}}\right) = 0,95$$

Innsatt for k blir da konfidensintervallet

$$\left[\frac{n\widetilde{\theta}}{(n+1)\sqrt[n]{0.975}}, \frac{n\widetilde{\theta}}{(n+1)\sqrt[n]{0.025}}\right].$$

Med $n = 10$ og $y = 0,46$ får vi

$$\widetilde{\theta} = \frac{10+1}{10} \cdot 0,46 = 0,506,$$

og konfidensintervallet blir

$$\left[\frac{10 \cdot 0,506}{11 \cdot \sqrt[10]{0.975}}\,,\ \frac{10 \cdot 0,506}{11 \cdot \sqrt[10]{0.025}}\right]$$
$$= [0,461\ ,\ 0,665].$$