Opgave 6

Opstiller $y_i = \begin{cases} 1 \\ 0 \end{cases}$. Dette er en bernoulli-fordeling.

Dette er en Bernoulli fordeling. Da man enten er ryger, eller også er man ikke. Taget ud fra stikprøven altså. Vi antager, at alle har samme ssh for at ryge, og alle observationer er uafhængige. Yderligere antager vi, at alle observationer har samme sandsynlighed, samt ligefordelt.

Uafhængighed er realt ikke en god antagelse, da respondenter kender og fester med hinanden. Ergo ryger de nok sammen, hvorved ens rygning ikke er 100% uafhængig fra en anden respondents.

1)

Dette er en Bernoulli fordeling. Da man enten er ryger, eller også er man ikke. Taget ud fra stikprøven altså. Vi antager, at alle har samme ssh for at ryge, og alle observationer er uafhængige. Uafhængighed er realt ikke en god antagelse, da respondenter kender og fester med hinanden. Ergo ryger de nok sammen, hvorved ens rygning ikke er 100% uafhængig fra en anden respondents.

2)

Sandsynligheden for et givet udfald for en bernoulli-fordeling er givet ved:

Dette kan skrives, som $f_{Y_i}(y|\phi) = \begin{cases} P(Y_i = 1) = \phi \\ P(Y_i = 0) = 1 - \phi \end{cases}$

Som også kan skrives om til $p(y) = \phi^y (1-\phi)^{1-y}$

Dette kan vi opstille en likelihood-funktion på.

$$L(\phi|y_i) = f_{Y_i}(y_i|\phi) = \phi^{y_i}(1-\phi)^{1-y_i}$$

Det vides, at observationerne er uafhængige. Derfor:

$$L(\phi|y_i) = \prod_{i=1}^{n} l(\phi|y_i) = \prod_{i=1}^{n} \phi^{y_i}(1-\phi)^{(1-y_i)}$$

Herved kan log-likelihood opstilles:

$$\ln L(\phi|y_i) = \ln \prod_{i=1}^{n} \phi^{y_i}(1-\phi)^{(1-y_i)}$$

$$\ln L(\phi|y_i) = \sum_{i=1}^{n} y_i \cdot \ln \phi + (1-y_i) \cdot \ln 1 - \phi$$

3)

Finder scoren ved $\frac{\partial \ln L(\phi|y_i)}{\partial} = \left(\sum_{i=1}^{n} y_i \cdot \ln \phi + (1-y_i) \cdot \ln 1 - \phi\right)'$

$$S(\phi) = \frac{\partial \ln L(\phi|y_i)}{\partial \phi} = \sum_{i=1}^{n} \left(\frac{y_i}{\phi} - \frac{1-y_i}{1-\phi}\right)$$

$$S(\phi) = \sum_{i=1}^{n} \left(\frac{y_i}{\phi} - \frac{1-y_i}{1-\phi}\right) = \frac{\sum_{i=1}^{n} y_i}{\phi} - \frac{n - \sum_{i=1}^{n} y_i}{1-\phi}$$

Isolerer for ϕ:

$$\frac{\sum_{i=1}^{n} y_i}{\phi} - \frac{n - \sum_{i=1}^{n} y_i}{1-\phi}$$

$$\Leftrightarrow \sum_{i=1}^{n} y_i \cdot (1-\phi) = \left(n - \sum_{i=1}^{n} y_i\right) \cdot \phi$$
$$\Leftrightarrow \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} y_i \cdot \phi = n \cdot \phi - \sum_{i=1}^{n} y_i \cdot \phi$$
$$\Leftrightarrow \sum_{i=1}^{n} y_i = n \cdot \phi \Leftrightarrow \frac{\sum_{i=1}^{n} y_i}{n} = \phi$$
$$\phi = \frac{realisationer}{observationer}$$

4)

Finder estimatet. Vi ved, $\sum_{i=1}^{n} y_i = 51$ og $n = 144$
Herved fås estimatet til:

$$\frac{\sum_{i=1}^{n} y_i}{n} = \phi$$
$$\frac{51}{144} = \phi$$
$$0{,}354 = \phi$$

5)

Den nye betingede model indeholder to fremfor en parameter, $\theta_1 \wedge \theta_2$. Hvis personen bruger og bruger ikke cykelhjelm.
Likelihood-funktionen

$$L(\theta_1, \theta_2 | Y_i, X_i) = \prod_{i=1}^{n} \left(\theta_1^{X_i}, \theta_2^{1-X_i}\right)^{Y_i} \left((1-\theta_1)^{X_i}(1-\theta_2)^{1-X_I}\right)^{1-Y_i}$$

Log-likelihood lyder derfor:

$$\ln L(\theta_1, \theta_2 | Y_i, X_i) = \ln \prod_{i=1}^{n} \left(\theta_1^{X_i}, \theta_2^{1-X_i}\right)^{Y_i} \left((1-\theta_1)^{X_i}(1-\theta_2)^{1-X_I}\right)^{1-Y_i}$$
$$= \sum_{i=1}^{n} Y_i \cdot \ln\left(\theta_1^{X_i}\theta_2^{1-X_i}\right) + \sum_{i=1}^{n} (1-Y_i) \cdot \ln\left((1-\theta_1)^{X_i}(1-\theta_2)^{1-X_I}\right)$$
$$= \sum_{i=1}^{n} Y_i \left(X_1 \cdot \ln\theta_1 + (1-X_i) \cdot \ln\theta_2\right) + \sum_{i=1}^{n} (1-Y_i) \left(X_1 \cdot \ln(1-\theta_1) + (1-X_i) \cdot \ln(1-\theta_2\right.$$
$$= \ln\theta_1 \cdot \sum_{i=1}^{n} Y_i X_i + \ln(1-\theta_1) \sum_{i=1}^{n} (1-Y_i) X_i + \ln\theta_2 \sum_{i=1}^{n} Y_i (1-X_i) + (1 - \theta_2) \sum_{i=1}^{n} (1-Y_i)(1-X_i)$$

Scoren bliver derfor:

$$S(\theta_i) = \begin{cases} \dfrac{\sum_{i=1}^{n} Y_i X_i}{\theta_1} - \dfrac{\sum_{i=1}^{n} (1-Y_i) X_i}{1-\theta_1} \\ \dfrac{\sum_{i=1}^{n} Y_i (1-X_i)}{\theta_2} - \dfrac{\sum_{i=1}^{n} (1-Y_i)(1-X_i)}{1-\theta_2} \end{cases}$$

Ud fra dette kan man isolere θ_i.

$$\frac{\sum_{i=1}^{n} Y_i X_i}{\theta_1} - \frac{\sum_{i=1}^{n}(1-Y_i)X_i}{1-\theta_1}$$

$$\frac{1-\theta_1}{\theta_1} = \frac{\sum_{i=1}^{n}(1-Y_i)X_i}{\sum_{i=1}^{n} Y_i X_i}$$

$$\frac{1}{\theta_1} - 1 = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} X_i\, Y_i}{\sum_{i=1}^{n} Y_i X_i}$$

$$\frac{1}{\theta_1} = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} X_i\, Y_i}{\sum_{i=1}^{n} Y_i X_i} + \frac{\sum_{i=1}^{n} Y_i X_i}{\sum_{i=1}^{n} Y_i X_i} \Leftrightarrow \frac{1}{\theta_1} = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} X_i\, Y_i + \sum_{i=1}^{n} Y_i X_i}{\sum_{i=1}^{n} Y_i X_i}$$

$$\theta_1 = \frac{\sum_{i=1}^{n} Y_i X_i}{\sum_{i=1}^{n} X_i}$$

Hermed må θ_2 estimatoren blive noget ala:

$$\theta_2 = \frac{\sum_{i=1}^{n} Y_i(1-X_i)}{\sum_{i=1}^{n}(1-X_i)}$$

Estimaterne bliver hermed:

$$\theta_1 \wedge \theta_2 = 0{,}44 \wedge 0{,}23$$

6)

Likelihood funktionen bliver ved udvidelsen:

$$L(\gamma, \delta, X_i) = \prod_{i=1}^{n} \big(\gamma + \delta(1-X_i)\big)^{Y_i} \big(1 - \gamma - \delta(1-X_i)\big)^{1-Y_i}$$

I denne udvidelse kan δ fortolkes som mersandsynligheden for at være ryger, taget man ikke bruger cykelhjelm.

- $\delta < 0$ er ssh for at være ryger lavere for personer, der ikke bruger cyjkelhjelm
- $\delta = 0$ er ssh for at være ryger den samme for personer, der bruger og ikke bruger cykelhjelm
- $\delta > 0$ er ssh for at være ryger højere for personer, der ikke bruger hjelm

γ kan beskrives som vores tidligere θ.