Lec 1/25

Wednesday, January 25, 2017 14:57

Method of Moments: $m_k' = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ $\mu_k' = E(X_i^k)$

No guarantees on bias, variance, sufficiency, etc.

Solving System of eqns:

$m_1' = \mu_1', \ldots, m_r' = \mu_r'$ for r parameters.

Ex: let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. find estimators for μ, σ^2.

$m_1' = \mu_1' \Rightarrow^{*} \bar{X} = \mu$

$m_2' = \mu_2' \Rightarrow \frac{\sum_{i=1}^{n} X_i^2}{n} = E(X_1^2) = Var(X_1) + E(X_1)^2 = \sigma^2 + \mu^2.$

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

So $*$ and $**$ are estimators for μ and σ^2.

10.6 Robustness

An estimator is robust if it's not seriously affected by violations of assumptions. (eg. assume a wrong model).

e.g. lets say we want to estimate mean μ of a pop. given a sample $x_1, \ldots, x_n$. from the population. If we assume $X_i \overset{iid}{\sim}$ Poisson(λ), then $\bar{X}$ and S^2 both estimate $\lambda = E(X_i)$. Which is more robust?

$\bar{X}$ will still estimate $E(X)$ if X_i are not Poisson.

S^2 may not.

So $\bar{X}$ is more robust.

10.8 Maximum Likelihood.

Ex: Suppose 4 boxes of cereal containing a toy are purchased.

The toy is spiderman, venom, or sandman.

Let K = number of boxes containing spiderman.

Now suppose 1 box is lost on the way home.

Of the remaining 3 boxes, 2 had spiderman.

How to estimate K?

~~$P(K=3) = \frac{1}{3}$, $P(K=2) = \frac{2}{3}$ So should estimate $K=2$~~

If $K = 2$, the prob. of "observed data" $= \dfrac{\binom{2}{2}\binom{2}{1}}{\binom{4}{3}} = \dfrac{2}{4}$

" 3 " " $= \dfrac{\binom{3}{2}\binom{1}{1}}{\binom{4}{3}} = \dfrac{3}{4}$

So 3 is the best estimator for K.

Idea of ML estimation:

Pick values of parameters that give the highest probability of observing what was observed.

Def (Likelihood function)

If $x_1, \dots, x_n$ are the values of a random sample $X_1, \dots, X_n$ from a population with parameter θ, then the likelihood function of the sample is given by:

$$L(\theta) = f(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

where f is the joint pmf/pdf of X_i s.

ML Estimation:

maximize $L(\theta)$ wrt θ. The resulting θ is $\hat{\theta}$, the maximum likelihood estimator. (MLE)

$$\hat{\theta} = \underset{\theta}{\arg\max}\left(L(\theta)\right)$$

$$\hat{\theta} = \underset{\theta \in \Omega}{\operatorname{argmax}} \left(L(\theta) \right)$$

$\nwarrow$ all possible values for θ. e.g if $\theta = \sigma^2$, $\Omega = [0, \infty)$

Ex: $X_1, \ldots, X_n \overset{iid}{\sim} \text{Poisson}(\lambda)$ find MLE of λ.

$$L(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \lambda^{\sum x_i} \cdot e^{-n\lambda} \cdot \prod_{i=1}^{n} \frac{1}{x_i!}$$

So maximize $\lambda^c \cdot e^{-n\lambda} \cdot d \Leftrightarrow$ maximize $\overset{(*)}{c \log(\lambda) - n\lambda + \log(d)}$

$$\frac{\partial}{\partial \lambda}(*) = \frac{c}{\lambda} - n = 0 \Leftrightarrow \frac{c}{n} = \lambda \Leftrightarrow \frac{\sum x_i}{n} = \overline{X} = \lambda.$$

$\overline{X}$ is MLE.

$$\frac{\partial^2}{\partial \lambda^2}(*) = \frac{-c}{n^2} < 0$$ so that pt was a max. if not $(x_1, \ldots, x_n) = (0, \ldots, 0)$.

if $x_i = 0 \ \forall i$, then $L(\lambda) = e^{-n\lambda} < 1 \ \forall \lambda > 0$ so there's no maximizer.

MLE DNE.