Parametric Inference

sample

$\underline{X = (X_1, \dots, X_n)} \overset{i.i.d}{\sim} F_\theta(\cdot)$ ← family of distr., known

θ ← parameter, unknown

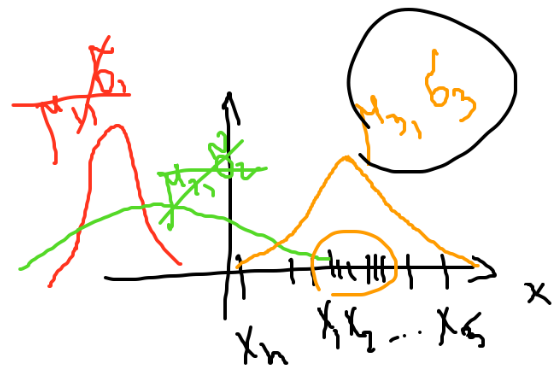
$\downarrow$

$\hat{\theta}$

$\Rightarrow \hat{\Theta}$ - estimate of paran

$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}$ ← plug-in

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



I Plug-in estimator

II Maximum Likelihood Estimator

III Method of Moments Estimator

Likelihood of sample? = pro

$p(X|\Theta)$ → density

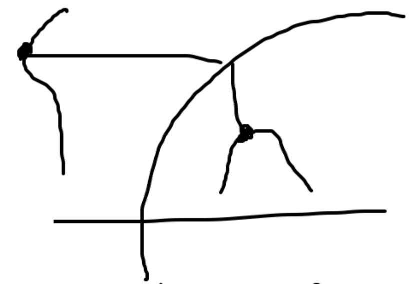
→ PMF for di

1000000

$p(X|\theta) = \prod^{n} p(X_i|\theta)$

likelihood $\boxed{p(X|\theta)}$ "how likely it is to have our sample w/ these pa

log-likelihood $\mathcal{L}(X,\theta) = \log p(X|\theta) = \log \prod_{i=1}^{n} p(x_i|\theta) = \underbrace{\sum_{i=1}^{n} \log \underbrace{p(x_i|\theta)}_{\in[0,1]}}_{\text{large negative value}}$

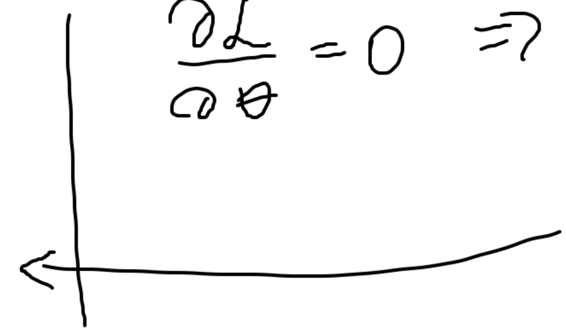


$\arg\max_{\theta} p(X|\theta) = \hat{\theta}_{MLE}$

$\boxed{\arg\max_{\theta} \mathcal{L}(X,\theta)} = \arg\max_{\theta} p(X|\theta) = \hat{\theta}_{MLE}$

$\frac{\partial \mathcal{L}}{\partial \theta} = 0 \Rightarrow$ solve

et



$X = (x_1, \dots, x_n) \sim \mathcal{N}(\mu, \sigma^2)$

unknown (μ), known (σ^2)

$\frac{\partial}{\partial \mu} \mathcal{L}(X,\mu) = \frac{\partial}{\partial \mu} \sum \log p(x_i|\mu)$

$(x-\mu)^2$

$$p_i = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\}$$

$$l = \sum \log p_i = \sum \log\left(\frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\}\right) =$$

$\hat{\mu} = 0 \rightarrow \ldots$

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

use gradient descent (or other opt.)

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

$$\boxed{\frac{\partial L}{\partial \mu} = \cancel{\frac{2}{2\sigma^2}}\sum_{i=1}^{n}(x_i-\hat{\mu}) = 0} \Rightarrow \sum_{i=1}^{n}(x_i-\hat{\mu}) = 0 \Rightarrow \sum_{i=1}^{n} x_i - n\hat{\mu} = 0$$

$$\Downarrow$$

III. Method of moments

$$\underbrace{\mathbb{E}[X^k]}_{\text{theoretical (with unk. params)}} = \underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i^k}_{\text{practical (number)}}$$

k? $\rightarrow$ $k = 1, \ldots, m$

m is the num of unknown params

$X_1, \ldots, X_n \sim N(?, \sigma^2)$

$$\hat{\mu}_{MoM} = \mathbb{E}X = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$X_1, \ldots, X_n \sim N(?, ?)$

$$\hat{\mu}_{MoM} = \mathbb{E}X = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\hat{\sigma}^2_{MoM} + \hat{\mu}^2_{MoM} = \mathbb{E}X^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

2nd sample moment

$$\underbrace{Var(X)}_{\sigma^2} = \mathbb{E}X^2 - (\underbrace{\mathbb{E}X}_{\mu})^2 \Rightarrow \mathbb{E}X^2 = \sigma^2 + \mu^2$$

$$\hat{\sigma}^2_{мом} = \frac{1}{n}\sum X_i^2 - \left(\frac{1}{n}\sum X_i\right)^2 = \underbrace{\frac{1}{n}\sum X_i^2} - \bar{X}^2 =$$

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

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

$$= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 + 2\bar{X}\,\cancel{\sum\left(X_i - \bar{X}\right)}^{=0} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

$$\sum X_i - n\bar{X}$$

$$\sum X_i - \cancel{n}\left(\cancel{\tfrac{1}{n}}\sum X_i\right)$$

$$\sum X_i - \sum X_i = 0$$