Nov. 3rd.

1. Law of large numbers

2. Point estimation

   a. consistency

   b. sufficiency

   c. efficiency.

Khinchin's WLLN:

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} X$, $EX = \mu < \infty$, $S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

Then $S_n \xrightarrow{P} \mu$.

Kolmogorov SLLN.

Under the same condition as above, $S_n \xrightarrow{a.s.} \mu$.

CLT $\quad \mathrm{Var}\, X = \sigma^2 < \infty$.

$$\frac{S_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1)$$

Consistency.

$\hat{\mu}_n$ be estimator of μ,

$P(|\hat{\mu}_n - \mu| > \varepsilon) \to 0$ as $n \to \infty$. i.e. $\hat{\mu}_n \xrightarrow{P} \mu$.

9.23 Refer to Exercise 9.21. Suppose that $Y_1, Y_2, \ldots, Y_n$ is a random sample of size n from a population for which the first four moments are finite. That is, $m_1' = E(Y_1) < \infty$, $m_2' = E(Y_1^2) < \infty$, $m_3' = E(Y_1^3) < \infty$, and $m_4' = E(Y_1^4) < \infty$. (*Note:* This assumption is valid for the normal and Poisson distributions in Exercises 9.21 and 9.22, respectively.) Again, assume that $n = 2k$ for some integer k. Consider

$$\hat{\sigma}^2 = \frac{1}{2k}\sum_{i=1}^{k}(Y_{2i} - Y_{2i-1})^2.$$

a Show that $\hat{\sigma}^2$ is an unbiased estimator for σ^2.

b Show that $\hat{\sigma}^2$ is a consistent estimator for σ^2.

c Why did you need the assumption that $m_4' = E(Y_1^4) < \infty$?

$X_i \overset{def}{=} Y_{2i} - Y_{2i-1}$, $i = 1, \cdots, k$, then X_i are i.i.d.

$\mathbb{E}X_i = \mathbb{E}(Y_{2i} - Y_{2i-1}) = 0$, $Var\, X_i = Var\, Y_{2i} + Var\, Y_{2i-1}$

$= 2\sigma^2$.

$\hat{\sigma}^2 = \frac{1}{2k}\sum_{i=1}^{k} X_i^2 = \frac{1}{2} \times \frac{1}{k}\sum_{i=1}^{k} X_i^2 \xrightarrow{P} \frac{1}{2} \times 2\sigma^2 = \sigma^2$.

$\mathbb{E}\,\hat{\sigma}^2 = \mathbb{E}\,\frac{1}{2k}\sum_{i=1}^{k} X_i^2 = \frac{1}{2k}\cdot k\,\mathbb{E}X_1^2 = \frac{1}{2k}\cdot k \cdot 2\sigma^2 = \sigma^2$.

9.32 Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from the probability density function

$$f(y) = \begin{cases} \dfrac{2}{y^2}, & y \geq 2, \\ 0, & \text{elsewhere.} \end{cases}$$

Does the law of large numbers apply to $\overline{Y}$ in this case? Why or why not?

No. $\mathbb{E}Y = \int_2^{\infty} y \cdot \frac{2}{y^2}\, dy$

$= \int_2^{\infty} \frac{2}{y}\, dy = 2\ln y \Big|_2^{\infty} = \infty$.

$MSE(\hat{\theta}) = Var\, \hat{\theta}$ $\quad$ $\hat{\theta}^* = \underset{\hat{\theta}}{\operatorname{argmin}}\, Var(\hat{\theta})$ MVUE complete UMVUE

Sufficiency.

Factorization theorem.

If $p(x|\theta) = h_\theta(T(X))\, g(X)$, where g doesn't involve θ,

then $T(X)$ is SS.

9.45 Suppose that $Y_1, Y_2, \ldots, Y_n$ is a random sample from a probability density function in the (one-parameter) exponential family so that

$$f(y \mid \theta) = \begin{cases} a(\theta)b(y)e^{-[c(\theta)d(y)]}, & a \le y \le b, \\ 0, & \text{elsewhere,} \end{cases}$$

where a and b do not depend on θ. Show that $\sum_{i=1}^{n} d(Y_i)$ is sufficient for θ.

$$f(Y_1,\ldots,Y_n|\theta) = \prod_{i=1}^{n} f(Y_i|\theta)$$
$$= a^n(\theta) \prod_{i=1}^{n} b(Y_i)\, e^{-c(\theta)\sum_{i=1}^{n} d(Y_i)}$$
$$= \left[a^n(\theta)\, e^{-c(\theta)\sum_{i=1}^{n} d(Y_i)}\right] \prod_{i=1}^{n} b(Y_i)$$
$$= h_\theta\left(\sum_{i=1}^{n} d(Y_i)\right) g(Y_1,\ldots,Y_n).$$

***9.51** Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from the probability density function

$$f(y \mid \theta) = \begin{cases} e^{-(y-\theta)}, & y \ge \theta, \\ 0, & \text{elsewhere.} \end{cases}$$

Show that $Y_{(1)} = \min(Y_1, Y_2, \ldots, Y_n)$ is sufficient for θ.

$$f(Y_1,\ldots,Y_n|\theta) = \prod_{i=1}^{n}\left(e^{-y_i}\cdot e^{\theta}\right)\cdot 1_{\{y_i \geq \theta\}}$$
$$= e^{-\sum_{i=1}^{n} y_i}\cdot e^{n\theta}\cdot 1_{\{y_1,\ldots,y_n \geq \theta\}}$$
$$= e^{-\sum_{i=1}^{n} y_i}\cdot e^{n\theta}\cdot 1_{\{Y_{(1)} \geq \theta\}}.$$

So $Y_{(1)}$ is SS.

Efficiency.

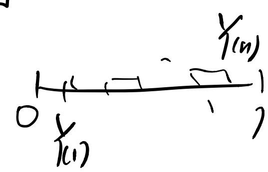
Restrict to unbiased estimators.

9.4 Let $Y_1, Y_2, \dots, Y_n$ denote a random sample of size n from a uniform distribution on the interval $(0, \theta)$. If $Y_{(1)} = \min(Y_1, Y_2, \dots, Y_n)$, the result of Exercise 8.18 is that $\hat{\theta}_1 = (n+1)Y_{(1)}$ is an unbiased estimator for θ. If $Y_{(n)} = \max(Y_1, Y_2, \dots, Y_n)$, the results of Example 9.1 imply that $\hat{\theta}_2 = [(n+1)/n]Y_{(n)}$ is another unbiased estimator for θ. Show that the efficiency of $\hat{\theta}_1$ to $\hat{\theta}_2$ is $1/n^2$. Notice that this implies that $\hat{\theta}_2$ is a markedly superior estimator.

$Y_{(1)} = \min(Y_1, \dots, Y_n)$, $\quad Y_1, \dots, Y_n \overset{i.i.d}{\sim} \mathcal{U}((0,\theta))$.

WLOG. $\boxed{\theta = 1}$.

$f_{Y_{(1)}}(y) = n(1-y)^{n-1}$.



$\mathbb{E}\, Y_{(1)} = \frac{1}{n+1}$. $\quad \hat{\theta}_1 = (n+1)\, Y_{(1)}$

$\mathbb{E}\, Y_{(1)}^2$

$\mathrm{Var}\, Y_{(1)} = \int_0^1 n y^2 (1-y)^{n-1}\, dy - (\mathbb{E}\, Y_{(1)})^2$

$1 - Y_1 \sim \mathcal{U}([0,1])$

$\vdots$

$1 - Y_n$

$= n \cdot \frac{\Gamma(3) \cdot \Gamma(n)}{\Gamma(n+3)} - \left(\frac{1}{n+1}\right)^2$.

$1 - Y_{(1)} \overset{d}{\sim} Y_{(n)}$

$= \mathrm{Var}\, Y_{(n)}$.

Most efficient one? $\underline{\text{MVUE}}$!

How to find it? Rao-Blackwell!

9.56 Refer to Exercise 9.38(b). Find an MVUE of σ^2.

9.38 Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a normal distribution with mean μ and variance σ^2.

a If μ is unknown and σ^2 is known, show that $\overline{Y}$ is sufficient for μ.

b If μ is known and σ^2 is unknown, show that $\underline{\sum_{i=1}^{n}(Y_i - \mu)^2}$ is sufficient for σ^2.

$\hat{\theta}$ is unbiased estimator of θ,

$T(X)$ is SS for θ.

$\hat{\theta}^* = \mathbb{E}[\hat{\theta} \mid T(X)]$ is MVUE.

$\hat{\theta}^* = f(T(X))$, $\mathbb{E}\hat{\theta}^* = \theta \Rightarrow \hat{\theta}^*$ is MVUE.

$T(Y) = \sum_{i=1}^{n}(Y_i - \mu)^2$ is SS for σ^2.

$$\frac{T(Y)}{\sigma^2} \sim \chi^2_{n-1}$$

$\mathbb{E}\,T(Y) = \sigma^2 \cdot \mathbb{E}\chi^2_{n-1} = \underline{(n-1)}\,\sigma^2$.

$\boxed{S_Y^2 = \frac{T(Y)}{n-1}}$ $\boxed{\mathbb{E}S_Y^2 = \sigma^2.}$ S_Y^2 is MVUE.