Introductory Statistics Lectures

Testing a claim about two means

Two sample hypothesis test of the mean: independent & dependent samples

Anthony Tanbakuchi
Department of Mathematics
Pima Community College

(Compile date: Tue May 19 14:51:00 2009)

Contents

1 Testing a claim about two means

1.1 Introduction

Example 1. A student wants to test the hypothesis that male college students are taller (on average) than female college students. The student records the heights of his fellow statistics class mates (in inches):

Males: 71 77 70 67 73 68 70

Females: 65 68 66 68 65 62 68 62 69 65 63

Question 1. What is our hypothesis for this problem?

In hypothesis testing we assume the null hypothesis is true and calculate the probability of observing a sample statistic at least as extreme as the one we observe.

Question 2. If the null hypothesis is true, how is it possible that the sample statistic could differ from the null value?

1.2 Dependent vs. independent samples

Independent vs dependent samples

DEFINITION 1.1 INDEPENDENT SAMPLES.

The samples from one population are not related to or paired with the samples values in from the other population.

Example 2 (Independent samples). Taking 100 study subjects and randomly placing half in a group receiving a placebo and the other half in a group that receives an aspirin a day to study its effect of heart attacks. (There is no relationship between individuals in the group receiving the placebo and the group receiving the aspirin.)

DEFINITION 1.2 DEPENDENT SAMPLES (MATCHED PAIRS / PAIRED SAMPLES).

When the samples from the first population have some relationship or pairing to the second population that is sampled.

Example 3 (Dependent samples). Randomly selecting 100 individuals and measuring their blood pressure (as sample 1) and then giving them a blood pressure reducing drug for two weeks and measuring their blood pressure again (as sample 2). (The two samples **are related**, it involves the sample individuals and we could even **pair** the before and after measurements of the blood pressure for each individual.)

Determine if the following two examples consist of independent or dependent samples.

Question 3. The effectiveness of Prilosec for treating heartburn is tested by measuring gastric acid secretion in patients before and after the drug treatment. The data consist of the before/after measurements for each patient.

Question 4. The effect of sugar as an ingredient is tested with a sample of cans of regular Coke and another sample of cans of diet Coke.

1.3 2 population means: indep. samples

USE

Often used to help answer:

1. Is the mean of x the same in the two populations?
2. Is the mean of x the same in the two populations?
3. Is process 1 equivalent to process 2?
4. Is the new process better than the current process?
5. Is the new process better than the current process by at least some predetermined threshold amount?

HYPOTHESIS TEST

TWO SAMPLE MEAN HYPOTHESIS TEST. DEFINITION 1.3

requirements (1) simple **independent** random samples, (2) CLT applies for **both** samples.

null hypothesis $\mu_1 = \mu_2$ or $\Delta\mu = 0$

alternative hypothesis (1) $\mu_1 \neq \mu_2$, (2) $\mu_1 < \mu_2$, or (3) $\mu_1 > \mu_2$ (**specify which is** μ_1 **&** μ_2**.**)

(We are not assuming $\sigma_1 = \sigma_2$.)

Test statistic

$$t = \frac{\Delta\bar{x} - \Delta\mu_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \qquad \Delta\bar{x} = \bar{x_1} - \bar{x_2},\ \Delta\mu_0 = \mu_1 - \mu_2 \tag{1}$$

where the degrees of freedom for the t distribution is:

$$df = \frac{(A+B)^2}{\frac{A^2}{n_1-1} + \frac{B^2}{n_2-1}} \tag{2}$$

where

$$A = \frac{s_1^2}{n_1} \qquad B = \frac{s_2^2}{n_2} \tag{3}$$

The df can be roughly approximated using:

$$df \approx \min\left(n_1 - 1,\ n_2 - 1\right) \tag{4}$$

Question 5. What distribution do you use to calculate the p-value?

R COMMAND

TWO SAMPLE MEAN HYPOTHESIS TEST:
`t.test(x1, x2, alternative="two.sided")`

- **x1 vector** of sample 1 data.
- **x2 vector** of sample 2 data.
- **alternative** H_a $\neq$:`"two.sided"`, $<$:`"less"`, $>$:`"greater"`)

Can only use `t.test` if you have the raw data!

DEFINITION 1.4 CONFIDENCE INTERVAL FOR $\Delta\mu$ (INDEPENDENT SAMPLES).

$$\text{CI for } \Delta\mu: \qquad \Delta\bar{x} \pm t_{\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \tag{5}$$

where

$$df \approx \min\left(n_1 - 1,\ n_2 - 1\right)$$

Back to our original example

Example 4. A student wants to test the hypothesis that the male college students are taller (on average) than female college students. The student records the heights of his fellow statistics class mates (in inches):

```
R: males
[1] 71 77 70 67 73 68 70
R: females
 [1] 65 68 66 68 65 62 68 62 69 65 63
```

Solving the problem: step-by-step

Step 0 known information: (I will make the males group 1)

Step 1 Test: two sample mean.

Step 2 Requirements: (1) simple independent random samples, (2) CLT applies to both samples.

Question 6. Have we satisfied the requirements?

Step 3 Hypothesis: $H_0 : \mu_1 = \mu_2$, $H_a : \mu_1 > \mu_2$. (H_a : mean height of males is greater than mean height of females.)

Step 4 Significance: $\alpha = 0.05$

Step 5 p-value: using R

```
R: res = t.test(males, females, alternative = "greater")
R: res
        Welch Two Sample t-test

data:  males and females
t = 3.6124, df = 10.277, p-value = 0.002272
alternative hypothesis: true difference in means is greater than ↩
    0
95 percent confidence interval:
 2.6539     Inf
sample estimates:
mean of x mean of y
   70.857    65.545
```

Question 7. What is the formal decision

Question 8. What is the final conclusion

Question 9. What are the point estimates for the males and female mean heights?

Question 10. How is our use of `t.test` different than the 1-sample case?

Question 11. If you are only given n, $\bar{x}$ and s for both samples, can you use the `t.test` ?

Find the 95% confidence interval for $\Delta\mu$:

CI for $\Delta\mu$: $\quad \Delta\bar{x} \pm t_{\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}, \qquad df \approx \min(n_1 - 1,\ n_2 - 1)$

```
R: x1.bar = mean(males)
R: s1 = sd(males)
R: n1 = length(males)
```

```
R: x2.bar = mean(females)
R: s2 = sd(females)
R: n2 = length(females)
R: delta.x.bar = x1.bar - x2.bar
R: df = min(n1 - 1, n2 - 1)
R: t.crit = qt(1 - 0.05/2, df)
R: E = t.crit * sqrt(s1^2/n1 + s2^2/n2)
```

	x1.bar	n1	s1	x2.bar	n2	s2	delta.x.bar	df	t.crit
1	70.86	7	3.34	65.55	11	2.50	5.31	6.00	2.45

```
R: delta.x.bar
[1] 5.3117
R: E
[1] 3.5979
R: CI = c(delta.x.bar - E, delta.x.bar + E)
R: CI
[1] 1.7137 8.9096
```

Question 12. How does the CI compare with our hypothesis tests conclusions?

REQUIRED SAMPLE SIZE

R COMMAND

t-TEST REQUIRED SAMPLE SIZE:
`power.t.test(delta=`h`, sd =`σ`, sig.level=`α`, power=1` $-$ β`, type ="two.sample", alternative="two.sided")`

- **delta** minimum effect size of interest
- **sd** estimated standard for both groups (assumed equal).
- **power** desired power
- **alternative** set to either "one.sided" or "two.sided"

1.4 2 population means: dep. samples / matched pairs

HYPOTHESIS TEST

Notation

If you have two dependent samples with sample 1 in x_i and sample 2 in y_i where x and y are **ordered vectors**:

$$\text{paired differences:} \quad d_i = x_i - y_i \tag{6}$$
$$\text{mean paired difference of sample:} \quad \bar{d} = \texttt{mean(d)} \tag{7}$$
$$\text{mean paired difference of population:} \quad \mu_d \tag{8}$$
$$\text{standard deviation of differences:} \quad s_d = \texttt{sd(d)} \tag{9}$$
$$\text{number of pairs:} \quad n_{\text{pairs}} \tag{10}$$

MATCHED PAIRS HYPOTHESIS TEST. DEFINITION 1.5

Think of the data as **one sample of differences**. Thus, it is equivalent to the one sample mean test.

requirements (1) simple **dependent / matched pair** random samples, (2) CLT applies to the **pairs** (population of paired differences are normally dist or $n > 30$).

null hypothesis $\mu_d = 0$

alternative hypothesis (1) $\mu_d \neq 0$, (2) $\mu_d < 0$, or (3) $\mu_d > 0$

Test statistic

$$t = \frac{\bar{d} - \mu_{d0}}{s_d/\sqrt{n}}, \qquad df = n - 1 \tag{11}$$

Note: n is the number of **pairs**.

Question 13. What distribution do you use to calculate the p-value?

TWO SAMPLE MEAN HYPOTHESIS TEST: R COMMAND

```
t.test(x, y, paired=TRUE, alternative="two.sided")
```

x ordered vector of sample 1 data.

y ordered vector of sample 2 data.

alternative H_a $\neq$:`"two.sided"`, $<$:`"less"`, $>$:`"greater"`)

Only difference from independent samples is the `paired=TRUE` **.**

Can only use `t.test` **if you have the raw data!**

CONFIDENCE INTERVAL FOR μ_d (INDEPENDENT SAMPLES). DEFINITION 1.6

$$\text{CI for } \mu_d: \qquad \bar{d} \pm t_{\alpha/2} \cdot \frac{s_d}{\sqrt{n}}, \qquad df = n - 1 \tag{12}$$

Is Friday the 13th Unlucky? Use the following data collected on the number of hospitals admissions resulting from motor vehicle crashes, and results are given below for Fridays on the 6th of a month and Fridays on the following 13th of the **same** month. Use a 0.05 significance level to test the claim that when the 13th day of a month falls on Friday, the numbers of hospital admissions from motor vehicle crashes are not affected. (Assume population differences are normally distributed.)

Friday the 6th	9	6	11	11	3	5
Friday the 13th	13	12	14	10	4	12

Question 14. What are the hypothesis?

Question 15. Conduct the hypothesis test. What is the p-value?

Question 16. What is the final conclusion?

Let Friday the 6th be y and the 13th be x, then positive d indicates more crashes on the 6th.

```
R: x = c(9, 6, 11, 11, 3, 5)
R: y = c(13, 12, 14, 10, 4, 12)
R: t.test(x, y, paired = TRUE)
        Paired t-test

data:  x and y
t = -2.7116, df = 5, p-value = 0.04219
alternative hypothesis: true difference in means is not equal to ↩
    0
95 percent confidence interval:
 -6.49328 -0.17339
sample estimates:
mean of the differences
                -3.3333
```

Conclusion: The sample data supports the claim that when the 13th day of the month falls on a Friday, the number of hospital admissions from motor vehicle crashes **is** affected.

Observation: since $\bar{d}$ is **negative** the sample data indicates **more** admissions on the 13th! What does this mean?

REQUIRED SAMPLE SIZE

R Command

t-TEST REQUIRED SAMPLE SIZE:
`power.t.test(delta=`h`, sd =`σ`, sig.level=`α`, power=1` $-$ β`, type ="paired", alternative="two.sided")`

- **delta** minimum effect size of interest
- **sd** estimated standard for both groups (assumed equal).
- **power** desired power
- **alternative** set to either "one.sided" or "two.sided"

1.5 Summary

Two sample mean hypothesis test (indep.)

requirements (1) simple **independent** random samples, (2) CLT applies for **both** samples.

null hypothesis $\mu_1 = \mu_2$ (**specify which is** μ_1 **&** μ_2**.**)

alternative hypothesis (1) $\mu_1 \neq \mu_2$, (2) $\mu_1 < \mu_2$, or (3) $\mu_1 > \mu_2$

test in R : `t.test(x1, x2 , alternative="two.sided")`

Where `x1, x2` are vectors of the sample 1 and sample 2 data.

CI :

$$\text{CI for } \Delta\mu: \qquad \Delta\bar{x} \pm t_{\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}, \qquad df \approx \min(n_1 - 1,\ n_2 - 1)$$

(We are not assuming $\sigma_1 = \sigma_2$.)

Matched pairs hypothesis test (dep.)

requirements (1) simple **dependent / matched pair** random samples, (2) CLT applies to the **pairs** (population of paired differences are normally dist or $n > 30$).

null hypothesis $\mu_d = 0$

alternative hypothesis (1) $\mu_d \neq 0$, (2) $\mu_d < 0$, or (3) $\mu_d > 0$

test in R : `t.test(x, y, paired=TRUE, alternative="two.sided")`

Where `x, y` are **ordered** vectors of the sample 1 and sample 2 data.

CI :

$$\text{CI for } \mu_d: \qquad \bar{d} \pm t_{\alpha/2} \cdot \frac{s_d}{\sqrt{n}}, \qquad df = n - 1$$

Note: n is the number of **pairs**.

1.6 Additional examples

Question 17. Use Data Set 1 (in Appendix B) to test the hypothesis that the mean height of males is greater than the mean height of females.

Question 18. Try book problem 9.4 question 14.