Math 218 Mathematical Statistics

Second Test Answers
April 2016

Scale. 89–106 A, 73–88 B, 65–72 C. Median 87.

1. [25] A British study (*Sleep*, Jan 1989) measured systematically the effect of sleep loss on creative faculties and the ability to deal with unfamiliar situations. In the study, 12 healthy college students, deprived of one night's sleep, received an array of tests intended to measure thinking time, fluency, flexibility, and originality of thought. The overall test scores of the sleep-deprived students were compared to the average score one would expect from students who received their accustomed sleep. Suppose that the overall scores of the 12 sleep-deprived students had a mean of $\overline{x} = 63$ and a standard deviation of 17. (Lower scores are associated with a decreased ability to think creatively.)

a. [5] Consider the hypothesis H_0 that the true mean score of sleep-deprived subjects is 80, the same mean score of subjects who received sleep prior to taking the test, as opposed to the alternative hypothesis H_1 that sleep-deprived subjects have depressed scores. Use $\alpha = .05$. First, what kind of test is this? Circle one.

(i). upper one-sided

(ii). lower one-sided. *Yes, that's it.*

(iii). two-sided

b. [5] Second, what kind of statistic will you use? Circle one.

(i). Z-statistic (i.e., normal) *No*, since n isn't large.

(ii). T-statistic. *Yes*, n is small so a t-stat is needed.

(iii). χ^2

(iv). F-statistic

c. [10] Carry out the test to see if the data are enough to reject H_0. (Show your work.)

We can reject H_0 if

$$\overline{x} < \mu_0 - t_{n-1,\alpha} \frac{s}{\sqrt{n}},$$

that is, if

$$63 < 80 - 1.796 \frac{17}{\sqrt{12}} = 71.2,$$

which is so. Therefore we reject H_0.

d. [5] What assumptions are required for the hypothesis test of part a to be valid?

That the X_i are independent form a random sample from a normally distribution with the same mean and variance.

2. [25] Refer to exercise 6 in chapter 7, page 263. Coffee cans are to filled with 16 oz. of coffee. The mean content of cans filled on a production line is monitored. It is known from past experience that the standard deviation of the contents is 0.1 oz. A sample of 9 cans is taken every hour and their mean content is measured.

a. [5] Set up the hypotheses to test whether the mean content is 16 oz. Should the alternative be one-sided or two-sided? Why?

Clearly the null hypothesis should be $H_0 : \mu = 16$, but what should the alternative hypothesis be?

1. $H_1 : \mu < 16$. A lower one-sided test.
2. $H_1 : \mu \neq 16$. A two-sided test.
3. $H_1 : \mu > 16$. An upper one-sided test.

You would use a lower one-sided test if you're worried that the coffee cans might not be filled enough; a two-sided test if you're worried that that they might not be filled enough or they might be filled too much; and an upper one-sided test if you're worried that the coffee cans might be filled too much.

It is not clear from the description what the worry is. You'll have to use your interpretation of the situation. I think it's quite reasonable to say that the company should be worried if the cans contain either too much or too little coffee. However, you could argue that the company needs to be very careful that the cans contain at least 16 oz. because of labelling laws and possible litigation or bad press if it's found that the company is falsely labelling 15 oz. cans as 16 oz. You could even argue that the company has to cut costs and make sure that no more than 16 oz. goes in the cans, but that would be a pretty weak argument.

I'll answer parts b and c as if the two-sided test is selected for part a.

b. [5] Give a decision rule in terms of the sample mean $\overline{x}$ for a 0.05-level test. Describe your rule in the form: Reject H_0 if $\overline{x}$ does not lie in some interval.

Assuming we're doing the two sided test we reject H_0 if $\overline{x}$ if

$$|\overline{x} - \mu_o| > z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

Here, $\mu_0 = 16$, $\sigma = 0.1$, $z_{\alpha/2} = z_{0.025} = 1.96$, and $n = 9$, so the rejection condition is

$$|\overline{x} - 16| > 1.96 \frac{0.1}{\sqrt{9}} = 0.065.$$

Thus, we reject H_0 if $\overline{x}$ does not lie in the interval

$$[15.935, 16.065].$$

c. [7] If the true mean content during a particular period is 16.1 oz., what is the probability that the test derived in part b will correctly detect this deviation from the target value of 16 oz.?

Method I. We need to compute the probability

$$P(\overline{X} \notin [15.935, 16.065]).$$

The easiest way to do this is to standardize $\overline{X}$. Since

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X} - 16.1}{0.1/\sqrt{9}} = (\overline{X} - 16.1)\,30$$

has a standard normal distribution, we can rewrite the probability as

$$\begin{aligned} & P(Z \notin [(15.935 - 16.1)\,30, (16.065 - 16.1)\,30]) \\ = \ & P(Z \notin [-4.95, -1.05]) \\ = \ & P(Z < -4.95) + P(Z > -1.05) \\ = \ & 0 + 0.85. \end{aligned}$$

Thus, the probability that a mean of $\mu = 16.1$ will be detected by the test is approximately 0.85.

Method II. The power $\pi(\mu)$ is defined to be the probability we're looking for. According to the table on page 247, a formula for computing it is

$$\begin{aligned} & \pi(\mu) \\ = \ & \Phi\left(-z_{\alpha/2} + \frac{\mu_0 - \mu)\sqrt{n}}{\sigma}\right) + \Phi\left(-z_{\alpha/2} + \frac{\mu - \mu_0)\sqrt{n}}{\sigma}\right) \\ = \ & \Phi\left(\frac{-1.96 - 0.1\sqrt{9}}{0.1}\right) + \Phi\left(\frac{-1.96 + 0.9\sqrt{9}}{0.9}\right) \\ = \ & \Phi(-4.96) + \Phi(1.04) \\ = \ & 0 + 0.85 \end{aligned}$$

d. [8] How many cans should be sampled to assure 90% power in part c?

We need to find n so that $\pi(\mu) = 0.9$. From part b, we know n will have to be larger than 9. We need to solve

$$\begin{aligned} & 0.9 = \pi(\mu) \\ = \ & \Phi\left(-z_{\alpha/2} + \frac{\mu_0 - \mu)\sqrt{n}}{\sigma}\right) + \Phi\left(-z_{\alpha/2} + \frac{\mu - \mu_0)\sqrt{n}}{\sigma}\right) \\ = \ & \Phi\left(\frac{-1.96 - 0.1\sqrt{n}}{0.1}\right) + \Phi\left(\frac{-1.96 + 0.9\sqrt{n}}{0.9}\right) \\ = \ & \Phi(-1.96 - \sqrt{n}) + \Phi(1.96 + \sqrt{n}) \end{aligned}$$

The first term, $\Phi(-1.96 - \sqrt{n})$ will be zero, so our equation reduces to $0.9 = \Phi(-1.96 - \sqrt{n})$. Since $0.9 = \Phi(1.28)$, we need to solve the equation $1.28 = -1.96 - \sqrt{n}$. Then $\sqrt{n} = 3.24$, and $n = 10.5$, so 11 cans are needed.

3. [25] Tax assessors. In response to a complaint that a particular tax assessor (X) was biased, an experiment was conducted to compare the assessor named in the complaint with another assessor (Y) in the same office. Eight properties were selected, and each was assessed by both assessors. The assessments (in thousands of dollars) are shown in the table

Property	Assessor X	Assessor Y
1	76.3	75.1
2	88.4	86.8
3	80.2	77.3
4	94.7	90.6
5	68.7	69.1
6	82.8	81.0
7	76.1	75.3
8	79.0	79.1

With this data, the mean assessment for X is $\overline{x} = 80.775$, the mean assessment for Y is $\overline{y} = 79.287$, so the mean difference is $\overline{d} = \overline{x} - \overline{y} = 1.487$. Also the sample standard deviation for X is $s_X = 7.994$, that for Y is $s_Y = 6.851$, and the sample standard deviation for the difference is $s_d = 1.491$.

a. [5] What are the assumptions for this experiment to have a matched pairs design?

That $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ are all normally distributed, the X_is with the same mean, the Y_is with the same mean, all with the same variance, all independent except that X_i and Y_i need not be independent but their correlations are the same for all i.

Assume in parts b and c that those assumptions are met.

b. [10] Determine a 2-sided confidence interval for the mean of the population difference μ_D which is $\mu_X - \mu_Y$ at the 95% confidence level. (Show your work and use standard notations.)

According to the top of page 284, the endpoints of the intervals are

$$\begin{aligned} & \overline{d} \pm t_{n-1,\alpha/2}\,\frac{s_d}{\sqrt{n}} \\ = \ & 1.487 \pm 2.365\,\frac{1.491}{\sqrt{8}} \\ = \ & 1.487 \pm 1.247 \end{aligned}$$

Thus, the interval is $[0.24, 2.73]$.

c. [10] Does the data allow you to reject the null hypotheses that the two assessors tend to give the same average assessments? Explain in a short sentence why.

Since 0 is not in the interval, we reject H_0.

4. [25] On a family of conjugate priors for Bayesian statistics.

a. [10] Let $\alpha > 1$ and $\beta > 0$. Show that the function

$$f_{\alpha\beta}(x) = \begin{cases} \frac{(\alpha-1)\beta^{\alpha-1}}{x^\alpha} & \text{if } x \geq \beta \\ 0 & \text{if } x < \beta \end{cases}$$

is a density function.

You just have to show that the integral of f equals 1.

$$\begin{aligned} \int_\beta^\infty \frac{(\alpha-1)\beta^{\alpha-1}}{x^\alpha}\,dx &= (\alpha-1)\beta^{\alpha-1}\int_\beta^\infty x^{-\alpha}\,dx \\ &= (\alpha-1)\beta^{\alpha-1}\frac{x^{1-\alpha}}{1-\alpha}\bigg|_\beta^\infty \\ &= (\alpha-1)\beta^{\alpha-1}\left(0 - \frac{\beta^{1-\alpha}}{1-\alpha}\right) = 1 \end{aligned}$$

b. [15] Consider the family of uniform distributions on $[0, \theta)$ parameterized by θ. The density function for such a distribution is

$$f(x) = \begin{cases} 1/\theta & \text{if } x \in [0, \theta] \\ 0 & \text{if } x > \theta \end{cases}$$

Suppose that the prior distribution on the parameter θ has density $f_{\alpha\beta}$ described in part **a**. As described in our notes on Bayesian statistics, the posterior distribution on θ is

$$f(\theta|\mathbf{x}) \propto f(\mathbf{x}|\theta) f_{\alpha\beta}(\theta).$$

As mentioned in example 15.4, page 617, of our text

$$f(\mathbf{x}|\theta) = \begin{cases} 1/\theta^n & \text{if } x_{\max} \in [0, \theta] \\ 0 & \text{otherwise} \end{cases}$$

where $x_{\max}$ is the maximum of the values $x_1, \ldots, x_n$ in the sample. Show that the posterior distribution $f(\theta|\mathbf{x})$ is another distribution of the form found in part **a**, and determine what the new α and new β for this posterior distribution.

The computations are easier if you use proportions rather than equations because you can drop all the messy constants.

The prior distribution is $f_{\alpha\beta}(\theta) = \dfrac{(\alpha-1)\beta^{\alpha-1}}{\theta^\alpha}$ for $\theta \geq \beta$ and 0 otherwise. In other words, $f_{\alpha\beta}(\theta)$ is proportional to $\dfrac{1}{\theta^\alpha}$ on $[\beta, \infty)$. Also, $f(\mathbf{x}|\theta)$ is proportional to $\dfrac{1}{\theta^n}$ on $[x_{\max}, \infty)$ and 0 otherwise. Therefore, the posterior distribution is proportional to their product

$$f(\theta|\mathbf{x}) = \frac{1}{\theta^{\alpha+n}}.$$

so long as θ is greater than both β and $x_{\max}$. So the posterior distribution is $f_{\alpha'\beta'}$ where $\alpha' = \alpha + n$ and β' is the maximum of β and $x_{\max}$.

In other words, α keeps track of the number of trials, and β keeps track of the largest observed value. By the way, these are sufficient statistics for the experiment.