Oppg. 1

a) Har en Poisson-fordeling og definerer da

$$L(\mathbf{x}, \lambda_R) = f(x_1, \lambda_R) \cdots f(x_n, \lambda_R) \tag{1}$$

med

$$f(x_i, \lambda_R) = \frac{\lambda_R^{x_i}}{x_i!} e^{-\lambda_R} \tag{2}$$

og finner så

$$\frac{\partial \ln L}{\partial \lambda_R} = \frac{\partial}{\partial \lambda_R} \left\{ \ln \left[\frac{\lambda_R^{\sum_{i=1}^n x_i}}{\prod_{j=1}^n x_j!} e^{-n\lambda_R} \right] \right\} = \frac{\partial}{\partial \lambda_R} \left\{ \ln \lambda_R \sum_{i=1}^n x_i - \ln \prod_{j=1}^n x_j! - n\lambda_R \right\} = \frac{1}{\lambda_R} \sum_{i=1}^n x_i - n \tag{3}$$

setter vi så likning (3) lik 0 finner vi

$$\widehat{\lambda}_R = \frac{1}{n} \sum_{i=1}^n x_i \tag{4}$$

ser også at $\widehat{\lambda}_R = \overline{X}$ og har da

$$\mathrm{E}\left[\widehat{\lambda}_R\right] = \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^n x_i\right] = \frac{1}{n}\mathrm{E}\left[\sum_{i=1}^n x_i\right] = \frac{1}{n} \cdot n\lambda_R = \lambda_R \tag{5}$$

og ser så på

$$\mathrm{Var}\left[\widehat{\lambda}_R\right] = \mathrm{Var}\left[\frac{1}{n}\sum_{i=1}^n x_i\right] = \frac{1}{n^2}\mathrm{Var}\left[\sum_{i=1}^n x_i\right] = \frac{1}{n^2} \cdot n\lambda_R = \frac{\lambda_R}{n} \tag{6}$$

ved sentralgrenseteoremet er dette normalfordelt.

b) Setter da opp hypotesene

$$H_0 : \lambda_R = 10, \quad H_1 : \lambda_R < 10 \tag{7}$$

og kan sette opp ved sentralgrenseteoremet

$$Z = \frac{\widehat{\lambda}_R - \lambda_R}{\sqrt{\frac{\lambda_R}{n}}} \approx -1.34164 \tag{8}$$

for H_0 og kan så finne p-verdien

$$p = \mathrm{P}\left[Z \leq -1.34164\right] = 0.0901 \tag{9}$$

og har dermed med $\alpha = 0.05$ at vi ikke forkaster hypotesen.

c) Antar $\lambda_R = 9$ og fortsetter å operere med $\alpha = 0.05$ og vil da finne

$$\mathrm{P}\left[\text{forkast } H_0 | \lambda_R = 9\right] = 0.9 \Rightarrow \mathrm{P}\left[p < 0.05 | \lambda_R = 9\right] = 0.9 \tag{10}$$

som da gir $Z \leq -2.58$ og siden vi har

$$\widehat{\lambda}_R = -2.58\sqrt{\frac{10}{n}} + 10 \tag{11}$$

Oppg. 2

a) Ser på $Y_i = ax_i(1-x_i) + \epsilon_i$ for $i = 1, 2, \ldots, n$ og har

$$\mathrm{E}\left[Y_i|x_i\right] = ax_i(1-x_i), \quad \mathrm{Var}\left[Y_i|x_i\right] = \sigma_0^2 = 0.025^2 \tag{12}$$

ser på sannsynlighetsmaksimeringsestimatoren for a og har $Y_i \sim N(ax_i(1-x_i), \sigma^2)$ og dermed

$$L(a) = f(\mathbf{y}; a) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left\{-\frac{\left[y_i - ax_i(1-x_i)\right]^2}{2\sigma_0^2}\right\} \tag{13}$$

og har dermed

$$\ln L(a) = \sum_{i=1}^{n} \left\{-\ln\sqrt{2\pi} - \ln\sigma_0 - \frac{\left[y_i - ax_i(1-x_i)\right]^2}{2\sigma_0}\right\} \tag{14}$$

og videre

$$\begin{aligned} \partial_a \ln L(a) = \frac{1}{\sigma_0} \sum_{i=1}^{n} \left[y_i - \widehat{a}x_i(1-x_i)\right]\left[x_i(1-x_i)\right] = 0 \\ \sum_{i=1}^{n} \left[x_i(1-x_i)\right] y_i = \widehat{a} \sum_{i=1}^{n} \left[x_i(1-x_i)\right]^2 \\ \widehat{a} = \frac{\sum_{i=1}^{n} x_i(1-x_i)y_i}{\sum_{i=1}^{n} \left[x_i(1-x_i)\right]^2} \end{aligned} \tag{15}$$

b) Skriver om

$$\widehat{a} = \sum_{i=1}^{n} \frac{x_i(1-x_i)y_i}{\sum_{j=1}^{n}\left[x_j(1-x_j)\right]^2} = \sum_{i=1}^{n} a_i y_i, \quad a_i = \frac{x_i(1-x_i)}{\sum_{j=1}^{n}\left[x_j(1-x_j)\right]^2} \tag{16}$$

der a_i er en konstant og kan videre se

$$\mathrm{E}\left[\widehat{a}\right] = \mathrm{E}\left[\sum_{i=1}^{n} a_i y_i\right] = \sum_{i=1}^{n} \mathrm{E}\left[y_i a_i\right] = \sum_{i=1}^{n} ax_i(1-x_i)a_i = a\frac{\sum_{i=1}^{n}\left[x_i(1-x_i)\right]^2}{\sum_{j=1}^{n}\left[x_j(1-x_j)\right]^2} = a \tag{17}$$

og ser videre på variansen

$$\mathrm{Var}\left[\widehat{a}\right] = \mathrm{Var}\left[\sum_{i=1}^{n} a_i y_i\right] = \sum_{i=1}^{n} \mathrm{Var}\left[a_i y_i\right] = \sum_{i=1}^{n} a_i^2\sigma_0^2 = \frac{\sigma_0^2\sum_{i=1}^{n}\left[x_i(1-x_i)\right]^2}{\left\{\sum_{j=1}^{n}\left[x_j(1-x_j)\right]^2\right\}^2} = \frac{\sigma_0^2}{\sum_{j=1}^{n}\left[x_j(1-x_j)\right]^2} \tag{18}$$

og kan se at $\widehat{a}$ er normalfordelt da det er en lineærkombinasjon av uavhengige normalfordelinger.

c) Setter opp hypotesetesten

$$H_0 : a = 0, \quad H_1 : a \neq 0 \tag{19}$$

og kan fra informasjonen sette opp

$$Z = \frac{\widehat{a} - a}{\sqrt{\mathrm{Var}\left[a\right]}} \Rightarrow Z_0 = \frac{\widehat{a}}{\sqrt{\mathrm{Var}\left[a\right]}} \tag{20}$$

og har at

$$\begin{aligned} \widehat{a} = \frac{\sum_{i=1}^{9} x_i(1-x_i)y_i}{\sum_{j=1}^{9}\left[x_j(1-x_j)\right]^2} \approx -0.283528 \\ \mathrm{Var}\left[a\right] = \frac{\sigma_0^2}{\sum_{j=1}^{9}\left[x_j(1-x_j)\right]^2} \approx 0.0018752 \end{aligned} \tag{21}$$

og med $z_{\alpha/2} = 1.960$

$$Z_0 = -6.54746 < -z_{\alpha/2} \tag{22}$$

og må dermed forkaste hypotesen.

Oppg. 3

a) Har den stokastiske variablen $Y_i = \beta x_i + \epsilon_i$ med $\epsilon_i \sim N(0, \sigma^2)$ og kan se på differansen

$$\delta = \sum_{i=1}^{n} (y_i - \widehat{\beta} x_i)^2 \tag{23}$$

og vil finne $\widehat{\beta}$ slik at α blir minst mulig, og ser da på

$$\partial_\beta \delta = -2 \sum_{i=1}^{n} x_i (y_i - \widehat{\beta} x_i) = 0 \Rightarrow \sum_{i=1}^{n} x_i y_i = \widehat{\beta} \sum_{i=1}^{n} x_i^2 \tag{24}$$

og har dermed

$$\widehat{\beta} = \frac{\sum_{i=1}^{11} x_i y_i}{\sum_{i=1}^{11} x_i^2} \tag{25}$$

og ved å sette inn verdier får vi

$$\beta = 0.0567 \tag{26}$$

ser videre også at

$$\mathrm{E}\left[\widehat{\beta}\right] = \mathrm{E}\left[\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}\right] = \frac{1}{\sum_{i=1}^{n} x_i^2} \sum_{i=1}^{n} \mathrm{E}\left[x_i y_i\right] = \frac{1}{\sum_{i=1}^{n} x_i^2} \sum_{i=1}^{n} x_i \mathrm{E}\left[y_i\right] \tag{27}$$

og kan se at

$$\mathrm{E}\left[y_i\right] = \mathrm{E}\left[\beta x_i\right] + \mathrm{E}\left[\epsilon_i\right] = \beta x_i \tag{28}$$

og kan dermed se

$$\mathrm{E}\left[\widehat{\beta}\right] = \frac{\beta \sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^2} = \beta \tag{29}$$

ser videre

$$\mathrm{Var}\left[\widehat{\beta}\right] = \frac{1}{\left[\sum_{i=1}^{n} x_i^2\right]^2} \sum_{i=1}^{n} x_i^2 \mathrm{Var}\left[y_i\right] \tag{30}$$

og kan se at

$$\mathrm{Var}\left[y_i\right] = \mathrm{Var}\left[\beta x_i\right] + \mathrm{Var}\left[\epsilon_i\right] = \sigma^2 \tag{31}$$

og dermed

$$\mathrm{Var}\left[\widehat{\beta}\right] = \frac{\sigma^2 \sum_{i=1}^{n} x_i^2}{\left[\sum_{i=1}^{n} x_i^2\right]^2} = \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2} \tag{32}$$

b) Ser først at

$$y_0 = \widehat{\beta} x_0 = 51.03 \tag{33}$$

ser så på $Y_0 - \widehat{Y}_0$ som er normalfordelt og har også $\mathrm{E}\left[Y_0 - \widehat{Y}_0\right] = 0$ og

$$\mathrm{Var}\left[Y_0 - \widehat{Y}_0\right] = \mathrm{Var}\left[Y_0\right] + \mathrm{Var}\left[\widehat{\beta} x_0\right] = \sigma^2 \left(1 + \frac{x_0^2}{\sum_{i=1}^{n} x_i^2}\right) \tag{34}$$

og kan se på standardiseringen

$$Z = \frac{Y_0 - \widehat{Y}_0}{\sqrt{\sigma^2 \left(1 + \frac{x_0^2}{\sum_{i=1}^{n} x_i^2}\right)}} \sim N(0, 1) \tag{35}$$

men siden vi ikke kjenner σ^2 kan vi heller bruke

$$T = \frac{Y_0 - \widehat{Y}_0}{\sqrt{S^2 \left(1 + \frac{x_0^2}{\sum_{i=1}^{n} x_i^2}\right)}} \sim t_{n-1}, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2 \tag{36}$$

og kan dermed finne intervallet slik at

$$\mathrm{P}\left[-t_{n-1,\alpha/2} \leq T \leq t_{n-1,\alpha/2}\right] = 0.95 \tag{37}$$

med $\alpha = 0.05$ kan vi finne prediksjonsintervallet

$$Y_0 \in \left[\widehat{Y}_0 - t_{10,0.025}\sqrt{S^2\left(1 + \frac{x_0^2}{\sum_{i=1}^n x_i^2}\right)}, \widehat{Y}_0 + t_{10,0.025}\sqrt{S^2\left(1 + \frac{x_0^2}{\sum_{i=1}^n x_i^2}\right)}\right] = [48.50, 53.56] \tag{38}$$