Final Exam

MATH 60062/70062: Mathematical Statistics II

May 5, 2022

- Please turn off your phone.
- Print your name clearly at the top of this page.
- This is a closed-book and closed-notes exam.
- This exam contains 4 questions. There are 100 points in total.
- You have 75 minutes to complete the exam.
- Please show your work and explain all of your reasoning.
- You must work by yourself. Do not communicate in any way with others.

1. (15 points) Give full definitions for the following concepts:

   a. Coverage probability

   b. Confidence coefficient

   c. Pivotal quantity

   d. Consistent estimator

   e. Asymptotic relative efficiency

   Solution:

   a. For an interval estimator $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ of a parameter $\theta \in \Theta$, the coverage probability of the interval is
   $$P_\theta(L(\boldsymbol{X}) \le \theta \le U(\boldsymbol{X})).$$

   b. For an interval estimator $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ of a parameter $\theta \in \Theta$, the confidence coefficient of the interval is
   $$\inf_{\theta \in \Theta} P_\theta(L(\boldsymbol{X}) \le \theta \le U(\boldsymbol{X})).$$

   c. A random variable $Q = Q(\boldsymbol{X}, \theta)$ is a pivotal quantity (or pivot) if the distribution of Q is independent of θ.

   d. A sequence of estimators W_n is consistent for a parameter θ if
   $$W_n \xrightarrow{p} \theta \text{ for all } \theta \in \Theta.$$
   That is, for all $\epsilon > 0$ and for all $\theta \in \Theta$,
   $$\lim_{n\to\infty} P_\theta(|W_n - \theta| \ge \epsilon) = 0,$$
   or equivalently,
   $$\lim_{n\to\infty} P_\theta(|W_n - \theta| < \epsilon) = 1.$$

   e. Suppose we have two competing sequences of estimators denoted by W_n and V_n that satisfy
   $$\sqrt{n}(W_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma_W^2)$$
   $$\sqrt{n}(V_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma_V^2).$$
   Both estimators are consistent estimators of θ. The asymptotic relative efficiency (ARE) is defined as
   $$\text{ARE}(W_n \text{ to } V_n) = \frac{\sigma_W^2}{\sigma_V^2}.$$

2. (35 points) Suppose that $X_1, \ldots, X_n$ are iid $\mathcal{N}(\mu, \sigma^2)$, where $-\infty < \mu < \infty$ and $\sigma^2 > 0$. Both parameters are unknown. Consider testing

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu \neq \mu_0.$$

Let $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$, where $\bar{X}$ is the sample mean. The size α one-sample two-sided t-test rejects H_0 when

$$|\bar{x} - \mu_0| \geq t_{n-1,\alpha/2}\sqrt{s^2/n}.$$

a. (20 points) Show that the test can be derived as a likelihood ratio test.

b. (15 points) Find a $1-\alpha$ confidence set for μ by inverting the two-sided t-test.

Solution:

a. Set $\boldsymbol{\theta} = (\mu, \sigma^2)$. The likelihood function is

$$L(\boldsymbol{\theta} \mid \boldsymbol{x}) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2},$$

and the log-likelihood function is

$$\log L(\boldsymbol{\theta} \mid \boldsymbol{x}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu)^2.$$

To find the unrestricted maximum likelihood estimator (MLE):

$$\frac{\partial}{\partial \mu}\log L(\boldsymbol{\theta} \mid \boldsymbol{x}) = \frac{1}{\sigma^2}\sum_{i=1}^n (x_i - \mu) = 0$$
$$\frac{\partial}{\partial \sigma^2}\log L(\boldsymbol{\theta} \mid \boldsymbol{x}) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^n (x_i - \mu) = 0.$$

Solving the above equations gives the unrestricted MLE

$$\hat{\boldsymbol{\theta}} = \begin{pmatrix} \bar{X} \\ \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2 \end{pmatrix}.$$

To find the restricted MLE (under $H_0 : \mu = \mu_0$):

$$-\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^n (x_i - \mu_0) = 0$$

$$\implies \hat{\boldsymbol{\theta}}_0 = \begin{pmatrix} \mu_0 \\ \frac{1}{n}\sum_{i=1}^n (X_i - \mu_0)^2 \end{pmatrix}.$$

Therefore, the likelihood ratio test (LRT) statistic is

$$\lambda(\boldsymbol{x}) = \frac{L(\hat{\boldsymbol{\theta}}_0 \mid \boldsymbol{x})}{L(\hat{\boldsymbol{\theta}} \mid \boldsymbol{x})} = \left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \mu_0)^2}\right]^{n/2}$$

Noting that $\sum_{i=1}^n (x_i - \mu_0)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2$, the condition that the LRT rejects H_0,

$$\lambda(\boldsymbol{x}) = \left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \mu_0)^2}\right]^{n/2} \leq c, \text{ for } 0 \leq c \leq 1,$$

can be written as

$$\frac{\sum_{i=1}^n (x_i-\bar{x})^2}{\sum_{i=1}^n (x_i-\bar{x})^2 + n(\bar{x}-\mu_0)^2} \le c^{2/n},$$

which is equivalent to

$$\frac{1}{1+\frac{n(\bar{x}-\mu_0)^2}{\sum_{i=1}^n (x_i-\bar{x})^2}} \le c^{2/n} \iff \frac{n(\bar{x}-\mu_0)^2}{\sum_{i=1}^n (x_i-\bar{x})^2} \ge c^{-2/n}-1.$$

By defining $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i-\bar{X})^2$, the rejection condition is equivalent to

$$\frac{(\bar{x}-\mu_0)^2}{s^2/n} = \frac{n(\bar{x}-\mu_0)^2}{\frac{1}{n-1}\sum_{i=1}^n (x_i-\bar{x})^2} \ge (n-1)(c^{-2/n}-1)$$

That is,

$$\lambda(\boldsymbol{x}) \le c \iff \left|\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\right| \ge c',$$

where c' satisfies (because of the size α condition)

$$\begin{aligned}\alpha &= \sup_{\boldsymbol{\theta}\in\Theta_0} P_{\boldsymbol{\theta}}\left(\left|\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\right| \ge c'\right) \\ &= \sup_{\boldsymbol{\theta}\in\Theta_0} P_{\boldsymbol{\theta}}\left(\frac{\bar{X}-\mu}{S/\sqrt{n}} \ge c' + \frac{\mu_0-\mu}{S/\sqrt{n}} \text{ or } \frac{\bar{X}-\mu}{S/\sqrt{n}} \le -c' + \frac{\mu_0-\mu}{S/\sqrt{n}}\right).\end{aligned}$$

Since $X_1,\ldots,X_n$ are iid $\mathcal{N}(\mu,\sigma^2)$, $\frac{\bar{X}-\mu}{S/\sqrt{n}}$ has a t_{n-1} distribution. Thus, the critical value c' is chosen to satisfy

$$\alpha = P(|T_{n-1}| \ge c'),$$

which gives $c' = t_{n-1,\alpha/2}$. Therefore, the LRT rejects H_0 when

$$\left|\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\right| \ge t_{n-1,\alpha/2} \iff |\bar{x}-\mu_0| \ge t_{n-1,\alpha/2}\sqrt{s^2/n}.$$

This is the size α one-sample two-sided t-test.

b. Consider testing $H_0:\mu=\mu_0$ versus $H_1:\mu\neq\mu_0$. The test function of the level α two-sided t-test is

$$\phi(\boldsymbol{x}) = \begin{cases} 1 & \frac{|\bar{x}-\mu_0|}{s/\sqrt{n}} \ge t_{n-1,\alpha/2} \\ 0 & \text{otherwise.}\end{cases}$$

The acceptance region for the test is

$$A_{\mu_0} = \left\{\boldsymbol{x}\in\mathcal{X} : \frac{|\bar{x}-\mu_0|}{s/\sqrt{n}} < t_{n-1,\alpha/2}\right\},$$

where

$$P_{\mu_0}(\boldsymbol{X}\in A_{\mu_0}) = P_{\mu_0}\left(\frac{|\bar{X}-\mu_0|}{S/\sqrt{n}} < t_{n-1,\alpha/2}\right) = 1-\alpha.$$

This is true for all $-\infty < \mu_0 < \infty$. Thus, by inverting the acceptance region for the test, a $1-\alpha$ confidence set for μ is expressed as

$$
\begin{aligned}
C(\boldsymbol{x}) &= \{\mu \in \mathbb{R} : \boldsymbol{x} \in A_\mu\} \\
&= \left\{\mu : -t_{n-1,\alpha/2} < \frac{\bar{x}-\mu}{s/\sqrt{n}} < t_{n-1,\alpha/2}\right\} \\
&= \left\{\mu : -t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} < \bar{x}-\mu < t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}\right\} \\
&= \left\{\mu : \bar{x} - t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}\right\}.
\end{aligned}
$$

The random version of this confidence set (interval) is

$$
\left(\bar{X} - t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}, \bar{X} + t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}\right).
$$

3. (35 points) Suppose $X_1, \ldots, X_n$ are iid Beta$(\theta, 1)$, where $\theta > 0$.

   a. (5 points) Find the method of moments estimator of θ, $\hat{\theta}_{\text{MOM}}$.

   b. (10 points) Show that $\hat{\theta}_{\text{MOM}}$ satisfies

   $$\sqrt{n}(\hat{\theta}_{\text{MOM}} - \theta) \xrightarrow{d} \mathcal{N}\left(0, \frac{\theta(\theta+1)^2}{\theta+2}\right).$$

   Hint: Use Central Limit Theorem and Delta Method. **Useful fact:** For $Y \sim \text{Beta}(\alpha, \beta)$,

   $$f_Y(y \mid \theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}.$$

   The mean and variance of Y are $E[Y] = \frac{\alpha}{\alpha+\beta}$ and $\text{Var}[Y] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$, respectively.

   c. (5 points) Find the maximum likelihood estimator of θ, $\hat{\theta}_{\text{MLE}}$.

   d. (10 points) Show that $\hat{\theta}_{\text{MLE}}$ satisfies

   $$\sqrt{n}(\hat{\theta}_{\text{MLE}} - \theta) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

   Hint: Use large sample results for MLEs.

   e. (5 points) What is the asymptotic relative efficiency (ARE) of $\hat{\theta}_{\text{MOM}}$ to $\hat{\theta}_{\text{MLE}}$? Graph the ARE as a function of θ, and summarize the graph in 1-3 sentences.

Solution:

   a. The method of moments estimator equalizes the first sample moment $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and the first population moment $\frac{\theta}{\theta+1}$,

   $$\bar{X} = \frac{\theta}{\theta+1}.$$

   Thus, the method of moments estimator of θ is given by

   $$\hat{\theta}_{\text{MOM}} = \frac{\bar{X}}{1-\bar{X}}.$$

   b. The population mean and variance of Beta$(\theta, 1)$ are $\frac{\theta}{\theta+1}$ and $\frac{\theta}{(\theta+1)^2(\theta+2)}$, respectively. Therefore, by the Central Limit Theorem (CLT),

   $$\sqrt{n}\left(\bar{X} - \frac{\theta}{\theta+1}\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\theta}{(\theta+1)^2(\theta+2)}\right).$$

   Consider a function $g(x) = \frac{x}{1-x}$, which is differentiable

   $$g'(x) = \frac{1}{(1-x)^2}.$$

   and $g'(\bar{X}) \neq 0$. Using the Delta Method, we have

   $$\sqrt{n}\left(g(\bar{X}) - g\left(\frac{\theta}{\theta+1}\right)\right) \xrightarrow{d} \mathcal{N}\left(0, [g'(\bar{X})]^2 \frac{\theta}{(\theta+1)^2(\theta+2)}\right),$$

where $g(\bar{X}) = \hat{\theta}_{\text{MOM}}$, $g(\frac{\theta}{\theta+1}) = \theta$, and

$$[g'(\bar{X})]^2 \frac{\theta}{(\theta+1)^2(\theta+2)} = \left[\frac{1}{(1-\frac{\theta}{\theta+1})^2}\right]^2 \frac{\theta}{(\theta+1)^2(\theta+2)} = \frac{\theta(\theta+1)^2}{\theta+2}.$$

Therefore,

$$\sqrt{n}(\hat{\theta}_{\text{MOM}} - \theta) \xrightarrow{d} \mathcal{N}\left(0, \frac{\theta(\theta+1)^2}{\theta+2}\right).$$

c. The likelihood function is

$$L(\theta \mid \boldsymbol{x}) = \prod_{i=1}^{n} \frac{\Gamma(\theta+1)}{\Gamma(\theta)\Gamma(1)} x_i^{\theta-1}(1-x_i)^{1-1} = \theta^n \prod_{i=1}^{n} x_i^{\theta-1}.$$

The log-likelihood function is

$$\log L(\theta \mid \boldsymbol{x}) = n \log \theta + (\theta - 1) \sum_{i=1}^{n} \log x_i.$$

To find the MLE, set the first derivative

$$\frac{\partial}{\partial \theta} \log L(\theta \mid \boldsymbol{x}) = \frac{n}{\theta} + \sum_{i=1}^{n} \log x_i = 0$$

$$\implies \hat{\theta}_{\text{MLE}} = -\frac{n}{\sum_{i=1}^{n} \log X_i}.$$

d. Under certain regularity conditions, an MLE $\hat{\theta}_{\text{MLE}}$ satisfies

$$\sqrt{n}(\hat{\theta}_{\text{MLE}} - \theta) \xrightarrow{d} \mathcal{N}(0, v(\theta)),$$

where $v(\theta) = 1/I_1(\theta)$. The PDF of X is

$$f_X(x \mid \theta) = \theta x^{\theta-1},$$

and the derivatives of $\log f_X(x \mid \theta)$ are

$$\frac{\partial}{\partial \theta} \log f_X(x \mid \theta) = \frac{1}{\theta} + \log x$$
$$\frac{\partial^2}{\partial \theta^2} \log f_X(x \mid \theta) = -\frac{1}{\theta^2}$$

Therefore,

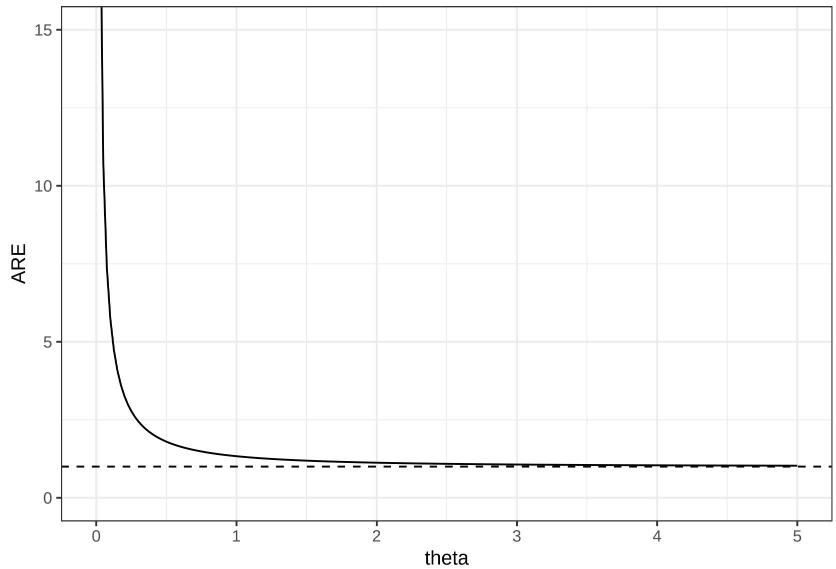
$$I_1(\theta) = -E_\theta\left[\frac{\partial^2}{\partial \theta^2} \log f_X(X \mid \theta)\right] = \frac{1}{\theta^2},$$

and

$$\sqrt{n}(\hat{\theta}_{\text{MLE}} - \theta) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

e. The ARE ($\hat{\theta}_{\text{MOM}}$ to $\hat{\theta}_{\text{MLE}}$) is

$$\frac{(\theta+1)^2}{\theta(\theta+2)}.$$



The MOM estimator is not as efficient as the MLE (ARE is always greater than unity).

4. (15 points) Suppose $X_1, \ldots, X_n$ are iid Bern(p), where $0 < p < 1$. Derive a $1 - \alpha$ Wald confidence interval for

$$g(p) = \log\left(\frac{p}{1-p}\right),$$

the log odds of p.

Solution:

The likelihood function of p is

$$L(p \mid \boldsymbol{x}) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}.$$

The log-likelihood function is

$$\log L(p \mid \boldsymbol{x}) = \sum_{i=1}^{n} x_i \log p + (n - \sum_{i=1}^{n} x_i)\log(1-p).$$

To find the MLE, set the first derivative

$$\frac{\partial}{\partial p}\log L(p \mid \boldsymbol{x}) = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n - \sum_{i=1}^{n} x_i}{1-p} = 0$$

$$\hat{p} = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X}.$$

Under certain regularity conditions, the MLE $\hat{p}$ satisfies

$$\sqrt{n}(\hat{p} - p) \xrightarrow{d} \mathcal{N}(0, v(p)),$$

where $v(p) = 1/I_1(p)$. The derivatives of $\log f_X(x \mid p) = x\log p + (1-x)\log(1-p)$ are

$$\frac{\partial}{\partial p}\log f_X(x \mid p) = \frac{x}{p} - \frac{1-x}{1-p}$$
$$\frac{\partial^2}{\partial p^2}\log f_X(x \mid p) = -\frac{x}{p^2} - \frac{1-x}{(1-p)^2}.$$

Therefore

$$I_1(p) = -E_p\left[-\frac{X}{p^2} - \frac{1-X}{(1-p)^2}\right] = \frac{1}{p(1-p)}.$$

Thus,

$$\sqrt{n}(\hat{p} - p) \xrightarrow{d} \mathcal{N}(0, p(1-p)).$$

Now, consider the function $g(p) = \log[p/(1-p)]$, which is a differentiable and

$$g'(p) = \frac{1}{p(1-p)} \neq 0.$$

The Delta Method gives

$$\sqrt{n}\left[\log\left(\frac{\hat{p}}{1-\hat{p}}\right) - \log\left(\frac{p}{1-p}\right)\right] \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{p(1-p)}\right).$$

Since $\hat{p}$ is a consistent estimator of p and $g(p)$ is a continuous function, the asymptotic variance $1/[p(1-p)]$ can be consistently estimated by $1/[\hat{p}(1-\hat{p})]$. By Slutsky's Theorem, we have

$$\frac{\log\left(\frac{\hat{p}}{1-\hat{p}}\right) - \log\left(\frac{p}{1-p}\right)}{\sqrt{\frac{1}{n\hat{p}(1-\hat{p})}}} \xrightarrow{d} \mathcal{N}(0,1).$$

Therefore,

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) \pm z_{\alpha/2}\sqrt{\frac{1}{n\hat{p}(1-\hat{p})}}$$

is an approximate $1-\alpha$ Wald confidence interval for the log odd $g(p) = \log[p/(1-p)]$.