20180502 Kom ihåg

$X_1, X_2, \ldots, X_n$ är ett stickprov & vi vill skatta parametern θ.

Def då θ är okänd skriver vi $f(x_1, x_2, \ldots, x_n | \theta)$ för den gemensamma tfkn/slf, för s.v. $X_1, \ldots, X_n$.

Anm: 1) i det diskreta fallet är kanske $p(x_1, x_2, \ldots, x_n | \theta)$ en mer naturlig notation. Det blir mkt opraktiskt att använda olika notation för diskret/kontinuerlig.

2) om $X_1, \ldots, X_n$ är ober. får vi att $f(x_1, \ldots, x_n | \theta) =$
$= f(x_1|\theta) f(x_2|\theta) \ldots f(x_n|\theta)$.

3) vi skriver $f(\cdot|\theta)$ för att poängtera att $\theta = (\theta_1, \ldots, \theta_m)$ är okänd.

Def givet observerade data $x_1, \ldots, x_n$ (dvs. $X_1 = x_1, \ldots, X_n = x_n$) def. vi likelihooden av θ som $lik(\theta) = f(x_1, \ldots, x_n | \theta)$

Anm: 1) observera att $lik(\theta)$ ses som en funk. av θ, ej av de uppmätta värdena $x_1, \ldots, x_n$.

2) om $X_1, \ldots, X_n$ är diskreta & ober så blir $lik(\theta) = f(x_1, \ldots, x_n | \theta)$
$= P(X_1 = x_1 | \theta) \ldots P(X_n = x_n | \theta)$.

Alltså är $lik(\theta)$ sanno. att få de uppmätta värdena $x_1, x_2, \ldots, x_n$ som funk. av θ.

Def ML-skattaren (Maximum Likelihood Estimator, MLE) av θ är det värde $\hat{\theta}$ som maximerar $lik(\theta)$.

Anm: 1) i det diskreta fallet är $\hat{\theta}$ det värde på θ som gör de observerade värdena (data) mest troligt (most likely)

2) det kont. fallet är analogt om än ngt mer svårtolkat

Vi kommer vilja maximera $lik(\theta)$. Ofta är det enklast att maximera $l(\theta) = \log(lik(\theta)) \overset{\text{ober.}}{=} \sum_{i=1}^{n} \log(f(x_i|\theta))$

Anm: om $X_1 = x_1, \ldots, X_n = x_n$ är alltså $lik(\theta) = \prod_{i=1}^{n} f(x_i|\theta) = lik(\theta, x_1, \ldots, x_n)$. Man skriver ofta $lik(\theta) \overset{\text{ober.}}{=} \prod_{i=1} f(X_i|\theta) = lik(\theta, X_1, \ldots, X_n)$. Detta är analogt med hur $E[X|Y]$ definierades av att den blev $E[X|Y=y]$ om händelsen $Y=y$ inträffade.

Antar numera alltid ober.

Tal låt $X_1, \ldots, X_n$ vara iid $Poi(\lambda)$

a) hitta MME'n $\hat{\lambda}$ b) hitta MLE'n $\hat{\lambda}$

L a) då $X_i \sim Poi(\lambda)$ är $E[X_i] = \lambda$.

MM: $\bar{x} = \hat{\lambda}$ (så trivialt) $\hat{\lambda} = \bar{X}$

b) vi har att $lik(\lambda) = f(k_1, \ldots, k_n|\lambda) = \prod_{i=1}^{n} f(k_i|\lambda) = \prod_{i=1}^{n} P(X_i = k_i|\lambda) = \prod_{i=1}^{n} \frac{\lambda^{k_i}}{k_i!} e^{-\lambda}$

alt. $lik(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{X_i}}{X_i!} e^{-\lambda}$.

$\Rightarrow l(\lambda) = \log(lik(\lambda)) = \sum_{i=1}^{n} (X_i \log(\lambda) - \log X_i! - \lambda) = \log(\lambda) \sum_{i=1}^{n} X_i - n\lambda - \sum_{i=1}^{n} \log X_i!$.

Detta ska maximeras: $l'(\lambda) = \frac{1}{\lambda} \sum_{i=1}^{n} X_i - n$, $l''(\lambda) = \frac{-1}{\lambda^2} \sum_{i=1}^{n} X_i < 0$ om $\neq 0$

i) Så $l'(\lambda) = 0 \Rightarrow \lambda = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}$ om $\neq 0$

Dessutom är då $lik''(\lambda) < 0$ s.a. $\hat{\lambda} = \bar{X}$ maximerar $l(\lambda)$

ii) Om nu $\sum_{i=1}^{n} X_i = 0 \Rightarrow lik(\lambda) = e^{-n\lambda}$ maximeras av $\lambda = 0$ (ty $\lambda \geq 0$) s.a. återigen blir $\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}$.

Anm: om bytet $k_i \to X_i$ <u>ej</u> genomförts:

om $X_1 = k_1, \ldots, X_n = k_n \Rightarrow \hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} k_i$

Ta1 Låt $X_1, X_2, \dots, X_n$ vara iid $N(\mu, \sigma^2)$

a) hitta MME'n för μ & σ^2

b) hitta MLE'n för μ & σ (ej σ^2)

L a) Vi har att $E[X] = \mu$ & $E[X^2] = Var(X) + E[X]^2 = \sigma^2 + \mu^2$.

Enl. momentmetoden ska vi sätta $\overline{X} = \hat{\mu}$ & $\overline{X^2} = \widehat{\sigma^2} + \hat{\mu}^2$.

Vi får att $\widehat{\sigma^2} = \overline{X^2} - \hat{\mu}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 = \frac{1}{n}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)$

$= \frac{1}{n}\sum_{i=1}^{n}(X_i^2 + \overline{X}^2 - 2X_i\overline{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ //

b) Vi har att om $X_i \sim N(\mu, \sigma^2)$

$f(x_1,..x_n \mid \mu, \sigma) = \prod_{i=1}^{n}(\sigma\sqrt{2\pi})^{-1} exp\left(-\frac{1}{2}\frac{(x_i-\mu)^2}{\sigma^2}\right)$

$\Rightarrow L(\mu, \sigma) = \sum_{i=1}^{n} \log\left(\frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\frac{(x_i-\mu)^2}{\sigma^2}\right)\right) =$

$= -n\log\sigma - n\log\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$. Vi hittar max:

$\frac{\partial L}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)$, $\frac{\partial L}{\partial \sigma} = \frac{-n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2$

villkoret $\frac{\partial L}{\partial \mu} = 0 \Rightarrow 0 = \sum_{i=1}^{n}(x_i-\mu) = -n\mu + \sum_{i=1}^{n} x_i$

$\Rightarrow \mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$.

Vidare är $\frac{\partial L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2 = \{\mu = \overline{x}\} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\overline{x})^2$,

om $\frac{\partial L}{\partial \sigma} = 0 \Rightarrow \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2}$.

Övning: har vi hittat maximat? (Ja)

Om $X_1 = x_1, \dots, X_n = x_n$ så är $\hat{\mu} = \overline{x}$, $\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2}$

alt. $\hat{\mu} = \overline{X}$ $\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}$

Anm: $\mathbb{E}[\hat{\sigma}^2] = \mathbb{E}[\frac{n-1}{n} \cdot \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2] =$

$= \mathbb{E}[\frac{n-1}{n} s^2] = \frac{n-1}{n} \sigma^2$ så ej VVR.

Man justerar sina skattare!

Observera även att om $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ är VVR, så har ni att

$\mathbb{E}[s] = \mathbb{E}[\sqrt{s^2}] \neq \sqrt{\mathbb{E}[s^2]} = \sqrt{\sigma^2} = \sigma$

så s är ej en VVR-skattare av σ.

I allmänhet $\mathbb{E}[\sqrt{X}] \leq \sqrt{\mathbb{E}[X]}$ om $X \geq 0$