<u>Kom ihåg:</u> $X_1, \dots, X_n$ är ett stickprov och vi vill skatta parametern θ.

<u>Def:</u> Då θ är okänd skriver vi

$f(x_1, \dots, x_n \mid \theta)$ för den gemensamma slf/tfkn

för s.v. $X_1, \dots, X_n$.

<u>Anm:</u> 1) I det diskreta fallet är kanske

$p(x_1, \dots, x_n \mid \theta)$ mer naturlig notation. Det blir dock opraktiskt med olika notationer för diskret/kont.

2) Om $X_1, \dots, X_n$ är oberoende får vi att

$$f(x_1, \dots, x_n \mid \theta) = f(x_1 \mid \theta) \cdots f(x_n \mid \theta).$$

3) Vi skriver $f(\cdot \mid \theta)$ för att poängtera att $\theta = (\theta_1, \dots, \theta_m)$ är okänd.

<u>Def:</u> Givet observerade data $x_1, \dots, x_n$ (dvs $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$) definierar vi likelihooden av θ som

$$lik(\theta) = f(x_1, \dots, x_n \mid \theta)$$

<u>Anm:</u> 1) Observera att $lik(\theta)$ ses som en funktion av θ, ej $x_1, \ldots, x_n$.

2) Om X_i är diskreta så blir

$$lik(\theta) = f(x_1, \ldots, x_n \mid \theta) = \mathbb{P}(X_1 = x_1 \mid \theta) \cdots \mathbb{P}(X_n = x_n \mid \theta).$$

Alltså är $lik(\theta)$ sann. att få de givna värdena $x_1, \ldots, x_n$ som funktion av θ.

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

<u>Anm:</u> 1) I det diskreta fallet är $\hat{\theta}$ det värde på θ som gör de observerade data mest troligt (most likely).

2) Det kont. fallet är mer svårtolkat.

Vi kommer alltså vilja maximera $lik(\theta)$. Ofta är det enklast att maximera

$$l(\theta) = \log(lik(\theta)) = \sum_{i=1}^{n} \log(f(x_i \mid \theta))$$

Anm: Om $X_1 = x_1, \ldots, X_n = x_n$ är alltså

$$\mathrm{lik}(\theta) = \prod_{i=1}^{n} f(x_i \mid \theta) = \mathrm{lik}(\theta, x_1, \ldots, x_n)$$

Man skriver ofta

$$\mathrm{lik}(\theta) = \prod_{i=1}^{n} f(X_i \mid \theta) = \mathrm{lik}(\theta, X_1, \ldots, X_n).$$

Detta är analogt med $\mathbb{E}[X \mid Y]$ def. av $\mathbb{E}[X \mid Y = y]$ $\forall y$.

Tal: Låt $X_1, \ldots, X_n$ vara iid $\mathrm{Poi}(\lambda)$.

a) Hitta MME $\hat{\lambda}$.

b) Hitta MLE $\hat{\lambda}$.

L: a) Då $X_i \sim \mathrm{Poi}(\lambda)$ är $\mathbb{E}[X_i] = \lambda$

MME: Lös $\bar{X} = \hat{\lambda}$ (trivialt) så $\hat{\lambda} = \bar{X}$.

b) Vi har att

$$\mathrm{lik}(\lambda) = l(k_1, \ldots, k_n \mid \lambda) = \prod_{i=1}^{n} l(k_i \mid \lambda) = \prod_{i=1}^{n} \frac{\lambda^{k_i}}{k_i!} e^{-\lambda},$$

alternativt

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

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

$$= \log \lambda \sum_{i=1}^{n} X_i - n\lambda - \sum_{i=1}^{n} \log X_i!$$

Detta skall maximeras:

$$\ell'(\lambda) = \frac{1}{\lambda}\sum_{i=1}^{n} X_i - n$$

$$\ell''(\lambda) = -\frac{1}{\lambda^2}\sum_{i=1}^{n} X_i < 0 \quad \text{om} \quad \sum_{i=1}^{n} X_i \neq 0$$

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

och $\ell''(\lambda) < 0$ ger att $\hat{\lambda} = \overline{X}$ maximerar $\ell(\lambda)$.

Om nu $\sum_{i=1}^{n} X_i = 0 \Rightarrow \ell(\lambda) = e^{-n\lambda}$ maximeras av

$\lambda = 0$, så att återigen blir $\hat{\lambda} = \overline{X}$.

Tal: Låt $X_1, X_2, \ldots$ vara i.i.d. $N(\mu, \sigma^2)$.

a) Hitta MME för μ, σ^2

b) Hitta MLE för μ, σ. ← subtilt men viktigt!

1: Vi har att $\mathbb{E}[X] = \mu$ och

$$\mathbb{E}[X^2] = \mathrm{Var}(X) + \mathbb{E}[X]^2 = \sigma^2 + \mu^2.$$

Enligt Momentmetoden skall vi sätta

$$\overline{X} = \hat{\mu} \quad \text{och}$$

$$\overline{X^2} = \widehat{\sigma^2} + \hat{\mu}^2$$

$$\Rightarrow \quad \hat{\sigma}^2 = \overline{X^2} - \hat{\mu}^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - \left(\frac{1}{n}\sum_{i=1}^n X_i\right)^2$$

$$= \frac{1}{n}\left(\sum_{i=1}^n X_i^2 - n\bar{X}^2\right) = \frac{1}{n}\sum_{i=1}^n (X_i^2 + \bar{X}^2 - 2X_i\bar{X})$$

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

b)

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

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

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

$$= -n\log\sigma - n\log\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum_{i=1}^n (X_i-\mu)^2.$$

Vi hittar maximum:

$$\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^n (X_i - \mu)$$

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

S.a. $\frac{\partial \ell}{\partial \mu} = 0 \Rightarrow 0 = \sum_{i=1}^n (X_i - \mu) = -n\mu + \sum_{i=1}^n X_i$

$$\Rightarrow \quad \mu = \bar{X}$$

Vidare är

$$0 = \frac{\partial l}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (X_i - \mu)^2 = \{ \mu = \bar{X} \}$$

$$= -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (X_i - \bar{X})^2 \Rightarrow \sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2}$$

D.v.s. $\hat{\mu} = \bar{X}$ och $\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2}$.

<u>Anm:</u> 1) MLE för σ^2 (dvs $\widehat{\sigma^2}$) blir

$\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$ dvs $(\hat{\sigma})^2$. Detta följer

av att $\frac{\partial l}{\partial \sigma^2} = \frac{\partial l}{\partial \sigma} \cdot \frac{\partial \sigma}{\partial \sigma^2} = \frac{\partial l}{\partial \sigma} \cdot \frac{1}{\frac{\partial \sigma^2}{\partial \sigma}} = \frac{\partial l}{\partial \sigma} \cdot \frac{1}{2\sigma}$

$\ulcorner$ σ^* s.a. $\left(\frac{\partial l}{\partial \sigma} \cdot \frac{1}{2\sigma}\right)(\sigma^*) = 0$ samma σ^* s.a. $\frac{\partial l}{\partial \sigma^2}(\sigma^*) = 0$

$\Rightarrow \frac{\partial l}{\partial \sigma^2}((\sigma^*)^2) = 0$ där vi ser $\frac{\partial l}{\partial \sigma^2}$ som fkn av σ^2! $\lrcorner$

2) MME $\hat{\sigma}$ för σ är samma som $\sqrt{\widehat{\sigma^2}}$

3) $E[\widehat{\sigma^2}] = E\left[\frac{n-1}{n} \cdot s^2\right] = \frac{n-1}{n} \sigma^2$ så ej VVR!

Observera även att (med $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$)

$$E[s] = E[\sqrt{s^2}] \neq \sqrt{E[s^2]} = \sqrt{\sigma^2} = \sigma \quad \text{så}$$

s är ej en VVR skattare för σ.

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