SOLUTIONS

Homework 06: Due 10/25 (completion based)

Stat061-F23
Prof Amanda Luby

---

1. Let's explore (through a few examples) the *efficiency* property of large-sample MLEs. Recall that the large-sample normal approximation for the MLE is $\hat{\theta}_{MLE} \sim N(\theta, \frac{1}{nI(\theta)})$.

   (a) Explain why the normal approximation for $\hat{\theta}_{MLE}$ implies that the MLE for large samples is efficient.
   (b) Confirm the normal sample approximation for the MLE of the binomial distribution. (There's an example in Notes04 that may be helpful).
   (c) In Homework01, you showed that the MLE for p in the *geometric distribution* is $\frac{1}{\bar{X}}$. Find the normal approximation of $\hat{p}_{MLE}$. Why is it useful to use the normal approximation instead of finding $V(\hat{p}_{MLE})$ directly in this case?
   (d) Also in Homework01, you showed that the MLE for β in the Pareto pdf $f_x = \frac{\beta}{x^{\beta+1}}$ is $\hat{\beta}_{MLE} = \frac{n}{\sum \ln x_i}$. Find the approximate variance of $\hat{\beta}_{MLE}$.

2. Suppose we have an unbiased estimator $\hat{\theta}$. Explain how the Rao-Blackwell theorem, taken together with the Cramer-Rao Lower Bound, implies that an estimator must be *sufficient* before it can be *efficient*.

Delta Method (again)

A less general, but perhaps more useful, version of the delta method is:
Suppose that $\frac{\sqrt{n}(Y_n - \mu)}{\sigma} \to_d N(0,1)$ and suppose that g is a differentiable function with $g'(\mu) \neq 0$. Then,

$$\frac{\sqrt{n}(g(Y_n) - g(\mu))}{|g'(\mu)|\sigma} \to_d N(0,1).$$

Stated in another way, if $Y_n \approx N(\mu, \frac{\sigma^2}{n})$ then $g(Y_n) \approx N(g(\mu), (g'(\mu))^2 \frac{\sigma^2}{n})$.

3. Suppose that $X_1, ..., X_n \sim N(0, \sigma^2)$.

   (a) Determine the asymptotic distribution of the statistic $T = \frac{1}{\frac{1}{n}\sum X_i^2}$.
   (b) Find a variance stabilizing transformation for the statistic $T^{-1} = \frac{1}{n}\sum X_i^2$.

1. Let's explore (through a few examples) the *efficiency* property of large-sample MLEs. Recall that the large-sample normal approximation for the MLE is $\hat{\theta}_{MLE} \sim N(\theta, \frac{1}{nI(\theta)})$.

   (a) Explain why the normal approximation for $\hat{\theta}_{MLE}$ implies that the MLE for large samples is efficient.

   (b) Confirm the normal sample approximation for the MLE of the binomial distribution. (There's an example in Notes04 that may be helpful).

   (c) In Homework01, you showed that the MLE for p in the *geometric distribution* is $\frac{1}{\bar{X}}$. Find the normal approximation of $\hat{p}_{MLE}$. Why is it useful to use the normal approximation instead of finding $V(\hat{p}_{MLE})$ directly in this case?

   (d) Also in Homework01, you showed that the MLE for β in the Pareto pdf $f_x = \frac{\beta}{x^{\beta+1}}$ is $\hat{\beta}_{MLE} = \frac{n}{\sum \ln x_i}$. Find the approximate variance of $\hat{\beta}_{MLE}$.

(a) The CRLB tells us the MVUE has variance $\frac{1}{nI(\theta)}$, which is the asymptotic variance of $\hat{\theta}_{MLE}$ by the formula. Since $\underbrace{E(\hat{\theta}_{MLE}) = \theta}_{\text{for large samples}}$, $\hat{\theta}_{MLE}$ is unbiased for large samples. So CRLB applies.

(b) Recall $I(p) = \frac{1}{p(1-p)}$ from Notes 04.

Approximation: $\bar{x} \sim N(p, \frac{p(1-p)}{n})$

$E(\bar{x}) = \frac{1}{n} \sum E(X_i) = \frac{1}{n} \cdot n \cdot p = p$ ✓

$V(\bar{x}) = \frac{1}{n^2} \sum V(X_i) = \frac{1}{n^2} \cdot n \cdot p(1-p) = \frac{p(1-p)}{n}$ ✓

(c) $I(p) = -E\left[\frac{\partial^2}{\partial p^2} \cdot \ln f_x\right] \longrightarrow f_x = (1-p)^{y-1} p$

$\ell(p) = (y-1)\ln(1-p) + \ln p$

$\frac{\partial \ell}{\partial p} = \frac{y-1}{1-p} \cdot -1 + \frac{1}{p}$

$= \frac{1-y}{1-p} + \frac{1}{p}$

$\frac{\partial^2 \ell}{\partial p^2} = \frac{(1-y)}{(1-p)^2} \cdot -1 - \frac{1}{p^2}$

$= \frac{y-1}{(1-p)^2} - \frac{1}{p^2}$

$= -E\left[\frac{y-1}{(1-p)^2} - \frac{1}{p^2}\right]$

$= -\frac{1}{(1-p)^2}\left[E(y) - 1\right] + \frac{1}{p^2}$

$= \frac{1}{(1-p)^2}\left[\frac{1}{p} - 1\right] + \frac{1}{p^2}$

$= \frac{(1-p)}{p(1-p)^2} + \frac{1}{p^2}$

$= \frac{1}{p(1-p)} + \frac{1}{p^2}$

$\hat{p}_{MLE} \sim N(p, \frac{1}{nI(p)})$

$\sim N\left(p, \frac{1}{n\left(\frac{1}{p(1-p)} + 1/p^2\right)}\right)$

Since $V(\frac{1}{\bar{x}}) \neq \frac{1}{V(\bar{x})}$, would be hard to find $V(\frac{1}{\bar{x}})$ directly.

(d) Also in Homework01, you showed that the MLE for β in the Pareto pdf $f_x = \frac{\beta}{x^{\beta+1}}$ is $\hat{\beta}_{MLE} = \frac{n}{\sum \ln x_i}$. Find the approximate variance of $\hat{\beta}_{MLE}$.

$$f_x = \frac{\beta}{x^{\beta+1}}$$

$$l(\beta) = \ln \beta - (\beta+1) \ln x = \ln \beta - \beta \ln x - \ln x$$

$$\frac{\partial l}{\partial \beta} = \frac{1}{\beta} - \ln x$$

$$\frac{\partial^2 l}{\partial \beta^2} = \frac{-1}{\beta^2}$$

$$I(\beta) = -E\left(\frac{\partial^2 l}{\partial \beta^2}\right) = -E\left(-\frac{1}{\beta^2}\right) = \frac{1}{\beta^2} E(1) = \frac{1}{\beta^2}$$

$$\hat{\beta}_{MLE} \,\dot\sim\, N\left(\beta, \frac{1}{n(1/\beta^2)}\right) = N\left(\beta, \frac{\beta^2}{n}\right)$$

2. Suppose we have an unbiased estimator $\hat{\theta}$. Explain how the Rao-Blackwell theorem, taken together with the Cramer-Rao Lower Bound, implies that an estimator must be *sufficient* before it can be *efficient.*

From Rao-Blackwell:

$\theta^* = E(\hat{\theta} \mid T=t)$, where $\hat{\theta}$ is any estimator and T is sufficient. If $\hat{\theta}$ is not a function of T,

$$MSE(\theta^*, \theta) < MSE(\hat{\theta}, \theta) \leftarrow \text{strict if } \hat{\theta} \neq f(T)$$

$$V(\theta^*) + bias(\theta^*, \theta)^2 < V(\hat{\theta}) + bias(\hat{\theta}, \theta)^2$$

if $\hat{\theta}$ is unbiased, $MSE(\hat{\theta}, \theta) = V(\hat{\theta})$

$$\Longrightarrow V(\theta^*) + bias(\theta^*, \theta)^2 < V(\hat{\theta})$$

But, $E(\theta^*) = E[E(\hat{\theta} \mid T=t)] = E(\hat{\theta})$] Law of total Expectation from Stat 51. See Blitzstein & Hwang ch 9

So θ^* is also unbiased. So

$$V(\theta^*) < V(\hat{\theta})$$

So θ^* is unbiased and has smaller variance than $\hat{\theta}$, so $\hat{\theta}$ cannot be MVUE and therefore can't meet CRLB and be efficient. θ^* could be, so only have a chance if T is sufficient (since otherwise inequality wouldn't be strict)

3. Suppose that $X_1, \ldots, X_n \sim N(0, \sigma^2)$.

(a) Determine the asymptotic distribution of the statistic $T = \frac{1}{\frac{1}{n}\sum X_i^2}$.

(b) Find a variance stabilizing transformation for the statistic $T^{-1} = \frac{1}{n}\sum X_i^2$.

(a) Let $Y_n = \frac{1}{n}\sum X_i^2$. Note that $E(X_i)=0$, can find w/ MGF

$E(X_i^2) = V(X_i) + E(X_i)^2 = \sigma^2$

$V(X_i^2) = E(X_i^4) - E(X_i)^4 = E(X_i^4) = 2\sigma^4$

Then, by CLT, $Y_n \,\dot\sim\, N\left(\sigma^2, \frac{2\sigma^4}{n}\right)$

Then, let $g(x) = \frac{1}{x}$. So $g'(x) = \frac{-1}{x^2}$ By delta method,

$$T = g(Y_n) \approx N\left(g(\sigma^2), (g'(\sigma^2))^2 \frac{2\sigma^4}{n}\right)$$
$$\approx N\left(1/\sigma^2, \frac{1}{\sigma^8} \cdot \frac{2\sigma^4}{n}\right)$$
$$\approx N\left(1/\sigma^2, 2/n\sigma^4\right)$$

(b) From above, $T^{-1} \,\dot\sim\, N\left(\sigma^2, \frac{2\sigma^4}{n}\right)$ want to "undo" the σ^4 in the asymptotic variance in order to stabilize

$\Rightarrow$ want to find $h(x)$ such that $[h'(\sigma^2)]^2 = \frac{1}{2\sigma^4}$

$$h'(\sigma^2) = \frac{1}{\sqrt{2}\sigma^2}$$
$$\int \frac{1}{\sqrt{2}\sigma^2}\, d\sigma^2 = \frac{1}{\sqrt{2}}\log(\sigma^2)$$
$$= \frac{2}{\sqrt{2}}\log\sigma$$

let $h(x) = \frac{1}{\sqrt{2}}\log(x)$.

Then $h(Y_n) \approx N\left(\sqrt{2}\log\sigma, \frac{1}{2\sigma^4} \cdot \frac{2}{n\sigma^4}\right)$

$$\frac{1}{\sqrt{2}}\log\frac{1}{n}\sum X_i^2 \approx N\left(\sqrt{2}\log\sigma, \frac{1}{n}\right)$$

doesn't depend on σ, so 'stabilized'