Statistics with Spa  ows

Lecture 6

Julia Schroeder

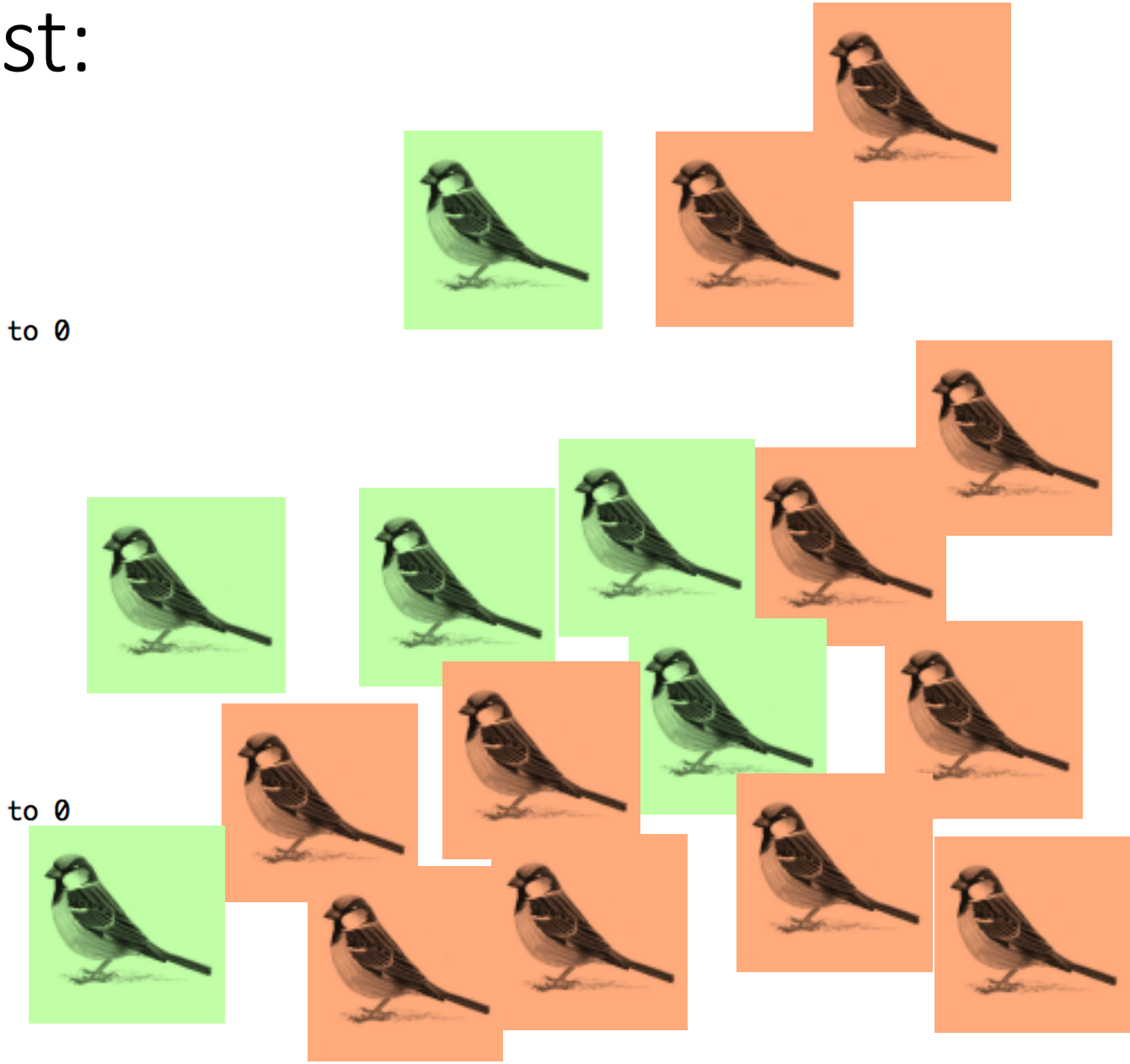
Julia.schroeder@imperial.ac.uk

Outline

- Influence of df on test
- Hypothesis testing

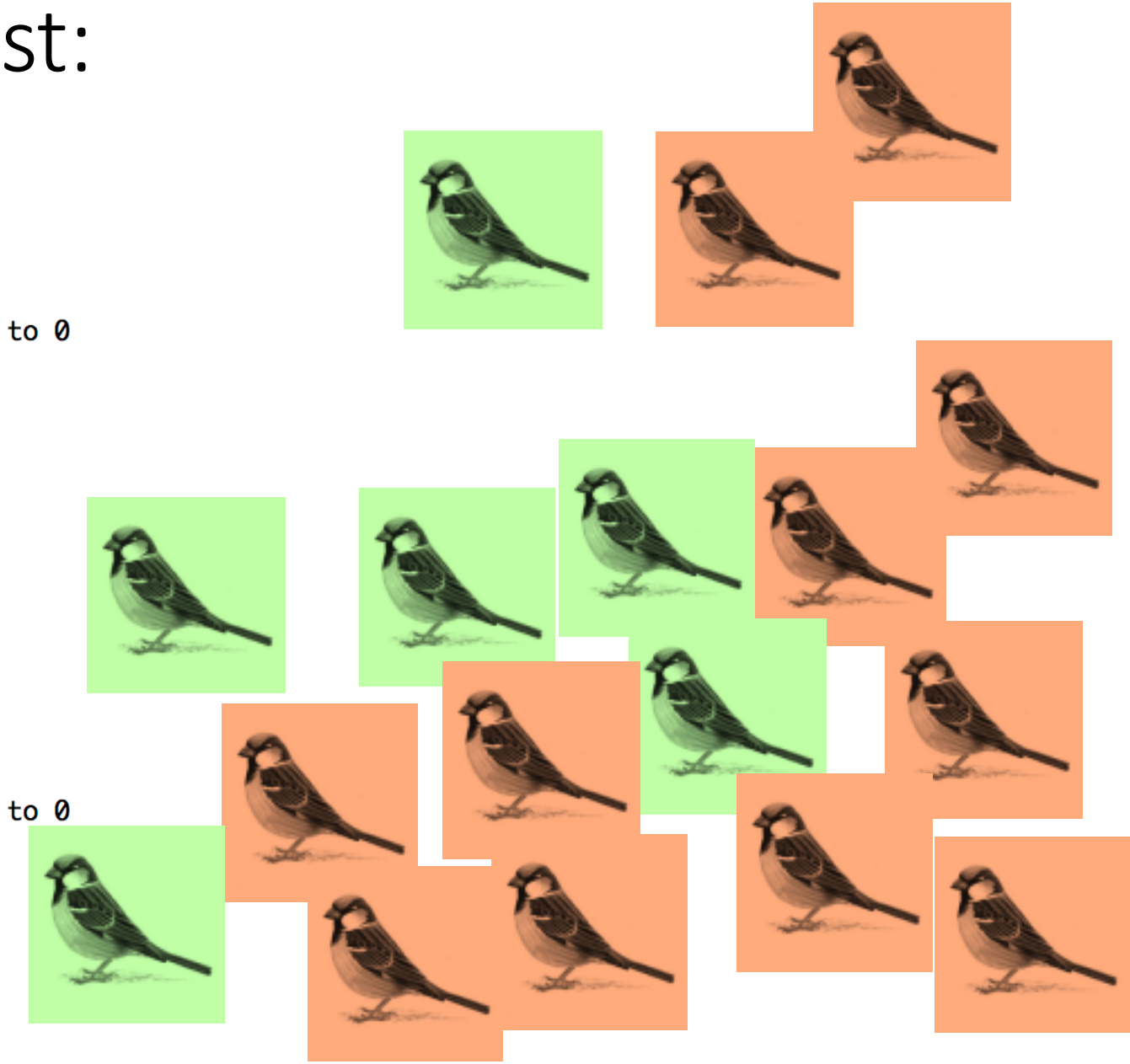
Influence of *df* on t-test:

```
> t.test(d1$Tarsus~d1$Sex,na.rm=TRUE)

        Welch Two Sample t-test

data:  d1$Tarsus by d1$Sex
t = 1.2257, df = 139.07, p-value = 0.2224
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1012318  0.4314949
sample estimates:
mean in group 0 mean in group 1
       18.27763        18.11250

> t.test(d$Tarsus~d$Sex,na.rm=TRUE)

        Welch Two Sample t-test

data:  d$Tarsus by d$Sex
t = -3.7382, df = 1677.4, p-value = 0.0001916
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.23814658 -0.07424028
sample estimates:
mean in group 0 mean in group 1
       18.44317        18.59936

>
```



Influence of *df* on t-test:

```
> t.test(d1$Tarsus~d1$Sex,na.rm=TRUE)

        Welch Two Sample t-test

data:  d1$Tarsus by d1$Sex
t = 1.2257, df = 139.07, p-value = 0.2224
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1012318  0.4314949
sample estimates:
mean in group 0 mean in group 1
       18.27763        18.11250

> t.test(d$Tarsus~d$Sex,na.rm=TRUE)

        Welch Two Sample t-test

data:  d$Tarsus by d$Sex
t = -3.7382, df = 1677.4, p-value = 0.0001916
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.23814658 -0.07424028
sample estimates:
mean in group 0 mean in group 1
       18.44317        18.59936

>
```



Influence of *df (n)* on t-test:

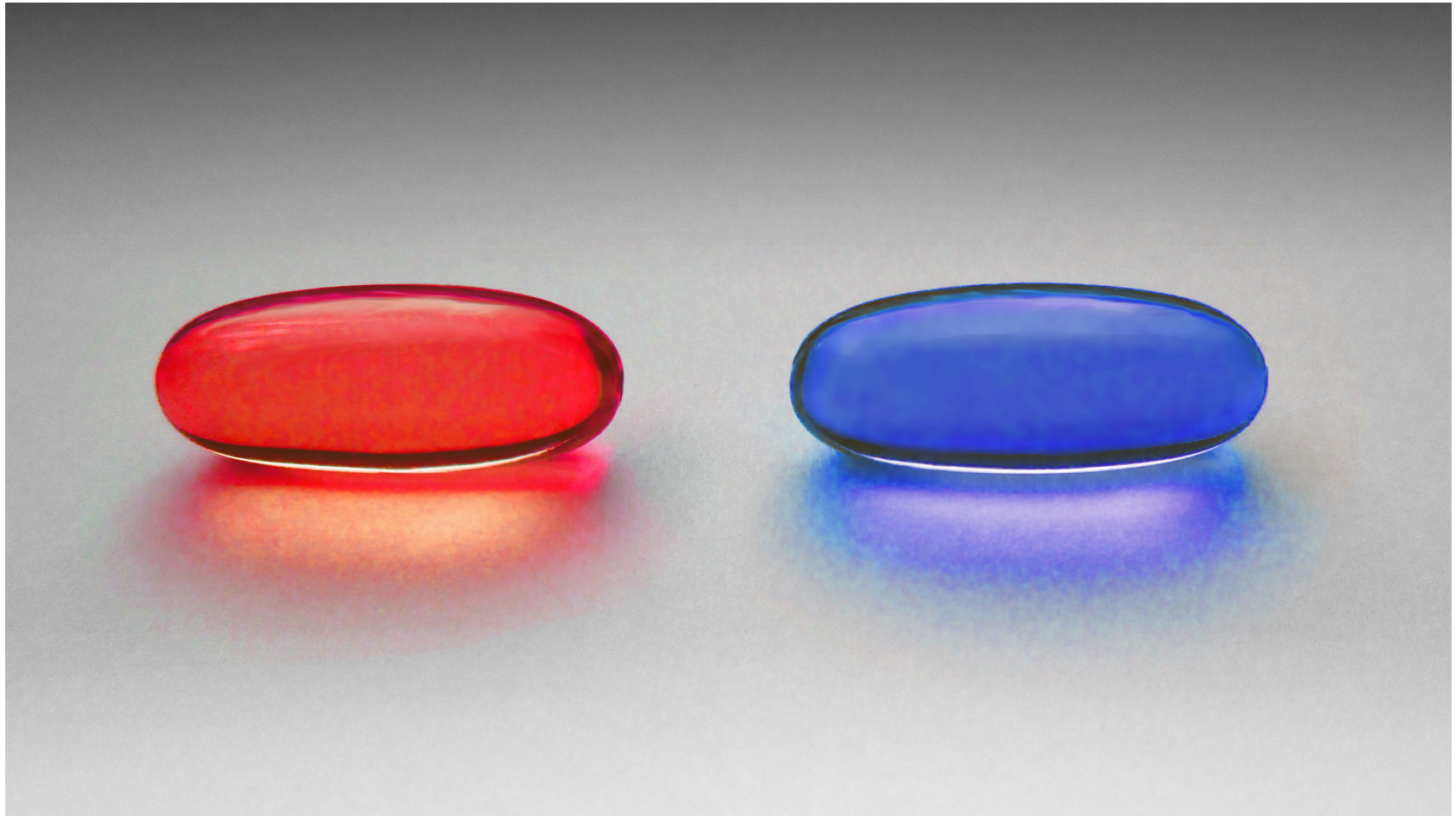
```
> t.test(d1$Tarsus~d1$Sex,na.rm=TRUE)

        Welch Two Sample t-test

data:  d1$Tarsus by d1$Sex
t = 1.2257, df = 139.07, p-value = 0.2224
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1012318  0.4314949
sample estimates:
mean in group 0 mean in group 1
       18.27763        18.11250

> t.test(d$Tarsus~d$Sex,na.rm=TRUE)

        Welch Two Sample t-test

data:  d$Tarsus by d$Sex
t = -3.7382, df = 1677.4, p-value = 0.0001916
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.23814658 -0.07424028
sample estimates:
mean in group 0 mean in group 1
       18.44317        18.59936

>
```

Wait, what?

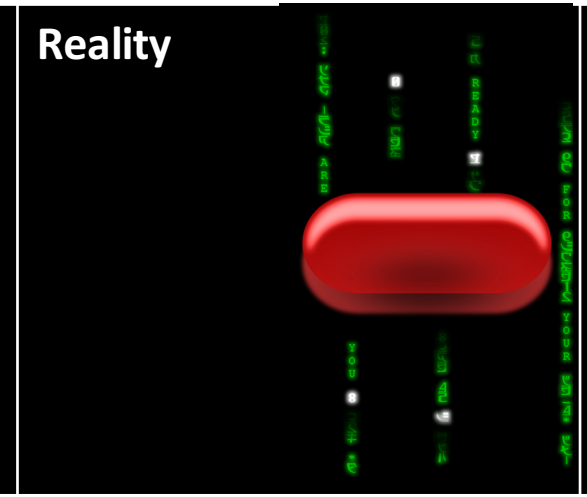
- Let’s digress for a moment…

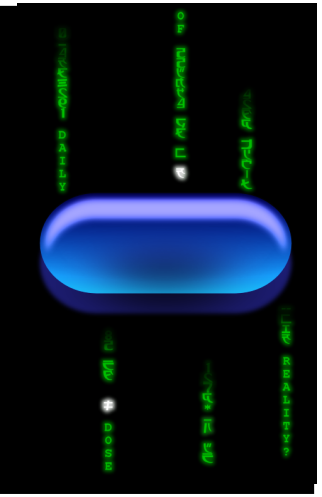
What does this mean?

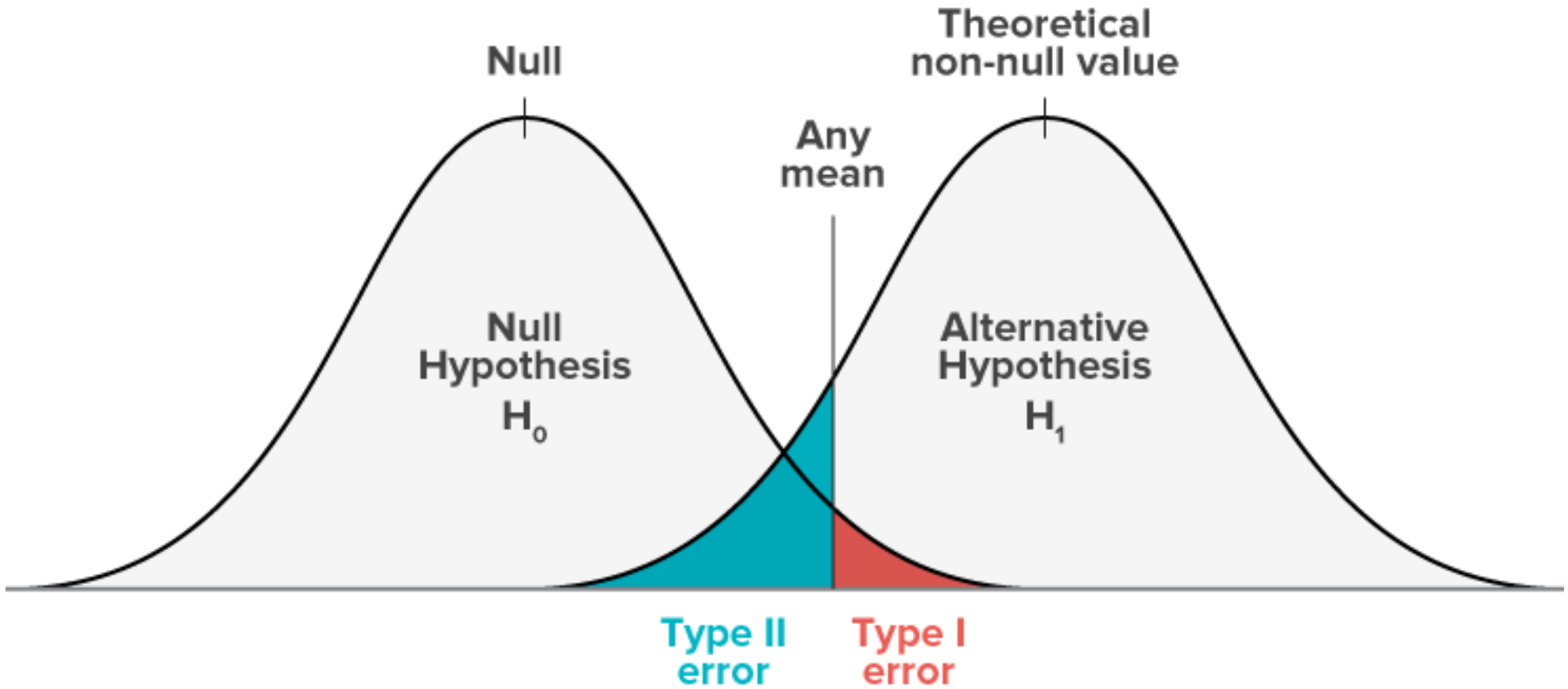
- What is real and what is an illusion?
- Decision errors

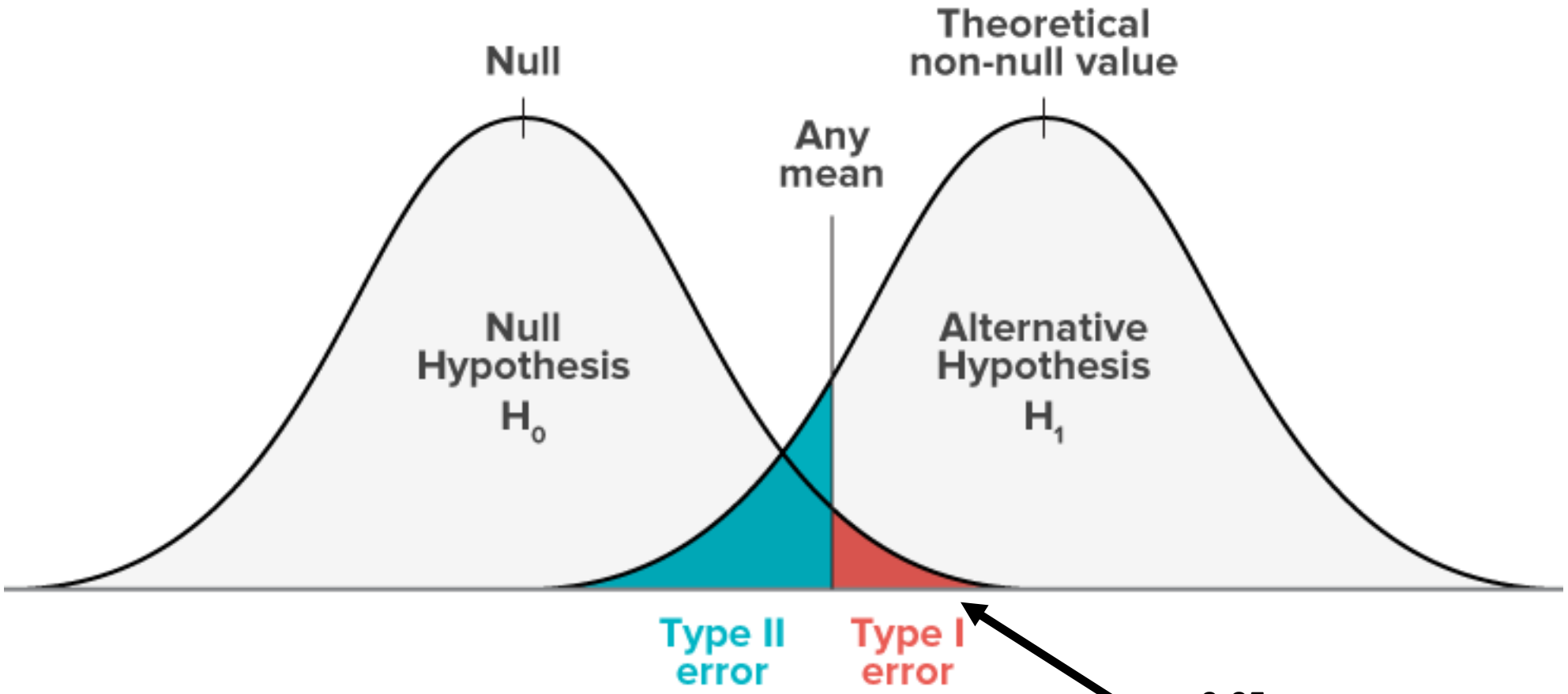


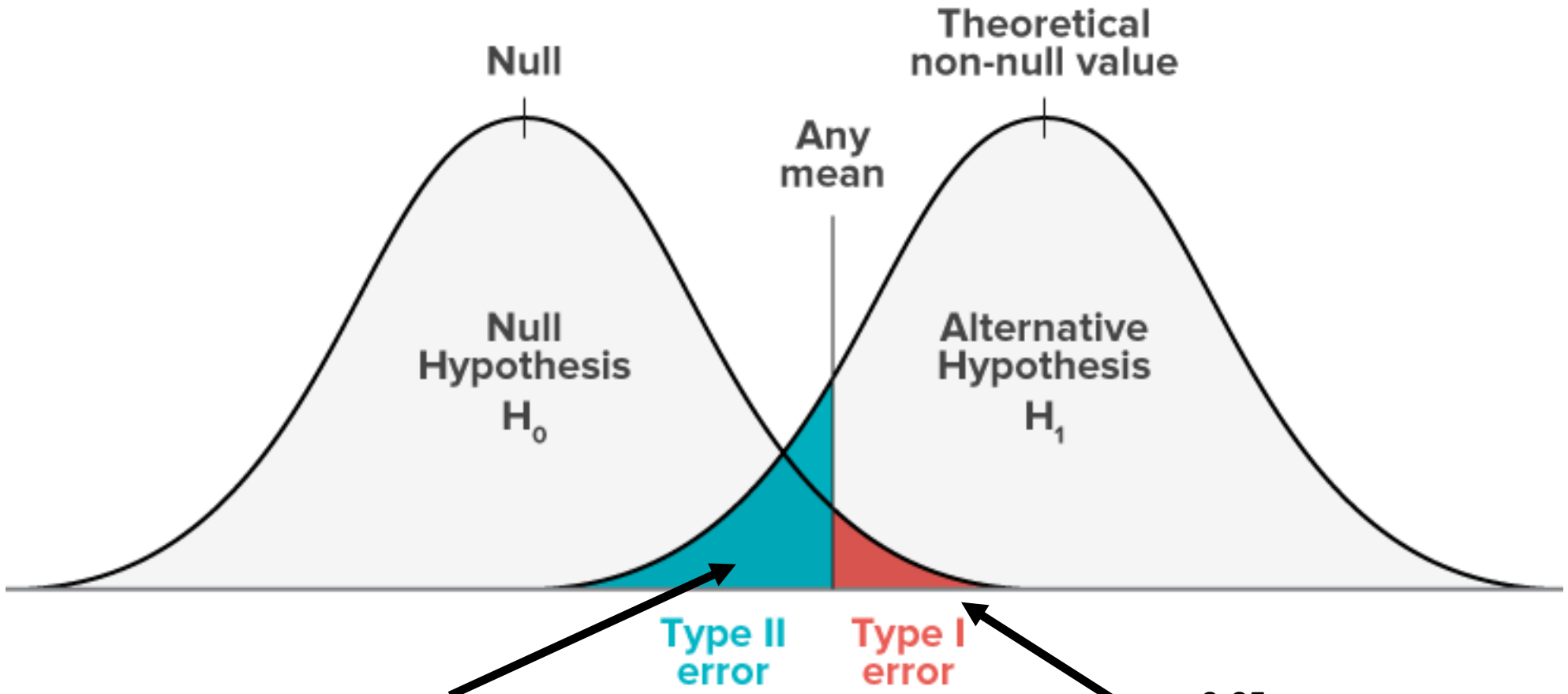
H0	Reality	Illusion
rejected	Effect detected Effect exists	Type I error Effect detected, but does not exist
accepted	No effect detected No effect exists	Type II error No effect detected, but it exists






Null
Theoretical non-null value
Any mean
Null Hypothesis H_0
Alternative Hypothesis H_1
Type II error
Type I error


Null
Theoretical non-null value
Any mean
Null Hypothesis H_0
Alternative Hypothesis H_1
Type II error
Type I error
0.05
5% chance of detecting an effect that does not exist



Type II effects depend on statistical power

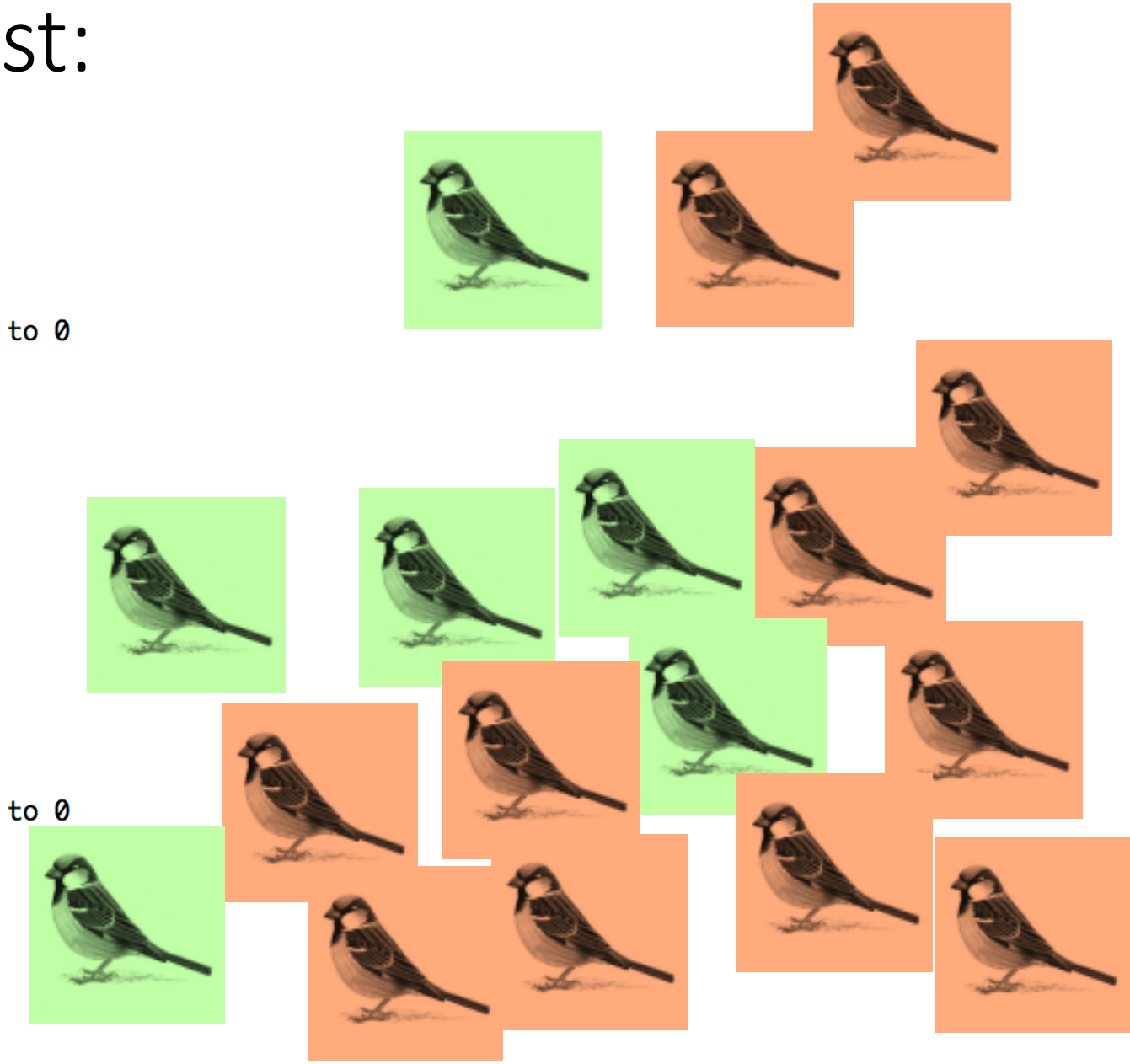
The bigger the sample size, the smaller the chance for type II errors

0.05

5% chance of detecting an effect that does not exist

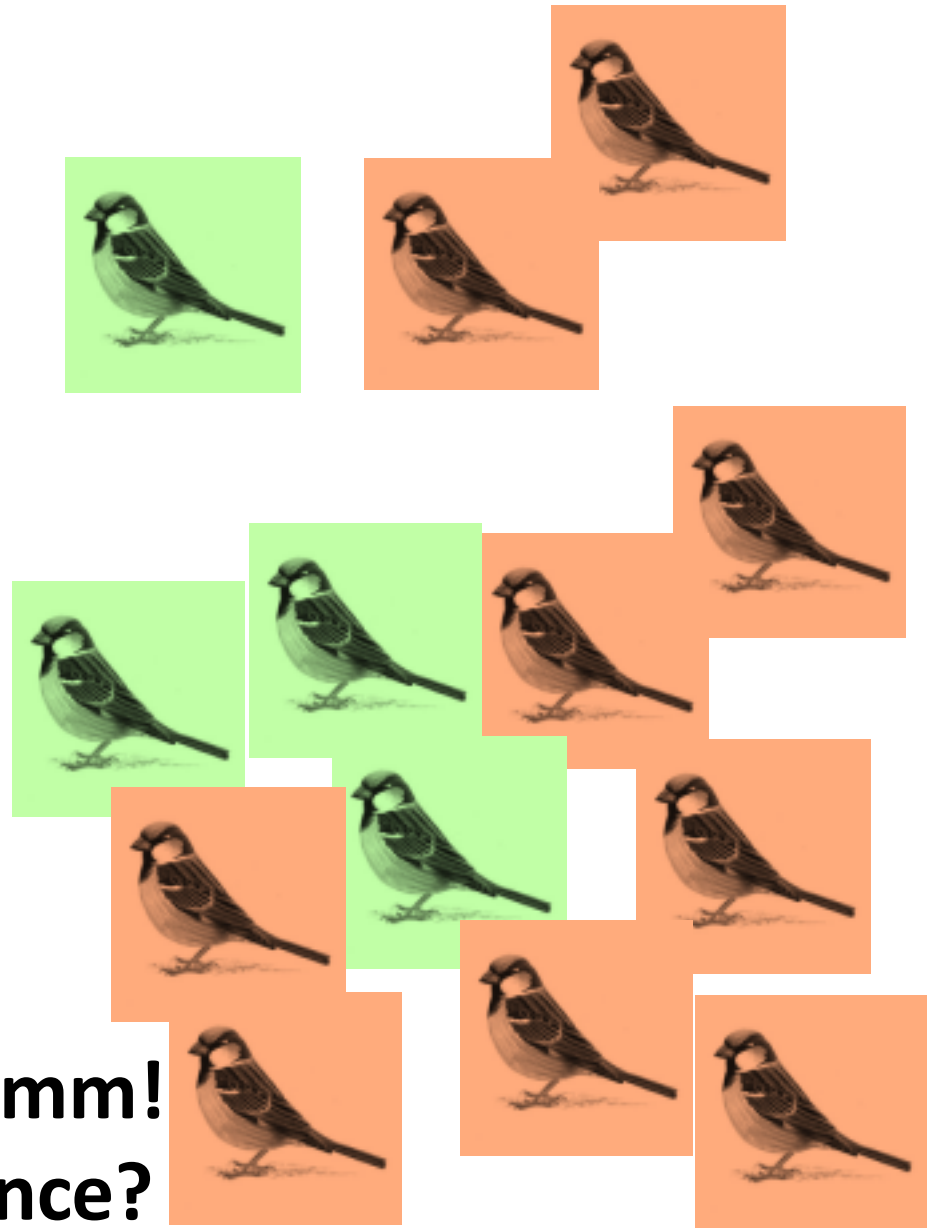
Influence of *df* on t-test:

```
> t.test(d1$Tarsus~d1$Sex,na.rm=TRUE)

        Welch Two Sample t-test

data:  d1$Tarsus by d1$Sex
t = 1.2257, df = 139.07, p-value = 0.2224
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1012318  0.4314949
sample estimates:
mean in group 0 mean in group 1
       18.27763        18.11250

> t.test(d$Tarsus~d$Sex,na.rm=TRUE)

        Welch Two Sample t-test

data:  d$Tarsus by d$Sex
t = -3.7382, df = 1677.4, p-value = 0.0001916
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.23814658 -0.07424028
sample estimates:
mean in group 0 mean in group 1
       18.44317        18.59936

>
```



Influence of *df* on t-test:

```
> t.test(d1$Tarsus~d1$Sex,na.rm=TRUE)

        Welch Two Sample t-test

data:  d1$Tarsus by d1$Sex
t = 1.2257, df = 139.07, p-value = 0.2224
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1012318  0.4314949
sample estimates:
mean in group 0 mean in group 1
       18.27763        18.11250

> t.test(d$Tarsus~d$Sex,na.rm=TRUE)

        Welch Two Sample t-test

data:  d$Tarsus by d$Sex
t = -3.7382, df = 1677.4, p-value = 0.0001916
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.23814658 -0.07424028
sample estimates:
mean in group 0 mean in group 1
       18.44317        18.59936

>
```



Difference of 0.5mm!
Biological relevance?

Statistical power

- Probability to detect an effect of specific size

- What effect size do you need?
- What effect size is biological meaningful?

Statistical power

- To calculate statistical power you need
- Mean in each group (make one 0, the other difference. 0.16)
- N (sample size – we want to find that one out)
- Sd of combined groups 0.96
- Power level (usually 80% is ok)

Statistical power

- To calculate statistical power you need
- Mean in each group (make one 0, the other difference. 0.16)
- N (sample size – we want to find that one out)
- Sd of combined groups 0.96
- Power level (usually 80% is ok = type II error chance 20%)

```
> library(pwr)
> pwr.t.test(d=(0-0.16)/0.96,power=.8,sig.level=.05,type="two.sample",alternative="two.sided")

     Two-sample t test power calculation

              n = 566.0799
              d = 0.1666667
      sig.level = 0.05
          power = 0.8
    alternative = two.sided

NOTE: n is number in *each* group
```

Statistical power

- To calculate statistical power you need
- Mean in each group (make one 0, the other difference. 0.16)
- N (sample size – we want to find that one out)
- Sd of combined groups 1
- Power level (usually 80% is ok = type II error chance 20%)

```
> library(pwr)
> pwr.t.test(d=(0-0.16)/0.96,power=.8,sig.level=.05,type="two.sample",alternative="two.sided")

     Two-sample t test power calculation

              n = 566.0799
              d = 0.1666667
      sig.level = 0.05
          power = 0.8
    alternative = two.sided

NOTE: n is number in *each* group
```

DO IT NOW – no handout!

- Download package “pwr”
- Run a power analysis to find out how large a sample of wing length data must be to detect a difference of an effect size of 5mm!

DO IT NOW – no handout!

- Download package “pwr”
- Run a power analysis to find out how large a sample of wing length data must be to detect a difference of an effect size of 5mm!
- effect size d = effect size you want to detect (differnce between two mean) = 5/sd