Oppg. 1

Har en normalfordeling med ukjent μ og $\sigma^2 = 0.060^2$

a) Antar $\mu = 6.8$ og har da

$$\mathrm{P}\left[X < 6.74\right] = \mathrm{P}\left[\frac{X-\mu}{\sigma} < \frac{6.74-\mu}{\sigma}\right] = \mathrm{P}\left[Z < -1\right] = 0.1587 \tag{1}$$

ved å slå opp i tabell. Kan så finne $\mathrm{P}\left[6.74 < X < 6.86\right]$, men siden normalfordelingen er lik på begge sider av μ har vi

$$\mathrm{P}\left[6.74 < X < 6.86\right] = 1 - 2\mathrm{P}\left[X < 6.74\right] = 0.6826 \tag{2}$$

til slutt ser vi på $\mathrm{P}\left[|X-\mu| > 0.06\right]$ og har da

$$\mathrm{P}\left[|X-\mu| > 0.06\right] = \mathrm{P}\left[X < 6.74\right] + \mathrm{P}\left[X > 6.86\right] = 1 - \mathrm{P}\left[6.74 < X < 6.86\right] = 0.3174 \tag{3}$$

b) Ser på $Y = \sum_{i=1}^{5} X_i/5$ og vet ikke lenger hva μ er, kan dermed se på

$$\mathrm{Var}\left[Y\right] = \frac{1}{25}\left[5\sigma^2\right] = \frac{1}{5}\sigma^2 \tag{4}$$

vet at Y er normalfordelt og at det ikke har noe å si hvor μ er og kan dermed anta $\mu = 0$ og da

$$\mathrm{P}\left[|Y| > 0.06\right] = 2\mathrm{P}\left[Y > 0.06\right] = 2\mathrm{P}\left[\frac{\sqrt{5}Y}{\sigma} > \frac{0.06\sqrt{5}}{\sigma}\right] = 0.025 \tag{5}$$

ved oppslag i tabell og ser videre på

$$\begin{aligned} \overline{x} - z_{0.025}\frac{\sigma}{\sqrt{n}} &= 6.707 \\ \overline{x} + z_{0.025}\frac{\sigma}{\sqrt{n}} &= 6.813 \end{aligned} \tag{6}$$

og har dermed konfidensintervallet på 95%

$$\mu \in (6.707, 6.813) \tag{7}$$

Oppg. 2

a) Ser på poissonfordelingen

$$f(x) = \frac{\lambda^x}{x!}e^{-\lambda}, x \in \mathbb{N} \tag{8}$$

antar $\lambda = 15$ og ser da at vi har

$$\mathrm{E}\left[X\right] = \mu = \lambda = 15 \tag{9}$$

og har dermed

$$\mathrm{P}\left[X > 20\right] = 1 - \mathrm{P}\left[X \leq 20\right] = 0.083 \tag{10}$$

og

$$\mathrm{P}\left[10 \leq X < 20\right] = \mathrm{P}\left[X < 20\right] - \mathrm{P}\left[X < 10\right] = 0.8053 \tag{11}$$

b) Antar så at vi ikke kjenner λ og har $\hat{\lambda} = \overline{X} = 359/30$ og har da fra sentralgrenseteoremet at dette kan tilnærmes en standard normalfordeling og kan se på

$$Z = \frac{\overline{X} - \lambda}{\sigma}\sqrt{n} \tag{12}$$

og vi bruker variansen i en poissonfordeling som σ og dermed

$$\sigma^2 = \mathrm{Var}\left[X\right] = \overline{X} \tag{13}$$

vil ha et konfidensintervall med 99% sannsynlighet og må dermed ha

$$\begin{aligned} \mathrm{P}\left[-t_{0.005} < T < t_{0.005}\right] = 0.99 \\ \mathrm{P}\left[\overline{X} - t_{0.005}\frac{\sigma}{\sqrt{n}} < \lambda < \overline{X} + t_{0.005}\frac{\sigma}{\sqrt{n}}\right] = 0.99 \end{aligned} \tag{14}$$

slår opp i tabell og finner $t_{0.005} = 2.750$ med frihetsgrad $\nu = 30$ og har dermed

$$\lambda \in (10.34, 13.59) \tag{15}$$

c) Ser så på sannsynlighetsmaksimeringsestimatoren for λ og definerer da

$$\begin{aligned} L(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m; \lambda) &= \prod_{i=1}^{n} f(x_i; \lambda) \cdot \prod_{j=1}^{m} f(y_j; \lambda) \\ &= \lambda^{\sum_{i=1}^{n} x_i} \frac{e^{-n\lambda}}{\prod_{i=1}^{n} x_i!} \cdot \left(\frac{\lambda}{2}\right)^{\sum_{j=1}^{m} y_j} \frac{e^{-\frac{m\lambda}{2}}}{\prod_{j=1}^{m} y_j!} \end{aligned} \tag{16}$$

og finner så logaritmen av L

$$\ln L = \sum_{i=1}^{n} x_i \ln\lambda - n\lambda - \ln\prod_{i=1}^{n} x_i! + \sum_{j=1}^{m} y_j \ln\frac{\lambda}{2} - \frac{m\lambda}{2} - \ln\prod_{j=1}^{m} y_j! \tag{17}$$

og deriverer så med hensyn på λ

$$\begin{aligned} \frac{\partial L}{\partial \lambda} = 0 &= \sum_{i=1}^{n} x_i \frac{1}{\lambda} - n + \sum_{j=1}^{m} y_j \frac{1}{\lambda} - \frac{m}{2} \\ \lambda\left(n + \frac{1}{2}m\right) &= \sum_{i=1}^{n} x_i + \sum_{j=1}^{m} y_j \\ \lambda &= \frac{\sum_{i=1}^{n} x_i + \sum_{j=1}^{m} y_j}{n + \frac{1}{2}m} \\ \lambda &= \frac{\overline{x}n + \overline{y}m}{n + \frac{1}{2}m} \end{aligned} \tag{18}$$

og har dermed

$$\hat{\lambda} = \frac{\overline{x}n + \overline{y}m}{n + \frac{1}{2}m} \tag{19}$$

Oppg. 3

a) Ser at X er definert slik at for hver del av DNA-strukturen så er det to muligheter, at det er en match eller ikke. Siden det sjekkes for n deler som er uavhengige er X binomisk fordelt

$$f(x) = b(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}, x \in \mathbb{N} \cup \{0\} \tag{20}$$

kan dermed finne ved $n = 5$

$$\mathrm{P}\left[X = 2\right] = f(2) = \frac{5!}{2!(5-2)!} \cdot 0.15^2 \cdot (1 - 0.15)^{5-2} = 0.138 \tag{21}$$

og

$$\mathrm{P}[X \geq 2] = 1 - \mathrm{P}[X = 1] - \mathrm{P}[X = 0] = 1 - 5 \cdot 0.15 \cdot 0.85^4 - 0.85^5 = 0.165 \tag{22}$$

og

$$\mathrm{P}[X = 2|X \geq 2] = \frac{\mathrm{P}[X = 2]}{\mathrm{P}[X \geq 2]} = 0.836 \tag{23}$$

b) Sjansen for å begå en type 1 feil vil være at et binomisk forsøk med $p = 0.15, n = 5$ gir 5 og har da sjansen for type 1 feil gitt ved

$$P(X = 5) = p^n = 0.15^5 = 7.594 \cdot 10^{-5} \tag{24}$$

og siden H_1 har $p = 1$ så er det ingen mulighet for at utfallet av forsøket ikke er 5 og dermed er sjansen for en type 2 feil lik 0. Ser så på hvor mange deler av DNA-strukturen som må sammenliknes for å få sjansen for en type 1 feil lik 10^{-6} som er gitt ved

$$10^{-6} = 0.15^n \Rightarrow n = \log_{0.15} 10^{-6} = 7.282 \tag{25}$$

og ved å runde opp får vi at en må sammenlikne $n = 8$ deler av DNA-strukturen for å få ønsket sannsynlighet for type 1 feil.

Oppg. 4

Har X med normalfordeling med $\mu = 80, \sigma^2 = 18^2$ og $Y = \sum_{n=1}^{16} X_n$ ser på

$$\mathrm{P}[X_1 > 90] = \mathrm{P}\left[\frac{X_1 - \mu}{\sigma} > \frac{90 - \mu}{\sigma}\right] = 0.2877 \tag{26}$$

og vet at

$$\mu_Y = 16\mu_X, \quad \sigma_Y^2 = 16\sigma_X^2 \tag{27}$$

har da

$$\mathrm{E}[Y] = 16\mu_X = 1280 \tag{28}$$

og

$$\mathrm{Var}[Y] = 16\sigma_X^2 = 4^2 \cdot 18^2 = 72^2 \tag{29}$$

og til slutt fordi Y også er normalfordelt

$$\mathrm{P}[Y > 16 \cdot 90] = \mathrm{P}\left[\frac{Y - \mu_Y}{\sigma_Y}\sqrt{n} > \frac{16 \cdot 90 - \mu_Y}{\sigma_Y}\sqrt{n}\right] = \mathrm{P}[Z > 2.22] = 0.013 \tag{30}$$

ved oppslag i tabell. Dersom vi hadde sett bortifra uavhengighet kunne vi ikke brukt regnereglene for varians og forventningsverdi og de tre siste utregningene ville ikke vært rett. Vi trenger så vite sannsynlighetsfordelingen for å finne sannsynlighetene $\mathrm{P}[X_1 > 90], \mathrm{P}[Y > 16 \cdot 90]$, men variansen og forventningsverdien er ikke avhengig av dette og disse vil da fortsatt stemme.

b) Ser at vi kan skrive

$$\overline{x} = \frac{1591}{20} = 79.55, S^2 = \frac{1}{19}\sum_{i=1}^{20}(x_i - \overline{x})^2 = 192.471 \tag{31}$$

finner så et 90% konfidensintervall

$$\begin{aligned} \overline{x} - t_{0.05}\frac{S}{\sqrt{n}} &= 74.20 \\ \overline{x} + t_{0.05}\frac{S}{\sqrt{n}} &= 84.90 \end{aligned} \tag{32}$$

og har dermed

$$\mu \in (74.20, 84.90) \tag{33}$$

c) Finner sannsynligheten for at vi har $\sigma^2 \geq 15^2$ og ser på

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = 6.283 \tag{34}$$

vi må ha $\chi^2 < \chi^2_{0.01} = 36.191$ med 19 frihetsgrader og fra det vi ser stemmer dette og vi kan ikke forkaste H_0.