

Unit 7 – Hypothesis Testing
Week #11 - Practice Problems

SOLUTIONS

1. An independent testing agency was hired prior to the November 2004 election to study whether or not the work output is different for construction workers employed by the state and receiving prevailing wages versus construction workers in the private sector who are paid rates determined by the free market. A sample of 100 private sector workers reveals an average output of 74.3 parts per hour with a sample standard deviation of 16 parts per hour. A sample of 100 state workers reveals an average output of 69.7 parts per hour with a standard deviation of 18 parts per hour.

- (a) Is there evidence of a difference in productivity at the 0.10 level of significance?
- (b) Is there evidence of a difference in productivity at the 0.05 level of significance?
- (c) What is the achieved level of significance?

ANSWER

- a. The p-value is less than 0.1, so it is significant at 0.1 level.**
- b. The p-value is bigger than 0.05, so it is not significant at 0.05 level.**
- c. The achieved level is 0.058.**

SOLUTION -

This question is asking for a hypothesis test of the equality of two means in the setting of two independent groups (state v private)

Research Question. Is the work output of state workers is different from that of workers in the private sector?

Assumptions.

Let subscript “1” reference the group of state employees, “2” the private sector employees.

\bar{X}_1 is distributed Normal (μ_1 , $\sigma^2/100$) and \bar{X}_2 is distributed Normal (μ_2 , $\sigma^2/100$)

H_0 and H_A .

$$H_0 : \mu_1 = \mu_2$$

$$H_A : \mu_1 \neq \mu_2$$

Test statistic is a t-score.

$$t_{\text{score}} = \left[\frac{(\bar{X}_1 - \bar{X}_2) - E[(\bar{X}_1 - \bar{X}_2) | H_0 \text{ true}]}{\hat{SE}[(\bar{X}_1 - \bar{X}_2) | H_0 \text{ true}]} \right]$$

The estimated SE formula assumes equality of unknown variances

$$\hat{SE}(\bar{X}_1 - \bar{X}_2) = \sqrt{\frac{S_{\text{pool}}^2}{n_1} + \frac{S_{\text{pool}}^2}{n_2}} \quad \text{where } S_{\text{pool}}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

For these data:

$$\hat{\sigma}^2 = S_{\text{pool}}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{(100 - 1)18^2 + (100 - 1)16^2}{(100 - 1) + (100 - 1)} = 290$$

$$\hat{SE}(\bar{X}_1 - \bar{X}_2) = \sqrt{\frac{S_{\text{pool}}^2}{n_1} + \frac{S_{\text{pool}}^2}{n_2}} = \sqrt{\frac{290}{100} + \frac{290}{100}} = 2.4083$$

$$\text{Degrees of freedom} = (n_1 - 1) + (n_2 - 1) = (100 - 1) + (100 - 1) = 198.$$

“Evaluation” rule.

The likelihood of these findings or ones more extreme if H_0 is true is

$$\text{p-value} = \Pr[(\bar{X}_1 - \bar{X}_2) \geq |(69.7 - 74.3)| | H_0 \text{ true}].$$

Calculations.

$$\text{p-value} = (2)\Pr[(\bar{X}_1 - \bar{X}_2) \geq |(69.7 - 74.3)|] \quad \text{note – The (2) is in front because this is two sided}$$

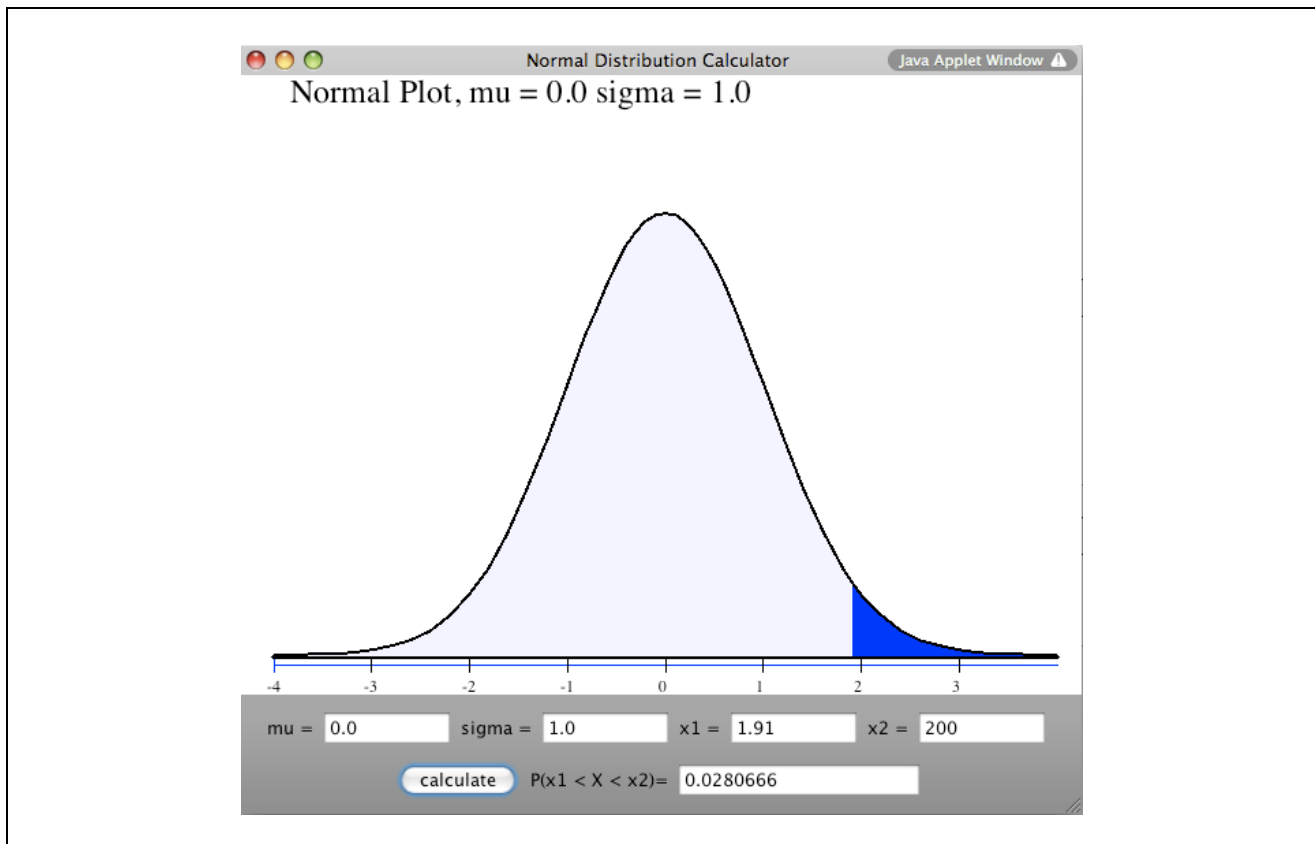
$$= 2\Pr\left[\frac{(\bar{X}_1 - \bar{X}_2) - (0)}{\hat{SE}(\bar{X}_1 - \bar{X}_2)} \geq \frac{(69.7 - 74.3) - (0)}{2.4083}\right]$$

$$= (2)\Pr[t_{\text{score}} \geq 1.91] \quad \text{where degrees of freedom} = 198$$

$= (2)(.028) = .056$ note – I used the Normal(0,1) table as degrees of freedom is so large

“Evaluate”.

Under the null hypothesis H_0 (worker output is the same in both groups) the chances that the average work outputs differ by a magnitude greater than $|69.7 - 74.3|$ is about 6 in 100. This is a borderline suggestion that the two groups differ in their work output.



2. For the data in Exercise 1, what level of significance is achieved by the data if the sample means and sample standard deviations are unchanged but the within group sample sizes are

- (a) both equal to 10
- (b) both equal to 200
- (c) Comment on the role of sample size in the probability of a type I error.

ANSWER

- a. p-value = .554
- b. p-value = .007
- c. All other things equal, a larger sample size reduces type I error.

SOLUTION

The solution involves substitution of the new values of the sample sizes into the formulae shown in the solution for Exercise 1.

a. n=10 in each group**Two-Sample T-Test and CI**

Sample	N	Mean	StDev	SE Mean
1	10	74.3	16.0	5.1
2	10	69.7	18.0	5.7

Difference = mu (1) - mu (2)

Estimate for difference: 4.60000

95% CI for difference: (-11.46788, 20.66788)

T-Test of difference = 0 (vs not =): T-Value = 0.60 P-Value = 0.554 DF = 17

b. n=200 in each group**Two-Sample T-Test and CI**

Sample	N	Mean	StDev	SE Mean
1	200	74.3	16.0	1.1
2	200	69.7	18.0	1.3

Difference = mu (1) - mu (2)

Estimate for difference: 4.60000

95% CI for difference: (1.25196, 7.94804)

T-Test of difference = 0 (vs not =): T-Value = 2.70 P-Value = 0.007 DF = 392

3. Halcion is a sleeping pill that is relatively rapidly metabolized by the body and therefore having fewer hangover effects the next morning, compared to other sleeping pills. Opponents of Halcion argue that, because this agent is so rapidly metabolized by the body, patients do not sleep as long with this drug as with Dalmane. Data on 10 insomniacs, each of whom took Dalmane on one occasion and Halcion on a second, is collected. The variable measured is number of hours of sleep:

Patient	Number of Hours Sleep with	
	Dalmane	Halcion
1	4.58	3.97
2	5.19	4.88
3	3.94	4.09
4	6.32	5.87
5	7.68	6.93
6	3.48	4.00
7	5.72	5.08
8	7.04	6.95
9	5.27	4.96
10	5.84	5.13

Do these data suggest that Halcion is not as effective as Dalmane with respect to number of hours of sleep? Carry out an appropriate statistical test and interpret your findings.

ANSWER

Yes, a paired t-test suggests that the average difference in hours slept (Dalmane – Halcion) = 0.32 is statistically significant (one sided p-value = .018).

SOLUTION

This question is asking for a hypothesis test of the equality of two means in the setting of paired data. The data are paired because each participant was measured on two occasions, once on Dalmane and once on Halcion

Research Question. Are sleep durations shorter on Dalmane than on Halcion?

Assumptions.

\bar{d} is distributed Normal ($\mu_d, \sigma_d^2/10$)

Differences are calculated as (Dalmane – Halcion)

For these 10 paired measurements, we have

Obs	dalmane	halcion	diff
1	4.58	3.97	0.61
2	5.19	4.88	0.31
3	3.94	4.09	-0.15
4	6.32	5.87	0.45
5	7.68	6.93	0.75
6	3.48	4.00	-0.52
7	5.72	5.08	0.64
8	7.04	6.95	0.09
9	5.27	4.96	0.31
10	5.84	5.13	0.71

H_0 and H_A .

$$H_0 : \mu_d = 0$$

$H_A : \mu_d > 0$ ("Dalmane is better than Halcion) – one sided

Test statistic is a t-score.

To obtain the t-score for the variable DIFF, it is necessary to obtain the sample mean, sample variance, and sample standard error. SAS gives

The MEANS Procedure			
Analysis Variable : diff diff: Dalmane - Halcion			
N	Mean	Std Dev	Std Error
10	0.3200000	0.4109609	0.1299573

$$t_{\text{score}} = \left[\frac{(\bar{d}) - E[\bar{d}]|H_0 \text{ true}}{\hat{S}\hat{E}[(\bar{d})|H_0 \text{ true}]} \right] \text{ where } \hat{S}\hat{E}(\bar{d}) = \frac{S_d}{\sqrt{n}}$$

$$= \left[\frac{0.32 - 0}{0.4109609/\sqrt{10}} \right]$$

$$= [2.46]$$

Degrees of freedom = $(n-1) = (10-1) = 9$.

“Evaluation” rule.

The likelihood of these findings or ones more extreme if H_0 is true is

$$p\text{-value} = \Pr\left[\left(\bar{d}\right) \geq 0.32 \mid H_0 \text{ true}\right].$$

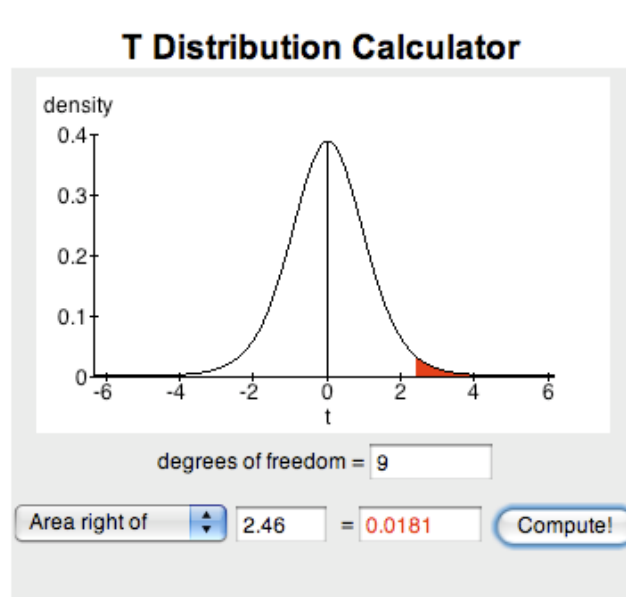
Calculations.

$$p\text{-value} = \Pr\left[t_{\text{score}} \geq 2.46\right] \quad \text{where degrees of freedom} = 9$$

$$=.018$$

If you want to use a student’s t-distribution calculator on the internet, one choice is

<http://www.stat.tamu.edu/~west/applets/tdemo.html>



How it works: The calculator above takes the place of the traditional textbook table. First, enter the appropriate number of degrees of freedom in the top box. Then, the calculator can be used in two ways. To find t critical values, enter a probability in the "=" box and hit "Compute!". The answer is displayed in the "Area right of" box. To Find tail probabilities (or p-values), enter the t value in the "Area right of" box and hit "Compute!". The probability will be displayed in the "=" box. In either case, the probability is represented graphically.

“Evaluate”.

Under the null hypothesis H_0 (duration of sleep is the same with both drugs) the chance that the difference in average hours slept is as great or greater than 4.6 is about 2 in 100. This is statistically significant.

For SAS Users: note – this fragment is a more extensive than what I actually needed – cb.

```
data temp;
  input patid dalmane halcion;
  diff = dalmane - halcion;
  label diff="diff: Dalmane - Halcion";
  cards;
  1 4.58 3.97
  2 5.19 4.88
  3 3.94 4.09
  4 6.32 5.87
  5 7.68 6.93
  6 3.48 4.00
  7 5.72 5.08
  8 7.04 6.95
  9 5.27 4.96
  10 5.84 5.13
  ;
run;
quit;
proc print data=temp;
  var dalmane halcion diff;
  title "Week 12 #3";
  title2 "Differences Dalmane - Halcion";
run;
quit;

proc means data=temp n mean var std stderr;
  var diff;
  title "Statistics on diff=[dalmane] - [halcion]";
run;
quit;

proc ttest data=temp alpha=.01;
  Title "Paired T-test";
  title2 "Difference is [Dalmane] - [Halcion]";
  paired dalmane*halcion;
run;
quit;
```


For Minitab Users**STAT > BASIC STATISTICS > PAIRED t****Paired T-Test and CI: Dalmane, Halcion**

Paired T for Dalmane - Halcion

	N	Mean	StDev	SE Mean
Dalmane	10	5.50600	1.31104	0.41459
Halcion	10	5.18600	1.09882	0.34748
Difference	10	0.320000	0.410961	0.129957

95% lower bound for mean difference: 0.081774

T-Test of mean difference = 0 (vs > 0): T-Value = 2.46 P-Value = 0.018

4. For the Halcion versus Dalmane data in Exercise 3, construct a 99% confidence interval estimate of discrepancy in the efficacies of the two drugs. Compare this to the acceptance region that would have been obtained had you constructed a statistical test with type I error pre-specified at 0.01.

ANSWER The 99% confidence interval is (-0.10, 0.74).

The acceptance region is $\bar{d} < 0.3666$.

SOLUTION

Solution for the 99% CI is as follows

$$\bar{d} = 0.32 \quad \hat{SE}(\bar{d}) = \frac{S_d}{\sqrt{n}} = 0.1299573$$

$$df = (n-1) = 9 \quad t_{1-\alpha/2; df} = t_{.995; 9} = 3.25 \text{ from the calculator on the web above.}$$

$$99\% \text{ CI for } \mu_d = \bar{d} \pm (t_{.995; DF=9}) \hat{SE}(\bar{d})$$

$$= (0.32) \pm (3.25)(0.1299573) = (-0.1024, +0.7424)$$

Solution for the acceptance region of a one sided test with alpha = .01 is obtained by reasoning as follows

Rejection occurs for t-score \geq the 99th percentile of a student's t on df=9 \rightarrow

Rejection occurs for t-score $\geq t_{.99; df=9} \rightarrow$

Acceptance occurs for t-score $< t_{.99; df=9} \rightarrow$

Substituting in the definition of a t-score allows us to write this equivalently as

Acceptance occurs for $\frac{\bar{d}-0}{\hat{SE}(\bar{d})} < t_{.99,df=9}$ By plugging in numbers for the SE and the percentile \rightarrow

Acceptance occurs for $\frac{\bar{d}}{0.1299573} < 2.821$ where I used the calculator on the web as before \rightarrow

Acceptance occurs for $\bar{d} < 0.3666$

Comparison

These two regions overlap but are not identical. They are not identical because the confidence interval is two sided whereas the acceptance region is one sided.