Math 3339

Section 27204
MWF 10-11:00am AAAud 2

Bekki George
bekki@math.uh.edu
639 PGH

Office Hours:
M & Th noon – 1:00 pm & T 1:00 – 2:00 pm
and by appointment

Matched Pairs t – Test ← Dependent Data → $\bar{x}_p \pm t^* \frac{s_p}{\sqrt{n}}$

In a paired-sample test, the two groups being tested are **not independent** and may even be the same group (e.g. before and after).

Example: Ten engineers' knowledge of basic statistical concepts was measured on a scale of 100 before and after a short course in statistical quality control. The engineers were selected at random. The table below shows the results of the tests:

Engineer	Before, y	After, x	improv $(x-y)$
1	43	51	8
2	82	84	2
3	77	74	-3
4	39	48	9
5	51	53	2
6	66	61	-5
7	55	59	4
8	61	75	14
9	79	82	3
10	43	48	5

W = Improv. amount

$\bar{W} = 3.9$

$S_w = \sqrt{31.21}$

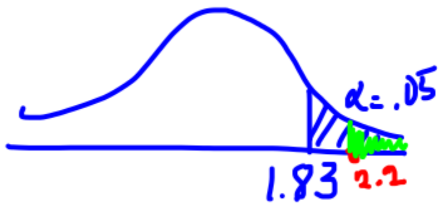
$\alpha = .05$

$H_0: \mu_w = 0$

$H_a: \mu_w > 0$

Let $w = x - y$. Then $\bar{w} = 3.9$ and $s_w^{\,2} = 31.21$. Test the claim that there is an improvement (use a 0.05 significance level).

$n = 10$, $df = 9$



$\Delta X \rightarrow \overline{w} = 3.9$ $\quad s_w = \sqrt{31.21}$

$H_0: \mu_w = 0$

$H_a: \mu_w > 0$

$$t = \frac{3.9 - 0}{\sqrt{31.21}/\sqrt{10}} = 2.2076$$

pvalue: $P(t > 2.2076) = 1 - pt(2.2076, 9) = .0273 < \alpha$

Based on $\alpha = .05$, we will reject the null hypothesis (no change in test score) in favor of saying there is an improvement in test scores.

The guidance office of a school wants to test the claim of an SAT test preparation company that students who complete their course will improve their SAT Math score by at least 50 points. Ten members of the junior class who have had no SAT preparation but have taken the SAT once were selected at random and agreed to participate in the study. All took the course and re-took the SAT at the next opportunity. The results of the testing indicated:

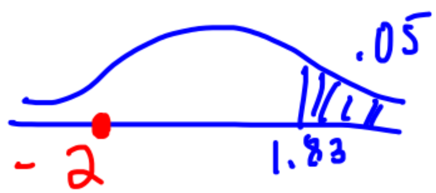
Student	1	2	3	4	5	6	7	8	9	10
Before	475	512	492	465	523	560	610	477	501	420
After	500	540	512	530	533	603	691	512	489	458

Is there sufficient evidence to support the prep course company's claim that scores will improve by at least 50 points at the 5% level of significance?

$\alpha = .05$

$H_0: \mu_D = 50 \quad (\leq)$

$H_a: \mu_D > 50$



$\bar{D} = 33.3 \qquad s_D = 26.36 \qquad t = \dfrac{33.3 - 50}{26.36/\sqrt{10}} = -2.001$

Fail to reject (pvalue ≈ .96)

But if $H_a: \mu_D > 0$ we would reject H_0

```
> t.test(d,alternative="greater")

        One Sample t-test

data:  d
t = 3.9902, df = 9, p-value = 0.001578
alternative hypothesis: true mean is greater than 0
95 percent confidence interval:
 18.00196      Inf
sample estimates:
mean of x
    33.3

> t.test(d,alternative="greater",mu=50)

        One Sample t-test

data:  d
t = -2.0011, df = 9, p-value = 0.9618
alternative hypothesis: true mean is greater than 50
95 percent confidence interval:
 18.00196      Inf
sample estimates:
mean of x
    33.3
```

Two (Independent) Sample Mean t – Test

The **two-sample t confidence interval for $\mu_1 - \mu_2$** with confidence level $100(1-\alpha)\%$ is then

t^*

$$(\bar{x} - \bar{y}) \pm t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$

$s^2/n = (s/\sqrt{n})^2$

A one-sided confidence bound can be calculated as described earlier.

The **two-sample t test** for testing $H_0\colon \mu_1 - \mu_2 = \Delta_0$ is as follows:

$$\text{Test statistic value: } t = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

0

$H_0\colon \mu_1 = \mu_2$

$H_a\colon \begin{cases} \mu_1 \neq \mu_2 \\ \mu_1 > \mu_2 \\ \mu_1 < \mu_2 \end{cases}$

Where the degrees of freedom is:

$$\nu = \text{df} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$ and round down to nearest integer

$$\frac{(SE_1^2 + SE_2^2)^2}{\frac{SE_1^4}{n_1 - 1} + \frac{SE_2^4}{n_2 - 1}}$$

st. error: $s/\sqrt{n} = se$ $\qquad s^2/n = se^2$

Example:

The void volume within a textile fabric affects comfort, flammability, and insulation properties. Permeability of a fabric refers to the accessibility of void space to the flow of a gas or liquid. The article "The Relationship Between Porosity and Air Permeability of Woven Textile Fabrics" (*J. of Testing and Eval.*, 1997: 108–114) gave summary information on air permeability (cm³/cm²/sec) for a number of different fabric types. Consider the following data on two different types of plain-weave fabric:

```
> se1=.79/sqrt(10)
> se2=3.59/sqrt(10)
> (se1^2+se2^2)^2/(se1^4/9+se2^4/9)
[1] 9.869602
```

Fabric Type	Sample Size	Sample Mean	Sample Standard Deviation	
Cotton	10 n_1	51.71 $\bar{x}_1$	.79 s_1	$SE_1 = .79/\sqrt{10}$
Triacetate	10 n_2	136.14 $\bar{x}_2$	3.59 s_2	$SE_2 = 3.59/\sqrt{10}$

Determine a 95% CI for the difference in mean.

$V = 9.8696 \Rightarrow df = 9$

$\rightarrow (51.71 - 136.14) \pm 2.262\sqrt{se1^2 + se2^2}$

$= 84.43 \pm 2.6296$

$= (-87.06, -81.80)$

Analysis of a random sample consisting of 20 specimens of cold-rolled steel (R) to determine yield strengths resulted in a sample average strength of 29.8 ksi. A second random sample of 25 two-sided galvanized (G) steel specimens gave a sample average strength of 34.7 ksi. It is also found that the samples gave standard deviations of 4.0 ksi (for cold-rolled) and 5.0 ksi (for two-sided galvanized). Assume that the two yield-strength distributions are normal. Does the date indicate that the corresponding true average yield strengths of the two methods are different? (Use 0.01 significance level) (α)

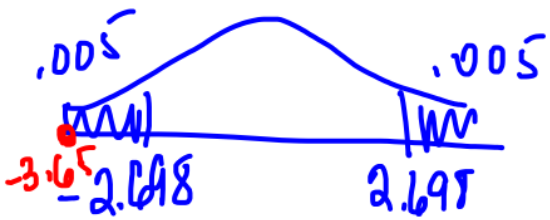
$\left(2/\sqrt{5}\right)^2 = \frac{4}{5}$

$H_0: \mu_R = \mu_G$

$H_a: \mu_R \neq \mu_G$

$\bar{X}_R = 29.8 \quad n_R = 20 \quad S_R = 4 \quad SE_R = 4/\sqrt{20} = 2/\sqrt{5}$

$\bar{X}_G = 34.7 \quad n_G = 25 \quad S_G = 5 \quad SE_G = 5/\sqrt{25} = 1$



$$df = \frac{\left(4/5 + 1\right)^2}{\left(\frac{(4/5)^2}{19} + \frac{(1)}{24}\right)} \doteq 42.999$$

$\downarrow = 42$

$$t = \frac{29.8 - 34.7}{\sqrt{\left(2/\sqrt{5}\right)^2 + (1)^2}} = -3.65$$

pvalue = .0007 < .01 $\Rightarrow$ Rejecting H_0

The true average tread lives of two competing brands of radial tires (brand X and brand Y) are known to be normally distributed. A sample of 45 brand X tires results in a sample mean of 42,500 and sample standard deviation of 2450. A sample of 45 brand Y tires results in a sample mean of 40,400 and sample standard deviation of 2150. Find a 95% confidence interval for the difference in the true means, mean of X minus mean of Y.

Popper 26

Suppose we compare the class averages for two classes on the same exams and get the following data:

Class	n	$\bar{x}$	s
A	25	88.4	4.3
B	36	86.7	1.9

1. What degrees of freedom will we use?
   a. 25
   b. 36
   c. 24
   d. 30
   e. none of these

2. Find the margin of error for a 90% CI
   a. 1.89
   b. 1.70
   c. 2.10
   d. 2.06
   e. none of these

3. Suppose we claim that class A has a higher average than class B. What is the alternate hypothesis?

a. H_a: $\mu_A \neq \mu_B$
b. H_a: $\mu_A > \mu_B$
c. H_a: $\mu_A < \mu_B$
d. H_a: $\mu_A = \mu_B$
e. none of these

4 + 5 = C