STAT 513 fa 2018 Exam I (take-home)

Karl B. Gregory

assigned: Tuesday, Oct 2nd, 2018
due: Thursday, Oct 4th, 2018

Instructions:

- Looking at course notes IS allowed.
- Working with others IS NOT allowed. Asking a friend to help you puts the friend in the uncomfortable position of wanting to be nice but not wanting to break a rule. I recommend not doing that to your friends, because friendships are more important than exam grades.
- Write solutions on blank sheets of paper and turn these sheets in WITH A BLANK SHEET ON TOP (to cover answers) which has only your name on it. You do not need to turn in a copy of the exam itself.
- I expect answers to by very neatly written, since you have the time. Partial credit will be given only for legible work.

The table below gives some values of the function $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$:

z	1.282	1.645	1.96	2.576
$\Phi(z)$	0.9	0.95	0.975	0.995

Let $X_1, \dots, X_n$ be a random sample with likelihood function $L(\theta; X_1, \dots, X_n)$. Then for hypotheses of the form H_0: $\theta = \theta_0$ versus H_1: $\theta \neq \theta_0$ the likelihood ratio takes the form

$$\text{LR}(X_1, \dots, X_n) = \frac{L(\theta_0; X_1, \dots, X_n)}{L(\hat{\theta}; X_1, \dots, X_n)},$$

where $\hat{\theta}$ is the maximum likelihood estimator of θ.

1. Let $X_1, \ldots, X_n$ be a random sample from the Normal$(\mu, 1)$ distribution, where μ is unknown. A researcher plans to test H_0: $\mu \geq 5$ versus H_1: $\mu < 5$ with the test

$$\text{Reject } H_0 \text{ iff } \sqrt{n}(\bar{X}_n - 5) < -1.96.$$

(a) Give the size of the test.

Solution: We have $z_{0.25} = 1.96$, so the size is 0.25. This can be shown by the following, where $\gamma(\mu)$ is the power function:

$$\sup_{\mu \geq 5} \gamma(\mu) = \gamma(5) = P_{\mu=5}(\sqrt{n}(\bar{X}_n - 5) < -1.96) = P(Z < -1.96), \quad Z \sim \text{Normal}(0, 1),$$

which is equal to 0.025.

(b) Make a sketch which gives the shape of the power curve, indicating its height at $\mu = 5$.

Solution: Sketch should show a decreasing function taking values in $[0, 1]$ and crossing through the point $(5, 0.025)$.

(c) Give an expression for the power $\gamma(\mu)$ at any value of μ.

Solution: The power $\gamma(\mu)$ of the test is given by

$$\begin{aligned}
\gamma(\mu) &= P_\mu(\sqrt{n}(\bar{X}_n - 5) < -1.96) \\
&= P_\mu(\sqrt{n}(\bar{X}_n - \mu + \mu - 5) < -1.96) \\
&= P_\mu(\sqrt{n}(\bar{X}_n - \mu) < -1.96 - \sqrt{n}(\mu - 5)) \\
&= P(Z < -1.96 - \sqrt{n}(\mu - 5)), \quad Z \sim \text{Normal}(0, 1) \\
&= \Phi(-1.96 - \sqrt{n}(\mu - 5)).
\end{aligned}$$

(d) Which of the following numbers, when rounded up, is equal to the smallest sample size under which the test will reject H_0 with probability at least 0.90 for all $\mu \leq 4$? *Hint: You must do some calculations involving the power function.* **No points will be awarded if you do not show your work, even if you select the right answer.**

Solution: We want the power to be greater than 0.90 for all $\mu \leq 4$. Since the power function is decreasing, it will suffice to ensure that $\gamma(4) \geq 0.90$. The smallest sample size which ensures this is found by taking the smallest value of n which satisfies the following:

$$\begin{aligned} \gamma(4) &\geq 0.90 \\ \iff \Phi(-1.96 - \sqrt{n}(4-5)) &\geq 0.90 \\ \iff -1.96 - \sqrt{n}(4-5) &\geq \Phi^{-1}(0.90) \\ \iff -\sqrt{n}(4-5) &\geq \Phi^{-1}(0.90) + 1.96 \\ \iff \sqrt{n} &\geq z_{0.10} + 1.96 \\ \iff n &\geq (1.282 + 1.96)^2. \end{aligned}$$

A. $(1.96 + 2.575)^2$
B. $(1.96 + 2.575)^2/(5)^2$
C. $(1.96 + 1.282)^2$
D. $(2.575 + 2.575)^2/4$
E. $(1.96 + 2.575)^2/(\sqrt{4})^2$
F. $(1.96 + 1.96)^2/(\sqrt{5})^2$
G. $2^2(1.96 + 1.282)^2$
H. $(1.96 - 1.282)^2/5$
I. $2^2(1.96)^2/5^2$
J. $2^2(1.282)^2$

(e) Assuming that the correct sample size from the previous part is used, sketch the power curve of the test; include vertical lines positioned at $\mu = 4$ and $\mu = 5$ and horizontal lines positioned at the heights 0.90 and 0.025.

Solution: Sketch should show a decreasing function taking values in $[0, 1]$ and crossing through the point $(5, 0.025)$ and approximately crossing through the point $(4, 0.90)$.

(f) Suppose that a random sample of size n results in $\sqrt{n}(\bar{X}_n - 5) = -2.65$. In which of the following intervals does the p-value lie? *Hint: Consult the table of values of the standard Normal cdf* $\Phi(z)$. **No points will be awarded if you do not show your reasoning, even if you select the right answer.**

A. $(0, .005)$
B. $[.005, .01)$
C. $[.01, .05)$
D. $[.05, .5)$
E. $[.5, 1)$

(g) Suppose that a random sample of size n results in $\sqrt{n}(\bar{X}_n - 5) = 2.65$. In which of the following intervals does the p-value lie? **No points will be awarded if you do not show your reasoning, even if you select the right answer.**

A. $(0, .005)$

B. $[.005, .01)$
C. $[.01, .05)$
D. $[.05, .5)$
E. $[.5, 1)$

2. Let n^* be the smallest sample size required to detect a deviation as small as δ^* from the null with probability at least γ^* using a test with size α. Fill in the blanks:

   (a) For __**smaller**__ (larger/smaller) δ^* a __**larger**__ (larger/smaller) n^* is required.

   (b) For __**larger**__ (larger/smaller) γ^* a __**larger**__ (larger/smaller) n^* is required.

   (c) For __**smaller**__ (larger/smaller) α a __**larger**__ (larger/smaller) n^* is required.

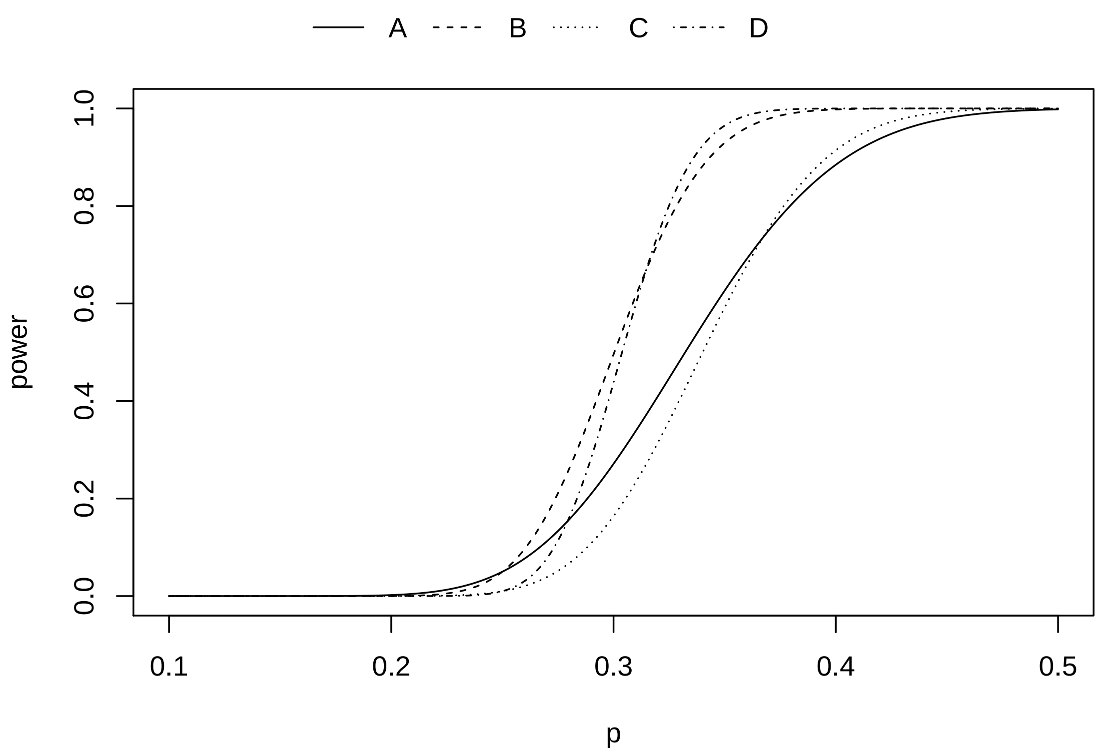
3. Let $X_1, \dots, X_n$ be a random sample from the Bernoulli(p) distribution, where p is unknown, and suppose it is of interest to test

$$H_0\colon\ p \leq 1/4 \text{ versus } p > 1/4.$$

Power curves of the test

$$\text{Reject } H_0 \text{ iff } \sqrt{n}(\hat{p}_n - 1/4)/\sqrt{1/4(1-1/4)} > z_\alpha$$

are shown under four settings, A, B, C, and D, which correspond to different values of α and n.



Answer the following by looking carefully at the power curves. **Explain answers to recieve credit!**

(a) Which settings have $\alpha = 0.05$?

> **Solution:** They are A and B: At $p = 1/4$, the heights of the curves under A and B have the height 0.05.

(b) Between settings C and D, which has the smaller sample size n?

> **Solution:** It is C: The value of α is the same under C and D, and the power curve under D climbs more steeply, so the sample size under D must be greater.

(c) Which of the four settings has the largest sample size n?

> **Solution:** It is D: For settings with the same value of α, the power curve under the larger sample size will climb more steeply. Therefore the answer must be either B or D. If B and D had equal sample sizes, the power under D would be smaller than under B over all values of p, since α under D is smaller. If the sample size under D were smaller than under B, the power under D could never exceed the power under B. However, the curve under D increases more steeply that that under B such that the power under D is greater than under B for any value of p greater than, say, 0.33. Therefore the sample size under D must be greater than under B.

(d) Under which setting is the probability of a Type II error the smallest when $p = 0.28$?

> **Solution:** It is B: When $p = 0.28$, the power curve under B is the highest, corresponding to the smallest Type II error probability.

4. Let $X_1, \dots, X_n$ be a random sample from the distribution with pdf given by

$$f(x) = \beta^{-1} \exp(-x\beta^{-1}) \mathbb{1}(x > 0),$$

where $\beta \geq 0$ is unknown, and suppose it is of interest to test the hypotheses

$$H_0\colon\ \beta = \beta_0 \text{ versus } H_1\colon\ \beta \neq \beta_0.$$

(a) Give the likelihood function $L(\beta; X_1, \dots, X_n)$ for $X_1, \dots, X_n$.

> **Solution:**
>
> $$L(\beta; X_1, \dots, X_n) = \prod_{i=1}^n \beta^{-1} \exp(-X_i\beta^{-1}) = \beta^{-n} \exp(-\beta^{-1} \sum_{i=1}^n X_i)$$

(b) Give the log-likelihood function $\ell(\beta; X_1, \dots, X_n)$ for $X_1, \dots, X_n$.

> **Solution:**
>
> $$\ell(\beta; X_1, \dots, X_n) = -n \log \beta - \beta^{-1} \sum_{i=1}^n X_i$$

(c) Find the maximum likelihood estimator $\hat{\beta}$ of β based on $X_1, \dots, X_n$.

Solution: The derivative of the log-likelihood function is

$$\frac{\partial}{\partial \beta}\ell(\beta; X_1, \ldots, X_n) = -n\beta^{-1} + \beta^{-2}\sum_{i=1}^{n} X_i,$$

which is equal to zero at the value $\hat{\beta} = \bar{X}_n$.

(d) Give the likelihood ratio $\text{LR}(X_1, \ldots, X_n)$ for testing the hypotheses H_0: $\beta = \beta_0$ versus H_1: $\beta \neq \beta_0$.

Solution:

$$\begin{aligned}
\text{LR}(X_1, \ldots, X_n) &= \frac{\sup_{\beta \in \{\beta_0\}} L(\beta; X_1, \ldots, X_n)}{\sup_{\beta \geq 0} L(\beta; X_1, \ldots, X_n)} \\
&= \frac{L(\beta_0; X_1, \ldots, X_n)}{L(\hat{\beta}; X_1, \ldots, X_n)} \\
&= \frac{\beta_0^{-n} \exp(-\beta_0^{-1} \sum_{i=1}^n X_i)}{\hat{\beta}^{-n} \exp(-\hat{\beta}^{-1} \sum_{i=1}^n X_i)} \\
&= \frac{\beta_0^{-n} \exp(-\beta_0^{-1} \sum_{i=1}^n X_i)}{\bar{X}_n^{-n} \exp(-\bar{X}_n^{-1} \sum_{i=1}^n X_i)} \\
&= (\bar{X}_n/\beta_0)^n \exp(-\beta_0^{-1} n \bar{X}_n + n) \\
&= [(\bar{X}_n/\beta_0) \exp(-\bar{X}_n/\beta_0)]^n \exp(n)
\end{aligned}$$

(e) Show that the rejection criterion $\text{LR}(X_1, \ldots, X_n) < c$ of the likelihood ratio test is equivalent to

$$\frac{\bar{X}_n}{\beta_0} \exp\left(-\frac{\bar{X}_n}{\beta_0}\right) < c^{1/n} e^{-1}$$

for any $c \in [0, 1]$.

Solution: We have

$$\begin{aligned}
&[(\bar{X}_n/\beta_0) \exp(-\bar{X}_n/\beta_0)]^n \exp(n) < c \\
\iff &[(\bar{X}_n/\beta_0) \exp(-\bar{X}_n/\beta_0)] < c^{1/n} e^{-1}.
\end{aligned}$$

(f) Since the function ze^{-z} is strictly increasing for $z < 1$ and strictly decreasing for $z > 1$, we have that rejecting H_0 when $\text{LR}(X_1, \ldots, X_n) < c$ is equivalent to rejecting H_0 when

$$\bar{X}_n < c_1 \text{ or } \bar{X}_n > c_2$$

for some c_1 and c_2. Explain in words how we can use the fact that

$$X_1, \ldots, X_n \overset{iid}{\sim} \text{Exponential}(\beta) \implies \bar{X}_n \sim \text{Gamma}(n, \beta/n)$$

to choose c_1 and c_2 such that the test has size α for any $\alpha \in (0, 1)$.

Solution: Choose c_1 and c_2 to be the $\alpha/2$ and $1 - \alpha/2$ quantiles of the Gamma$(n, \beta_0/n)$ distribution.