Des - 97

1a La $X_i = [\text{rødt lys dag } i] \sim \text{bernoulli}(p)$. Da er $E\overline{X} = p$, dvs $\overline{X}$ er en forv. rett estimator for p. Et estimat for p er da $\bar{x} = \frac{8}{20} = \underline{\underline{40\%}}$

b.) Fra $ES^2 = \sigma^2 = p\cdot(1-p)$ og $\text{Var}\,\overline{X} = \frac{1}{n}\sigma^2$ følger det at $\frac{1}{n}S^2 = \frac{1}{n-1}(\overline{X^2} - \overline{X}^2) = \frac{1}{n-1}(\overline{X} - \overline{X}^2)$ er en forv. rett estimator for $\text{Var}\,\overline{X}$. Et estimat for standardavviket til $\overline{X}$ er dermed

$$S/\sqrt{n} = \sqrt{\frac{1}{19}\cdot\frac{8}{20}\cdot\frac{12}{20}} \approx 0{,}11239 \approx \underline{\underline{11{,}2\%}}$$

c.) Det er 68% sannsynlighet for at intervallet $[\overline{X} - \text{SDEV}\,\overline{X}, \overline{X} + \text{SDEV}\,\overline{X}]$ inneholder p dersom* $\overline{X} \sim N(p, (\text{SDEV}\,\overline{X})^2)$ fordi $P(|\overline{X} - p| > \text{SDEV}\,\overline{X})$ $= 2\cdot P\left(\frac{\overline{X}-p}{\text{SDEV}\,\overline{X}} > 1\right) = 2\cdot(1 - \Phi(1)) = 32\%$.

*Begrunnes ved sentralgrenseteoremet. Et tilnærmet 68% intervallestimat er gitt ved grensene $\bar{x} \pm s/\sqrt{n}$, dvs $\underline{\underline{[28{,}8\%; 51{,}2\%]}}$

d.) Metoden på s.279 i Larsen & Marx gir at et eksakt intervall $[p_1, p_2]$ er gitt ved $P(Y \geq 8; p_1) = P(Y \leq 8; p_2) = 16\%$ hvor $Y \sim \text{binomisk}(20, p_i)$. Lineær interpolasjon gir p_i:

$P(Y \leq 7; p_1) = 84\%$, $p_1 = 0{,}25 + \frac{0{,}30 - 0{,}25}{0{,}772 - 0{,}898}(0{,}84 - 0{,}898) \approx 0{,}2730$

$P(20 - Y \leq 11; p_2) = 84\%$, $1 - p_2 = 0{,}40 + \frac{0{,}50 - 0{,}40}{0{,}748 - 0{,}943}(0{,}84 - 0{,}943)$ $\approx 0{,}4528$, $p_2 \approx 0{,}5472$): $\underline{\underline{[27{,}3\%; 54{,}7\%]}}$

2a - Avreisetidspunkt er tilfeldig.

- Kjøretiden varierer pga trafikken — Stor variasjon på sekundskala!
- Det er nærliggende å anta at Øyvind kommer til lyskrysset ved et tidspunkt som er normalfordelt pga en sum av tilfeldigheter. Dette kan benyttes til å sette opp en modell med 3 parametre. Denne modellen vil inneholde den uniforme modellen som et grensetilfelle tilsvarende stor varians i normalfordelingen.
- Den uniforme modellen foretrekkes her fordi den er enkel, og naturlig. Mot: For enkel, og derfor lite fleksibel.

b.) $P(Y_{maks} \leq y) = (y/a)^n$

$$E\,Y_{maks} = \int_0^a y \cdot n\,(y/a)^{n-1} \cdot 1/a \, dy = \frac{n}{n+1} \cdot a$$

$$W = \frac{n+1}{n} Y_{maks}, \quad w = \frac{9}{8} \cdot 29\,s \approx 32{,}625\,s \approx \underline{\underline{33\,s}}$$

c.) $E\,W = \frac{n+1}{n} \cdot E\,Y_{maks} \underline{\underline{= a}}$): $\underline{\underline{W \text{ er forventningsrett for } a.}}$

d.) $$f(\vec{y}\,; a) = \prod_{i=1}^{n} \frac{1}{a}[0 < y_i < a] = \underbrace{\left(\prod_{i=1}^{n}[0 < y_i]\right)}_{\text{uavh. av } a} \cdot \underbrace{\frac{1}{a^n}\left[W < \frac{n+1}{n} a\right]}_{\text{avh. av } \vec{y} \text{ via } W.}$$

gir at W er tilstrekkelig.

d.) $f(\vec{y}; a) = \prod_{i=1}^{n} \frac{1}{a}[0 < y_i < a] = \underbrace{\left(\prod_{i=1}^{n}[0 < y_i]\right)}_{\text{uavh. av } a} \cdot \underbrace{\frac{1}{a^n}\cdot[y_{maks} < a]}_{\text{avhenger av } a \text{ gjennom } y_{maks}}$

Dermed er y_{maks} tilstrekkelig.
Omskrivningen $[y_{maks} < a] = [w < \frac{n+1}{n} a]$ viser
at også w er tilstrekkelig.
(Strengt tatt: Y_{maks} og W er estimatorer
og det er disse, ikke estimatene y_{maks}, w,
som er tilstrekkelige.)

3a

$P(Y_{maks} \geq 29s) = 1 - P(Y_{maks} \leq 29s) = 1 - \left(\frac{29s}{30s}\right)^8$
$\approx 23{,}75\% > 5\%$

$):$ Alternativet er ikke bevist med signifikansnivå 5%

(Alternativt kan det kritiske området i b.) benyttes!)

b.) $5\% = 1 - \left(\frac{y}{30s}\right)^8$ gir $y = 30s \cdot \sqrt[8]{95\%} \approx 29{,}808$

$[29{,}81s; \infty)$ $[30s \cdot 95\%; \infty) \approx [28{,}50s; \infty)$

H_0 beholdes også for denne testen

c.) $P(Y_{maks} \geq 29{,}81s; a) = 1 - \left(\frac{29{,}81}{a}\right)^8 = \beta(a)$

$P(Y_1 \geq 28{,}50s; a) = 1 - \left(\frac{28{,}50s}{a}\right) = \gamma(a)$

a	30	33	40	50	75	100
β	4,96%	55,66%	90,48%	98,40%	99,94%	99,99%
γ	5,00%	13,64%	28,75%	43%	62%	71,5%

c.)

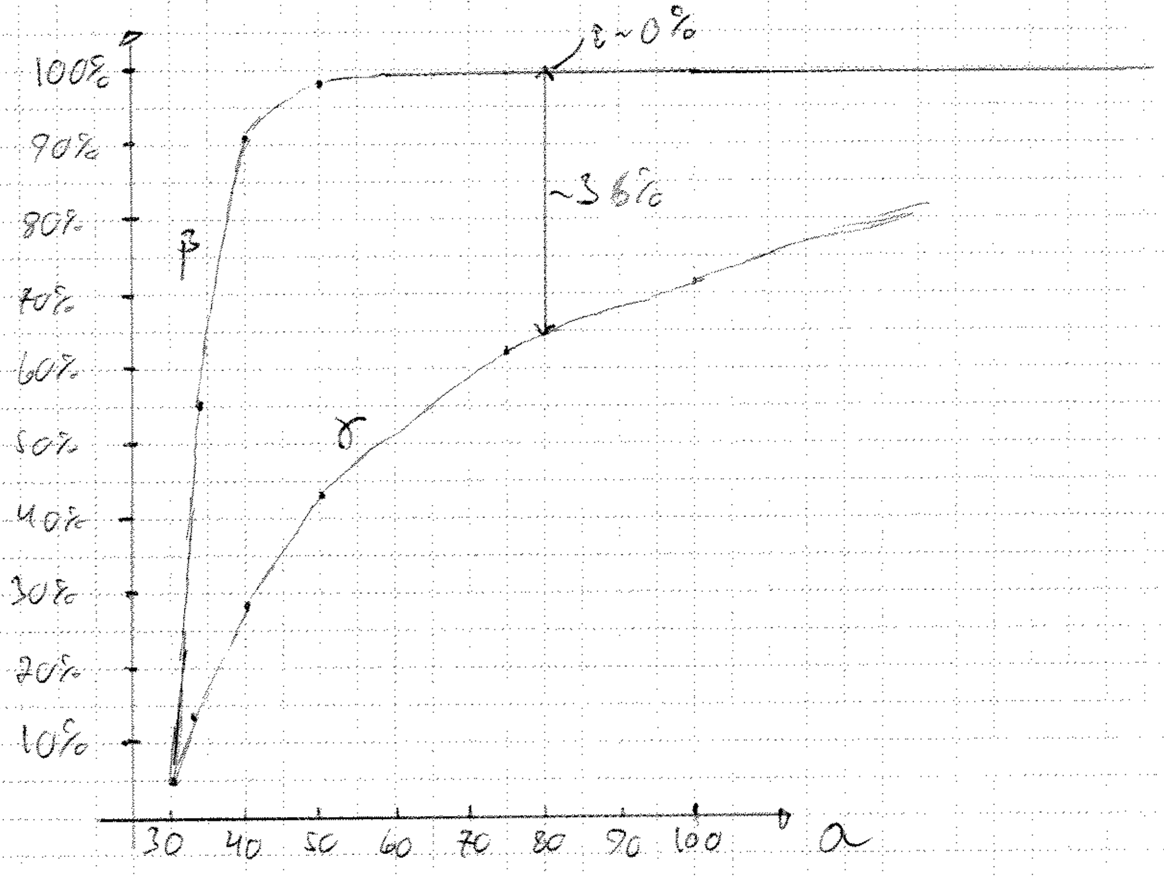


Fig. 1 Testen med testobservator Y_{maks} er uniformt bedre enn testen med testobservator Y_1.

d.) Sannsynligheten for type-2 feil er hhv (omtrent) 0% og 36% fra fig. 1.

e.) Statistisk tolkning: I det lange løp vil prosedyren gi type-2 feil i (høyst) 1% (34%) av tilfellene (relativ frekvens → 1% (34%)).
Generelt: Vår vurdering er at prosedyren gir type-2 feil med 1% (34%) sannsynlighet. Dette er en konsekvens av vurderingene av sannsynligheter som ligger til grunn for modellen.

4a $P(T \le t) = P([(T \le t) \cap R] \cup [(T \le t) \cap R^c])$

$\overset{\text{Kolmogorov}}{\underset{\text{Add. regel}}{=}} P((T \le t) \cap R) + P((T \le t) \cap R^c)$

$\overset{\text{Def. av bet.}}{\underset{\text{sannsynl.}}{=}} P(T \le t \mid R) \cdot P(R) + P(T \le t \mid R^c) \cdot P(R^c)$

$= P(T \le t \mid R) \cdot p + P(T \le t \mid R^c) \cdot (1-p)$

b.) $F(T \le t) = P(T \le t \mid R) \cdot p + P(T \le t \mid R^c) \cdot (1-p)$

$= \begin{cases} 0 & t < 0 \quad (\text{da er } P(T \le t) = 0!) \\ p \cdot t/a + (1-p) & 0 \le t \le a \quad (\text{unif. ved rødt og } T=0 \text{ ved grønt}) \\ 1 & t > a \quad (P(T \le a) = 1) \end{cases}$

T er ikke kontinuerlig fordelt fordi F ikke er deriverbar i 0.
T er ikke diskret fordelt fordi F ikke er en trappefunksjon.

5a Poissonparameteren tilsvarende 15 minutter estimeres av $\hat{\lambda} = \frac{16}{4\text{timer}} : 15\text{minutter} = 1$. Dette gir estimatet

$$\frac{\hat{\lambda}^0}{0!} e^{-\hat{\lambda}} = \underline{\underline{e^{-1}}} \approx 0{,}3679 \underline{\underline{\approx 37\%}}$$

b.) Del opp 15 min. intervallet i n like store intervall. La X_i = [Det kom en buss i intervall i], $i = 1, \ldots, n$. Anta at X_i'ene er et tilfeldig utvalg fra bernoulli(p)-fordelingen. Det følger da at $\Sigma X_i \sim$ binomisk(n,p). Dersom n er stor nok er det rimelig å anta at det maksimalt kommer én buss i hvert intervall, og dette gir at antall busser $= \Sigma X_i \sim$ binomisk(n,p). Vi har $E\Sigma X_i = np = \lambda$ ved $p = \lambda/n$. Ved $n \to \infty$ følger $\binom{n}{x} \cdot (\lambda/n)^x \cdot (1-\lambda/n)^{n-x} \to \frac{\lambda^x}{x!} \cdot e^{-\lambda}$ (Poisson-approksimasjonen) som gir $\Sigma X_i \xrightarrow{L} Y \sim$ Poisson(λ).

c.) Poissonantagelsen fører til at ventetiden er eksponensialfordelt med forventning (estimert lik 15 min) uavhengig av når Magnar kommer til bussholdeplassen, og spesielt uavhengig av om bussen har gått en gang i løpet av de 10 minuttene før Magnar kom til bussholdeplassen. Dette er i strid med Magnars erfaring i situasjonen gitt i oppgaven.

d.) Tettheten til avgangstiden T er $f(t) = 2 - 2t$, $0 \le t \le 1$, hvor $t = 1$ svarer til 15 minutter. Forv. ventetid dersom Magnar kommer ved tidspunkt $s = 0$ er $ET = \int_0^1 t \cdot (2-2t) = \frac{1}{3}$ (dvs 5 min.). La V være Magnars ventetid. Magnar når enten bussen i det intervallet han ankommer, eller så når han bussen en gang i neste intervall: $EV = E(V \mid T \ge s) \cdot P(T \ge s)$
$+ E(V \mid T < s) \cdot P(T < s)$.

5d (fortsatt) Fordelingen gitt $T \geq s$: $g(t) = f(t)/P(T \geq s)$, $s \leq t \leq 1$. $E(V|T \geq s) \cdot P(T \geq s) = \int_s^1 (2-2t) \cdot t\,dt = \frac{1}{3} - s^2 + \frac{2}{3}s^3$.

$P(T < s) = \int_0^s (2-2t)\,dt = 2s - s^2$. $E(V|T<s) = (1-s) + \frac{1}{3}$.

Tilsammen $EV = \frac{1}{3} + \frac{8}{3}s - \frac{13}{3}s^2 + \frac{5}{3}s^3$, $0 \leq s \leq 1$.

Tilfeldig tidspunkt tilsvarer uniform S, som gir

$E(E[V|S]) = \frac{1}{3} + \frac{4}{3} - \frac{13}{9} + \frac{5}{12} = \frac{23}{36} \approx 0{,}64$, dvs $\underline{\underline{9 \text{ min og } 35 \text{ sek.}}}$ ventetid ved tilfeldig valgt tidspunkt.

Forventet ventetid når Magnar ankommer ved tidspunkt u: $EV = 15\text{min} \cdot \left(\frac{1}{3} + \frac{8}{3} \cdot \frac{u}{15\text{min}} - \frac{13}{3} \cdot \left(\frac{u}{15\text{min}}\right)^2 + \frac{5}{3} \cdot \left(\frac{u}{15\text{min}}\right)^3\right)$

$$= \underline{\underline{\left(5 + 40 \cdot \frac{u}{15\text{min}} - 65 \cdot \left(\frac{u}{15\text{min}}\right)^2 + 25 \cdot \left(\frac{u}{15\text{min}}\right)^3\right)\text{min}}}$$

Tidspunkt for ekstremverdi til EV:

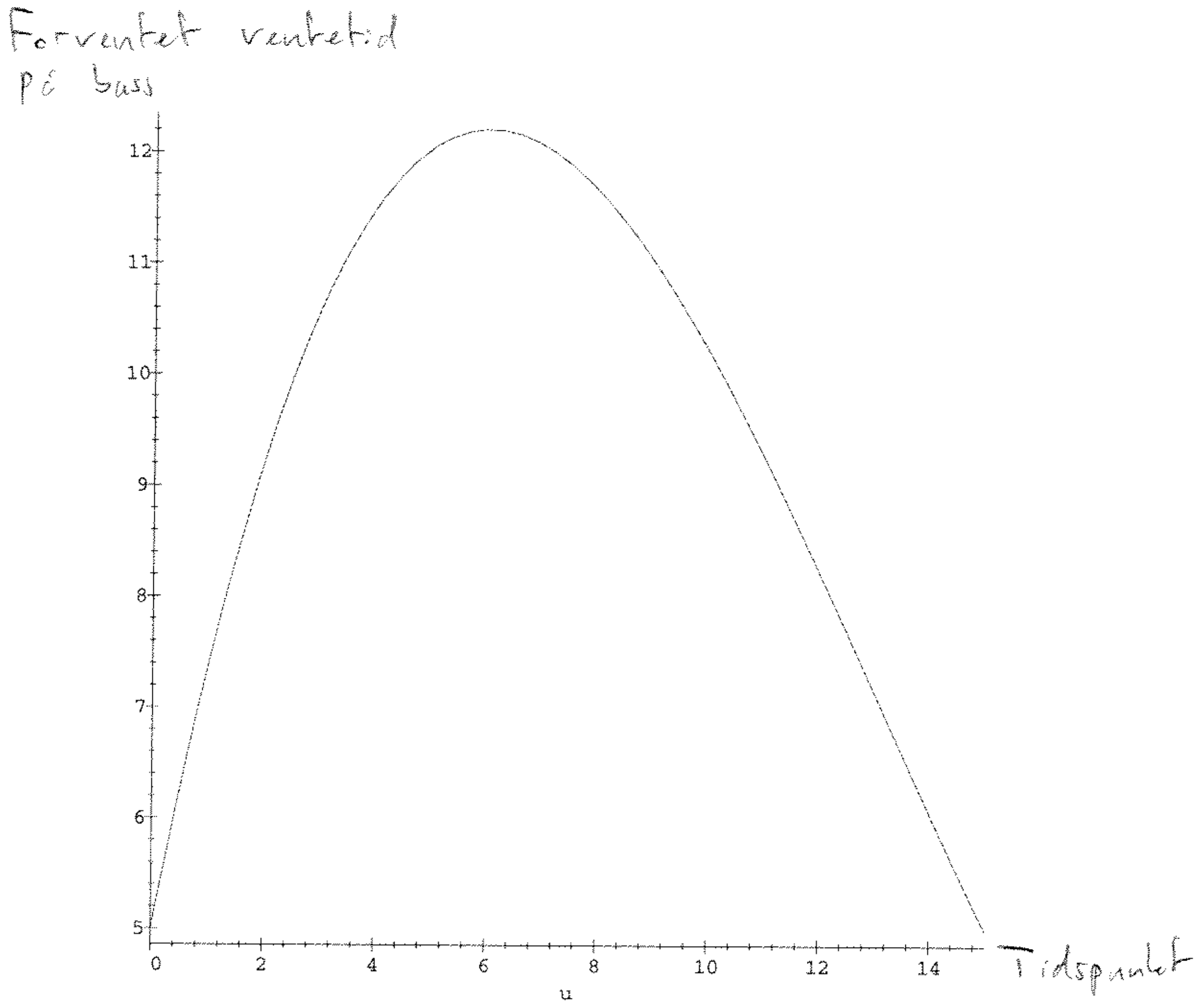
$0 = \frac{8}{3} - \frac{26}{3}s + \frac{15}{3}s^2$, $s = \frac{\cancel{26}^{13} \pm \sqrt{\cancel{26}^{13\,2} - \cancel{4} \cdot 15 \cdot 8}}{\cancel{2} \cdot 15} = \frac{13 \pm 7}{15} = \begin{cases} 2/5 \\ 4/3 \leftarrow \text{ikke aktuell} \end{cases}$

$EV_{s=2/5} = \frac{1}{3} + \frac{8}{3} \cdot \frac{2}{5} - \frac{13}{3} \cdot \left(\frac{2}{5}\right)^2 + \frac{5}{3} \cdot \left(\frac{2}{5}\right)^3 = \frac{61}{75}$

): Dersom du kommer 6 min etter tidspunktet hvor bussen skulle ha gått, så blir den forventede ventetid maksimal, og lik $\underline{12 \text{ min } 12 \text{ sek.}}$

<u>Det optimale tidspunktet er gitt ved $s=0$</u>, dvs Magnar konkluderer at han bør møte opp til det tidspunktet hvor bussen skal gå. Da er ventetiden minimal, og lik <u>5 min.</u>

```
> ?plot
> plot(5 + 40*(u/15) - 65*(u/15)^2 + 25*(u/15)^3, u=0..15);
```




```
>
```