

t-worksheet

(Session 22 Group Quiz)

July and August are the strongest months for monsoon rains in Tucson. Your friend suspects that more rain takes place in July.

(Do Not Repeat Solutions) You have already stated a hypothesis appropriate to your friend's suspicion.

(Do Not Repeat Solutions) For the years 2007 to 2016, here are the data for rainfall total in inches.

	2007	2008	2009	2010	2011	2012	2013	2104	2015	2016
July	4.38	4.03	2.15	1.44	1.69	0.71	1.44	4.58	2.06	3.32
August	4.27	1.69	0.86	2.89	1.03	0.00	1.85	3.32	1.80	1.09

You have already determined the test procedure that you will use in this case.

t-worksheet t-procedures

- Compute the appropriate t statistics for the hypothesis stated in the quiz. Show the calculation you would do by hand and show the command to do this in R.

Calculation by hand on next page

```
#The data has been entered for you!
july <- c(4.38,4.03,2.15,1.44,1.69,0.71,1.44,4.58,2.06,3.32)
august <- c(4.27,1.69,0.86,2.89,1.03,0.00,1.85,3.32,1.80,1.09)
```

```
# Command to perform this test in R
t.test(july, august, paired = TRUE, alternative = "greater")
```

```
##
## Paired t-test
##
## data: july and august
## t = 1.9099, df = 9, p-value = 0.04424
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
## 0.02813084 Inf
## sample estimates:
## mean of the differences
## 0.7
```

$j-a$	$\bar{x}-(j-a)$	$(\bar{x}-(j-a))^2$
0.11	0.59	0.349
2.34	-1.64	2.689
1.29	-0.59	0.3481
-1.45	2.15	4.622
0.66	0.04	0.0016
0.71	-0.01	0.0001
-0.41	1.11	1.2321
1.26	-0.56	0.3136
0.26	0.44	0.1936
2.23	-1.53	2.3409
7		12.0902

$$\bar{x} = \frac{7}{10} = 0.7 \quad sd = \sqrt{\frac{12.0902}{9}} = 1.159$$

$$t = \frac{\bar{x}_{j-a}}{s_{j-a}} \cdot \sqrt{n} = \frac{0.7}{1.15903} \cdot \sqrt{10}$$

$$= 1.9098$$

$$p\text{-value} = 1 - \Phi(t, df) = \boxed{0.05} \text{ approx}$$

will actually
be a bit
less

Figure 1: By hand

- b. What is the conclusion concerning your hypothesis? Use your answer in part (a) to explain your conclusion.

When we perform the test, we see that the p-value that we get is 0.04424. This means that we can reject the null hypothesis since the p-value is 0.044 which is less than our alpha which is 0.05

- c. If you incorrectly used the two-sample t procedure, how does the p -value change?

If we incorrectly use the two-sample t procedure, i.e., do not specify `paired = TRUE`, then the p value that we get will be significantly higher than the correct test. This means that it will be much harder to reject the null hypothesis. This is because we are not making use of the fact that the variables are actually matched. Since we are essentially discarding an important identity of the data, this makes the test much weaker. Therefore, since we are ignoring the fact that the variables are matched, the p -value we get will be significantly higher than before.

```
t.test(july, august, alternative = "greater")

##
##  Welch Two Sample t-test
##
## data:  july and august
## t = 1.1739, df = 17.894, p-value = 0.1279
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.3343955      Inf
## sample estimates:
## mean of x mean of y
##      2.58      1.88
```