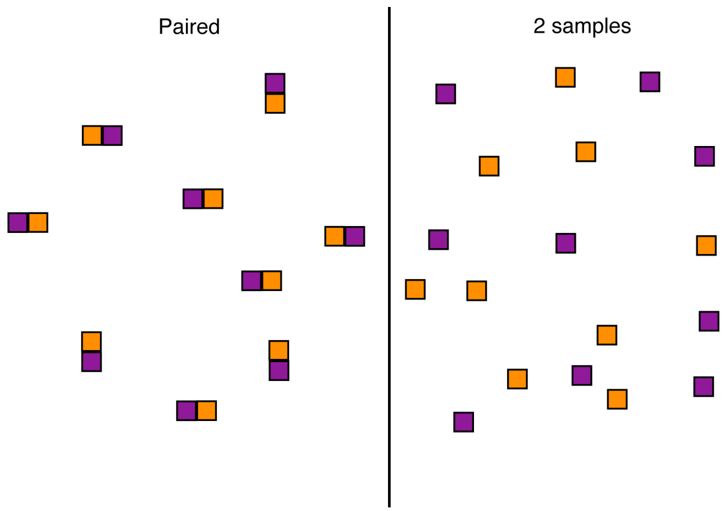
Comparing the means of two groups

Chapter 12

Comparing means

- Tests with one categorical and one numerical variable
- *Goal*: to compare the mean of a numerical variable for different groups.

Paired vs. 2 sample comparisons



Paired comparisons allow us to account for a lot of extraneous variation.

2-sample methods are sometimes easier to collect data for.

Paired design: Examples

- Before and after treatment
- Upstream and downstream of a power plant
- Identical twins: one with a treatment and one without
- Earwigs in each ear: how to get them out? Compare tweezers to hot oil

Paired designs

- Data from the two groups are paired
- Each member of the pair shares much in common with the other, *except* for the tested categorical variable
- There is a one-to-one correspondence between the individuals in the two groups

Paired comparisons

- We have many pairs
- In each pair, there is one member that has one treatment and another who has another treatment

("Treatment" can mean "group")

Paired comparisons

To compare two groups, we use the mean of the *difference* between the two members of each pair

Paired t test

- Compares the mean of the differences to a value given in the null hypothesis
- For each pair, calculate the difference.
- The paired t-test is simply a one-sample t-test on the differences.

Example: Emergency room admissions on 4/20

- Counted emergency room admissions in Vancouver on April 20
- Compared to average admissions one week before and after
- Each data point is a year

Staples, J.A., et al. 2020. Emergency department visits during the 4/20 cannabis celebration *Emergency Medicine Journal* 37:187-192.

Data

year	ERvisits420	ERvisitsControls
2009	269	306.5
2010	289	283
2011	292	301.5
2012	338	304
2013	326	308
2014	340	333.5
2015	429	367
2016	404	341.5
2017	402	383
2018	379	334

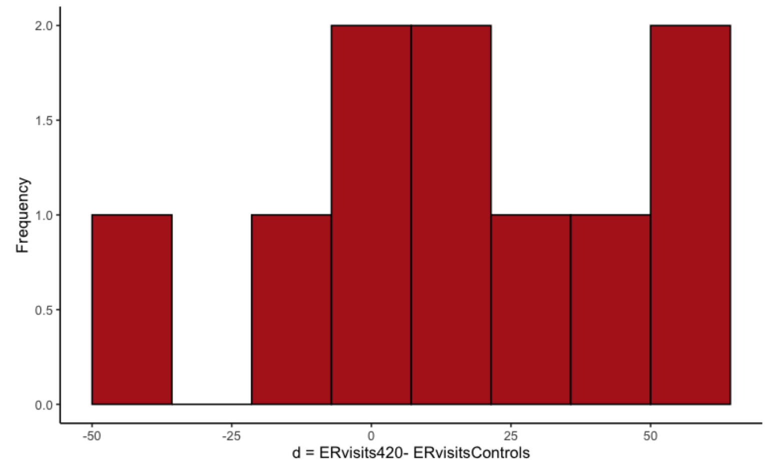
Hypotheses

H_0: ER admissions are the same on average on 4/20 as on control days.

H_A: ER admissions are different on average on 4/20 compared to control days.

Calculate differences

year	ERvisits420	ERvisitsControls	Difference (d)
2009	269	306.5	−37.5
2010	289	283	6
2011	292	301.5	−9.5
2012	338	304	34
2013	326	308	18
2014	340	333.5	6.5
2015	429	367	62
2016	404	341.5	62.5
2017	402	383	19
2018	379	334	45



Calculate t using d's

$$\bar{d} = 20.6$$
$$s_d^2 = 994.16$$
$$n = 10$$

$$t = \frac{20.6 - 0}{\sqrt{994.16/10}} = 2.06$$

CAUTION!

- The number of data points in a paired t test is the number of *pairs*. -- *Not* the number of individuals
- Degrees of freedom = Number of pairs - 1

Critical value of *t*

$$t_{0.05(2),9} = 2.26$$
$$t = 2.06 < 2.26$$

So we cannot reject the null hypothesis. There is insufficient evidence that ER visits increase on 4/20.

R for paired *t*-test

```
> t.test(ER420data$ERvisits420,
ER420data$ERvisitsControls, paired = TRUE)

      Paired t-test

data:  ER420data$ERvisits420 and
ER420data$ERvisitsControls
t = 2.066, df = 9, p-value = 0.0688
alternative hypothesis: true difference in means
is not equal to 0
95 percent confidence interval:
 -1.955369 43.155369
sample estimates:
mean of the differences
                   20.6
```

Assumptions of paired *t* test

- Pairs are chosen at random
- The differences have a normal distribution

It does *not* assume that the individual values are normally distributed, only the differences.

Comparing the means of two groups

Hypothesis test: 2-sample *t* test

Estimation:
Difference between two means

$$\bar{Y}_1 - \bar{Y}_2$$

Confidence interval: $\left(\bar{Y}_1 - \bar{Y}_2\right) \pm SE_{\bar{Y}_1 - \bar{Y}_2}\, t_{\alpha(2),df}$

Standard error of difference in means

$$SE_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$$

Pooled variance: $s_p^2 = \dfrac{df_1 s_1^2 + df_2 s_2^2}{df_1 + df_2}$

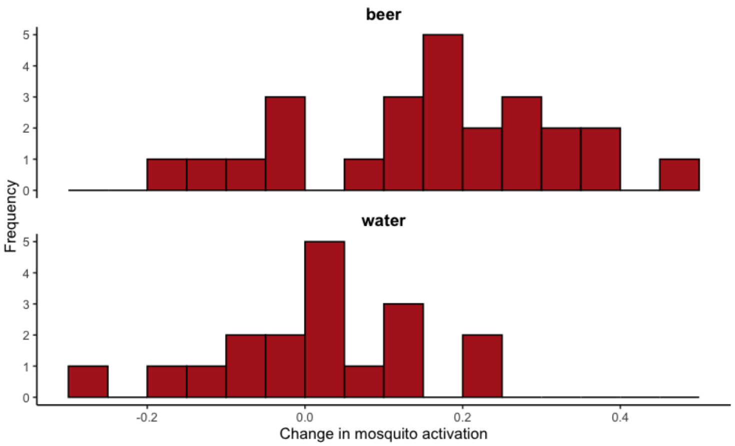
$df_1 = n_1 - 1$; $df_2 = n_2 - 1$

Malaria, mosquitoes and beer

Mosquitoes find prey by odor

Are mosquito biting rates affected by beer consumption?

Anopheles gambiae



Lefèvre, T. et al. 2010. Beer consumption increases human attractiveness to malaria mosquitoes. PLoS ONE 5: e9546.

https://commons.wikimedia.org/wiki/File:Zanzara_Anopheles_Gambiae.jpg

Data, summarized

	Beer	Water
Mean change in mosquitoes	0.1544	0.0078
SD of change in mosquitoes	0.1623	0.1269
Sample size	25	18

Both distributions are approximately normal.

Calculating the standard error

$df_1 = 25 - 1 = 24$; $df_2 = 18 - 1 = 17$

$$s_p^2 = \frac{df_1 s_1^2 + df_2 s_s^s}{df_1 + df_2} = \frac{24(0.1623)^2 + 17(0.1269)^2}{24 + 17} = 0.0221$$

$$SE_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = \sqrt{\frac{0.0221}{25} + \frac{0.0221}{18}} = 0.0459$$

Finding *t*

$df = df_1 + df_2 = n_1 + n_2 - 2$

$= 25 + 18 - 2$
$= 41$

$$t_{(0.05),41} = 2.02$$

The 95% confidence interval of the difference in the means

$$(\bar{Y}_1 - \bar{Y}_2) \pm SE_{\bar{Y}_1 - \bar{Y}_2} t_{\alpha(2),df}$$
$$= (0.1544 - 0.0078) \pm 0.0459(2.02)$$
$$= 0.1466 \pm 0.0990$$

$$0.0538 < \mu < 0.239$$

Testing hypotheses about the difference in two means

2-sample *t*-test

The *two sample* t-*test* compares the means of a numerical variable between two populations.

2-sample t-test

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{SE_{\bar{Y}_1 - \bar{Y}_2}}$$

Hypotheses

H_0: There is no difference between the mosquito activation between beer and water drinkers. ($\mu_1 = \mu_2$)

H_A: The beer and water drinkers differ in their mean mosquito activation. ($\mu_1 \neq \mu_2$)

Calculating *t*

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{SE_{\bar{Y}_1 - \bar{Y}_2}} = \frac{(0.1544 - 0.0078)}{0.0459} = 3.19$$

Drawing conclusions...

$t_{0.05(2),41} = 2.02$

$t > 2.02$, so we reject the null hypothesis.

These data suggest that beer drinkers attract more mosquitoes than water drinkers.

Assumptions of two-sample *t*-tests

- Both samples are random samples.
- Both populations have normal distributions
- The variance of both populations is equal.

2-sample *t*-test in R

```
t.test(change ~ drink, data = beerMosquitoData,
      var.equal = TRUE)
```

```
Two Sample t-test

data: change by drink
t = 3.1913, df = 41, p-value = 0.002717
alternative hypothesis: true difference in means is not
equal to 0
95 percent confidence interval:
0.05383517 0.23940928
sample estimates: mean in group beer mean in group water
                         0.154400000         0.007777778
```

Welch's *t*

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$

Relaxing equal variance assumption: Welch's *t*-test

Welch's *t*-test compares the means of two groups without requiring the assumption of equal variance.

Welch's *t* for mosquito and beer

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{0.1544 - 0.0078}{\sqrt{\frac{0.02632}{25} + \frac{0.01611}{18}}} = 3.32$$

Degrees of freedom

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}} = \frac{\left(\frac{0.0263}{25} + \frac{0.0161}{18}\right)^2}{\frac{\left(0.0263/25\right)^2}{25-1} + \frac{\left(0.0161/18\right)^2}{18-1}} = 40.66$$

Note that the *df* are not necessarily an integer here.

Reaching a conclusion

$t_{0.05(2),\ 40.6} = 3.32$

$t = 3.32 > 2.02$

So we can reject the null hypothesis with $P<0.05$.

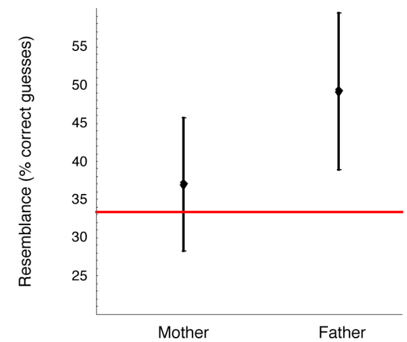
Beer attract mosquitoes more than water.

Welch's *t*-test in R

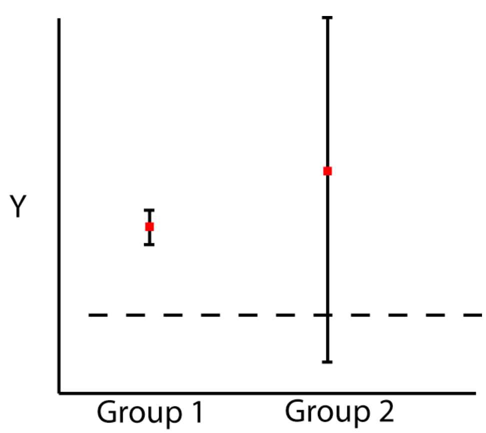
```
t.test(change ~ drink, data = beerMosquitoData,
      var.equal = FALSE)

      Welch Two Sample t-test

data:  change by drink
t = 3.3219, df = 40.663, p-value = 0.001897
alternative hypothesis: true difference in means is not
equal to 0
95 percent confidence interval:
 0.05746134 0.23578311
sample estimates:
 mean in group beer mean in group water
            0.154400000    0.007777778
```

The wrong way to make a comparison of two groups

"Group 1 is significantly different from a constant, but Group 2 is not. Therefore Group 1 and Group 2 are different from each other."



A more extreme case...



Comparing means when variances are not equal

Welch's *t* test

Welch's approximate t-test compares the means of two normally distributed populations that have unequal variances.

Comparing the variance of two groups

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_A : \sigma_1^2 \neq \sigma_2^2$$

Levene's test

Compares the variances of two (or more) groups

```
leveneTest(data = titanicData, age ~ 
survive, center = mean)
```

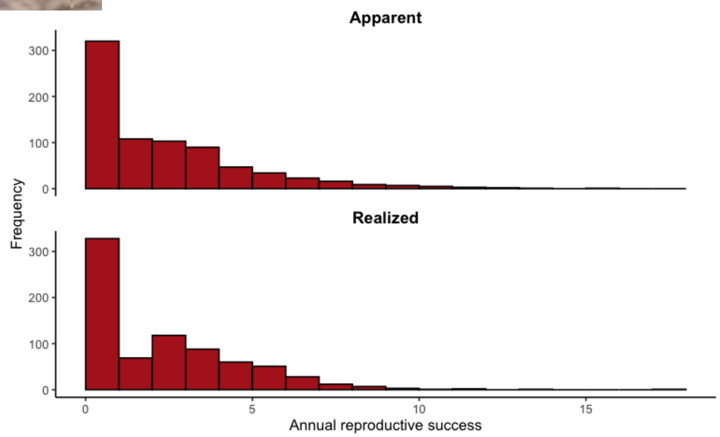
Example: Comparing apparent and realized reproductive success

- Variance is crucial to evolution
- Apparent reproductive success assigns offspring to "social males"
- Realized reproductive success assigns paternity by genotyping

Is there a lot more variance among males in reproductive success than we think?

© Eric Bégin

Data from song sparrows on Mandarte Island



Levene's test in R

```
leveneTest(data = RSData, reproductiveSuccess ~
measure, center = mean)
```

```
Levene's Test for Homogeneity of Variance
(center = mean)

      Df    F value     Pr(>F)
group   1   0.4773      0.4897
      1536
```

These data show no strong evidence for increased variance in reproductive success in the realized data.

Note about the *F* test

$$F = \frac{s_1^2}{s_2^2}$$

- The most commonly used test to compare variances
- The *F* test is very sensitive to its assumption that both distributions are normal.
- Therefore use Levene's test instead

Summary

- We can compare the means of two groups, using the mean of paired differences or the mean difference between two groups
- Paired data can be analyzed using the differences between two members of the pairs. (confidence intervals just like single sample on the differences; paired t-tests)

Summary

- Two-sample comparisons can be done with 2-sample t-test or Welch's t-test. Both assume normal distributed variables. 2-sample assumes equal variance; Welch's does not.
- The variances of two groups can be compared with a Levene's test.