Nov. 10th. Point estimation.

1. Review last lab

2. Special point estimations

a. Method of moments (MoM)

b. Maximinum likelihood estimator (MLE).

MoM.

$X_i \overset{i.i.d}{\sim} X$, $\mu_k = \mathbb{E}X^k$, $m_k = \dfrac{\sum_{i=1}^{n} X_i^k}{n}$

Simple, intuitive.

**9.75** Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from the probability density function given by

$$f(y \mid \theta) = \begin{cases} \dfrac{\Gamma(2\theta)}{[\Gamma(\theta)]^2}(y^{\theta-1})(1-y)^{\theta-1}, & 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the method-of-moments estimator for $\theta$.

$Y \sim Beta(\alpha, \beta)$, $\alpha = \beta = \theta$. $\mathbb{E}Y = \dfrac{\alpha}{\alpha+\beta} = \dfrac{1}{2}$,

$Var\,Y = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \dfrac{1}{4(2\theta+1)}$.

$\mathbb{E}Y^2 = Var\,Y + (\mathbb{E}Y)^2 = \dfrac{1}{4(2\theta+1)} + \dfrac{1}{4}$

$S_n^2 = \dfrac{\sum_{i=1}^{n} Y_i^2}{n}$

$\mathbb{E}Y^2 = S_n^2$, $\Rightarrow$ $\dfrac{1}{4(2\theta+1)} + \dfrac{1}{4} = S_n^2$,

$\hat{\theta} = \dfrac{1}{2}\left(\dfrac{1}{4S_n^2 - 1} - 1\right)$.

MLE.

$f(x;\theta)$ is pdf/pmf.

$f(X_1,\dots,X_n;\theta) = \prod_{i=1}^{n} f(X_i;\theta)$ joint pdf.

$L_n(\theta) \stackrel{def}{=} f(X_1,\dots,X_n;\theta)$ likelihood function. From a Bayesian viewpoint.

$\propto f(\theta|X_1,\dots,X_n)$ posterior density of $\theta$ given $\wedge$ uniform prior.

$\hat{\theta} = \arg\max_{\theta} L_n(\theta)$. MLE.

**9.80** Suppose that $Y_1, Y_2, \dots, Y_n$ denote a random sample from the Poisson distribution with mean $\lambda$.

- **a** Find the MLE $\hat{\lambda}$ for $\lambda$.
- **b** Find the expected value and variance of $\hat{\lambda}$.
- **c** Show that the estimator of part (a) is consistent for $\lambda$.
- **d** What is the MLE for $P(Y = 0) = e^{-\lambda}$?

a. $P(Y=y) = \frac{\lambda^y}{y!} e^{-\lambda}$ $\mathbb{E}Y=\lambda$, $Var\,Y=\lambda$.

$L_n(\lambda) = f(Y_1,\dots,Y_n;\lambda) = \prod_{i=1}^{n} f(Y_i;\lambda)$

$= \prod_{i=1}^{n} \frac{\lambda^{y_i}}{y_i!} e^{-\lambda}$.

$\ell_n(\lambda) = \underbrace{\log}_{} L_n(\theta) = \sum_{i=1}^{n} \log f(Y_i;\lambda)$

$= \sum_{i=1}^{n} (y_i \ln\lambda - \lambda - \underbrace{\ln y_i!}_{})$

$= (\sum_{i=1}^{n} y_i) \ln\lambda - n\lambda \underline{+c}$

$\frac{\partial \ln(\lambda)}{\lambda} = 0 \Leftrightarrow \frac{\sum_{i=1}^{n} y_i}{\lambda} - n = 0 \Leftrightarrow \hat{\lambda} = \frac{\sum_{i=1}^{n} y_i}{n} = \bar{Y}$.

b. $\mathbb{E}\hat{\lambda} = \mathbb{E}\bar{Y} = \lambda$, $\underbrace{Var\,\hat{\lambda} = \frac{Var\,Y}{n} = \frac{\lambda}{n}}_{}$.

c. WLLN.

d. $e^{-\lambda}$; $e^{-\hat{\lambda}}$ $\boxed{9.94}$

**9.97** The geometric probability mass function is given by

$$p(y \mid p) = p(1-p)^{y-1}, \qquad y = 1, 2, 3, \ldots.$$

A random sample of size $n$ is taken from a population with a geometric distribution.

**a** Find the method-of-moments estimator for $p$.

**b** Find the MLE for $p$.

a. $\underline{\mathbb{E}Y = \frac{1}{p}}$. $\quad \mathbb{E}Y = \sum_{y=1}^{\infty} y \cdot p(1-p)^{y-1}$.

$\frac{1}{\hat{p}} = \bar{Y} \qquad \hat{p} = \frac{1}{\bar{Y}}$.

b. $L_n(p) = \prod_{i=1}^{n} p(1-p)^{y_i - 1}$.

$\ln(p) = n \ln p + (\Sigma y_i - n) \ln(1-p)$.

$\frac{d \ln(p)}{dp} = 0 \iff \frac{n}{p} - \frac{\Sigma y_i - n}{1-p} = 0$

$(1-p)n - p(\Sigma y_i - n) = 0$.

$(\Sigma y_i) p = n$

$\therefore \hat{p} = \frac{n}{\Sigma y_i} = \frac{1}{\bar{Y}}$.

In general, $\hat{\theta}_{MOM} \neq \hat{\theta}_{MLE}$,

moreover, $\hat{\theta}_{MLE}$ is preferred.

(Additional part*)

Why MLE?

Enjoy nice property, under some mild conditions

① Consistency

② Asymptotic normality.

③ Asympotic efficiency.

Cramér - Theorem.

$\sqrt{n}(\hat{\theta}-\theta) \xrightarrow{d} N(0, I^{-1}(\theta))$. $I(\theta)$ is Fisher information,

i.e. $I(\theta) = \mathbb{E}\left[-\frac{\partial^2 p(X;\theta)}{\partial \theta^2}\right]$, hence $\hat{\theta}$ achieves

Cramer-Rao lower bound asymptotically.