STAT 512 sp 2020 Final Exam

Karl B. Gregory

This is a take-home test due to the COVID-19 suspension of face-to-face instruction. Do not communicate with classmates about the exam until after its due date/time. You may

- *Use your notes and the lecture notes.*
- *Use books.*
- *NOT work together with others.*

Write all answers on blank sheets of paper; then take pictures and merge to a PDF. Upload a single PDF to Blackboard.

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!}[F_X(x)]^{k-1}[1-F_X(x)]^{n-k}f_X(x)$$

$$X_1,\dots,X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu,\sigma^2) \implies \begin{cases} (n-1)S_n^2/\sigma^2 \sim \chi^2_{n-1} \\ \sqrt{n}(\bar{X}_n-\mu)/\sigma \sim \text{Normal}(0,1) \\ \sqrt{n}(\bar{X}_n-\mu)/S_n \sim t_{n-1} \end{cases}$$

$$W_1 \sim \chi^2_{\nu_1}, \quad W_2 \sim \chi^2_{\nu_2}, \quad W_1, W_2 \text{ indep.} \implies (W_1/\nu_1)/(W_2/\nu_2) \sim F_{\nu_1,\nu_2}$$

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	0.841	1.282	1.645	1.96	2.326	2.576
$\Phi(z)$	0.80	0.90	0.95	0.975	.990	0.995

1. Copy down this sentence on your answer sheet and put your signature underneath: *I have not collaborated with any other student on this exam. The work I have presented is my own.*

2. Let $X_1, \dots, X_n$ be a random sample from a population with pdf given by

$$f_X(x;\tau) = \sqrt{\frac{\tau}{\pi}} \exp(-\tau x^2), \quad -\infty < x < \infty,$$

for some $\tau > 0$, and note that the first two population moments are $\mu_1' = 0$ and $\mu_2' = 1/(2\tau)$.

(a) Set $m_2' = \mu_2'$, where m_2' is the second sample moment, to find the MoMs estimator for τ.

Solution: We have

$$m_2' = \mu_2' = \frac{1}{2\tau} \iff \bar{\tau} = \frac{1}{2m_2'},$$

so we may write

$$\bar{\tau} = \frac{n}{2\sum_{i=1}^n X_i^2}.$$

(b) Write down the likelihood function for τ based on $X_1, \dots, X_n$.

Solution: The likelihood function is given by

$$\mathcal{L}(\tau; X_1, \dots, X_n) = \prod_{i=1}^n \sqrt{\frac{\tau}{\pi}} \exp(-\tau X_i^2) = \left[\frac{\tau}{\pi}\right]^{n/2} \exp\left[-\tau \sum_{i=1}^n X_i^2\right].$$

(c) Identify a sufficient statistic for τ.

Solution: By looking at the likelihood function, which is nothing other than the joint pdf of the random variables $X_1, \dots, X_n$ in the sample, that $T(X_1, \dots, X_n) = \sum_{i=1}^n X_i^2$ is a sufficient statistic for τ by the factorization

$$f_{X_1,\dots,X_n}(x_1, \dots, x_n; \tau) = \underbrace{\left[\frac{\tau}{\pi}\right]^{n/2} \exp\left[-\tau \sum_{i=1}^n X_i^2\right]}_{g(\sum_{i=1}^n X_i^2;\tau)} \cdot \underbrace{1}_{h(x_1,\dots,x_n)}.$$

(d) Give the log-likelihood.

Solution: The log-likelihood is given by

$$\ell(\tau; X_1, \dots, X_n) = \frac{n}{2}\log\tau - \frac{n}{2}\log\pi - \tau\sum_{i=1}^n X_i^2.$$

(e) Find the MLE $\hat{\tau}$ for τ.

Solution: Setting the derivative of the log-likelihood with respect to θ equal to zero and solving for τ gives

$$\frac{d}{d\tau}\ell(\tau; X_1, \ldots, X_n) = \frac{n}{2\tau} - \sum_{i=1}^{n} X_i^2 = 0 \iff \tau = \frac{n}{2\sum_{i=1}^{n} X_i^2},$$

so the MLE is

$$\hat{\tau} = \frac{n}{2\sum_{i=1}^{n} X_i^2}.$$

(f) Let $Y = \sqrt{2\tau}X$, where $X \sim f_X(x; \tau)$. Use the transformation method to find the pdf of Y.

Solution: We have

$$y = \sqrt{2\tau}x = g(x) \iff x = y/\sqrt{2\tau}, \quad \text{and} \quad \frac{d}{dy}g^{-1}(y) = 1/\sqrt{2\tau},$$

and $-\infty < Y < \infty$. So by the transformation method we have

$$f_Y(y) = \sqrt{\frac{\tau}{\pi}} \exp\left[-\tau(y/\sqrt{2\tau})^2\right] \cdot 1/\sqrt{2\tau} = \frac{1}{\sqrt{2\pi}} \exp(-y^2/2).$$

(g) Identify the distribution of Y and give the distribution of $W = Y^2$.

Solution: By the answer to the previous part, $Y \sim \text{Normal}(0, 1)$, so $W = Y^2 \sim \chi_1^2$.

(h) Give the distribution of the random variable $S = \sum_{i=1}^{n}(\sqrt{2\tau}X_i)^2$.

Solution: Since $(\sqrt{2\tau}X_i)^2 \sim \chi_1^2$ for each $i = 1, \ldots, n$, and since $X_1, \ldots, X_n$ are independent, we have

$$S \sim \chi_n^2.$$

(i) Show that $\hat{\tau} = n\tau/S$.

Solution: We have

$$\hat{\tau} = \frac{n}{2\sum_{i=1}^{n} X_i^2} = \frac{n2\tau}{2\sum_{i=1}^{n}(\sqrt{2\tau}X_i)^2} = \frac{n\tau}{S}.$$

(j) We find that $\mathbb{E}[1/S] = 1/(n-2)$. Use this to find $\mathbb{E}\hat{\tau}$.

Solution: We have

$$\mathbb{E}\hat{\tau} = \mathbb{E}\left[\frac{n\tau}{S}\right] = n\tau\mathbb{E}\left[\frac{1}{S}\right] = \left(\frac{n}{n-2}\right)\tau.$$

As an added note, we find $\mathbb{E}[1/S]$ as follows:

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{S}\right] &= \int_0^\infty \frac{1}{s}\frac{1}{\Gamma(n/2)2^{n/2}} s^{n/2-1}e^{-s/2}ds \\
&= \int_0^\infty \frac{1}{\Gamma(n/2)2^{n/2}} s^{(n/2-1)-1}e^{-s/2}ds \\
&= \frac{\Gamma(n/2-1)2^{n/2-1}}{\Gamma(n/2)2^{n/2}} \underbrace{\int_0^\infty \frac{1}{\Gamma(n/2-1)2^{n/2-1}} s^{(n/2-1)-1}e^{-s/2}ds}_{=1,\text{ integral over Gamma}(n/2-1,2)\text{ pdf}} \\
&= \frac{\Gamma(n/2-1)}{2(n/2-1)\Gamma(n/2-1)} \\
&= 1/(n-2).
\end{aligned}$$

(k) Propose an estimator $\hat{\tau}_{\text{unbiased}}$ based on $\hat{\tau}$ that is unbiased for τ.

Solution: An unbiased estimator for τ is

$$\hat{\tau}_{\text{unbiased}} = \left(\frac{n-2}{n}\right)\hat{\tau} = \frac{n-2}{2\sum_{i=1}^n X_i^2}.$$

(l) Comment on whether there could exist an unbiased estimator besides $\hat{\tau}_{\text{unbiased}}$ with smaller variance.

Solution: Since the estimator $\hat{\tau}_{\text{unbiased}}$ is unbiased and is a function of the sufficient statistic $\sum_{i=1}^n X_i^2$, it is the MVUE, so there does not exist any other unbiased estimator with smaller variance.

3. Let $X_1, \ldots, X_n \overset{\text{ind}}{\sim} \text{Bernoulli}(p)$.

(a) Give the MoMs estimator $\bar{p}$ for p.

Solution: Setting $m_1' = \mu_1'$ gives $\bar{X}_n = p$, so we have

$$\bar{p} = \bar{X}_n.$$

(b) Give the variance of $\bar{p}$.

Solution: We have

$$\operatorname{Var}\bar{p} = \frac{p(1-p)}{n}.$$

(c) Give the MSE of $\bar{p}$.

Solution: Since $\bar{p}$ is an unbiased estimator for p, we have

$$\operatorname{MSE}\bar{p} = \operatorname{Var}\bar{p} = \frac{p(1-p)}{n}.$$

(d) Is the MoMs estimator a consistent estimator for p? Explain your answer.

Solution: Since $\operatorname{MSE}\bar{p}$ goes to zero as n goes to infinity, $\bar{p}$ is a consistent estimator of p.

(e) Write down the joint pmf of $X_1, \ldots, X_n$.

Solution: The joint pmf of $X_1, \ldots, X_n$ is given by

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n; p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}.$$

(f) Find a sufficient statistic for p.

Solution: By the factorization

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n; p) = \underbrace{p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}}_{g(\sum_{i=1}^n x_i; p)} \cdot \underbrace{1}_{h(x_1,\ldots,x_n)}$$

we see that $\sum_{i=1}^n X_i$ is a sufficient statistic for p.

(g) Is the MoMs estimator of p the MVUE? Explain your answer.

Solution: Since the MoMs estimator $\bar{p}$ is unbiased for p and since it is a function of $\sum_{i=1}^n X_i$, which is a sufficient statistic for p, it is the MVUE for p.

4. Let $Z_1, Z_2, Z_3, Z_4, Z_5 \overset{\text{ind}}{\sim} \text{Normal}(0, 1)$. Identify the distributions of the following.

(a) $Z_1^2 + Z_2^2$

Solution: This has the χ_2^2 distribution.

(b) $\dfrac{(Z_1^2 + Z_2^2 + Z_3^2)/3}{(Z_4^2 + Z_5^2)/2}$

Solution: This has the $F_{3,2}$ distribution.

(c) $\dfrac{1}{\sqrt{3}}(Z_1 + Z_2 + Z_3)$

> **Solution:** This has the Normal$(0, 1)$ distribution.

(d) $\dfrac{1}{2}(Z_1 + Z_2)\Big/\sqrt{\dfrac{1}{2}\left[\left(Z_1 - \left(\dfrac{Z_1 + Z_2}{2}\right)\right)^2 + \left(Z_2 - \left(\dfrac{Z_1 + Z_2}{2}\right)\right)^2\right]}$

> **Solution:** This has the t_1 distribution.

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

$$F_X(x; B, M, \nu) = \frac{1}{[1 + \exp(-B(x - M))]^{1/\nu}}, \quad -\infty < x < \infty,$$

for some parameters $B > 0$, $-\infty < M < \infty$, $\nu > 0$.

(a) Find the function F_X^{-1} such that the random variable $F_X^{-1}(U)$ has the cdf F_X if $U \sim$ Uniform$(0, 1)$.

> **Solution:** We have
>
> $$u = \frac{1}{[1 + \exp(-B(x - M))]^{1/\nu}} \iff M - \frac{1}{B}\log\left(\frac{1 - u^\nu}{u^\nu}\right) = x,$$
>
> so
>
> $$F_X^{-1}(U) = M - \frac{1}{B}\log\left(\frac{1 - U^\nu}{U^\nu}\right) \sim F_X.$$

(b) Give the pdf f_X corresponding to the cdf F_X.

> **Solution:** Taking the first derivative with respect to x of $F_X(x)$ gives
>
> $$f_X(x) = \frac{B}{\nu}\frac{\exp(-B(x - M))}{[1 + \exp(-B(x - M))]^{1/\nu + 1}}, -\infty < x < \infty.$$

(c) Write down the joint pdf of $X_1, \dots, X_n$.

> **Solution:** The joint pdf of $X_1, \dots, X_n$ is
>
> $$f_{X_1,\dots,X_n}(x_1, \dots, x_n; B, M, \nu) = \prod_{i=1}^{n} \frac{B}{\nu}\frac{\exp(-B(x_i - M))}{[1 + \exp(-B(x_i - M))]^{1/\nu + 1}}.$$

(d) Check whether $\sum_{i=1}^{n} X_i$ is a sufficient statistic for the parameter M. Explain your answer.

Solution: We cannot find any factorization of the joint pdf $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;B,M,\nu)$ into a function of $\sum_{i=1}^n x_i$ and the parameter M and a function of $x_1,\ldots,x_n$ which does not involve the parameters, so $\sum_{i=1}^n X_i$ is not a sufficient statistic for M.

(e) Give an interval based on $X_1,\ldots,X_n$ which will contain the first population moment with probability approaching 0.99 as $n \to \infty$.

Solution: The first population moment is the population mean; a large-n approximate 99% confidence interval for the mean is given by

$$\bar{X}_n \pm 2.576 \cdot S_n/\sqrt{n}.$$