University of Texas at Austin, Department of Mathematics
M358K - Applied Statistics

Practice Probems for In-Term III

True/false questions.

Problem 1.1. If a random variable X has a standard normal distribution, then X^2 has a chi-squared distribution with 1 degree of freedom. *True or false?*

Problem 1.2. Let X be a standard normal random variable, and let Y be a $\chi-$squared random variable with one degree of freedom. Assume that X and Y are independent. Then, X/Y is $t-$distributed. *True or false?*

Free-response problems.

Problem 1.3. (8 points) To write an article about Denver for a tourist magazine you would like to estimate the average nightly cost for a hotel room in the Denver area. You are willing to assume that the nighly cost of a room is normally distributed.

You open up the yellow pages and take a random sample of hotels. The sample of 16 hotels gives an average nightly cost of \$55.98 and a **sample** standard deviation of \$12. Estimate the mean nightly cost and include a 95% confidence interval.

Problem 1.4. (10 points)

It is claimed that the bags of chocolate chips available in Costco contain at least 4 pounds (64 ounces).

A random sample of 50 bag measurements resulted in the sample average of 62 and the sample standard deviation of 8. Please, test the hypothesis that the bags contain at least 64 ounces of delicious chocolate chips at the significance level of 5%.

Problem 1.5. An Airbus a330 can hold 277 passengers. The *Marginaire* airline knows from past experience that only about 90% of the passengers make it to their flights. So, they sell 300 tickets for every Airbus a330 flight.

Assume that passengers travel independently, i.e., the events that individual passengers make it to a flight are independent events. What is the approximate probability that a particular flight gets overbooked? *Note: Don't forget to use the continuity correction.*

Problem 1.6. (10 points) *Source: "Probability and Statistics for Engineers and Scientists" by Walpole, Myers, Myers, and Ye.*

An improvement is a process for manufacturing is being considered. Samples are taken both from parts manufactured using the existing method and the new method. It is found that 75 out of a sample of 1500 items produced using the existing method are defective. It is also found that 80 out of a sample of 2000 items produced using the new method are defective. The two samples are independent.

Find the 90%-confidence interval for the true difference in the proportions of defectives produced using the existing and the new method.

Problem 1.7. (15 points)

A casino game involves rolling three dice. The winnings are proportional to the total number of sixes rolled. Suppose a gambler plays the game 150 times, with the following observed counts:

Number of sixes	0	1	2	3
Count	72	51	21	6

Assuming that the die rolls are independent, test the null hypothesis that the dice are all fair. *Note: Keep five decimal places for your expected counts.*

Multiple-choice problems.

Problem 1.8. (5 points) *Source:"Probability and Statistics for Engineers and Scientists" by Walpole, Myers, Myers, and Ye.*

In a simple random sample of 500 households owning televisions in the city of Hamilton, Canada (pop. $536,915$), it is found that 340 subscribe to HBO. Find a 95% confidence interval for the true proportion of households with television which subscribe to HBO.

a. 0.68 ± 0.021
b. 0.68 ± 0.034
c. 0.68 ± 0.041
d. 0.68 ± 0.054
e. None of the above.

Problem 1.9. (5 points) *Source:"Probability and Statistics for Engineers and Scientists" by Walpole, Myers, Myers, and Ye.*

A commonly prescribed drug for relieving nervous tension is declared to be effective in 60% of patients. Experimental results with a **new** drug administered to a SRS of 100 patients show that 70 received relief. To answer the question whether the new drug is truly superior, you calculate the $p-$value. What do you get?

a. 0.0146
b. 0.0207.
c. 0.0292
d. 0.0414
e. None of the above.

Problem 1.10. (5 points) In 1956 Middletown, Lynd and Lynd conducted a sociological study in which questionnaires were administered to 784 white high school students. They were asked *"which* 2 *of the given* 10 *attributes where most desirable in their fathers."*

Among other things, and along with the students' genders, it was tallied how many of them mentioned *"being a college graduate"* as one of the 2 chosen desirable qualities. The following two-way table contains the resulting counts:

	Male	Female	Total
Mentioned	86	55	141
Not mentioned	283	360	643
Total	369	415	784

The question we can try to answer using the above data is whether males and females value this particular attribute differently. What is the conclusion of your hypothesis test of independence? *Note: When you calculate, keep four places after the decimal point for expected counts.*

a. The $p-$value is less than 0.001.
b. The $p-$value is between 0.001 and 0.005.
c. The $p-$value is between 0.005 and 0.01.
d. The $p-$value is between 0.01 and 0.02.
e. None of the above.

Problem 1.11. An experiment was designed to test whether people's reaction times to an orange light are different from their reaction times to a blue light. Upon being signaled with a light, the subjects would hit a button and the reaction time (in seconds) would be recorded. The reaction time (in seconds) of 16 subjects was recorded. The average reaction time for the blue light was 0.2025 seconds, and the average reaction time for the orange light was 0.1380 seconds. The sample standard deviation of the differences between reaction times was 0.0565. What is the 80%-confidence interval for the difference in mean reaction times? Assume the normal model for the reaction times.

(a) 0.0645 ± 0.0189
(b) 0.0645 ± 0.0122
(c) 0.2025 ± 0.0189
(d) 0.1380 ± 0.0122
(e) None of the above.

Problem 1.12. Twenty-five fortunate middle schoolers were put on an intense rope-jumping regimen in the hope of improving their times in the 40-yard dash. Assume that the distribution of the differences in the run times is normal. Let μ_d be the mean difference between the "pre-run" (before the regimen) and "post-run" (after the regimen). We want to test

$$H_0 : \mu_d = 0 \quad vs. \quad H_a : \mu_d > 0.$$

The observed average difference in run times was $\bar{x}_d = 0.0854$ while the sample standard deviation was $s_d = 0.2432$. What can you say about the $p-$value for this hypothesis test?

(a) It's below 0.01.
(b) It's between 0.01 and 0.02.
(c) It's between 0.02 and 0.025.
(d) It's between 0.025 and 0.05.
(e) None of the above.

Problem 1.13. *Source: "Probability and Statistical Inference" by Hogg, Tanis, and Zimmerman.*
Let p_m and p_f be the population proportions of male and female sparrows who return to their hatching site. You want to test whether the two proportions are different. The observed number of males who returned is 124 out of 894, while the observed number of females who returned is 70 out of 700. What is your decision for this hypothesis test?

(a) Reject at the 0.01 significance level.

(b) Fail to reject at the 0.01 significance level; reject at the 0.025 significance level.

(c) Fail to reject at the 0.025 significance level; reject at the 0.05 significance level.

(d) Fail to reject at the 0.05 significance level; reject at the 0.10 significance level.

(e) None of the above.