Spring 2023, MATH 408, Final Exam

Wednesday, May 3; 11am–1pm

Instructor — S. Lototsky (KAP 248D; x0–2389; lototsky@usc.edu)

Instructions:

- No notes, no books or other printed materials (including printouts from the web), no collaboration with anybody (or anything, like AI).
- You should have access to a calculator or some other computing device, and to the distribution tables (normal, t, χ^2, and F). Instead of the tables, you are welcome to use the statistical functions available on your computing device.
- Answer all questions and clearly indicate your answers; upload your solutions to GradeScope.
- **There are ten problems; each problem is worth 10 points.**

Problem 1. Let $X_1, \dots, X_{16}$ be an independent random sample from a normal population with unknown mean μ and unknown variance σ^2. It is known that

$$\bar{X}_{16} = \frac{160}{16} = 10. \qquad \sum_{k=1}^{16} X_k = 160, \quad \sum_{k=1}^{16} X_k^2 = 1840.$$

$S_{16}^2 = \frac{1}{15}\left[\sum_{k=1}^{16}(X_k - \bar X_{16})^2\right] = \frac{1}{15}\left[1840 - 100\cdot 16\right] = \frac{240}{15} = 16$

Construct the 99% confidence interval for the mean μ. $t_{15,\,0.005} = 2.947$ $\Rightarrow \bar X_{16} \pm \frac{S_{16}}{\sqrt{16}}\, t \approx \boxed{10 \pm 3}$

To get full credit, indicate the values of the sample mean, sample standard deviation, and the quantile of the corresponding distribution you need to construct the confidence interval.

Problem 2. Let $X_1, \dots, X_n$ be an independent random sample from the distribution with pdf

$$f(x;\theta) = \frac{1}{120\theta^6} x^5 e^{-x/\theta} 1(x > 0).$$

$\ln L_n \sim -6n \ln\theta - \frac{n}{\theta}\bar X_n$; $\frac{\partial}{\partial\theta} = 0 \Rightarrow \frac{6n}{\theta} = \frac{n}{\theta^2}\bar X_n$; $\boxed{\hat\theta_n = \frac{\bar X_n}{6}}$

$\frac{\partial \ln L_n}{\partial\theta} = -\frac{6n}{\theta} + \frac{n\bar X_n}{\theta^2}$; $+ \quad - \quad \hat\theta_n \Rightarrow \max$

Construct the MLE of θ. *Do not forget to check that the critical point is indeed the point of maximum.*

Problem 3. A study reports that freshmen at public universities work 11.2 hours a week for pay, on average, and the s_n is 8.5 hours; at private universities, the average is 9.3 hours and the s_n is 7.2 hours. Assume these data are based on two independent simple random samples, each of size 400. Is the difference between the averages due to chance? Explain your conclusion by stating the corresponding null and alternative hypotheses and computing the P-value.

$\frac{11.2-9.3}{\sqrt{\frac{8.5^2+7.2^2}{400}}} \approx 3.4$

P-value $2P(N(0,1) > 3.4)$

$0.0003 \cdot 2 = 0.0006$

< 0.01

not due to chance

Problems 4. Let $X_1, \dots, X_n$ be an independent random sample from the distribution with pdf

$$\mathrm{Gamma}(4,\theta) \Leftarrow f(x;\theta) = \frac{\theta^4}{6} x^3 e^{-\theta x} 1(x > 0).$$

$\frac{L_n(\bar X; 2)}{L_n(\bar X; 1)} = 2^{4n} e^{-n\bar X_n} < K_\alpha \Rightarrow$ Reject if $n\bar X_n > C_\alpha = \mathrm{Gam}(4n, 2)_{0.05}$

Construct the most powerful test with Type-I error equal to 0.05 for testing $H_0: \theta = 2$ against $H_1: \theta = 1$. 1) $EX = \frac{1}{4\theta}$. small $\bar X_n \Leftrightarrow$ big $\theta \Leftrightarrow$ reject H_0 if $\bar X_n$ is big 2) $n\bar X_n \sim \mathrm{Gamma}(4n, 2)$

Problem 5. For the first-year students at a certain university, the correlation between SAT scores and the first-year GPA was 0.86. Assume that the distribution of the scores is jointly normal. Predict the percentile rank on the first-year GPA for a student whose percentile rank on the SAT was 30%.

$z_x = \frac{X-\mu_x}{\sigma_x}$; $z_y = \frac{Y-\mu_y}{\sigma_y}$; $z_y = 0.86 z_x$

$z_x \approx -0.55$

$z_y \approx -0.473 \to 0.18 \to 32\%$

0.3

Problem 6. Below is part of a two-way ANOVA table for $b = 6$ blocks and $k = 5$ treatments. Fill out the rest of the table.

Source	SS	df	MS	F	Prob $> F$
Blocks	210	5	42	4.42	(0.005, 0.01) [0.0071243]
Treatments	83	4	20.75	2.18	> 0.1 [0.108250]
Error	190	20	9.5		
Total	483	29			

Problem 7.
To test whether a die is fair, 68 rolls were made, and the corresponding outcomes were as follows:

Face value	Observed frequency
1	8
2	11
3	15
4	15
5	11
6	8

exp. $= 68/6 = 11\frac{1}{3} = \frac{34}{3}$

$\varphi^* = 2 \cdot \frac{10^2 + 1^2 + 11^2}{9 \cdot 34/3} = \frac{222}{17 \cdot 3} \approx 4.35$

Estimate the P-value if the χ^2 test is used. $\mathbb{P}(\chi^2_5 > 4.35) > 0.1$ (≈ 0.5).

Problem 8. Assume that the following is an independent random sample from population X with a continuous cdf $F_X(x) = F(x)$:

14.1 15.3 13.6 11.5 12.8,

R 4 5 3 1 2

and assume that the following is an independent random sample from population Y with cdf $F_Y(x) = F(x + \theta)$:

8.6 5.5 6.1 10.3 9.1.

R 3 1 2 5 4

$n = 5$

$M = \sum_{K=1}^{5} 1(X_K > Y_K) = 5$

Compute the P-value of the sign test for the null hypothesis $\theta = 0$ against the alternative $\theta > 0$. *Note that the alternative means that the random variable X is more likely to be large, that is,* $\mathbb{P}(X > Y) > 1/2$.

P-value $= \mathbb{P}(B(5, \frac{1}{2}) \geq 5) = \frac{1}{32} = 0.03125$.

Problem 9. For the two samples in Problem 8, compute the Spearman rank correlation coefficient.

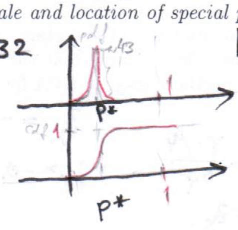
$\sum (R_x - R_y)^2 = 1 + 16 + 1 + 16 + 4 = 38$

$r_S = 1 - \frac{6 \cdot 38}{5 \cdot 24} = 1 - \frac{19}{10} = -\frac{9}{10}$

Problems 10. A coin-making machine produces pennies in such a way that, for each coin, the probability U to turn up heads is uniform on $[0, 1]$. A coin pops out of the machine, flipped 2500 times and lands heads 800 times. Sketch the graphs of the pdf and cdf of the posterior distribution of U. *Make sure to indicate the scale and location of special points on the horizontal axes.*

$p^* = \frac{800}{2500} = \frac{8}{25} = 0.32$

Beta(801, 1701)



Max height of pdf ≈ 43 (at p^*).