UNIVERSITY OF TEXAS AT AUSTIN, DEPARTMENT OF MATHEMATICS
M358K - Applied Statistics

PRACTICE FOR IN-TERM EXAM II

True/false questions.

Problem 1.1. (2 points) Freddie Threepwood conducts a hypothesis test. He calculates the observed value of the $z-$statistic to be 0.018 (under the null). At the significance level of 0.05, he should reject the null hypothesis. *True or false?*

Solution: FALSE

Problem 1.2. (2 points) The mean and median of any normal distribution are equal. *True or false?*

Solution: TRUE

Problem 1.3. (2 points) *Source: Problem 6.21 from the Moore/McCabe/Craig.*

Consider the following two scenarios:

- Take a simple random sample of 100 sophomore students at your college or university.
- Take a simple random sample of 100 students at your college or university.

For each of these samples you record the amount spent on textbooks used for classes during the fall semester. You should suspect that the first sample should have the smaller margin of error.

True or false?

Solution: TRUE

Problem 1.4. (2 points) Consider the normal distribution with mean μ and standard deviation σ. Then, the probability that this normal distribution takes a value within one standard deviation from the mean is approximately 68%. *True or false?*

Solution: TRUE

Free-response problems.

Problem 1.5. (10 points)
Source: "Probability and Statistics for Engineers and Scientists" by Walpole, Myers, Myers, and Ye.

A corrosion study was made in order to determine whether coating an aluminum metal with a corrosion retardation substance reduced the amount of corrosion. Also of interest is the influence of humidity on the amount of corrosion. Two levels of coating – no coating and chemical-corrosion coating – were used. In addition, there were two relative humidity levels at 20% relative humidity and at 80% relative humidity.

The coating is a protectant that is advertised to minimize fatigue damage in this type of material. A corrosion measurement can be expressed in thousands of cycles to failure.

There are eight aluminum specimens used.

(i) (5 points) What is the explanatory variable in the above experiment design? What are the possible values it can take? *Hint: Draw a table of possible treatment combinations!*

(ii) (2 points) What are the **experimental** units?

(iii) (3 points) How would you assign the experimental units to the treatments to ensure that you are not introducing bias in your results?

Solution:

(i) The explanatory variable is the combination of coating or no coating, and 20% and 80% relative humidity. There are 4 possible treatment combinations.

(ii) The eight aluminum specimens.

(iii) Randomize the assignment of speciments to different treatment combinations.

Multiple-choice problems.

Problem 1.6. (5 points) *Source: Ramachandran-Tsokos.*
A dendritic tree is a branched formation that originates from a nerve cell. In order to study brain development, researchers want to examine the brain tissues from adult guinea pigs. At least how many cells must the researchers select (randomly) so as to be 95% sure that the sample mean is within 3.4 cells of the population mean? Assume that a previous study has shown that the cells have the standard deviation of exactly 10 dendrites.

(a) 28
(b) 33
(c) 34
(d) 35
(e) None of the above.

Solution: (c) We require that the margin of error at the 95% confidence be at most 3.4. So, the sample size n must satisfy

$$1.96 \times \frac{10}{\sqrt{n}} \leq 3.4 \quad \Rightarrow \quad n \geq \left(\frac{1.96 \times 10}{3.4}\right)^2 = 33.23183.$$

Problem 1.7. (5 points) The mean area of the several thousand new apartments is advertized to be at least 1350 square feet. A tenant group thinks that the apartments are smaller than advertised. They hire an engineer to measure a sample of apartments, they should test

a.: $H_0 : \mu = 1350$ against $H_a : \mu > 1350$.
b.: $H_0 : \mu = 1350$ against $H_a : \mu < 1350$.
c.: $H_0 : \mu < 1350$ against $H_a : \mu = 1350$.
d.: $H_0 : \mu = 1350$ against $H_a : \mu \neq 1350$.
e.: $H_0 : \mu < 1350$ against $H_a : \mu > 1350$.

Solution: b.

Problem 1.8. (5 points) The manufacturer of the *Slim Steakburger* brand claims that the mean fat content of this grade of steakburger is at most 18%.

The *Fat Fighters* consumer group, concerned about the mean fat content of this grade of steakburger submits to an independent laboratory a random sample of 12 steakburgers for analysis.

Assuming the percentage fat content being normally distributed with a variance of 3, they carry out an appropriate hypothesis test in order to advise the consumer group as to the validity of the manufacturer's claim.

The rejection region for the significance level of 0.05 is $[18.8225, \infty)$. With the above significance level of 0.05, find the power of the test at the alternative population mean of 20.

a. 0.95
b. 0.96
c. 0.98
d. 0.99
e. None of the above.

Solution: d.
We need to find the probability that the random variable

$$\bar{X}_{12} \sim N(mean = 20, variance = 0.25)$$

falls above the value 18.8225. We get

$$\mathbb{P}[\bar{X}_{12} > 18.8225] = \mathbb{P}\left[\frac{\bar{X}_{20} - 20}{0.5} > \frac{18.8225 - 20}{0.5}\right] = 1 - \Phi\left(-2.355\right) = \Phi(2.355) \approx 0.99.$$

Problem 1.9. Let the population distribution be normal with mean μ and standard deviation σ. Let $\bar{X}$ denote the sample mean of a sample of size $n$from this population. Then, we know the following about the distribution of $\bar{X}$:

(a) $\bar{X} \sim Normal(mean = \mu, variance = \sigma^2)$

(b) $\bar{X} \sim Normal(mean = \mu, variance = \frac{\sigma^2}{n})$

(c) $\bar{X} \sim Normal(mean = \mu, variance = \frac{\sigma^2}{\sqrt{n}})$

(d) $\bar{X} \sim Normal(mean = \frac{\mu}{n}, variance = \frac{\sigma^2}{n})$

(e) None of the above.

Solution: (b)

Problem 1.10. Suppose a poll suggested the US President's approval rating is 45%. We would consider 45% to be ...

(a) ...the population mean.
(b) ...the point estimate.
(c) ...the sampling error.
(d) ...the bias.
(e) None of the above.

Solution: (b)

Problem 1.11. The *Cheesecake Manufacture and Dinery* claims that their famous cheesecake has at most 2000 per slice. You suspect the contrary and plan a study. You model the calorie content per slice using the normal distribution with un unknown mean μ and with a **known** standard deviation of 300.

You diligently study a random sample of 25 slices of cheesecake. The sample average turns out to be 2100.

What is the $p-$value corresponding to these data?

(a) 0.02375
(b) 0.0375
(c) 0.0475
(d) 0.095
(e) None of the above.

Solution: (c)
We are testing

$$H_0 : \mu = 2000 \quad vs. \quad H_a : \mu > 2000.$$

Under the null hypothesis, the $z-$score corresponding to the observed sample average is

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{2100 - 2000}{\frac{300}{\sqrt{25}}} = \frac{100}{60} \approx 1.67.$$

This test has a right-tailed alternative. So, from the standard normal tables, we get that the $p-$value equals

$$\mathbb{P}[Z > z] = 1 - \Phi(1.67) = 1 - 0.9525 = 0.0475.$$

Problem 1.12. A car manufacturer claims that the mean time until the car battery needs to be replaced is five years. From past experience, the lifetime of a car battery is modeled as normal with a **known** standard deviation of one year. An environmental institute wants to test the car manufacturer's claim. They collect the data from 49 cars and find the sample average of 4.7 years. What is their decision going to be at the 2% significance level?

(a) Reject the null hypothesis.
(b) Fail to reject the null hypothesis.
(c) Accept the null hypothesis.
(d) Reject the alternative hypothesis.
(e) None of the above.

Solution: (a)
The environmental institute is testing

$$H_0 : \mu = 5 \quad vs. \quad H_a : \mu < 5.$$

Under the null hypothesis, the $z-$score corresponding to the given sample average is

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{4.7 - 5}{\frac{1}{\sqrt{49}}} = -2.1.$$

The $p-$value is, according to the standard normal tables,

$$\mathbb{P}[Z < -2.1] = 0.0179.$$

where $Z \sim N(0,1)$. So, the null hypothesis is rejected at the 2% significance level.

Problem 1.13. In a hand sanitizer production facility, a machine is operated whose job is to fill the hand-sanitizer bottles with exactly 8 oz of the precious liquid. You are wondering whether the machine is correctly calibrated. From past experience, you know that you can model the amount in every bottle as normal with a known standard deviation of 1/4 oz. You are going to sample 100 bottles to test whether the machine is properly calibrated. If you choose that you are going to use a 1% significance level, what is the associated rejection region (in real units)?

(a) $[0, 7.9356] \cup [8.0644, \infty)$

(b) $(7.9356, 8.0644)$

(c) $[0, 7.951]$

(d) Not enough information is given.

(e) None of the above.

Solution: (a)
Let the unknown mean amount of hand sanitizer per bottle be denoted by μ. Then the distribution of the amount of hand sanitizer in a randomly chosen bottle can be written as

$$X \sim Normal(mean = \mu, sd = 1/4).$$

We need to test

$$H_0 : \mu = \mu_0 = 8 \quad vs. \quad H_a : \mu \neq \mu_0 = 8.$$

The rejection region for this two-sided test will be of the form (in our usual notation)

$$RR = \left(-\infty, \mu_0 + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)\right] \cup \left[\mu_0 + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right), \infty\right)$$

with $z_{\alpha/2} = \Phi^{-1}(0.005) = -2.576$. We have that

$$\mu_0 + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) = 8 - 2.576\left(\frac{1/4}{\sqrt{100}}\right) = 7.9356,$$
$$\mu_0 - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) = 8 + 2.576\left(\frac{1/4}{\sqrt{100}}\right) = 8.0644.$$

Of course, the amount in any bottle is at least 0. So, our rejection region is

$$RR = [0, 7.9356] \cup [8.0644, \infty).$$

We can say that the ∞ above is actually a placeholder for the capacity of a bottle.