

UNIVERSITY OF TEXAS AT AUSTIN

Problem Set # 17

The t -procedure: two means17.1. Two-sample t .

Problem 17.1. An instructor is teaching two sections of the same basic statistics course. The instructor is giving the same exams, homework assignments, and quizzes in both sections. Which t -procedure should be used to determine if there is a difference in the academic performance between the two course sections?

- (a) One-sample t -test.
- (b) Matched-pairs t -procedure.
- (c) Two-sample t -test.
- (d) None of the above.

Problem 17.2. This is an excerpt from findings of an educational study:

A study was done to determine whether there is a difference in the amount of time (in hours) that graduate students versus undergraduate students spend on the Internet per day. Five undergrads and five grad students were polled.

- (i) Is the alternative hypothesis one-sided or two-sided?
- (ii) A t -score for the data gathered was calculated to be 1.6664. Would you say that there is a significant difference in the amount of time that graduate and undergraduate students spend on the Internet?

Problem 17.3. (5 points)

There is a dispute about salaries of male versus female elves. The North Polar Bear collected the following data:

- the total number of male elves is 121;
- the total number of female elves is 100;
- the average salary of a male elf is 10,000 candy canes;
- the average salary of a female elf is 12,000 candy canes;
- the sample standard deviation of the salaries of male elves is 50;
- the sample standard deviation of the salaries of female elves is 132.

Assume independence between the salaries of individual elves.

Let μ_m denote the population mean for the distribution of the male elves' salaries and let μ_f denote the population mean for the distribution of the female elves' salaries. We wish to test:

$$H_0 : \mu_m = \mu_f \quad \text{vs.} \quad H_a : \mu_m \neq \mu_f.$$

What is the p -value associated with our data?

- a.: About 0.
- b.: About 0.01.
- c.: About 0.025
- d.: About 0.04.
- e.: None of the above.

Problem 17.4. Let the population distributions be normal with unknown parameters. Assume that sample data, based on two independent samples of size 25, give us $\bar{x}_1 = 505$, $\bar{x}_2 = 515$, $s_1 = 23$, and $s_2 = 28$.

- (i) What is a 95%-confidence interval (use the conservative value for the degrees of freedom) for the difference between the two population means?
 - (ii) Based on the confidence interval, we can conclude at the 5% significance level that there is no difference between the two population means. *True or false?*
 - (iii) The margin of error for the difference between the two sample means would be smaller if we were to take larger samples. *True or false?*
 - (iv) If a 99% confidence interval were calculated instead of the 95% interval, it would include more values for the difference between the two population means. *True or false?*
-

17.2. Pooled t .

Problem 17.5. The **pooled** two-sample t -procedure can be used when ...

- (a) you can assume the two populations have equal variances
- (b) you can assume the two populations have equal means
- (c) the sample sizes are equal
- (a) None of the above

Problem 17.6. Let n_1 and n_2 denote the sample sizes of each group. The pooled two-sample t -procedure is based how many degrees of freedom?

- (a) $n_1 + n_2 + 2$
- (b) $n_1 + n_2 - 2$
- (c) $n_1 + n_2 - 1$
- (d) $n_1 + n_2$
- (e) None of the above.

Problem 17.7. A study was done to determine if students learn better in an online basic statistics class versus a traditional face-to-face (f2f) course. A random sample of 12 students in an online course and 15 students in an f2f course was taken.

- (i) Let μ_{new} denote the population mean score for the online statistics class and let μ_{old} denote the population mean score for the face-to-face statistics class. What are the hypotheses being tested?
- (ii) We decide it is appropriate to use the pooled t -procedure. What is the number of degrees of freedom you are going to use?