Key

Statistics Chapter 10
Sections 10.3

1. Look at the following data below that compares Method A vs. Method B.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| A | 1 | 5 | 8 | 9 | 5 | 4 | 4 |
| B | 3 | 2 | 0 | 1 | 3 | | |

Is the above data related to two independent samples or one dependent sample? How do you know?

INDEPENDENT: Both of the sample sizes are different

$n_1 \neq n_2$

2. Can 2 Independent Populations yield samples that are the same size? yes

3. In a hypothesis test between 2 Independent Populations, the relation between the two populations is unimportant. Why?

The two populations should have no relation, therefore it is unimportant.

4. In a hypothesis test between 2 Independent Populations, the Null Hypothesis will usually be...

$H_o$: $\mu_1 = \mu_2$ or in this case $\mu_A = \mu_B$

What is the assumption between both population averages? Both populations are equal.

5. In a confidence interval between 2 Independent Populations, is it possible to get a "+" and "-" value for the results?

yes, it is possible to get,
(+, −), (−, +), (+, +), (−, −)

6. Suppose a calculation for a confidence interval between the average test scores for Period 1 Stats and Period 7 Stats produced the values,

(-5.35 to 15.21)

The manner of subtraction was (Period 1 average) – (Period 7 average). What does the interval indicate?

$P_1 - P_7$
−5.35 to 15.21

$P_1 \downarrow$ $P_1 \uparrow$

The difference between period 1 and period 7 is between 5.35 lower for 1 to 15.21 higher for 1.

7. What calculator function is used to conduct a hypothesis test between 2 Independent Populations?

2 sample t-test

8. What calculator function is used to conduct an estimation between 2 Independent Populations?

2 sample t-interval

9. There are two methods to find the Degree of Freedom between 2 Independent Populations.

a. Conservative Approach; df = Smaller n − 1

b. Calculator:

10. Referring to question #9, which method is the preferred method to find t*?

(b) – more accurate

11. A pharmaceutical company is interested in the time required for headache remedies to enter the blood stream. Below are the data comparing two brands.

| Medication Brand | $\bar{x}$ | $s_x$ | $n$ |
|---|---|---|---|
| A | 20.1 min | 8.7 min | 12 |
| B | 11.2 | 7.5 | 8 |

Brand B claims to be faster. Is there sufficient evidence to support Brand B's claim? Let the level of significance equal 0.05.

a. Are both brands of medication Independent? Why

yes, the sample sizes are different.

b. Is this problem considered a hypothesis test or confidence interval?

TEST

c. According to the problem what is the relation between the two brands of medication? (Define the relation by using inequalities)

$\mu_A > \mu_B$

d. Step 1.

more time ← $\mu_A > \mu_B$ → less time

$H_0: \mu_A = \mu_B$ ; $H_a: \mu_A > \mu_B$ ; $\alpha = .05$

e. What are the three necessary components of Step 2?

Random, Size, Distribution

f. In this example, which component of Step 2 cannot be evaluated?

The Distribution; you cannot make the mod. box plot. (no raw Data)

g. Is it possible that one sample is not biased, whereas the other sample is biased?

YES

h. Complete Step 2 for the problem.

| | med. A | med. B |
|---|---|---|
| Random | No | No |
| Size | n=12, small | n=8, small |
| Dist. | unknown | unknown. |

i. What are the 4 components necessary for Step 3?

curve, work, p-val, reject.

j. In the equation, what is $(\mu_1 - \mu_2)$ equal to?

0

k. What function in the TI 84 or 89 calculator will you use for this Hypothesis Test?

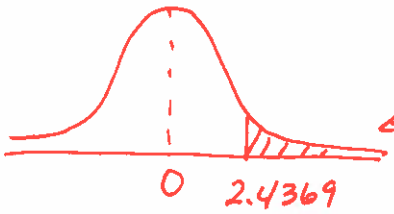
2-sample t-test

l. If you see a p-value = 2.56 E -4, what is that value equal to (in scientific notation or decimal format) $2.56 \times 10^{-4}$ or .000256

m. When can the Null Hypothesis be rejected?

when p-val $< \alpha$

n. Do all of Step 3




$$t = \frac{20.1 - 11.2}{\sqrt{\frac{8.7^2}{12} + \frac{7.5^2}{8}}}$$

$t = 2.4369$

p-val = .01318

p-val $< \alpha$ Reject

o. Do all of Step 4

I do have sufficient evidence to reject the null hypothesis. I can say that medication B does enter the blood stream faster than medication A. The test could be biased because both samples were not obtained randomly; they were both small; and their distributions are unknown.

12. A Medical Researcher would like to investigate the number of people that have Hay Fever in Kansas. Here are the data.

Rate of Hay Fever per 1000 population for people under 25

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 98 | 90 | 120 | 128 | 92 | 123 | 112 | 93 |
| 125 | 95 | 125 | 117 | 97 | 122 | 127 | 88 |

Rate of Hay Fever per 1000 population for people over 50

| | | | | | | |
|---|---|---|---|---|---|---|
| 95 | 110 | 101 | 97 | 112 | 88 | 110 |
| 79 | 115 | 100 | 89 | 114 | 85 | 96 |

Estimate with 95% confidence the average difference between the rate of Hay Fever for people under 25 years old and over 50 years old.

a. Step 1. Write up the statement

I am estimating with 95% confidence the average difference for the rate of Hay Fever between 25 year olds and 50 year olds

b. What are the 3 elements necessary to check off Step 2?

Random, Size, Distribution

c. Do you need to check the 3 elements for both samples?

yes

d. Do Step 2

| | 25 years or Under | 50 years or over |
|---|---|---|
| Random | No | No |
| Size | $n=16$, small | $n=14$ small |
| Dist. |  Not bad, median is to the right making it a bit skewed |  pretty symmetrical |

e. In Step 3 for estimation, what are the 2 necessary components.

work and results

f. What is the equation you will use for this problem?

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

g. Which way are you going to subtract the problem?

I will subtract 25years − 50years.

(There is no wrong method to subtract)

h. What are the possible degree's of freedom in this problem?

Conservative approach → 14−1 = 13 or Calculator 27.4252 2 sample tinv?

i. Which degree of freedom will you use for this problem to find t*?

27.4252 ... more accurate

(use 27 in the chart)

j. Suppose you ended up with the result of (-20.3 to 10.1) and you found the average difference by $\overline{x_{25}} - \overline{x_{50}}$. What do the values indicate? (Higher vs. Lower)

(I am going to use 25year olds)

25y − 50y

−20.3 to 10.1

25↓ 25↑

25year olds lower by 20.3

25year olds higher by 10.1

k. Do Step 3.

$$(109.5 - 99.3571) \pm 2.052 \cdot \sqrt{\frac{15.4056^2}{16} + \frac{11.5732^2}{14}}$$

25y − 50y

.0148 to 20.27

25↑ 25↑

l. Should you have any negative values when you write up Step 4?

No, report those "−" as lower

m. Do Step 4.

I am 95% confident that the average difference in the number of people that get hay fever between 25 years or younger ~~to~~ and 50 years or older, is between .0148 to 2027 higher for 25 years or younger (per 1000 people).

The estimation could be biased because both samples are small and they (the data) were not obtained randomly.