# Chapter 9: Comparing two paired populations
## (Ott & Longnecker Sections: 6.4)

https://dzwang91.github.io/stat324/

- Musculoskeletal disorders of the neck and shoulders are common in office workers because of repetitive tasks. Long periods of upper-arm elevation above 30 degrees have been shown to be related to disorders. It was thought that varying working conditions over the course of the day could alleviate some of these problems.
- "Can the diversification change the percentage of time where arm angle is below 30 degrees? "

- Musculoskeletal disorders of the neck and shoulders are common in office workers because of repetitive tasks. Long periods of upper-arm elevation above 30 degrees have been shown to be related to disorders. It was thought that varying working conditions over the course of the day could alleviate some of these problems.
- "Can the diversification change the percentage of time where arm angle is below 30 degrees? "
- 8 office workers were randomly selected. They were observed for one work day under the standard conditions, and the percentage of time that their dominant upper-arm was below 30 degrees was recorded. The next day, these same individuals had their work diversified, and again were observed.

- The results are:

| Participant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Diverse | 78 | 91 | 79 | 65 | 67 | 72 | 71 | 96 |
| Standard | 81 | 87 | 86 | 59 | 66 | 70 | 73 | 92 |

- Dotplots connected for each individual:
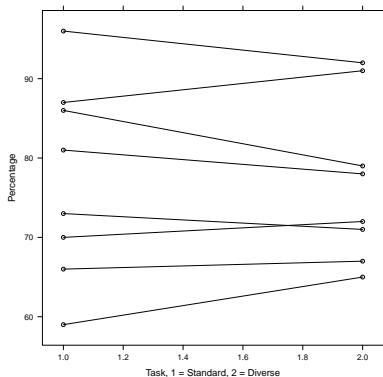


Some go up, some go down, and some stay the same. It's unlikely that we will see much difference.

- Let $\mu_{diverse}$ be true mean percent of time in the diverse condition and $\mu_{standard}$ be the true mean percent of time in the standard condition.

- Let $\mu_{diverse}$ be true mean percent of time in the diverse condition and $\mu_{standard}$ be the true mean percent of time in the standard condition.
- We want to test:

$$H_0 : \mu_{diverse} - \mu_{standard} = 0 \text{ vs. } H_A : \mu_{diverse} - \mu_{standard} \neq 0$$

- Let $\mu_{diverse}$ be true mean percent of time in the diverse condition and $\mu_{standard}$ be the true mean percent of time in the standard condition.
- We want to test:

$$H_0 : \mu_{diverse} - \mu_{standard} = 0 \text{ vs. } H_A : \mu_{diverse} - \mu_{standard} \neq 0$$

- Can we use two-sample t test or Welch t test?

- Let $\mu_{diverse}$ be true mean percent of time in the diverse condition and $\mu_{standard}$ be the true mean percent of time in the standard condition.
- We want to test:

$$H_0 : \mu_{diverse} - \mu_{standard} = 0 \text{ vs. } H_A : \mu_{diverse} - \mu_{standard} \neq 0$$

- Can we use two-sample t test or Welch t test?
  NO, because of the dependence. The data is clearly paired, since we are taking two measurements on each individual.

# Paired t test

- The data consists of paired observations. Let:
    - $X_{1,i}$ = i-th data point from population 1
    - $\mu_1$ = true mean of population 1
    - $X_{2,i}$ = i-th data point from population 2
    - $\mu_2$ = true mean of population 2
    - $D_i = X_{1,i} - X_{2,i}$ = the difference for pair $i$
    - $n$ = number of pairs
- We wish to test: $H_0 : \mu_1 - \mu_2 = \delta$ vs. $H_A : \mu_1 - \mu_2 \neq \delta$.

- If we can assume:
  - All of the *pairs* are indepdendent.
  - The differences follow a normal distribution. (how do we check?)

  Then the test statistic is:
  $$t = \frac{\bar{D} - \delta}{\frac{S_D}{\sqrt{n}}}$$

  Where $\bar{D}$ is the mean of the differences and $S_D$ is the sample standard deviation of the differences.

- Under $H_0$, test statistic is $T$ distribution with $n - 1$ degrees of freedom. (what is p-value and rejection region?)

- If we can assume:
    - All of the *pairs* are indepdendent.
    - The differences follow a normal distribution. (how do we check?)

  Then the test statistic is:
  $$t = \frac{\bar{D} - \delta}{\frac{S_D}{\sqrt{n}}}$$
  Where $\bar{D}$ is the mean of the differences and $S_D$ is the sample standard deviation of the differences.

- Under $H_0$, test statistic is $T$ distribution with $n-1$ degrees of freedom. (what is p-value and rejection region?)

- Paired t test is indeed one sample t test on the differences!

- diverse - standard: -3, 4, -7, 6, 1, 2, -2, 4.

# Back to example

- diverse - standard: -3, 4, -7, 6, 1, 2, -2, 4.
- check normality of the differences:



**QQ Plot of Differences**
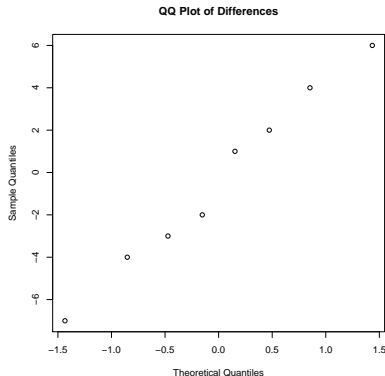
This looks quite straight, so we can assume normality.

- For the example, $\bar{D} = 0.63$ and $S_D = 4.34$, so observation of test statistic is

$$t_{obs} = \frac{0.63}{\frac{4.34}{\sqrt{8}}} = 0.41.$$

- For the example, $\bar{D} = 0.63$ and $S_D = 4.34$, so observation of test statistic is

$$t_{obs} = \frac{0.63}{\frac{4.34}{\sqrt{8}}} = 0.41.$$

- Compare to a $T_7$ distribution, p-value is about 0.702. Thus we would not reject. It seems that diversification of work has little effect on arm angle.

```
> rm(list=ls())
>
> #enter data#
> standard <- c(81, 87, 86, 59, 66, 70, 73, 92)
> diverse  <- c(78, 91, 79, 65, 67, 72, 71, 96)
>
> #reformat#
> all <- c(standard, diverse)
> task <- c(rep(1, 8), rep(2, 8))
> person <- c(1:8, 1:8)
```

Figure: make a dataset.

```
> #line graph#
> library(lattice)
>
> xyplot(all ~ task, group = person, type = 'o', col = 1, ylab = "Percentage",
xlab = "Task, 1 = Standard, 2 = Diverse")
```
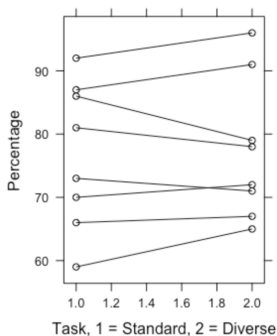


Figure: draw dotplot.

```
> #qq of diffs#
> diffs <- diverse - standard
> qqnorm(diffs, main = "QQ Plot of Differences")
```
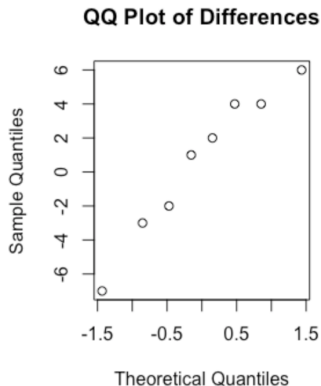


Figure: draw QQ plot.

```
> #Paired t test#
> #Output format is slightly different but these do the same thing#
> t.test(diverse, standard, paired = T)

        Paired t-test

data:  diverse and standard
t = 0.40728, df = 7, p-value = 0.696
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.003686  4.253686
sample estimates:
mean of the differences
                  0.625


> t.test(diffs)

        One Sample t-test

data:  diffs
t = 0.40728, df = 7, p-value = 0.696
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -3.003686  4.253686
sample estimates:
mean of x
    0.625
```

Figure: paired t test.

**1** Paired t test

**2** Paired sign test

- The data consists of paired observations. Let:
    - $X_{1,i}$ = i-th data point from population 1
    - $m_1$ = true median of population 1
    - $X_{2,i}$ = i-th data point from population 2
    - $m_2$ = true median of population 2
    - $D_i = X_{1,i} - X_{2,i}$ = the difference for pair $i$
    - $n$ = number of pairs
    - $\sigma_D^2$ = true variance of the differences
- We wish to test: $H_0 : m_1 - m_2 = 0$ vs. $H_A : m_1 - m_2 \neq 0$.

- The data consists of paired observations. Let:
  - $X_{1,i}$ = i-th data point from population 1
  - $m_1$ = true median of population 1
  - $X_{2,i}$ = i-th data point from population 2
  - $m_2$ = true median of population 2
  - $D_i = X_{1,i} - X_{2,i}$ = the difference for pair $i$
  - $n$ = number of pairs
  - $\sigma_D^2$ = true variance of the differences
- We wish to test: $H_0 : m_1 - m_2 = 0$ vs. $H_A : m_1 - m_2 \neq 0$. Equivalently, let $m_D = m_1 - m_2$,

$$H_0 : m_D = 0, \text{ vs. } H_A : m_D \neq 0$$

# Paired sign test

- If $D_1, \ldots, D_n$ are an i.i.d. sample, then test statistic is

  $B =$ number of data values greater than 0 (Ignore values tied with 0).

- Note that if $H_0$ is true, then $B \sim \text{Binomial}(n^*, 0.5)$, where $n^*$ is the number of data points not equal to 0.

- If $D_1, \ldots, D_n$ are an i.i.d. sample, then test statistic is

  $B =$ number of data values greater than 0 (Ignore values tied with 0).

- Note that if $H_0$ is true, then $B \sim \text{Binomial}(n^*, 0.5)$, where $n^*$ is the number of data points not equal to 0.

- P-value: Let $b$ be the observation of $B$, then for two-sided alternative,

$$\text{p-value} = 2\min\{P(B \geq b), P(B \leq b)\}.$$

For paired populations, take difference to remove the dependence and then use one sample tests.