Fisher-informasjon

- Eksempel: $X \sim \text{Bernoulli}(p)$. Vi vil finne Fisher-informasjonen $I(p)$ om p fra én observasjon X.

Vi har: $f(x;p) = p^x(1-p)^{1-x}$, $x = 0,1$.

Videre vet vi at $E(X) = p$ og $V(X) = p(1-p)$.

Vi har:

$$\log(f(x;p)) = \log\left(p^x(1-p)^{1-x}\right) = \log(p^x) + \log\left((1-p)^{1-x}\right)$$
$$= x\log(p) + (1-x)\log(1-p)$$

$$\longrightarrow \frac{d}{dp}\left(\log(f(x;p))\right) = \frac{d}{dp}\left(x\log(p) + (1-x)\log(1-p)\right)$$
$$= \frac{x}{p} - \frac{1-x}{1-p} = \frac{x(1-p)-(1-x)p}{p(1-p)} = \frac{x - \cancel{px} - p + \cancel{px}}{p(1-p)}$$
$$= \frac{x-p}{p(1-p)}$$

Det gir $U = \dfrac{X-p}{p(1-p)}$

og

$$I(p) = V(U) = V\left(\frac{X-p}{p(1-p)}\right) = V\left(\frac{1}{p(1-p)}X - \frac{p}{p(1-p)}\right)$$
$$= \left(\frac{1}{p(1-p)}\right)^2 V(X) = \frac{1}{p^{\cancel{2}}(1-p)^{\cancel{2}}}\cancel{p(1-p)}$$
$$= \frac{1}{p(1-p)}.$$

Videre er:

$$E(U) = E\left(\frac{X-p}{p(1-p)}\right) = \frac{1}{p(1-p)}E(X) - \frac{\cancel{p}}{\cancel{p}(1-p)} = \frac{\cancel{p}}{\cancel{p}(1-p)} - \frac{1}{1-p} = 0.$$

- Eksempel: $X \sim \text{Bernoulli}(p)$:

  Vi har: $f(x;p) = p^x(1-p)^{1-x}, \quad x = 0, 1$

  $\longrightarrow \{x : f(x;p) > 0\} = \{0, 1\}$ som ikke avhenger av p.

- Eksempel: $X \sim \text{Eksponensiell}(\lambda)$

  Vi har: $f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{ellers} \end{cases}$

  $\longrightarrow \{x : f(x;\lambda) > 0\} = (0, \infty)$ som ikke avhenger av λ.

- Eksempel: $X \sim U(0, \theta)$

  Vi har: $f(x;\theta) = \begin{cases} \frac{1}{\theta}, & 0 < x < \theta \\ 0, & \text{ellers} \end{cases}$

  $\{x : f(x;\theta) > 0\} = (0, \theta)$ som <u>avhenger</u> av θ.

• Eksempel : $X \sim \text{Bernoulli}(p)$

Vi vil finne $I(p)$ vha. alternativ formel:

Vi har:

$$\frac{d^2}{dp^2}\log f(x;p) = \frac{d}{dp}\cdot\Big(\overbrace{\frac{x-p}{p(1-p)}}^{\frac{d}{dp}\log f(x;p)}\Big) = \frac{d}{dp}\Big(\frac{x}{p} - \frac{1-x}{1-p}\Big) = -\frac{x}{p^2} - (1-x)\cdot(\cancel{-}1)\cdot\Big(\cancel{-}\frac{1}{(1-p)^2}\Big)$$

$$= -\frac{x}{p^2} - \frac{1-x}{(1-p)^2}$$

Vi får:

$$I(p) = -E\Big(\frac{d^2}{dp^2}\log\big(f(X;p)\big)\Big) = \cancel{-}E\Big(\cancel{-}\frac{X}{p^2} + \frac{1-X}{(1-p)^2}\Big)$$

$$= \frac{1}{p^2}E(X) + \frac{1}{(1-p)^2}\big(1-E(X)\big)$$

$$= \frac{1}{p^{\cancel{2}}}\cdot\cancel{p} + \frac{1}{(1-p)^{\cancel{2}}}\cancel{(1-p)}$$

$$= \frac{1}{p} + \frac{1}{1-p} = \frac{1\cancel{-p}\cancel{+p}}{p(1-p)} = \frac{1}{p(1-p)}$$

Vi antar at X har tetthet/punktsannsynlighet $f(x;\theta)$, som avhenger av θ, mens $\{x : f(x;\theta) > 0\}$ ikke avhenger av θ. For kontinuerlig X har vi da:

$$1 = \int f(x;\theta)\,dx \quad (*)$$

Videre merk at $\frac{\partial}{\partial\theta} \log f(x;\theta) = \frac{\frac{\partial f(x;\theta)}{\partial\theta}}{f(x;\theta)}$, slik at

$$\frac{\partial f(x;\theta)}{\partial\theta} = f(x;\theta)\cdot \frac{\partial}{\partial\theta}\log f(x;\theta).$$

Vi deriverer $(*)$ mhp θ på begge sider av likhetstegnet:

$$\begin{aligned}
0 = \frac{\partial}{\partial\theta}\int f(x;\theta)\,dx &= \int \frac{\partial}{\partial\theta} f(x;\theta)\,dx \\
&= \int \frac{\partial}{\partial\theta}\log f(x;\theta)\cdot f(x;\theta)\,dx \\
&= E\Big(\underbrace{\frac{\partial}{\partial\theta}\log f(X;\theta)}_{U}\Big) \\
&= E(U)
\end{aligned}$$

Vi partiellderiverer nok en gang mhp θ og får:

$$\begin{aligned}
0 &= \frac{\partial}{\partial\theta}\Big(\int \frac{\partial}{\partial\theta}\log f(x;\theta)\cdot f(x;\theta)\,dx\Big) \\
&= \int \frac{\partial}{\partial\theta}\Big(\frac{\partial}{\partial\theta}\log f(x;\theta)\cdot f(x;\theta)\Big)dx \\
&= \int\Big(\frac{\partial^2}{\partial\theta^2}\log f(x;\theta)\cdot f(x;\theta) + \frac{\partial}{\partial\theta}\log f(x;\theta)\cdot\frac{\partial f(x;\theta)}{\partial\theta}\Big)dx \\
&= \underbrace{\int \frac{\partial^2}{\partial\theta^2}\log f(x;\theta)\cdot f(x;\theta)\,dx}_{E\left(\frac{\partial^2}{\partial\theta^2}\log f(X;\theta)\right)} + \int \frac{\partial}{\partial\theta}\log f(x;\theta)\cdot\frac{\partial}{\partial\theta}\log f(x;\theta)\cdot f(x;\theta)\,dx \\
&= E\Big(\frac{\partial^2}{\partial\theta^2}\log f(X;\theta)\Big) + \underbrace{\int\Big(\frac{\partial}{\partial\theta}\log f(x;\theta)\Big)^2\cdot f(x;\theta)\,dx}_{E\left(\left(\frac{\partial}{\partial\theta}\log f(X;\theta)\right)^2\right)} \\
&= E\Big(\frac{\partial^2}{\partial\theta^2}\log f(X;\theta)\Big) + E\Big(\big(\underbrace{\tfrac{\partial}{\partial\theta}\log f(X;\theta)}_{U}\big)^2\Big) \\
&= E\Big(\frac{\partial^2}{\partial\theta^2}\log f(X;\theta)\Big) + E(U^2)
\end{aligned}$$

Da $E(U) = 0$, så er $V(U) = E(U^2) - \big(\underbrace{E(U)}_{=0}\big)^2 = E(U^2)$

Vi får:

$$0 = E\Big(\frac{\partial^2}{\partial\theta^2}\log f(X;\theta)\Big) + V(U) = E\Big(\frac{\partial^2}{\partial\theta^2}\log f(X;\theta)\Big) + I(\theta)$$

$$\longrightarrow \quad I(\theta) = -E\Big(\frac{\partial^2}{\partial\theta^2}\log f(X;\theta)\Big).$$

For diskret X får en akkurat samme resultater ved å bytte ut integraler med summer.

Fisher-informasjon i et tilfeldig utvalg

Vi har $X_1, \dots, X_n$ uif med fetthet/punktsannsynlighet $f(x;\theta)$.

Da er score-funksjonen:

$$\frac{\partial}{\partial\theta} \log f(X_1,\dots,X_n;\theta) \overset{\text{uif}}{=} \frac{\partial}{\partial\theta} \sum_{i=1}^{n} \log f(X_i;\theta) = \sum_{i=1}^{n} \underbrace{\frac{\partial}{\partial\theta} \log f(X_i;\theta)}_{U_i}$$

$$= \sum_{i=1}^{n} U_i$$

Fisher-informasjonen til utvalget blir:

$$I_n(\theta) = V\left(\frac{\partial}{\partial\theta} \log f(X_1,\dots,X_n;\theta)\right) = V\left(\sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(X_i;\theta)\right)$$

$$\overset{\text{uavh.}}{=} \sum_{i=1}^{n} V\Big(\underbrace{\frac{\partial}{\partial\theta} \log f(X_i;\theta)}_{U_i}\Big)$$

$$= \sum_{i=1}^{n} V(U_i)$$

$$= \sum_{i=1}^{n} I(\theta) = n \cdot I(\theta)$$

• Eksempel : $X_1, \dots, X_n \overset{uif}{\sim}$ Bernoulli(p)

Vi vet at for én observasjon er $I(p) = \frac{1}{p(1-p)}$

Da er nedre grense for variansen til forventningsrett estimator for p :

$$\frac{1}{n \cdot I(p)} = \frac{1}{n \cdot \frac{1}{p(1-p)}} = \frac{p(1-p)}{n}$$

Estimatoren $\hat{p} = \frac{\sum_{i=1}^{n} X_i}{n}$ er forventningsrett med varians

$V(\hat{p}) = \frac{p(1-p)}{n}$. Det betyr at $\hat{p}$ er effisient.

- Eksempel: $X_1, \dots, X_n \sim \text{Bernoulli}(p)$

Vi vet fra tidligere at MLE for p er $\hat{p} = \frac{\sum_{i=1}^{n} X_i}{n}$

og $I(p) = \frac{1}{p(1-p)}$. Det betyr at $\hat{p} \overset{\text{tiln.}}{\sim} N\left(p, \frac{p(1-p)}{n}\right)$ for store n.

- Eksempel: $X_1, \dots, X_n \overset{uif}{\sim} \text{Eksponensiell}(\lambda)$

$\longrightarrow$ MLE for λ er $\hat{\lambda} = \frac{1}{\overline{X}}$.

Vi har:

$$\log f(x;\lambda) = \log\left(\lambda e^{-\lambda x}\right) = \log(\lambda) - \lambda x$$

$$\longrightarrow \frac{\partial}{\partial \lambda} \log f(x;\lambda) = \frac{1}{\lambda} - x$$

$$\longrightarrow \frac{\partial^2}{\partial \lambda^2} \log f(x;\lambda) = -\frac{1}{\lambda^2}$$

$$\longrightarrow I(\lambda) = -E\left(\frac{\partial^2}{\partial \lambda^2} \log f(X;\lambda)\right) = -E\left(-\frac{1}{\lambda^2}\right) = \frac{1}{\lambda^2}.$$

Det betyr at $\hat{\lambda} \overset{\text{tiln.}}{\sim} N\left(\lambda, \frac{\lambda^2}{n}\right)$.