Biostatistics 602 - Statistical Inference
Lecture 26
Final Exam Review & Practice Problems for the Final

Hyun Min Kang

April 23rd, 2013

Review of the second half

Rao-Blackwell : If $W(\mathbf{X})$ is an unbiased estimator of $\tau(\theta)$, $\phi(T) = E[W(\mathbf{X})|T]$ is a better unbiased estimator for a sufficient statistic.

Uniqueness of MVUE : Theorem 7.3.19 - Best unbiased estimator is unique

MVUE and UE of zeros : Theorem 7.3.20 - Best unbiased estimator is uncorrelated with any unbiased estimators of zero

UMVE by complete sufficient statistics : Theorem 7.3.23 - Any function of complete sufficient statistic is the best unbiased estimator for its expected value

How to get UMVUE Strategies to obtain best unbiased estimators:
- Condition a simple unbiased estimator on complete sufficient statistics
- Come up with a function of sufficient statistic whose expected value is $\tau(\theta)$.

Bayesian Framework

Prior distribution $\pi(\theta)$

Sampling distribution $\mathbf{x}|\theta \sim f_{\mathbf{X}}(\mathbf{x}|\theta)$

Joint distribution $\pi(\theta)f(\mathbf{x}|\theta)$

Marginal distribution $m(\mathbf{x}) = \int \pi(\theta)f(\mathbf{x}|\theta)d\theta$

Posterior distribution $\pi(\theta|\mathbf{x}) = \frac{f_{\mathbf{X}}(\mathbf{x}|\theta)\pi(\theta)}{m(\mathbf{x})}$

Bayes Estimator is a posterior mean of θ : $E[\theta|\mathbf{x}]$.

Bayesian Decision Theory

Loss Function $L(\theta, \hat{\theta})$ (e.g. $(\theta - \hat{\theta})^2$)

Risk Function is the average loss : $R(\theta, \hat{\theta}) = \mathrm{E}[L(\theta, \hat{\theta})|\theta]$. For squared error loss $L = (\theta - \hat{\theta})^2$, the risk function is MSE

Bayes Risk is the average risk across all θ : $\mathrm{E}[R(\theta, \hat{\theta})|\pi(\theta)]$.

Bayes Rule Estimator minimizes Bayes risk $\iff$ minimizes posterior expected loss.

Asymptotics

Consistency: Using law of large numbers, show variance and bias converges to zero, for any continuous mapping function τ

Asymptotic Normality: Using central limit theorem, Slutsky Theorem, and Delta Method

Asymptotic Relative Efficiency: $\text{ARE}(V_n, W_n) = \sigma_W^2/\sigma_V^2$.

Asymptotically Efficient: ARE with CR-bound of unbiased estimator of $\tau(\theta)$ is 1.

Asymptotic Efficiency of MLE: Theorem 10.1.12 MLE is always asymptotically efficient under regularity condition.

Hypothesis Testing

Type I error: $\Pr(\mathbf{X} \in R|\theta)$ when $\theta \in \Omega_0$

Type II error: $1 - \Pr(\mathbf{X} \in R|\theta)$ when $\theta \in \Omega_0^c$

Power function: $\beta(\theta) = \Pr(\mathbf{X} \in R|\theta)$
$\beta(\theta)$ represents Type I error under H_0, and power (=1-Type II error) under H_1.

Size α test: $\sup_{\theta \in \Omega_0} \beta(\theta) = \alpha$

Level α test: $\sup_{\theta \in \Omega_0} \beta(\theta) \leq \alpha$

LRT: $\lambda(\mathbf{x}) = \frac{L(\hat{\theta}_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})}$ rejects H_0 when $\lambda(\mathbf{x}) \leq c$
$\iff -2\log\lambda(\mathbf{x}) \geq -2\log c = c^*$

LRT based on sufficient statistics: LRT based on full data and sufficient statistics are identical.

UMP

Unbiased Test: $\beta(\theta_1) \geq \beta(\theta_0)$ for every $\theta_1 \in \Omega_0^c$ and $\theta_0 \in \Omega_0$.

UMP Test: $\beta(\theta) \geq \beta'(\theta)$ for every $\theta \in \Omega_0^c$ and $\beta'(\theta)$ of every other test with a class of test $\mathcal{C}$.

UMP level α Test: UMP test in the class of all the level α test. (smallest Type II error given the upper bound of Type I error)

Neyman-Pearson: For $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$, a test with rejection region $f(\mathbf{x}|\theta_1)/f(\mathbf{x}|\theta_0) > k$ is a UMP level α test for its size.

MLR: $g(t|\theta_2)/g(t|\theta_1)$ is an increasing function of t for every $\theta_2 > \theta_1$.

Karlin-Rabin: If T is sufficient and has MLR, then test rejecting $R = \{T : T > t_0\}$ or $R = \{T : T < t_0\}$ is an UMP level α test for one-sided composite hypothesis.

Asymptotic Tests and p-Values

Asymptotic Distribution of LRT: For testing, $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$, $-2\log\lambda(\mathbf{x}) \xrightarrow{d} \chi_1^2$ under regularity condition.

Wald Test: If W_n is a consistent estimator of θ, and S_n^2 is a consistent estimator of $\text{Var}(W_n)$, then $Z_n = (W_n - \theta_0)/S_n$ follows a standard normal distribution

- Two-sided test : $|Z_n| > z_{\alpha/2}$
- One-sided test : $Z_n > z_{\alpha/2}$ or $Z_n < -z_{\alpha/2}$

p-Value: A p-value $0 \leq p(\mathbf{x}) \leq 1$ is valid if, $\Pr(p(\mathbf{X}) \leq \alpha|\theta) \leq \alpha$ for every $\theta \in \Omega_0$ and $0 \leq \alpha \leq 1$.

Constructing p-Value: Theorem 8.3.27 : If large $W(\mathbf{X})$ value gives evidence that H_1 is true, $p(\mathbf{x}) = \sup_{\theta \in \Omega_0} \Pr(W(\mathbf{X}) \geq W(\mathbf{x})|\theta)$ is a valid p-value

p-Value given sufficient statistics: For a sufficient statistic $S(\mathbf{X})$, $p(\mathbf{x}) = \Pr(W(\mathbf{X}) \geq W(\mathbf{x})|S(\mathbf{X}) = S(\mathbf{x}))$ is also a valid p-value.

Interval Estimation

Coverage probability $\Pr(\theta \in [L(\mathbf{X}), U(\mathbf{X})])$

Coverage coefficient is $1-\alpha$ if $\inf_{\theta\in\Omega} \Pr(\theta \in [L(\mathbf{X}), U(\mathbf{X})]) = 1-\alpha$

Confidence interval $[L(\mathbf{X}), U(\mathbf{X})])$ is $1-\alpha$ if $\inf_{\theta\in\Omega} \Pr(\theta \in [L(\mathbf{X}), U(\mathbf{X})]) = 1-\alpha$

Inverting a level α test If $A(\theta_0)$ is the acceptance region of a level α test, then $C(\mathbf{X}) = \{\theta : \mathbf{X} \in A(\theta)\}$ is a $1-\alpha$ confidence set (or interval).

Practice Problem 1 (continued from last week)

Problem

Let $f(x|\theta)$ be the logistic location pdf

$$f(x|\theta) = \frac{e^{(x-\theta)}}{(1+e^{(x-\theta)})^2} \qquad -\infty < x < \infty,\ -\infty < \theta < \infty$$

(a) Show that this family has an MLR

(b) Based on one observation X, find the most powerful size α test of $H_0 : \theta = 0$ versus $H_1 : \theta = 1$.

(c) Show that the test in part (b) is UMP size α for testing $H_0 : \theta \leq 0$ vs. $H_1 : \theta > 0$.

Solution for (a)

For $\theta_1 < \theta_2$,

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = \frac{\frac{e^{(x-\theta_2)}}{(1+e^{(x-\theta_2)})^2}}{\frac{e^{(x-\theta_1)}}{(1+e^{(x-\theta_1)})^2}} = e^{(\theta_1-\theta_2)}\left(\frac{1+e^{(x-\theta_1)}}{1+e^{(x-\theta_2)}}\right)^2$$

Let $r(x) = (1+e^{x-\theta_1})/(1+e^{x-\theta_2})$

$$r'(x) = \frac{e^{(x-\theta_1)}(1+e^{(x-\theta_2)}) - (1+e^{(x-\theta_1)})e^{(x-\theta_2)}}{(1+e^{(x-\theta_2)})^2} = \frac{e^{(x-\theta_1)} - e^{(x-\theta_2)}}{(1+e^{(x-\theta_2)})^2} > 0 \qquad (\because x-\theta_1 > x-\theta_2)$$

Therefore, the family of X has an MLR.

Solution for (b)

The UMP test rejects H_0 if and only if

$$\frac{f(x|1)}{f(x|0)} = e\left(\frac{1+e^x}{1+e^{(x-1)}}\right)^2 > k$$

$$\frac{1+e^x}{1+e^{(x-1)}} > k^*$$

$$\frac{1+e^x}{e+e^x} > k^{**}$$

$$X > x_0$$

Because under H_0, $F(x_0|\theta=0) = \frac{e^x}{1+e^x}$, the rejection region of UMP level α test satisfies

$$1 - F(x|\theta=0) = \frac{1}{1+e^{x_0}} = \alpha$$

$$x_0 \sim \log\left(\frac{1-\alpha}{\alpha}\right)$$

Solution for (c)

Because the family of X has an MLR, UMP size α for testing $H_0 : \theta \leq 0$ vs. $H_1 : \theta > 0$ should be a form of

$$\begin{aligned} X &> x_0 \\ \Pr(X > x_0 | \theta = 0) &= \alpha \end{aligned}$$

Therefore, $x_0 = \log\left(\frac{1-\alpha}{\alpha}\right)$, which is identical to the test defined in (b).

Practice Problem 2

Problem

Suppose $X_1, \cdots, X_n$ are iid random samples with pdf $f_X(x|\theta) = \theta\exp(-\theta x)$, where $x \geq 0, \theta > 0$

(a) Show that $\frac{n}{\sum_{x=1}^n X_i}$ is a consistent estimator for θ.

(b) Show that $\frac{n}{\sum_{x=1}^n X_i}$ is asymptotically normal and derive its asymptotic distribution

(c) Derive the Wald asymptotic size α test for $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$.

(d) Find an asymptotic $(1-\alpha)$ confidence interval for θ by inverting the above test

You may use the fact that $\mathrm{E}X = 1/\theta$ and $\mathrm{Var}(X) = 1/\theta^2$.

Solution (a) - Consistency

1. Obtain $\mathrm{E}X = 1/\theta$ (Derive yourself if not given)

$$\begin{aligned} \mathrm{E}X &= \int_0^\infty x f(x|\theta)\,dx = \int_0^\infty \theta x \exp(-\theta x)\,dx \\ &= [-x\exp(-\theta x)]_0^\infty + \int_0^\infty \exp(-\theta x)\,dx \\ &= 0 + \left[-\frac{1}{\theta}\exp(-\theta x)\right]_0^\infty = \frac{1}{\theta} \end{aligned}$$

2. By LLN (Law of Large Number), $\overline{X} \xrightarrow{P} \mathrm{E}X = 1/\theta$.
3. By Theorem of continuous map, $n / \sum_{i=1}^n X_i = 1/\overline{X} \xrightarrow{P} \theta$.

Solution (b) - Asymptotic Distribution

1. Obtain $\mathrm{Var}(X) = 1/\theta^2$ (Derive if needed, omitted here).
2. Apply CLT(Central Limit Theorem),

$$\overline{X} \sim \mathcal{AN}\left(\frac{1}{\theta}, \frac{1}{\theta^2 n}\right)$$

3. Apply Delta method. Let $g(y) = 1/y$, then $g'(y) = -1/y^2$.

$$\begin{aligned} \frac{\sum X_i}{n} &= 1/\overline{X} = g(\overline{X}) \sim \mathcal{AN}\left(g(1/\theta), \frac{[g'(1/\theta)]^2}{\theta^2 n}\right) \\ &= \mathcal{AN}\left(\theta, \frac{\theta^2}{n}\right) \\ \Longleftrightarrow \sqrt{n}\left(\frac{1}{\overline{X}} - \theta\right) &= \mathcal{N}(0, \theta^2) \end{aligned}$$

Solution (c) - Wald asymptotic size α test

1. Obtain a consistent estimator of θ :

$$W(\mathbf{X}) = \frac{\sum_{i=1}^n X_i}{n} \sim \mathcal{AN}\left(\theta, \frac{\theta^2}{n}\right)$$

2. Obtain a constant estimator of $\text{Var}(W)$

$$\frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2 \xrightarrow{\text{P}} \text{Var}(\mathbf{X}) = \frac{1}{\theta^2} \quad \text{(CLT)}$$

$$\frac{n-1}{\sum_{i=1}^n (X_i - \overline{X})^2} \xrightarrow{\text{P}} \theta^2 \quad \text{(Continuous Map Theorem).}$$

$$S^2 = \frac{n}{\sum_{i=1}^n (X_i - \overline{X})^2} \xrightarrow{\text{P}} \theta^2 \quad \text{(Slutsky's Theorem).}$$

Solution (c) - Wald Asymptotic size α test (cont'd)

3. Construct a two-sided asymptotic size α Wald test, whose rejection region is

$$\begin{aligned} |Z(\mathbf{X})| &= \left|\frac{W(\mathbf{X}) - \theta_0}{S/\sqrt{n}}\right| \\ &= \left|\frac{\frac{n}{\sum_{i=1}^n X_i} - \theta_0}{\frac{1}{\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2}}}\right| \\ &= \left|\frac{1}{\overline{X}} - \theta_0\right| \sqrt{\sum_{i=1}^n (X_i - \overline{X})^2} \geq z_{\alpha/2} \end{aligned}$$

Solution (d) - Asymptotic $1 - \alpha$ confidence interval

The acceptance region is

$$A = \left\{\mathbf{x} : \left|\frac{1}{\overline{x}} - \theta_0\right| \sqrt{\sum_{i=1}^n (x_i - \overline{x})^2} \leq z_{\alpha/2}\right\}$$

By inverting the acceptance region, the confidence interval is

$$C(\mathbf{X}) = \left\{\theta : \left|\frac{1}{\overline{X}} - \theta\right| \sqrt{\sum_{i=1}^n (X_i - \overline{X})^2} \leq z_{\alpha/2}\right\}$$

which is equivalent to

$$C(\mathbf{X}) = \left\{\theta \in \left[\frac{1}{\overline{X}} - \frac{z_{\alpha/2}}{\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2}}, \frac{1}{\overline{X}} + \frac{z_{\alpha/2}}{\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2}}\right]\right\}$$

Practice Problem 3

Problem

The independent random variables $X_1, \cdots, X_n$ have the following pdf

$$f(x|\theta, \beta) = \frac{\beta x^{\beta-1}}{\theta^\beta} \quad 0 < x < \theta, \ \beta > 0$$

1. Find the MLEs of β and θ
2. When β is a known constant β_0, construct a LRT testing $H_0 : \theta \geq \theta_0$ vs. $H_1 : \theta < \theta_0$.
3. When β is a known constant β_0, find the upper confidence limit for θ with confidence coefficient $1 - \alpha$.

(a) - MLE

$$L(\theta, \beta|\mathbf{x}) = \frac{\beta^n \left(\prod_{i=1}^n x_i\right)^{\beta-1}}{\theta^{n\beta}} I(x_{(n)} \leq \theta)$$

Because L is a decreasing function of θ and positive only when $\theta \geq x_{(n)}$

$$\begin{aligned} \hat{\theta} &= x_{(n)} \\ l(\theta, \beta|\mathbf{x}) &= n \log \beta + (\beta - 1) \sum \log x_i - n\beta \log \theta \\ \frac{\partial l}{\partial \beta} &= \frac{n}{\beta} + \sum \log x_i - n \log \theta = 0 \\ \hat{\beta} &= \frac{n}{n \log \hat{\theta} - \sum \log x_i} \\ &= \frac{n}{n x_{(n)} - \sum \log x_i} \end{aligned}$$

(b) - LRT

$$\begin{aligned} \lambda(\mathbf{x}) &= \frac{\sup_{\theta \in \Omega_0} L(\hat{\theta}|\mathbf{x})}{\sup_{\theta \in \Omega} L(\hat{\theta}|\mathbf{x})} \\ &= \begin{cases} 1 & \theta_0 < x_{(n)} \\ \frac{L(\theta_0|\mathbf{x})}{L(x_{(n)}|\mathbf{x})} & \theta_0 \geq x_{(n)} \end{cases} \\ &= \begin{cases} 1 & \theta_0 < x_{(n)} \\ \frac{\left(x_{(n)}\right)^{n\beta_0}}{\theta_0^{n\beta_0}} & \theta_0 \geq x_{(n)} \end{cases} \leq c \\ \frac{x_{(n)}}{\theta_0} &\leq c^* \end{aligned}$$

(b) - size α LRT

$$\begin{aligned} \alpha &= \Pr\left(\frac{x_{(n)}}{\theta_0} \leq c^*\right) \\ &= (c^*)^{n\beta_0} \\ c^* &= \alpha^{\frac{1}{n\beta_0}} \end{aligned}$$

Therefore, the rejection region for size α LRT is is

$$R = \left\{\mathbf{x} : x_{(n)} \leq \theta_0 \alpha^{\frac{1}{n\beta_0}}\right\}$$

(c) - Upper $1 - \alpha$ confidence limit

The acceptance region of size α LRT is

$$A(\theta_0) = \left\{\mathbf{x} : x_{(n)} > \theta_0 \alpha^{\frac{1}{n\beta_0}}\right\}$$

By inserting the acceptance region, the $1 - \alpha$ confidence interval becomes

$$\begin{aligned} C(\mathbf{X}) &= \left\{\theta : X_{(n)} > \theta \alpha^{\frac{1}{n\beta_0}}\right\} \\ &= \left\{\theta : \theta < X_{(n)} \alpha^{-\frac{1}{n\beta_0}}\right\} \end{aligned}$$

Therefore, the upper $1 - \alpha$ confidence limit is $X_{(n)} \alpha^{-\frac{1}{n\beta_0}}$.

Practice Problem 4

Problem

A random sample $X_1, \cdots, X_n$ is drawn from a population $\mathcal{N}(\theta, \theta)$ where $\theta > 0$.

(a) Find the $\hat{\theta}$, the MLE of θ

(b) Find the asymptotic distribution of $\hat{\theta}$.

(c) Compute $\text{ARE}(\hat{\theta}, \overline{X})$. Determine whether $\hat{\theta}$ is asymptotically more efficient than $\overline{X}$ or not.

You may use the following fact: $\text{Var}(X^2) = 4\theta^3 + 2\theta^2$.

(a) - MLE of θ

$$
\begin{aligned}
L(\theta|\mathbf{x}) &= (2\pi\theta)^{n/2} \exp\left[-\frac{\sum_{i=1}^n (x_i - \theta)^2}{2\theta}\right] \\
l(\theta|\mathbf{x}) &= \frac{n}{2}\log(2\pi) + \frac{n}{2}\log\theta - \frac{\sum_{i=1}^n (x_i-\theta)^2}{2\theta} \\
&= \frac{n}{2}\log(2\pi) + \frac{n}{2}\log\theta - \frac{\sum x_i^2}{2\theta} + \sum x_i - \frac{n\theta}{2} \\
l'(\theta|\mathbf{x}) &= \frac{n}{2\theta} + \frac{\sum x_i^2}{2\theta^2} - \frac{n}{2} = \frac{n\theta - \sum x_i^2 - n\theta^2}{2\theta^2} = 0 \\
n\theta^2 + n\theta - \sum x_i^2 &= 0 \\
\hat{\theta} &= \frac{-1 + \sqrt{1 + 4\sum x_i^2/n}}{2} \\
\frac{1}{n}\sum x_i^2 &= \hat{\theta}^2 + \hat{\theta}
\end{aligned}
$$

(b) - Asymptotic distribution of MLE

By CLT, Let $W = \frac{1}{n}\sum X_i^2$, then

$$
W \sim \mathcal{AN}\left(\mathrm{E}X^2, \frac{\text{Var}(X^2)}{n}\right) = \mathcal{AN}\left(\theta + \theta^2, \frac{4\theta^3 + 2\theta^2}{n}\right)
$$

The asymptotic distribution of MLE $\hat{\theta}$

$$
\hat{\theta} \sim \mathcal{AN}\left(\theta, \frac{\sigma^2(\theta)}{n}\right)
$$

for some function $\sigma^2(\theta)$ and we would like to find $\sigma^2(\theta)$ using the asymptotic distribution of W.

(b) - Asymptotic distribution of MLE (cont'd)

Let $g(y) = y^2 + y$, then $g'(y) = (2y+1)$ and $g(\hat{\theta}) = W$. Then by the Delta Method, the asymptotic distribution of W can be written as

$$
\begin{aligned}
W = g(\hat{\theta}) &\sim \mathcal{AN}\left(g(\theta), g'(\theta)\frac{\sigma^2(\theta)}{n}\right) \\
&= \mathcal{AN}\left(\theta^2 + \theta, \frac{(2\theta+1)^2\sigma^2(\theta)}{n}\right) \\
&= \mathcal{AN}\left(\theta^2 + \theta, \frac{4\theta^3 + 2\theta^2}{n}\right) \\
\sigma^2(\theta) &= \frac{4\theta^3 + 2\theta^2}{(2\theta+1)^2} = \frac{2\theta^2(2\theta+1)}{(2\theta+1)^2} = \frac{2\theta^2}{2\theta+1}
\end{aligned}
$$

(b) - Asymptotic distribution of MLE (cont'd)

The asymptotic distribution of MLE $\hat{\theta}$

$$\begin{aligned}\hat{\theta} &\sim \mathcal{AN}\left(\theta, \frac{\sigma^2(\theta)}{n}\right) \\ &= \mathcal{AN}\left(\theta, \frac{2\theta^2}{n(2\theta+1)}\right)\end{aligned}$$

Note that you cannot use CR-bound for the asymptotic variance of MLE because the regularity condition does not hold (open set criteria).

(c) - ARE of MLE compared to $\overline{X}$

By CLT, the asymptotic distribution of $\overline{X}$ is

$$\overline{X} \sim \mathcal{AN}\left(\theta, \frac{\theta}{n}\right)$$

Then, $\text{ARE}(\hat{\theta}, \overline{X})$ is

$$\begin{aligned}\text{ARE}(\hat{\theta}, \overline{X}) &= \frac{\theta}{\frac{2\theta^2}{2\theta+1}} \\ &= \frac{2\theta+1}{2\theta} = 1 + \frac{1}{2\theta} > 1\end{aligned}$$

Therefore, $\hat{\theta}$ is more efficient estimator than $\overline{X}$.

Wrapping Up

1. Many thanks for your attentions and feedbacks.
2. Please complete your teaching evaluations, which will be very helpful for further improvement in the next year.
3. Final exam will be Thursday April 25th, 4:00-6:00pm.
4. The last office hour will be held Wednesday April 24th, 4:00-5:00pm.
5. The grade will be posted during the weekend.
6. Don't forget the materials we have learned, because they are the key topics for your candidacy exam.