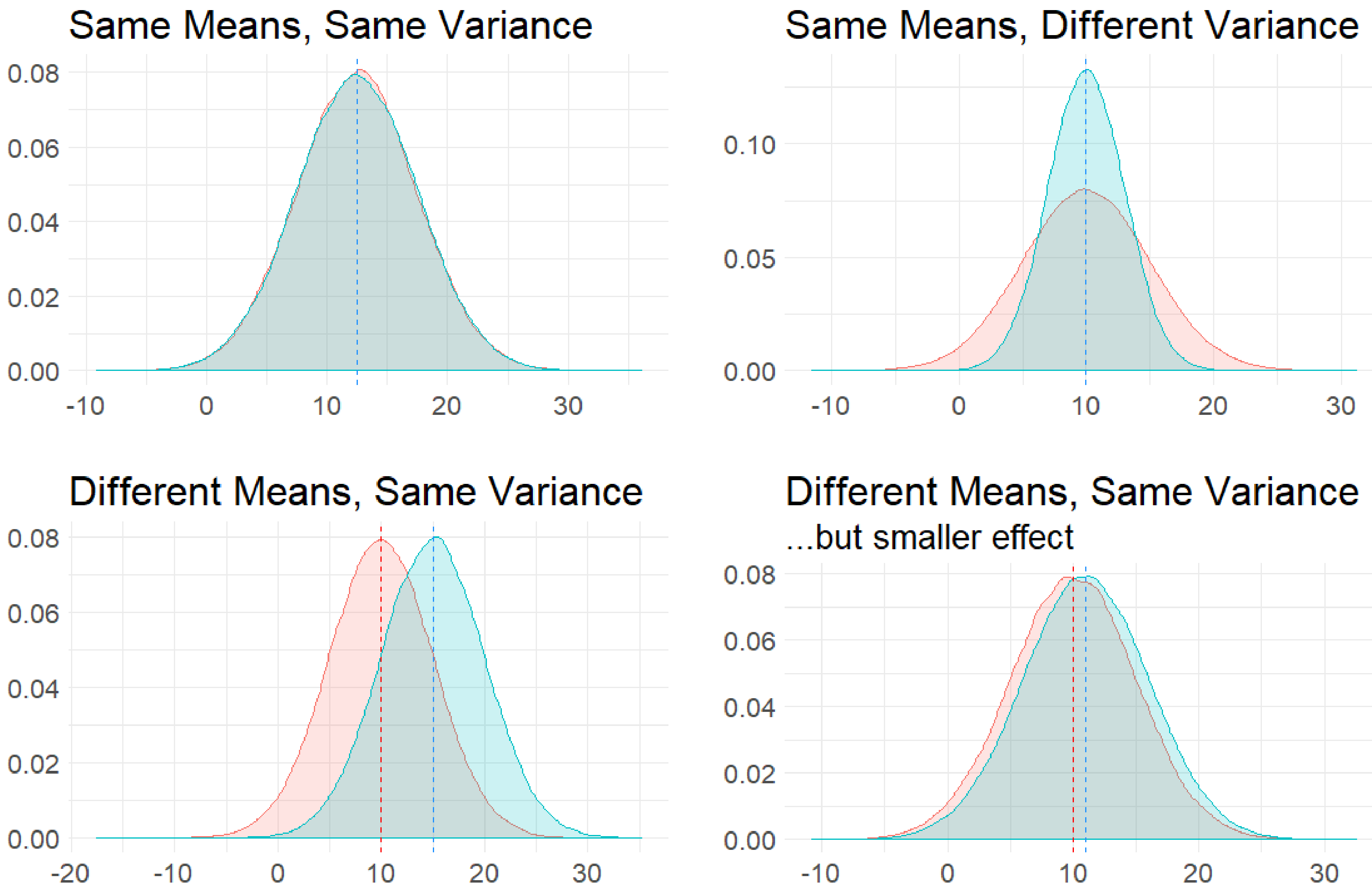
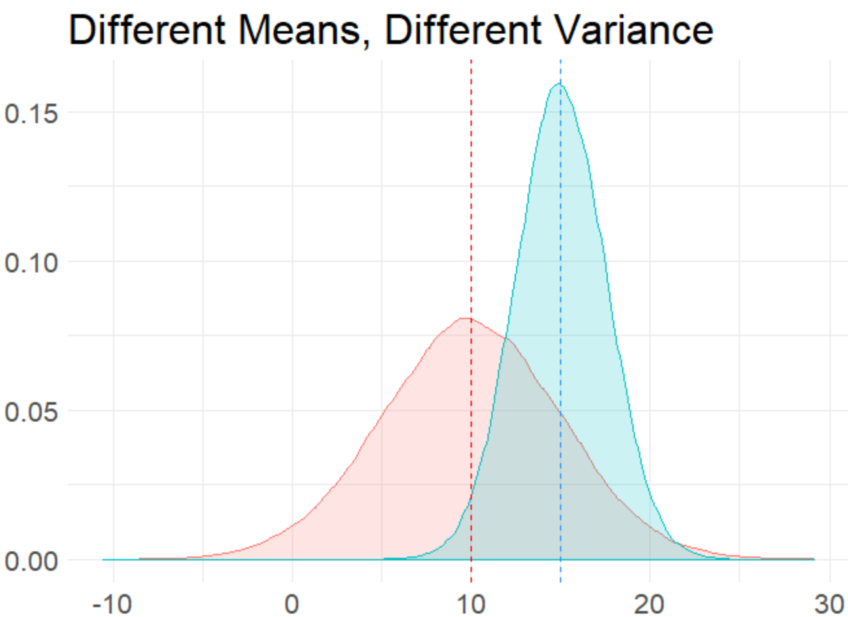
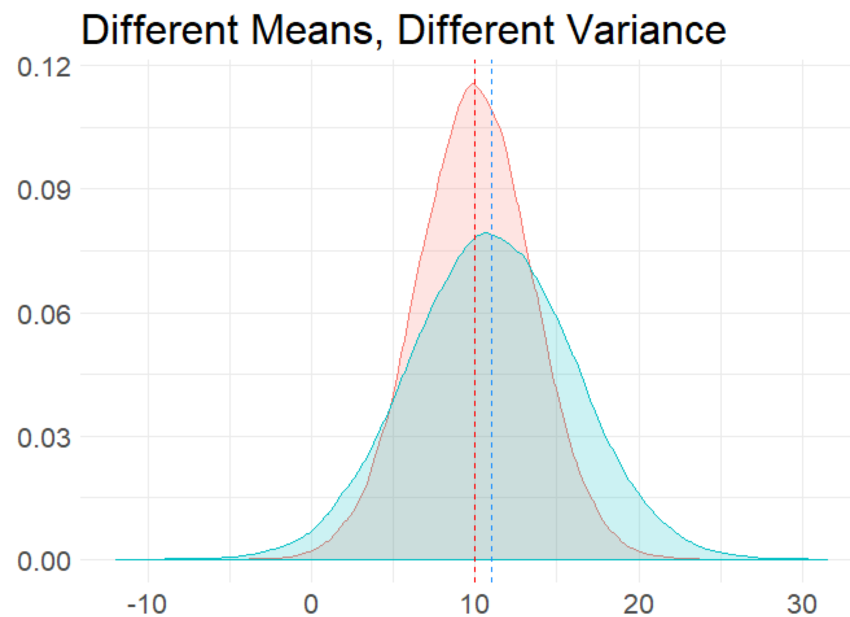
Comparing 2 samples



Same Means, Same Variance
0.08
0.06
0.04
0.02
0.00
-10
0
10
20
30
Same Means, Different Variance
0.10
0.05
0.00
-10
0
10
20
30
Different Means, Same Variance
0.08
0.06
0.04
0.02
0.00
-20
-10
0
10
20
30
Different Means, Same Variance
...but smaller effect
0.08
0.06
0.04
0.02
0.00
-10
0
10
20
30


Different Means, Different Variance
0.15
0.10
0.05
0.00
-10
0
10
20
30


Different Means, Different Variance
0.12
0.09
0.06
0.03
0.00
-10
0
10
20
30

e.g. Dr Harpo’s stats class has two TAs:

Anastasia (n=15)

65 74 73 83 76 65 86 70 80 55 78 78 90 77 68

Bernadette (n=18)

72 66 71 66 76 69 79 73 62 69 68 60 73 68 67 74 56 74

Are the scores of students different depending on which TA they got ?

Anastasia (n=15)

65 74 73 83 76 65 86 70 80 55 78 78 90 77 68

$$\overline{x} = 74.53$$

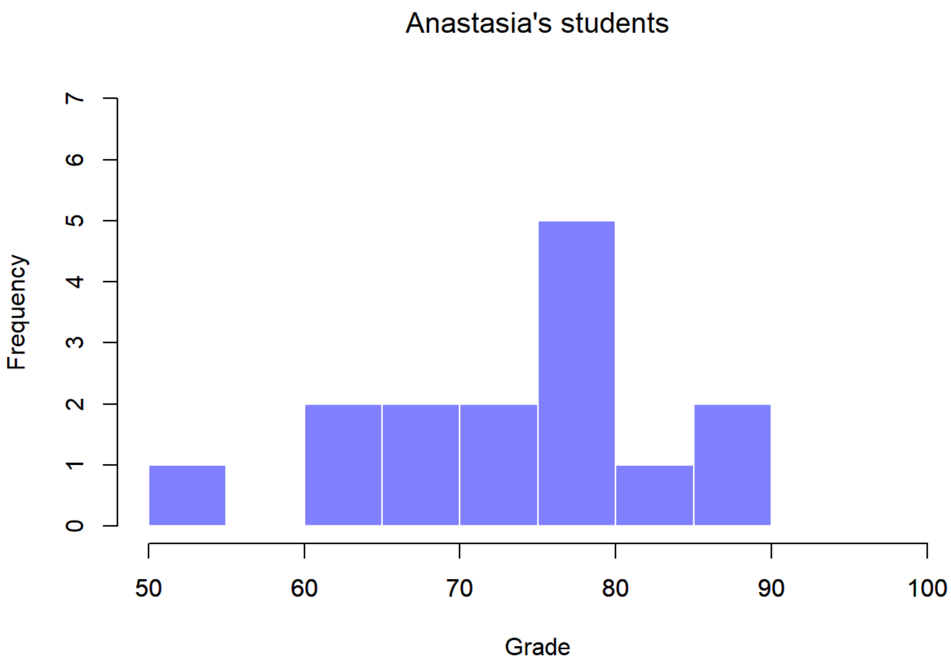
$$s = 9.00$$

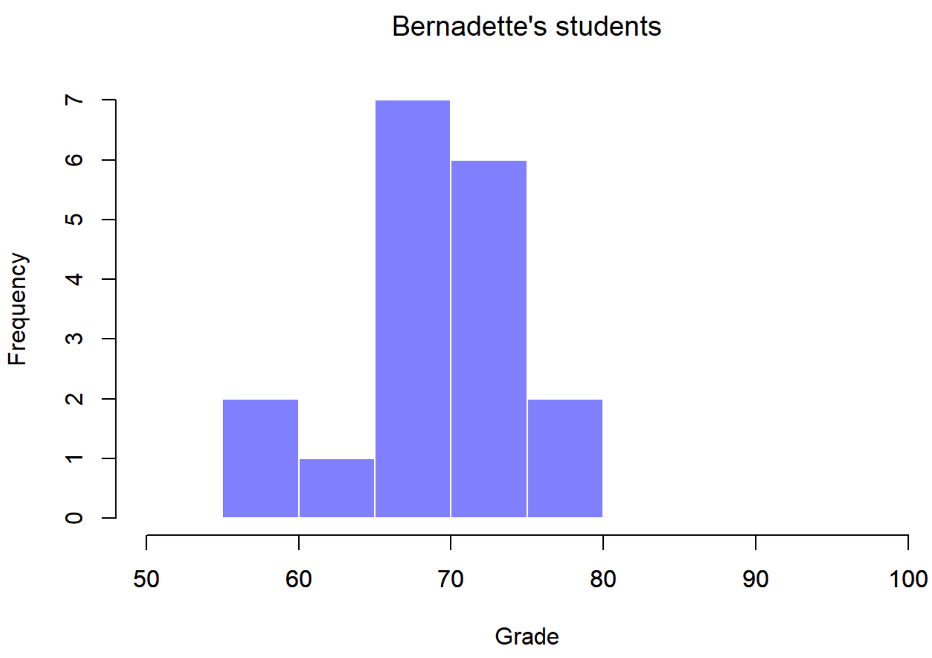
Bernadette (n=18)

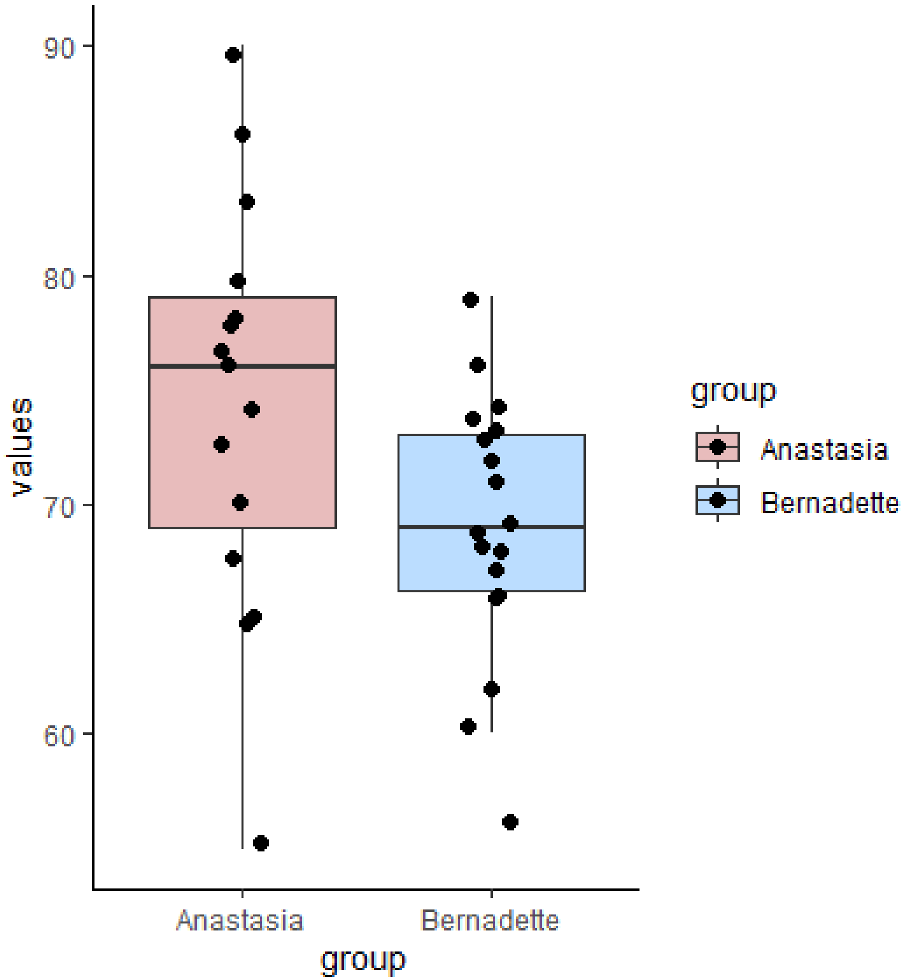
72 66 71 66 76 69 79 73 62 69 68 60 73 68 67 74 56 74

$$\overline{x} = 69.06$$

$$s = 5.77$$

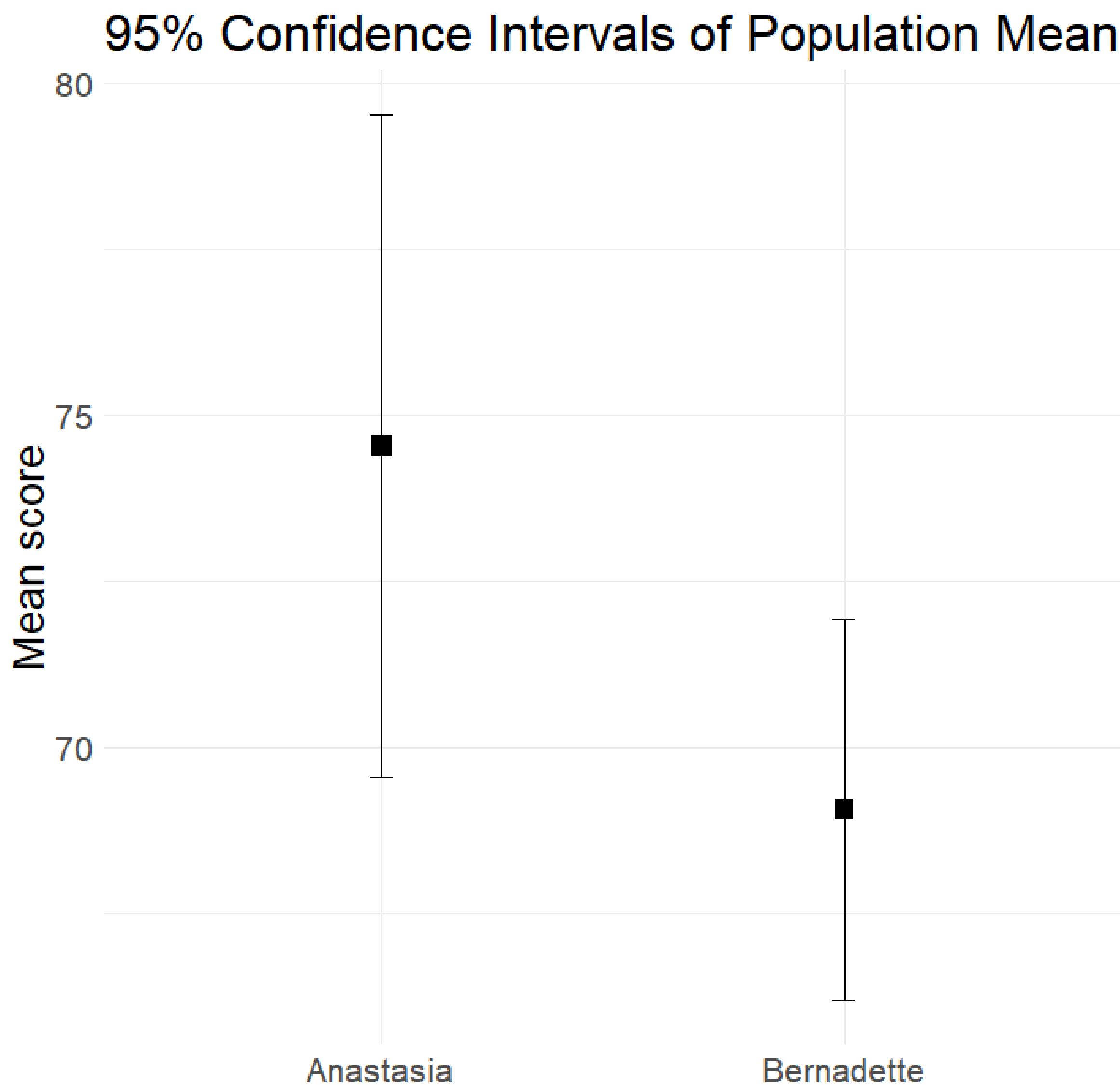

Anastasia's students
Frequency
0
1
2
3
4
5
6
7
50
60
70
80
90
100
Grade


Bernadette's students
Frequency
0
1
2
3
4
5
6
7
50
60
70
80
90
100
Grade



Is this a meaningful difference in the scores between the two groups?

Or, is it possible to find differences of this size due to chance alone?


95% Confidence Intervals of Population Mean
80
75
70
Mean score
Anastasia
Bernadette

We are imagining that these samples come from (are theoretically picked at random from) 'populations'...

Anastasia (n=15)

65 74 73 83 76 65 86 70 80 55 78 78 90 77 68

$$\overline{x} = 74.53$$

$$s = 9.00$$

Bernadette (n=18)

72 66 71 66 76 69 79 73 62 69 68 60 73 68 67 74 56 74

$$\overline{x} = 69.06$$

$$s = 5.77$$

Independent samples t***-test***

<u>2 versions:</u>

Student's

1. Test if there's a significant difference between the two groups in their population means

Welch's.

2. Estimating the difference in population means $(\mu_1 - \mu_2)$
with a confidence interval

Background to Student’s 2-Samples t-test

```
> anastasia
 [1] 65 74 73 83 76 65 86 70 80 55 78 78 90 77 68
> bernadette
 [1] 72 66 71 66 76 69 79 73 62 69 68 60 73 68 67 74 56 74
```

```
> t.test(anastasia,bernadette, var.equal = T)

        Two Sample t-test

data:  anastasia and bernadette
t = 2.1154, df = 31, p-value = 0.04253
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.1965873 10.7589683
sample estimates:
mean of x mean of y
 74.53333  69.05556
```

We are imagining that these samples come from (are theoretically picked at random from) ‘populations’…

Anastasia (n=15)

65 74 73 83 76 65 86 70 80 55 78 78 90 77 68

$$\overline{x} = 74.53$$

$$s = 9.00$$

Bernadette (n=18)

72 66 71 66 76 69 79 73 62 69 68 60 73 68 67 74 56 74

$$\overline{x} = 69.06$$

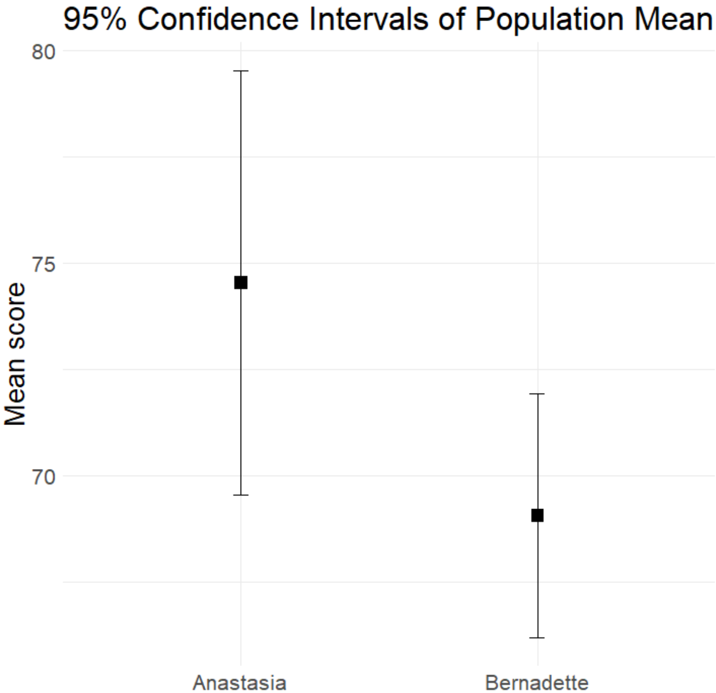
$$s = 5.77$$

Independent samples t*-test*

e.g. 2-sided test

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$



Are two "independent samples" of data drawn from populations with the same mean (the null hypothesis) or different means (the alternative hypothesis) ?

Assumptions of the independent t**-test**

Normality
Assumes that the true population distribution is normal

Independence
Observations are not correlated with each other

Equal variances (Homogeneity of variance)
Assumes equal population standard deviation between groups (same as saying equal population variances)

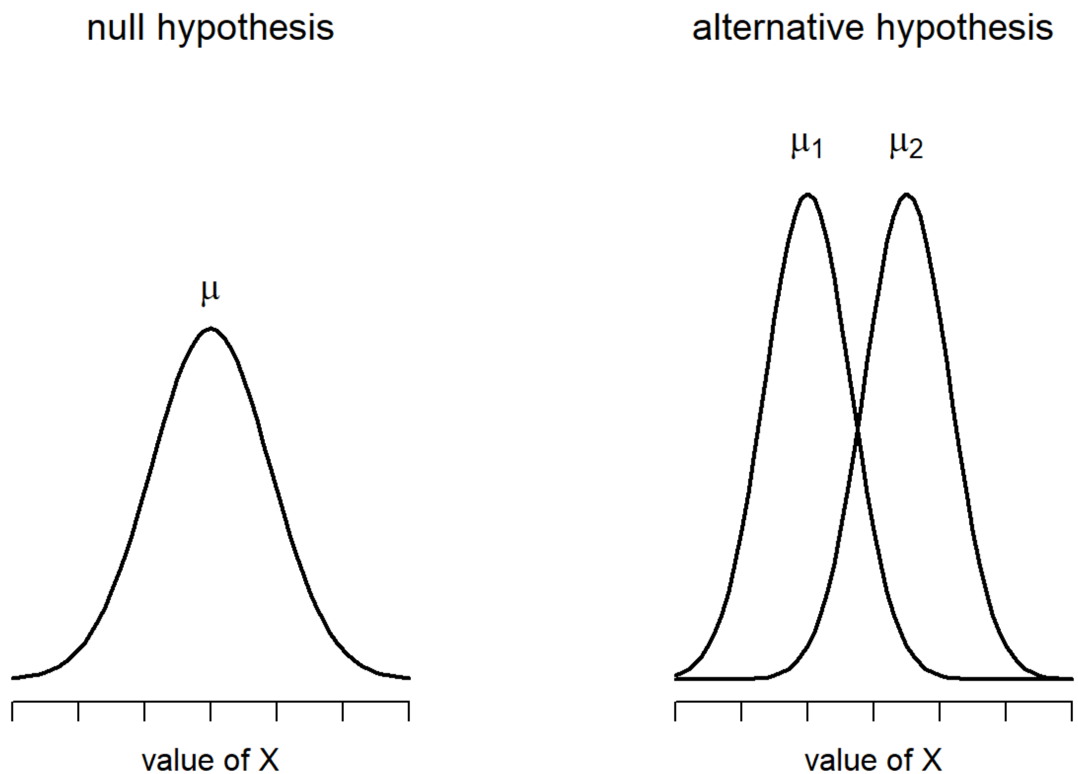
$$\sigma_1 = \sigma_2$$

Independent samples t*-test*

e.g. 2-sided test

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$



The null hypothesis assumes that both groups have the same mean μ, whereas the alternative assumes that they have different means μ_1 and μ_2.

Notice that it is assumed that the population distributions are normal, and that, although the alternative hypothesis allows the group to have different means, it assumes they have the same standard deviation (even though our samples don't)

Independent samples t***-test***

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$

$$\mu_1 - \mu_2 = 0$$

$$\bar{x}_1 - \bar{x}_2 \approx 0$$

$$74.53 - 69.06 = 5.47$$

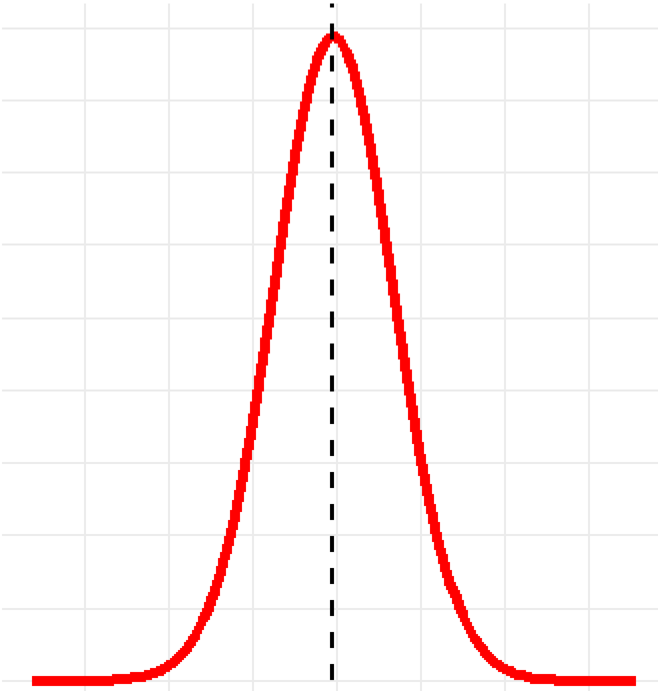
Assumption: $\sigma_1 = \sigma_2$

What if we took samples over and over again ....

Sample	$\bar{x}_1$	$\bar{x}_2$	d_i
1	74.53	69.06	5.47
2	71.33	68.66	2.67
3	69.55	69.91	-0.36
4	70.32	63.34	6.98
5	75.05	64.58	10.47
…	…	…	…
10000	74.31	67.72	6.59

We'd end up with a *sampling distribution* of the *difference in sample means*…

Sampling Distribution of the Difference in Sample Means



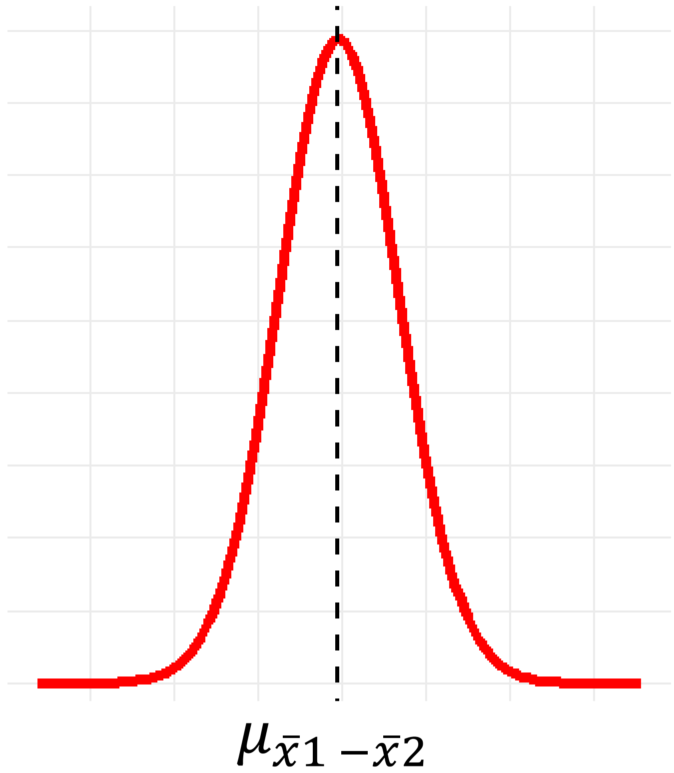
$\mu_{\bar{x}1 - \bar{x}2}$

μ_{d1}

We are testing whether our one observed difference in group means $\overline{x}_1 - \overline{x}_2$ is sufficiently different from 0 (i.e. no difference in sample means) -

d.f. = $n_1 + n_2 - 2$

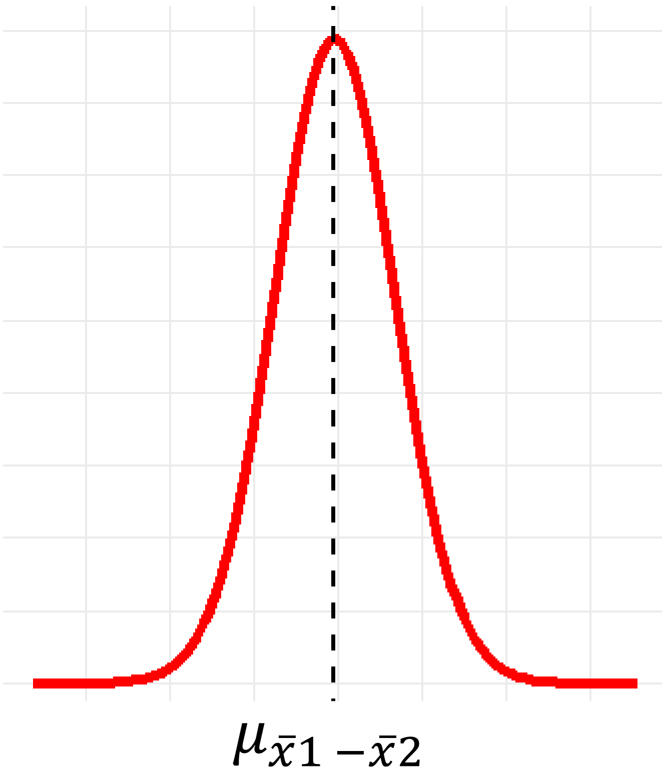
Sampling Distribution of the Difference in Sample Means



$$t = \frac{\bar{X}_1 - \bar{X}_2}{SE}$$

If we calculate the Standard Error of the Mean (SE) for the sampling distribution, we can get a t-statistic.

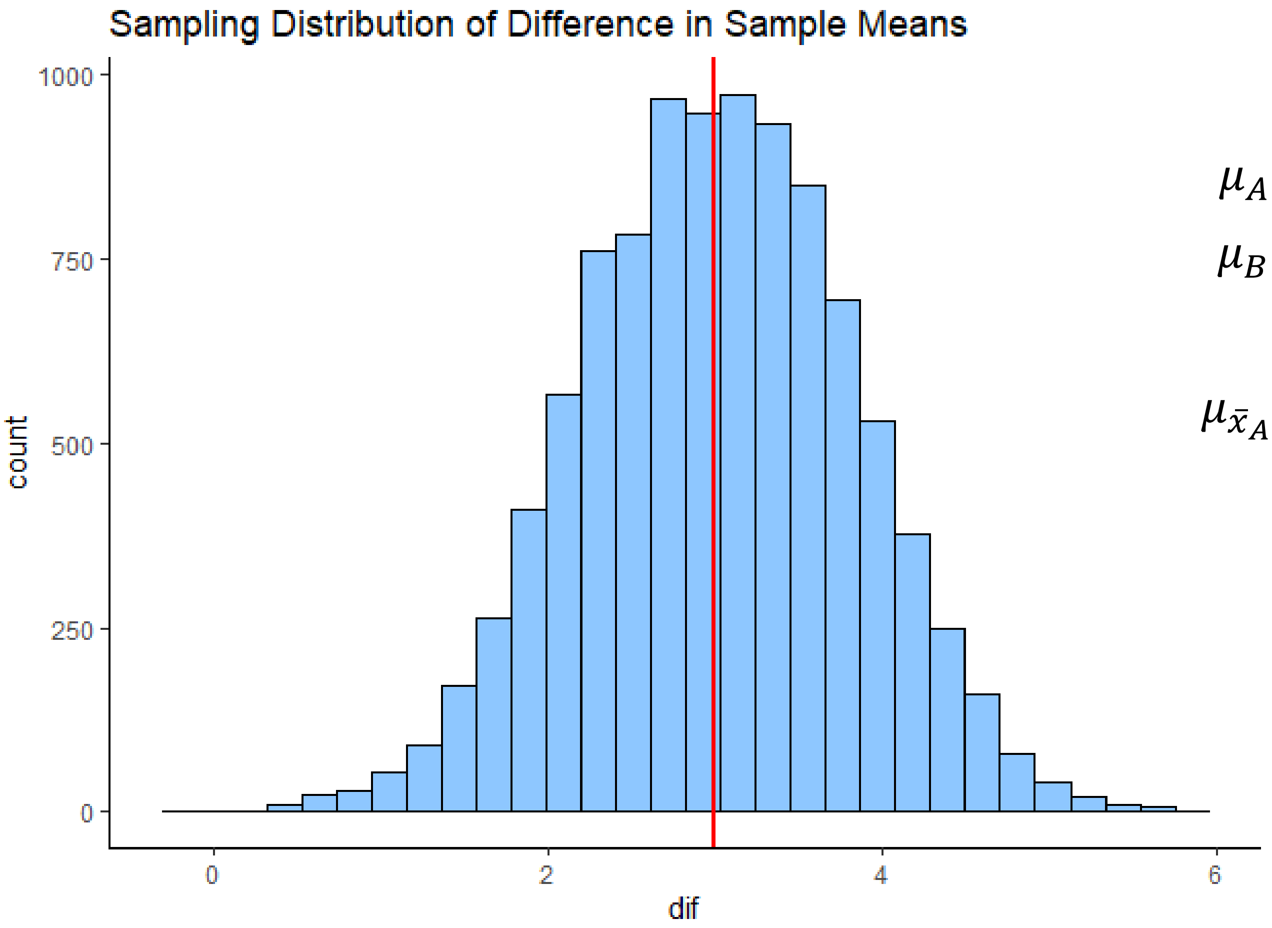
Sampling Distribution of the Difference in Sample Means



To do that we need to know what the standard deviation of this sampling distribution is?

i.e. the standard error of the mean difference scores

A simulated example:



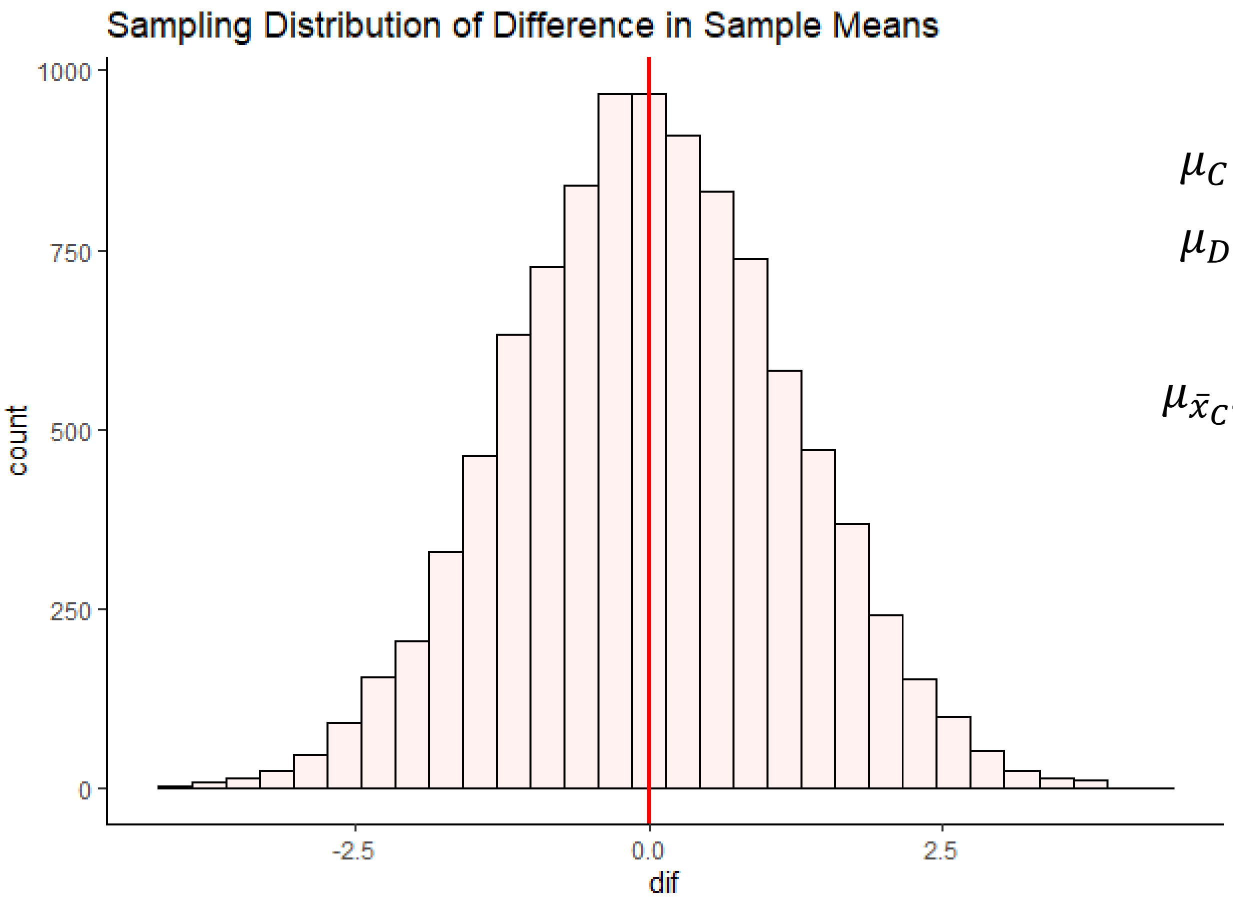
$\mu_A = 15 \quad \sigma_A = 2.5$

$\mu_B = 12 \quad \sigma_B = 2.5$

$\mu_{\bar{x}_A - \bar{x}_B} = 3 \quad \sigma_{\bar{x}_A - \bar{x}_B} = ?$

Taking samples of size n=17 For group A and n=20 for group B

Another simulated example:



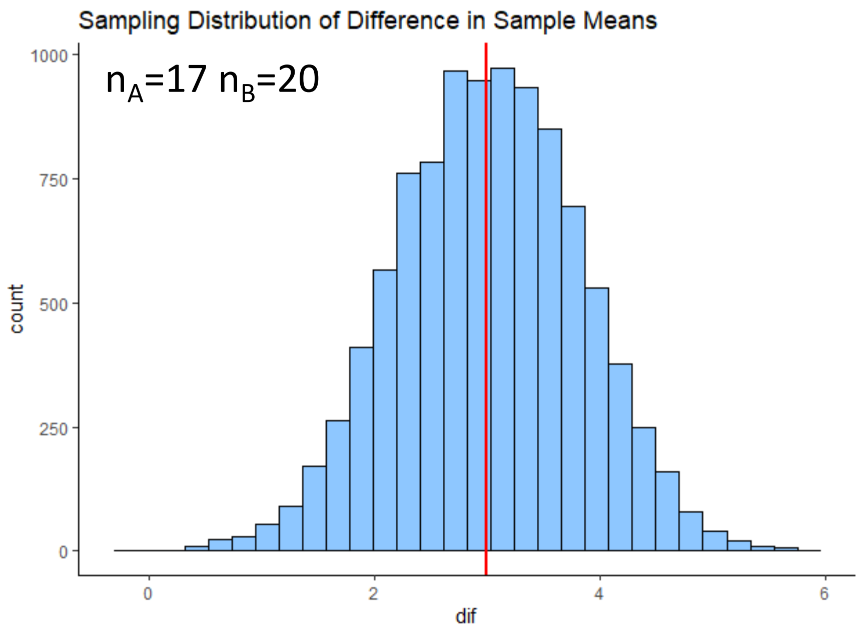
$$\mu_C = 10 \qquad \sigma_C = 3$$

$$\mu_D = 10 \qquad \sigma_D = 3$$

$$\mu_{\bar{x}_C - \bar{x}_D} = 0 \qquad \sigma_{\bar{x}_C - \bar{x}_D} = ?$$

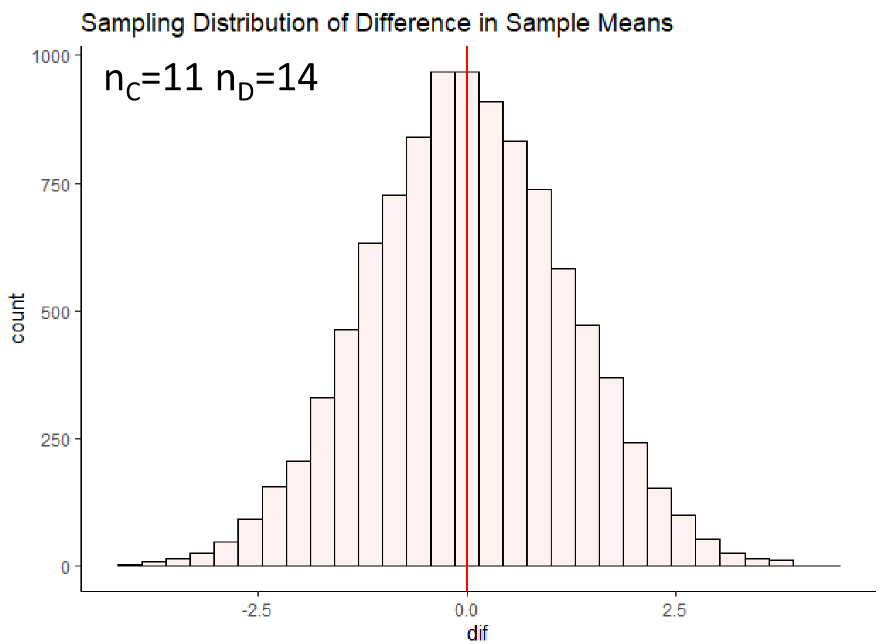
Taking samples of size n=11 For group C and n=14 for group D

Sampling Distribution of the Difference in Sample Means



$$\mu_A = 15 \qquad \sigma_A = 2.5$$
$$\mu_B = 12 \qquad \sigma_B = 2.5$$

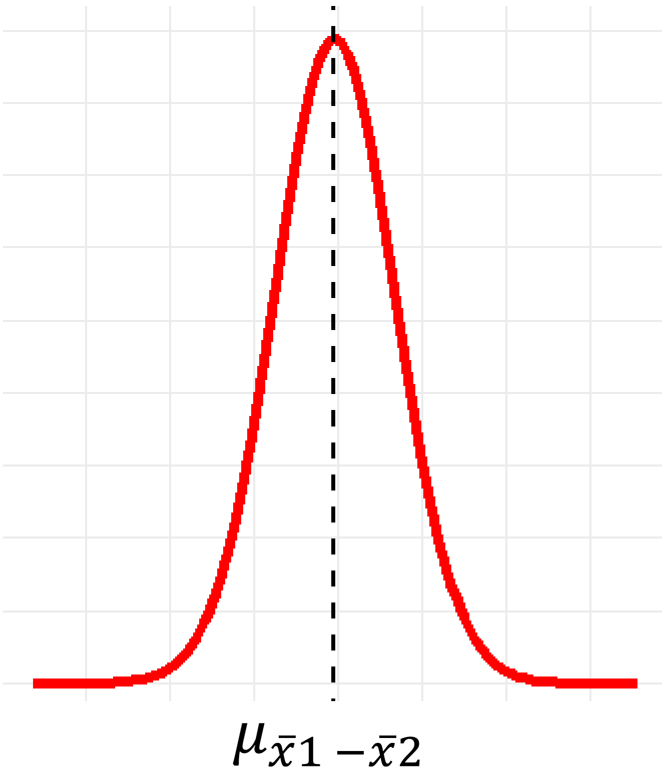
$$\mu_{\bar{x}_A - \bar{x}_B} = 3 \qquad \sigma_{\bar{x}_A - \bar{x}_B} = ?$$



$$\mu_C = 10 \qquad \sigma_C = 3$$
$$\mu_D = 10 \qquad \sigma_D = 3$$

$$\mu_{\bar{x}_C - \bar{x}_D} = 0 \qquad \sigma_{\bar{x}_C - \bar{x}_D} = ?$$

Sampling Distribution of the Difference in Sample Means



$$t = \frac{\bar{X}_1 - \bar{X}_2}{SE}$$

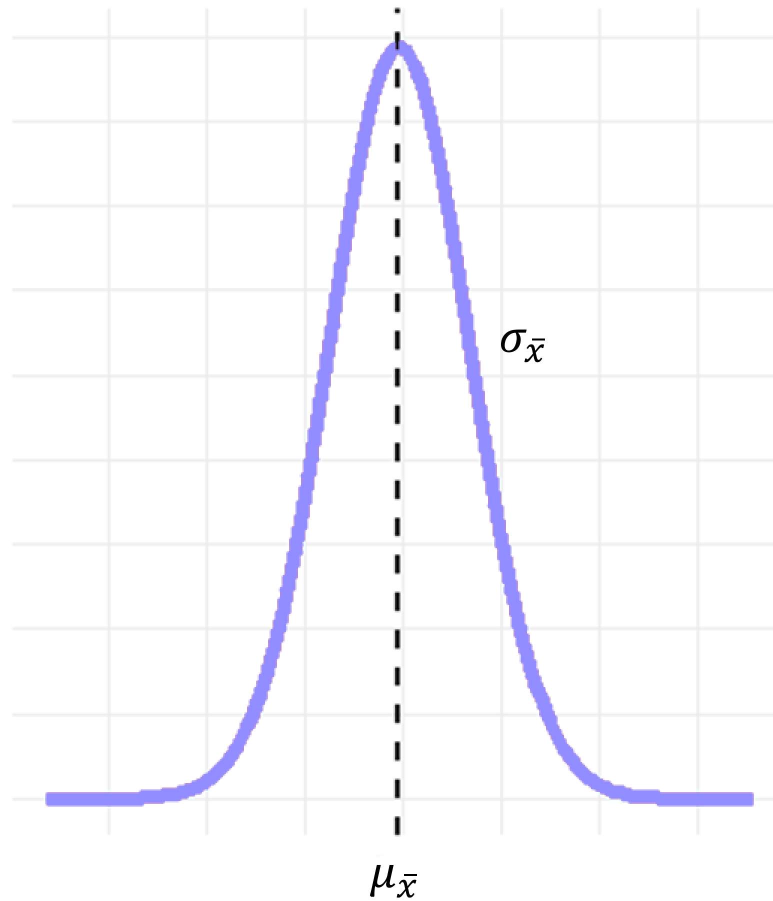
If we calculate the Standard Error of the Mean (SE) for the sampling distribution, we can get a t-statistic.

The Standard Deviation of the Sampling Distribution of Sample Means is the Standard Error of the Mean

When we only have one sample, we estimate it like this:

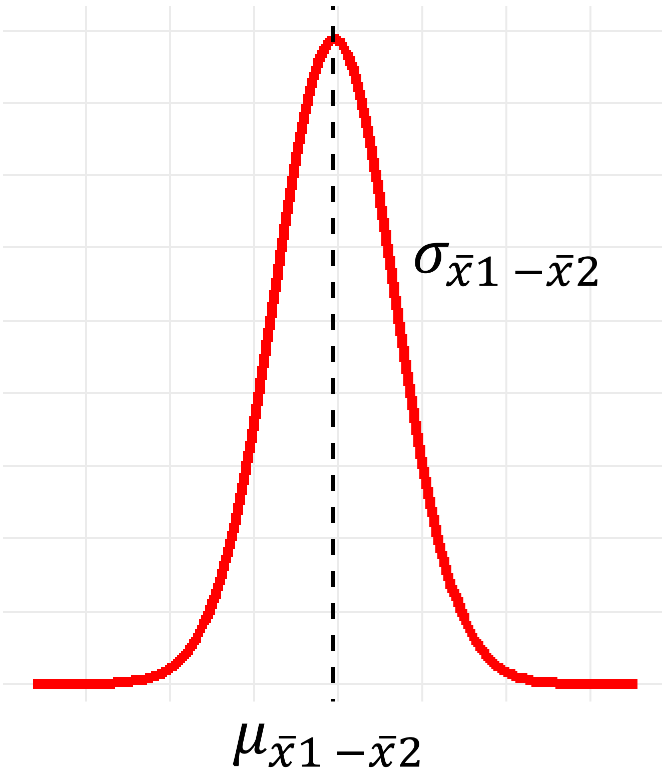
$$\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}$$

$$\sigma_{\bar{x}} = s \times \sqrt{\frac{1}{n}}$$



Sampling Distribution of Sample Means

Sampling Distribution of the Difference in Sample Means



How do we estimate the SE for the sampling distribution for the difference in sample means, when we have two samples that might also differ in sample sizes ?

We first calculate a **pooled standard deviation** between the two samples …. And use that to then estimate the SE

Anastasia (n=15)

65 74 73 83 76 65 86 70 80 55 78 78 90 77 68

$$\bar{x} = 74.53$$

$$s = 9.00$$

Bernadette (n=18)

72 66 71 66 76 69 79 73 62 69 68 60 73 68 67 74 56 74

$$\bar{x} = 69.06$$

$$s = 5.77$$

Our pooled standard deviation should be some kind of average of the two sample standard deviations – but adjusted in case one sample has a bigger sample size

Estimate of Pooled Standard Deviation

$$x_{ik} - \bar{x}_k$$

Get the difference of each score from its group mean

$$\hat{\sigma}_p{}^2 = \frac{\sum(x_{ik} - \bar{x}_k)^2}{n-2}$$

Get the difference of each score from its group mean, square them all, and sum. Then divide by N -2 to get an estimate of the pooled variance.

$$\hat{\sigma}_p = \sqrt{\frac{\sum(x_{ik} - \bar{x}_k)^2}{n-2}}$$

Square root this to get an estimate of the pooled standard deviation.

Anastasia (Mean = 74.53)			Bernadette (Mean = 69.06)	
65	-9.53		72	2.94
74	-0.53		66	-3.06
73	-1.53		71	1.94
83	8.47		66	-3.06
76	1.47		76	6.94
65	-9.53		69	-0.06
86	11.47		79	9.94
70	-4.53		73	3.94
80	5.47		62	-7.06
55	-19.53		69	-0.06
78	3.47		68	-1.06
78	3.47		60	-9.06
90	15.47		73	3.94
77	2.47		68	-1.06
68	-6.53		67	-2.06
			74	4.94
			56	-13.06
			74	4.94

Remember variance is just the average of all the squared deviations

So, add them all up and divide by …. N-2 (to get an estimate of the population variance)

Anastasia (Mean = 74.53)			Bernadette (Mean = 69.06)		
	$x_{ik} - \bar{x}_k$	$(x_{ik} - \bar{x}_k)^2$		$x_{ik} - \bar{x}_k$	$(x_{ik} - \bar{x}_k)^2$
65	-9.53	90.9	72	2.94	8.7
74	-0.53	0.3	66	-3.06	9.3
73	-1.53	2.4	71	1.94	3.8
83	8.47	71.7	66	-3.06	9.3
76	1.47	2.2	76	6.94	48.2
65	-9.53	90.9	69	-0.06	0.0
86	11.47	131.5	79	9.94	98.9
70	-4.53	20.6	73	3.94	15.6
80	5.47	29.9	62	-7.06	49.8
55	-19.53	381.6	69	-0.06	0.0
78	3.47	12.0	68	-1.06	1.1
78	3.47	12.0	60	-9.06	82.0
90	15.47	239.2	73	3.94	15.6
77	2.47	6.1	68	-1.06	1.1
68	-6.53	42.7	67	-2.06	4.2
			74	4.94	24.4
			56	-13.06	170.4
			74	4.94	24.4
	$\sum (x_{ik} - \bar{x}_k)^2$	1133.7		$\sum (x_{ik} - \bar{x}_k)^2$	566.9

Estimate of Pooled Standard Deviation

$x_{ik} - \bar{x}_k$ *Get the difference of each score from its group mean*

$$\hat{\sigma}_p{}^2 = \frac{\sum(x_{ik} - \bar{x}_k)^2}{n-2} = \frac{1133.7 + 566.9}{15 + 18 - 2} = 54.9$$

$$\hat{\sigma}_p = \sqrt{\frac{\sum(x_{ik} - \bar{x}_k)^2}{n-2}} = \sqrt{54.9} = 7.41$$

Anastasia (n=15)

65 74 73 83 76 65 86 70 80 55 78 78 90 77 68

$$\overline{x} = 74.53$$

$$s = 9.00$$

Bernadette (n=18)

72 66 71 66 76 69 79 73 62 69 68 60 73 68 67 74 56 74

$$\overline{x} = 69.06$$

$$s = 5.77$$

$$\hat{\sigma}_p = 7.41$$

```
> (9 + 5.77)/2
[1] 7.385
```

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

For one sample, the standard error of the mean

$$\sigma_{\bar{x}} = \sigma \times \sqrt{\frac{1}{n}}$$

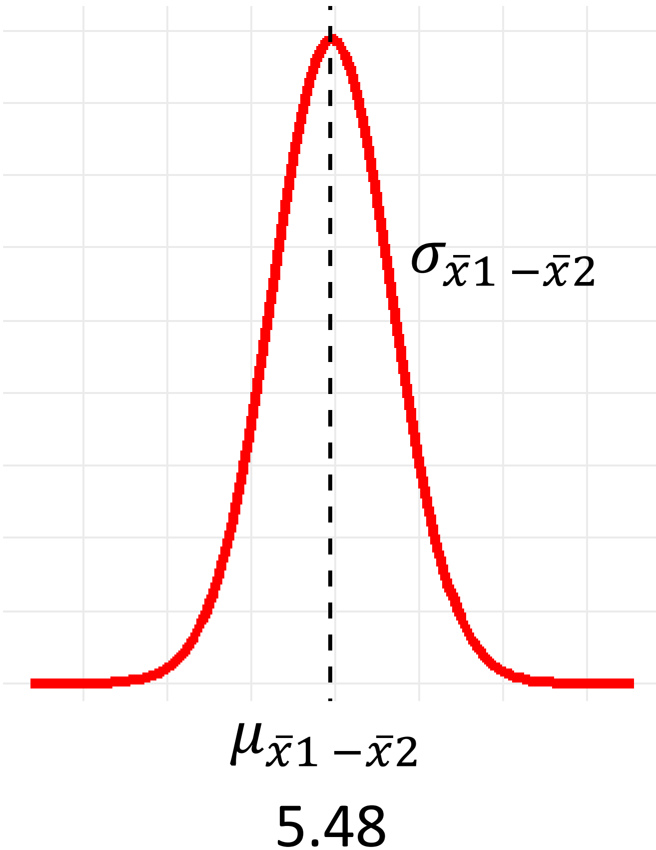
Which can be rewritten as (is the same as….)

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \hat{\sigma}_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

For the standard error of the mean difference between two samples, we use this formula.

Sampling Distribution of Difference in Sample Means

$$\sigma_{\bar{x}_1-\bar{x}_2} = \hat{\sigma}_p \times \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$



$$\sigma_{\bar{x}_1-\bar{x}_2} = 7.41 \times \sqrt{\frac{1}{15}+\frac{1}{18}} = 2.59$$

Confidence Interval for Difference in Means

Anastasia (n=15)

65 74 73 83 76 65 86 70 80 55 78 78 90 77 68

$$\overline{x} = 74.53$$

$$s = 9.00$$

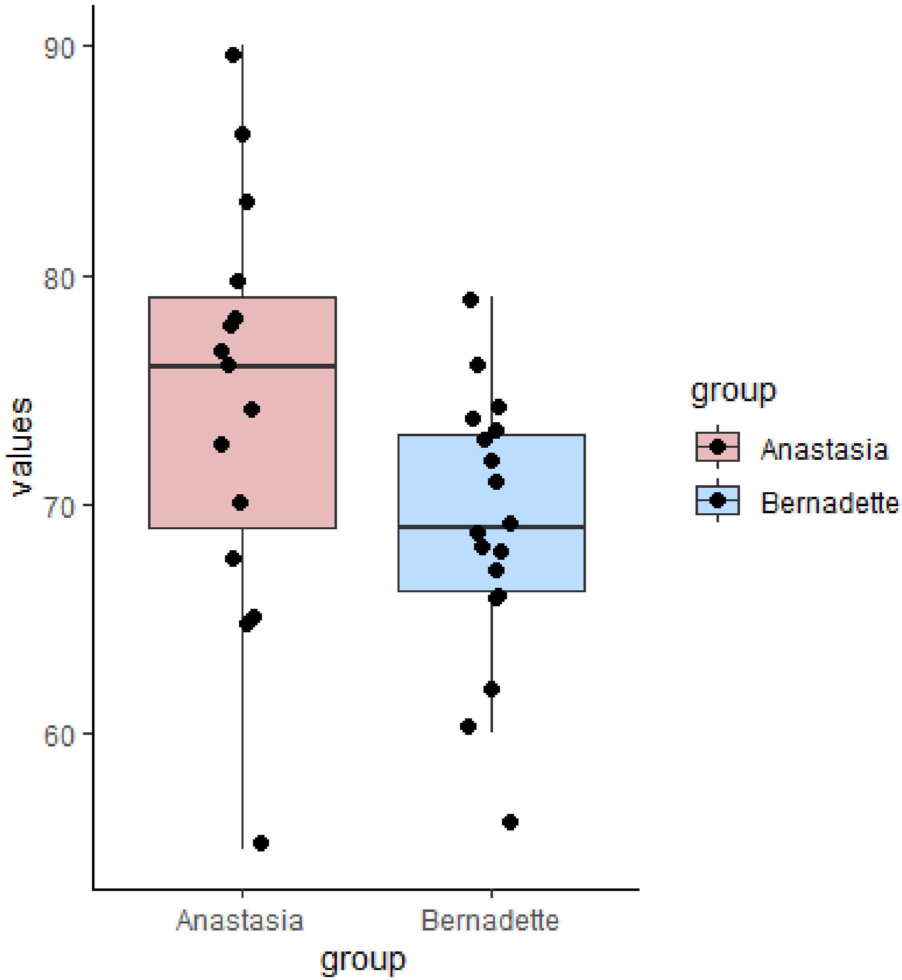
Bernadette (n=18)

72 66 71 66 76 69 79 73 62 69 68 60 73 68 67 74 56 74

$$\overline{x} = 69.06$$

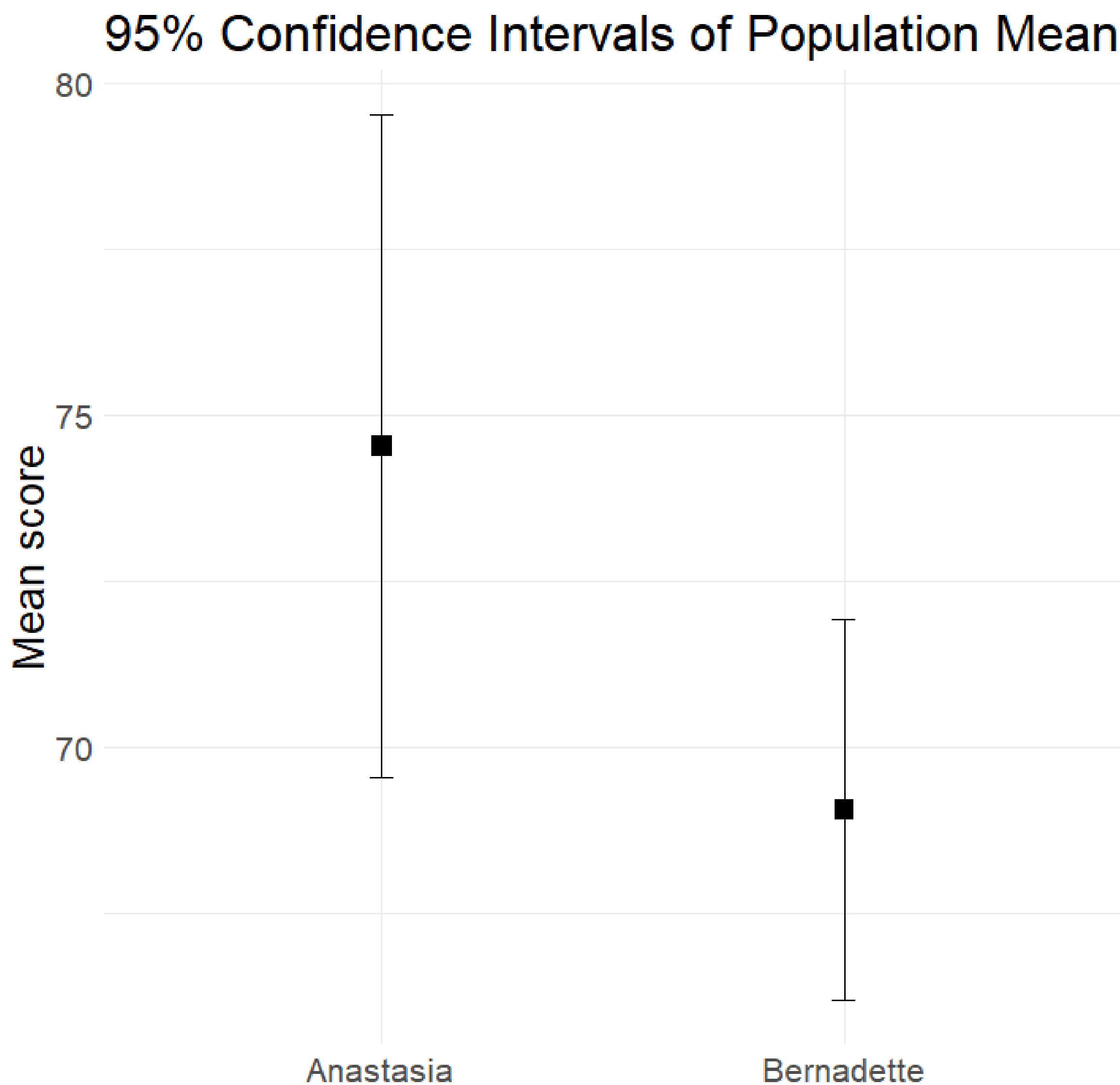
$$s = 5.77$$

$$\bar{x}1 - \bar{x}2 = 74.53 - 69.06 = 5.48$$



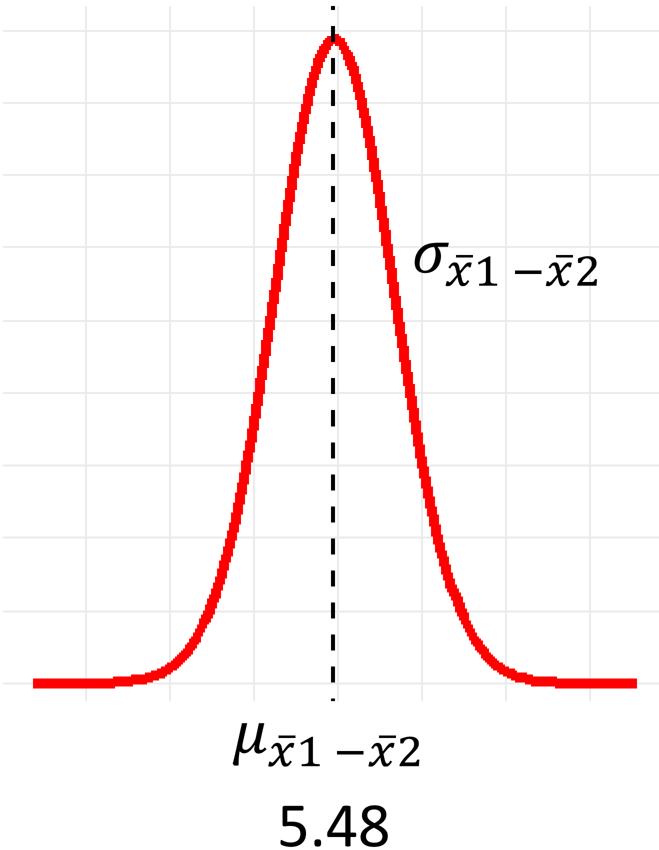
Is this a meaningful difference in the scores between the two groups?

Or, is it possible to find differences of this size due to chance alone?


95% Confidence Intervals of Population Mean
80
75
70
Mean score
Anastasia
Bernadette

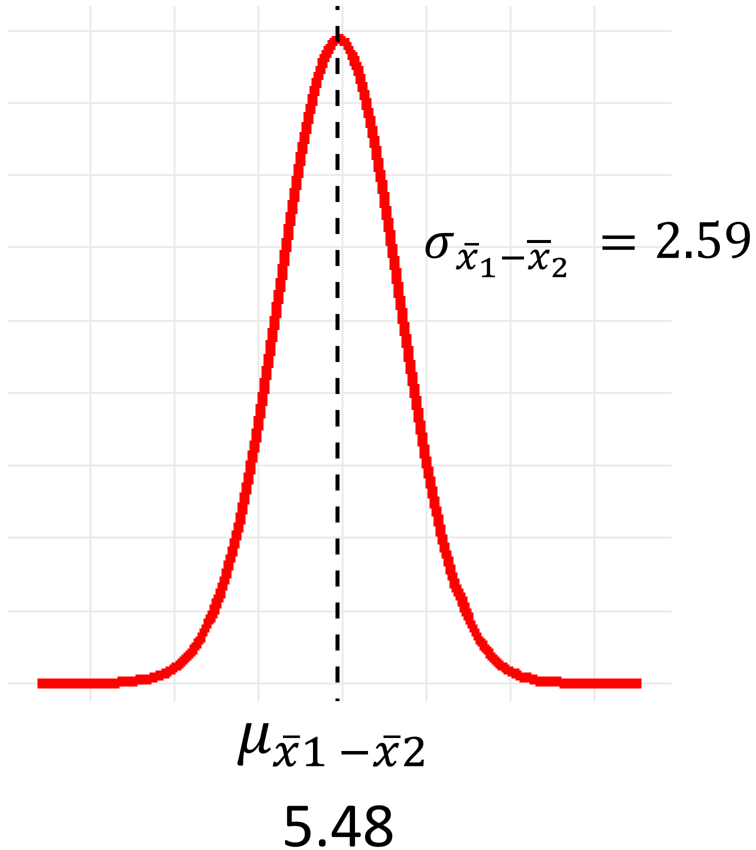
Sampling Distribution of Difference in Sample Means

$$\sigma_{\bar{x}_1-\bar{x}_2} = \hat{\sigma}_p \times \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$



$$\sigma_{\bar{x}_1-\bar{x}_2} = 7.41 \times \sqrt{\frac{1}{15}+\frac{1}{18}} = 2.59$$

Sampling Distribution of Difference in Sample Means

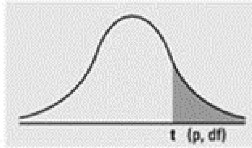


We can construct a CI based on the t-distribution

The d.f. for that distribution would be d.f. = $n_1 + n_2 - 2$

For a 95% CI, we need to find the values of 't' for that distribution that would leave 5% in the tails, (2.5% either side)

$$CI = \mu_{\bar{x}1 - \bar{x}2} \pm t \times \sigma_{\bar{x}_1 - \bar{x}_2}$$



df/p	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.54070	5.84091	12.9240
4	0.270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.943180	2.44691	3.14267	3.70743	5.9588
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079
8	0.261921	0.706387	1.396815	1.859548	2.30600	2.89646	3.35539	5.0413
9	0.260955	0.702722	1.383029	1.833113	2.26216	2.82144	3.24984	4.7809
10	0.260185	0.699812	1.372184	1.812461	2.22814	2.76377	3.16927	4.5869
11	0.259556	0.697445	1.363430	1.795885	2.20099	2.71808	3.10581	4.4370
12	0.259033	0.695483	1.356217	1.782288	2.17881	2.68100	3.05454	43178
13	0.258591	0.693829	1.350171	1.770933	2.16037	2.65031	3.01228	4.2208
14	0.258213	0.692417	1.345030	1.761310	2.14479	2.62449	2.97684	4.1405
15	0.257885	0.691197	1.340606	1.753050	2.13145	2.60248	2.94671	4.0728
16	0.257599	0.690132	1.336757	1.745884	2.11991	2.58349	2.92078	4.0150
17	0.257347	0.689195	1.333379	1.739607	2.10982	2.56693	2.89823	3.9651
18	0.257123	0.688364	1.330391	1.734064	2.10092	2.55238	2.87844	3.9216
19	0.256923	0.687621	1.327728	1.729133	2.09302	2.53948	2.86093	3.8834
20	0.256743	0.686954	1.325341	1.724718	2.08596	2.52798	2.84534	3.8495
21	0.256580	0.686352	1.323188	1.720743	2.07961	2.51765	2.83136	3.8193
22	0.256432	0.685805	1.321237	1.717144	2.07387	2.50832	2.81876	3.7921
23	0.256297	0.685306	1.319460	1.713872	2.06866	2.49987	2.80734	3.7676
24	0.256173	0.684850	1.317836	1.710882	2.06390	2.49216	2.79694	3.7454
25	0.256060	0.684430	1.316345	1.708141	2.05954	2.48511	2.78744	3.7251
26	0.255955	0.684043	1.314972	1.705618	2.05553	2.47863	2.77871	3.7066
27	0.255858	0.683685	1.313703	1.703288	2.05183	2.47266	2.77068	3.6896
28	0.255768	0.683353	1.312527	1.701131	2.04841	2.46714	2.76326	3.6739
29	0.255684	0.683044	1.311434	1.699127	2.04523	2.46202	2.75639	3.6594
30	0.255605	0.682756	1.310415	1.697261	2.04227	2.45726	2.75000	3.6460
z	0.253347	0.674490	1.281552	1.644854	1.95996	2.32635	2.57583	3.2905
CI	—	—	80%	90%	95%	98%	99%	99.9%

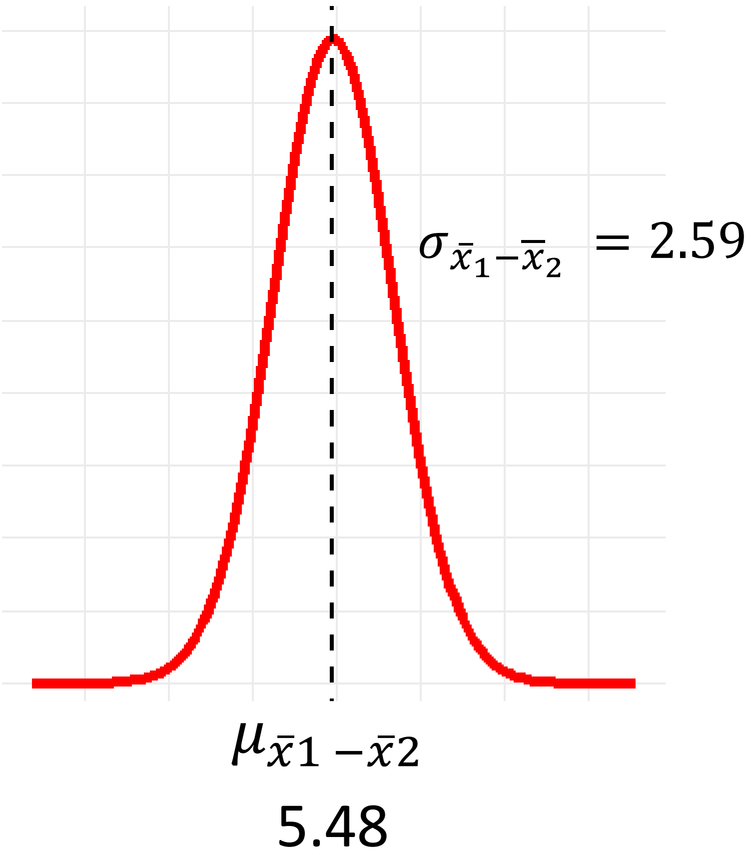
Most t-tables don't go past df = 30 !!!

df = 15 + 18 – 2 = 31

```
> qt(.975, df = 31)
[1] 2.039513
```

Our t value is 2.04

Sampling Distribution of Difference in Sample Means

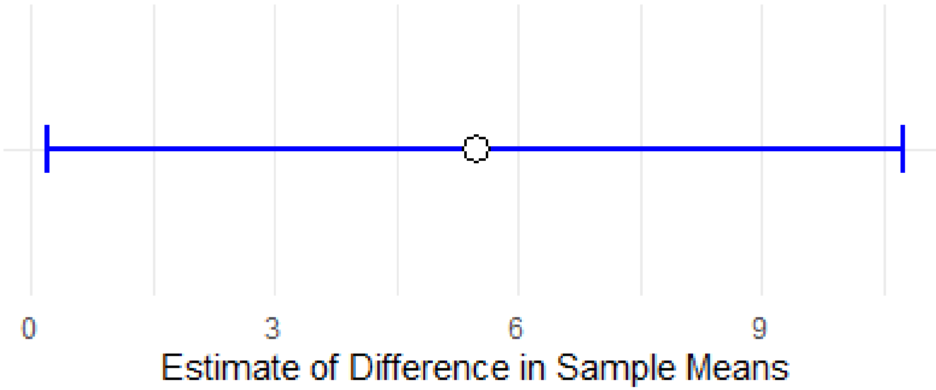


$$CI = \mu_{\bar{x}1-\bar{x}2} \pm t \times \sigma_{\bar{x}_1-\bar{x}_2}$$

$$CI = 5.48 \pm t \times 2.59$$

$$CI = 5.48 \pm 2.04 \times 2.59$$

CI = 5.48 [0.20, 10.76]



The 95% Confidence Interval for the true difference in means between the population of Anastasia students and the population of Bernadette students is:

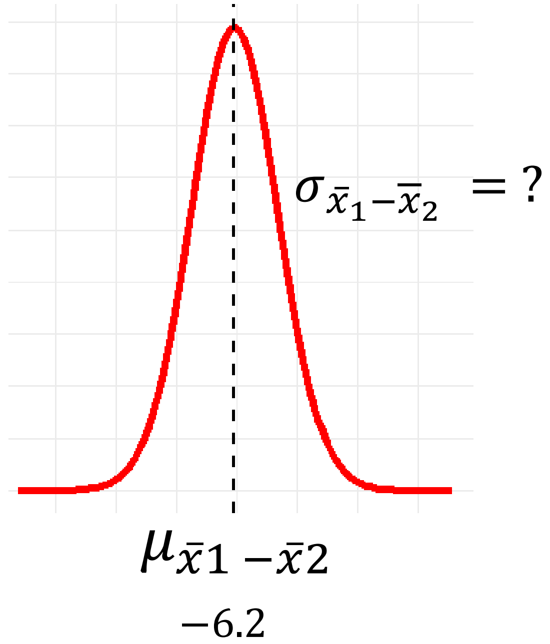
5.48 [0.20, 10.76]

Example 2 – Confidence Intervals for 2 samples.

Sometimes, we are not given the raw data, but just summary stats:

$\bar{x}_1 = 102.8 \qquad s_1 = 8.7 \qquad n_1 = 10$

$\bar{x}_2 = 109.0 \qquad s_2 = 8.3 \qquad n_2 = 11$

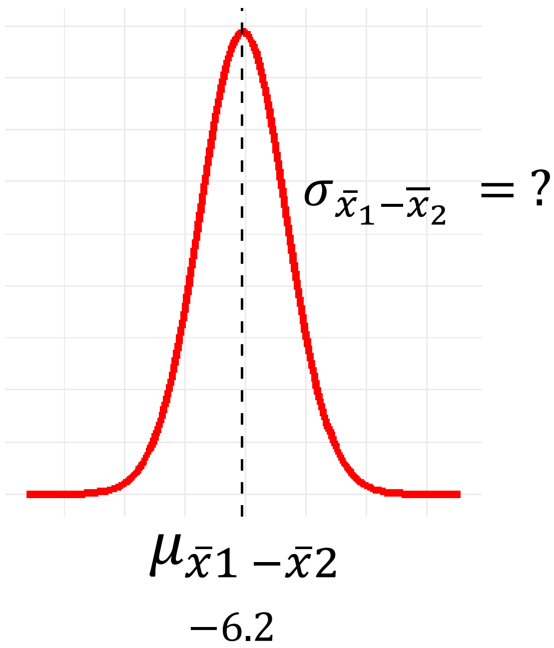


$$\sigma_{\bar{x}_1 - \bar{x}_2} = \hat{\sigma}_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Fortunately there is a shortcut to calculating $\hat{\sigma}_p$ if you don't have raw data:

$\bar{x}_1 = 102.8$ $s_1 = 8.7$ $n_1 = 10$

$\bar{x}_2 = 109.0$ $s_2 = 8.3$ $n_2 = 11$



$$\hat{\sigma}_p = \sqrt{\frac{w_1 {s_1}^2 + w_2 {s_2}^2}{w_1 + w_2}}$$

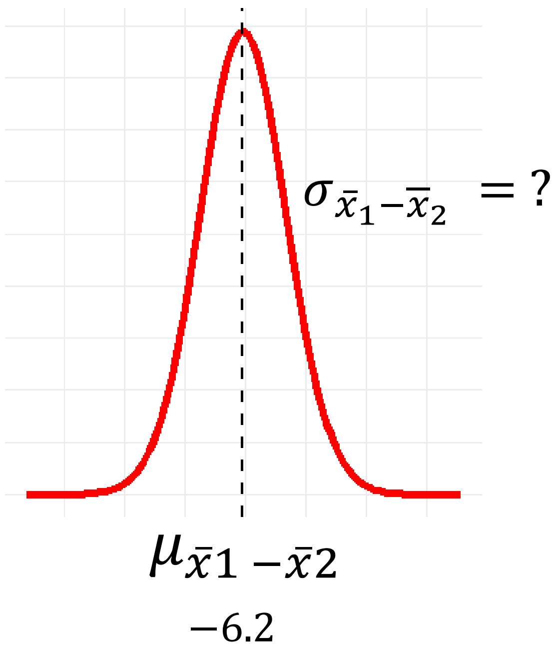
$w_1 = n_1 - 1$

$w_2 = n_2 - 1$

Fortunately there is a shortcut to calculating $\hat{\sigma}_p$ if you don't have raw data:

$\bar{x}_1 = 102.8$ $\quad s_1 = 8.7$ $\quad n_1 = 10$

$\bar{x}_2 = 109.0$ $\quad s_2 = 8.3$ $\quad n_2 = 11$



$$\hat{\sigma}_p = \sqrt{\frac{w_1{s_1}^2+w_2{s_2}^2}{w_1+w_2}} = 8.5$$

$w_1 = n_1 - 1$

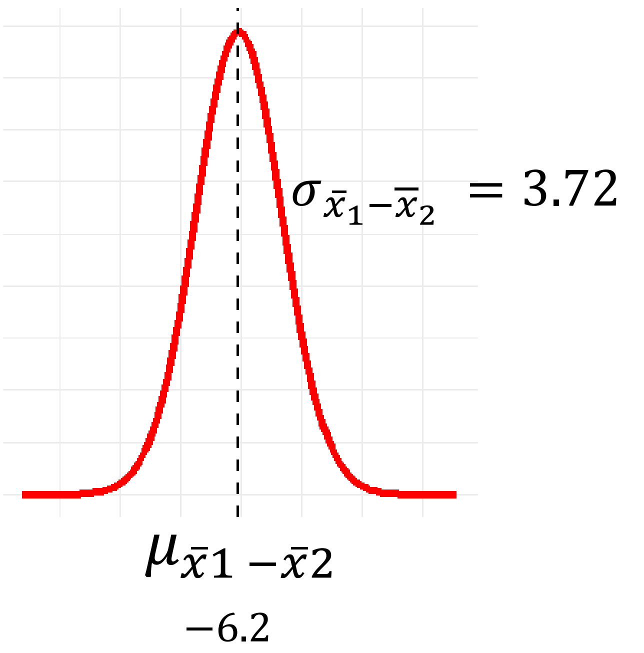
$w_2 = n_2 - 1$

We can then calculate $\sigma_{\bar{x}_1-\bar{x}_2}$ using the same formula as before:

$\bar{x}_1 = 102.8$ $\quad s_1 = 8.7$ $\quad n_1 = 10$

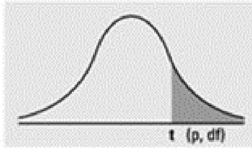
$\bar{x}_2 = 109.0$ $\quad s_2 = 8.3$ $\quad n_2 = 11$

$\hat{\sigma}_p = 8.5$



$$\sigma_{\bar{x}_1-\bar{x}_2} = \hat{\sigma}_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= 8.5 \times 0.44$$
$$= 3.72$$



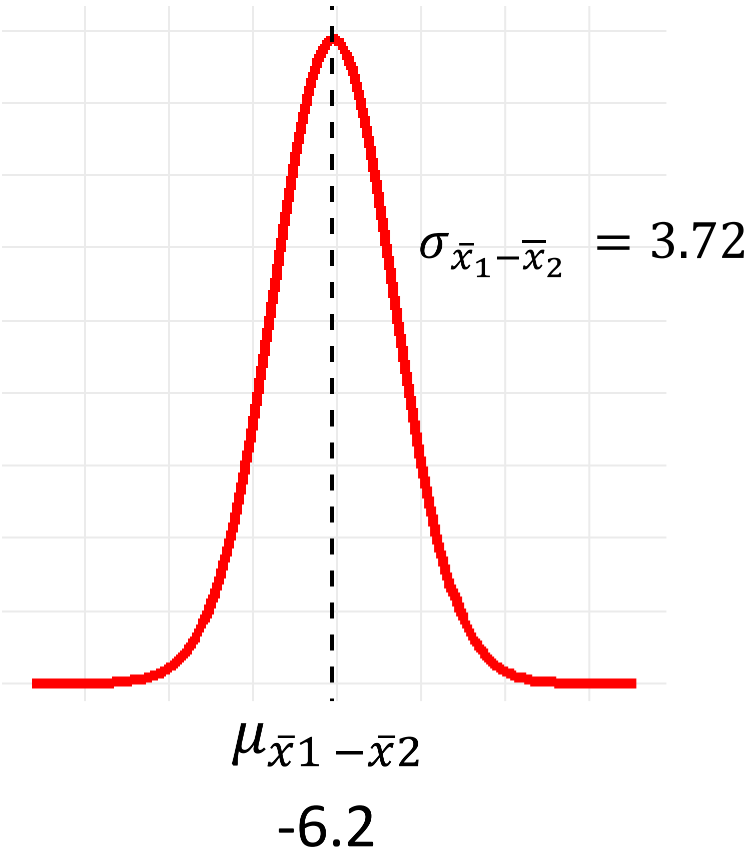
df/p	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.54070	5.84091	12.9240
4	0.270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.943180	2.44691	3.14267	3.70743	5.9588
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079
8	0.261921	0.706387	1.396815	1.859548	2.30600	2.89646	3.35539	5.0413
9	0.260955	0.702722	1.383029	1.833113	2.26216	2.82144	3.24984	4.7809
10	0.260185	0.699812	1.372184	1.812461	2.22814	2.76377	3.16927	4.5869
11	0.259556	0.697445	1.363430	1.795885	2.20099	2.71808	3.10581	4.4370
12	0.259033	0.695483	1.356217	1.782288	2.17881	2.68100	3.05454	43178
13	0.258591	0.693829	1.350171	1.770933	2.16037	2.65031	3.01228	4.2208
14	0.258213	0.692417	1.345030	1.761310	2.14479	2.62449	2.97684	4.1405
15	0.257885	0.691197	1.340606	1.753050	2.13145	2.60248	2.94671	4.0728
16	0.257599	0.690132	1.336757	1.745884	2.11991	2.58349	2.92078	4.0150
17	0.257347	0.689195	1.333379	1.739607	2.10982	2.56693	2.89823	3.9651
18	0.257123	0.688364	1.330391	1.734064	2.10092	2.55238	2.87844	3.9216
19	0.256923	0.687621	1.327728	1.729133	2.09302	2.53948	2.86093	3.8834
20	0.256743	0.686954	1.325341	1.724718	2.08596	2.52798	2.84534	3.8495
21	0.256580	0.686352	1.323188	1.720743	2.07961	2.51765	2.83136	3.8193
22	0.256432	0.685805	1.321237	1.717144	2.07387	2.50832	2.81876	3.7921
23	0.256297	0.685306	1.319460	1.713872	2.06866	2.49987	2.80734	3.7676
24	0.256173	0.684850	1.317836	1.710882	2.06390	2.49216	2.79694	3.7454
25	0.256060	0.684430	1.316345	1.708141	2.05954	2.48511	2.78744	3.7251
26	0.255955	0.684043	1.314972	1.705618	2.05553	2.47863	2.77871	3.7066
27	0.255858	0.683685	1.313703	1.703288	2.05183	2.47266	2.77068	3.6896
28	0.255768	0.683353	1.312527	1.701131	2.04841	2.46714	2.76326	3.6739
29	0.255684	0.683044	1.311434	1.699127	2.04523	2.46202	2.75639	3.6594
30	0.255605	0.682756	1.310415	1.697261	2.04227	2.45726	2.75000	3.6460
z	0.253347	0.674490	1.281552	1.644854	1.95996	2.32635	2.57583	3.2905
CI	—	—	80%	90%	95%	98%	99%	99.9%

df = 10 + 11 – 2 = 19

```
> qt(.975, df = 19)
[1] 2.093024
```

Our t value is 2.09

Sampling Distribution of Difference in Sample Means



$$CI = \mu_{\bar{x}1-\bar{x}2} \pm t \times \sigma_{\bar{x}_1-\bar{x}_2}$$

$$CI = -6.2 \pm t \times 3.72$$
$$CI = -6.2 \pm 2.09 \times 3.72$$

CI = -6.2 [-14.0, 1.6]

Conducting a Student t-test for 2 samples

Anastasia (n=15)

65 74 73 83 76 65 86 70 80 55 78 78 90 77 68

$$\overline{x} = 74.53$$

$$s = 9.00$$

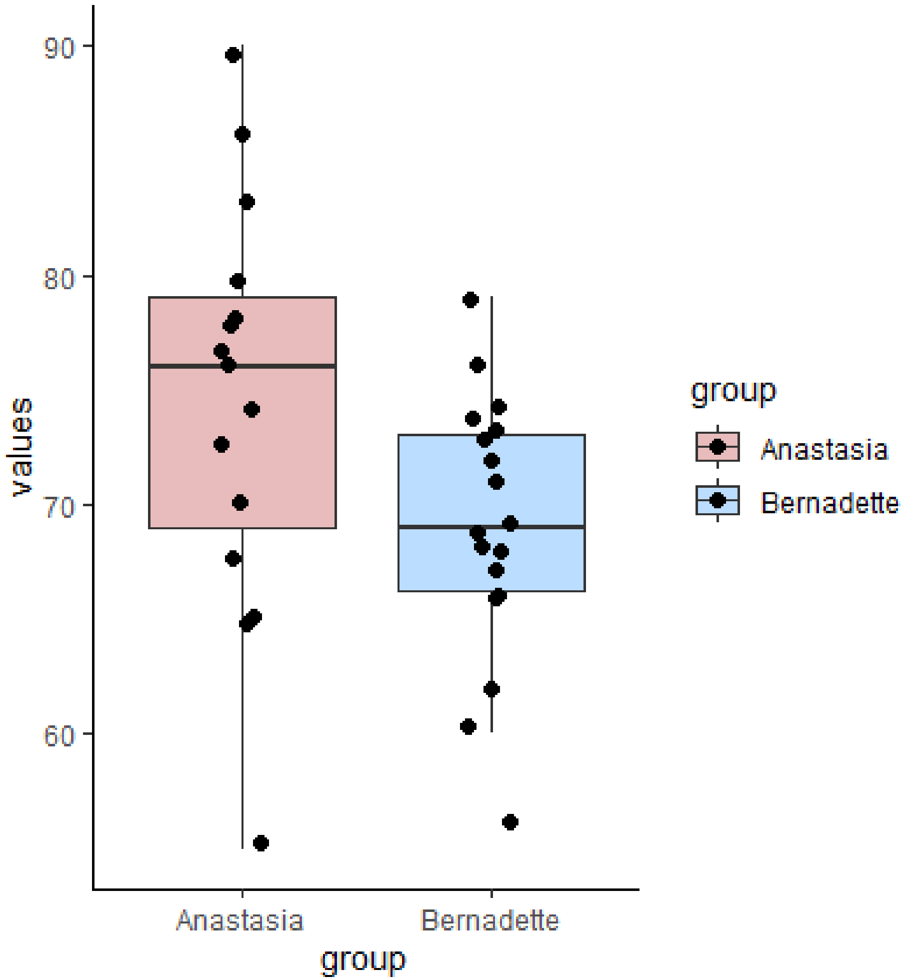
Bernadette (n=18)

72 66 71 66 76 69 79 73 62 69 68 60 73 68 67 74 56 74

$$\overline{x} = 69.06$$

$$s = 5.77$$

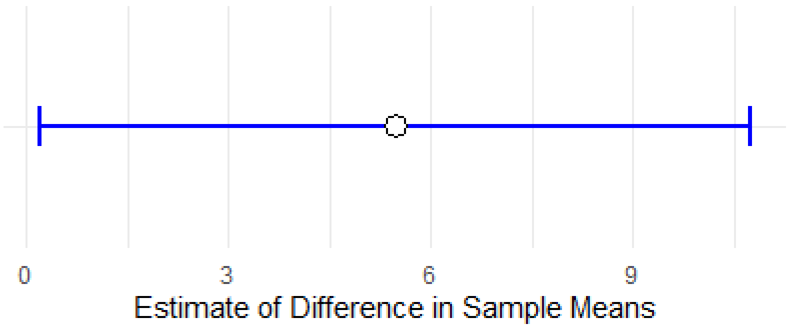
$$\bar{X}_1 - \bar{X}_2 = 5.48$$



$\bar{X}_1 - \bar{X}_2 = 5.47$

Is this a meaningful difference in the scores between the two groups?

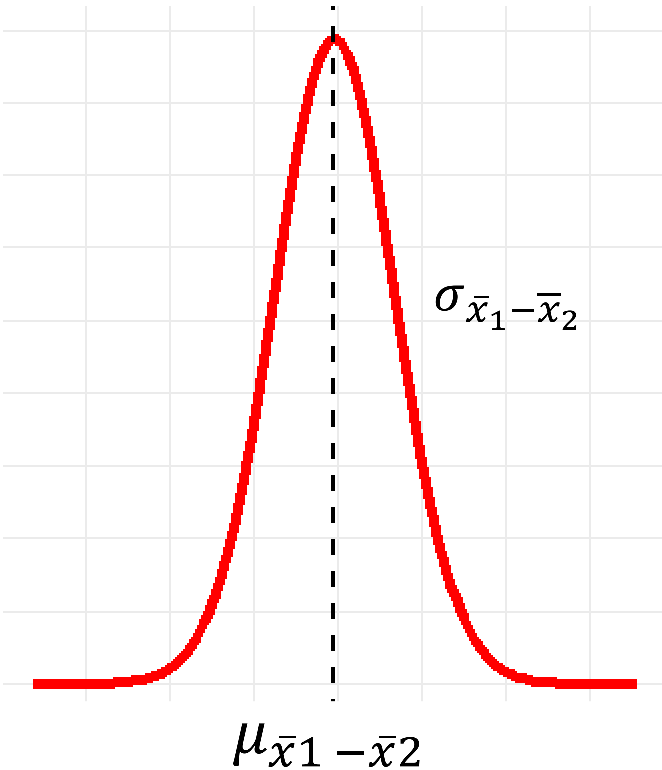
Or, is it possible to find differences of this size due to chance alone?



We already constructed a confidence interval based on the sampling distribution, that was:

5.48 [0.20, 10.76]

Sampling Distribution of the Difference in Sample Means



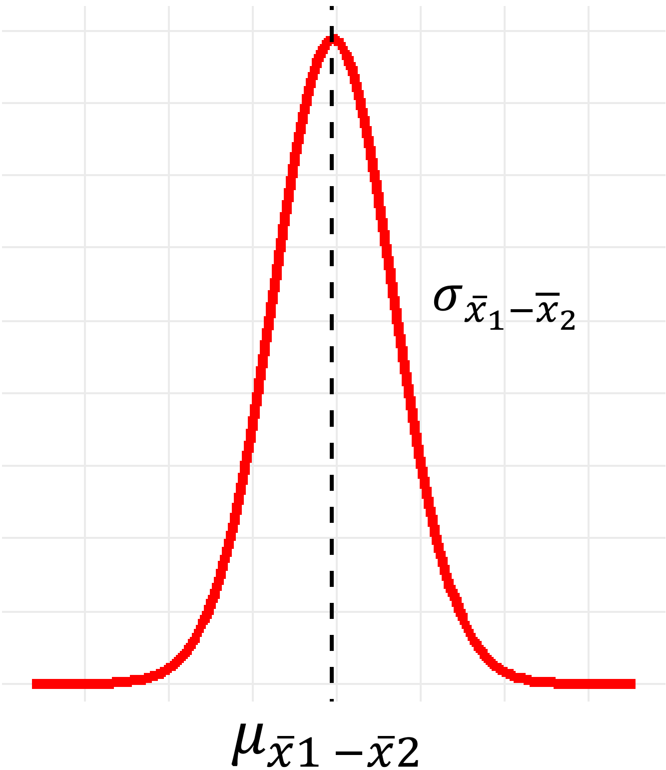
$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{x}1 - \bar{x}2}}$$

Using our two samples, we can estimate $\sigma_{\bar{x}_1 - \bar{x}_2}$

We're testing the hypothesis that $\mu_{\bar{x}1 - \bar{x}2} = 0$

We can then calculate t which is a measure of how many SD our observed difference in sample means $\bar{X}_1 - \bar{X}_2$ is from 0

Sampling Distribution of the Difference in Sample Means



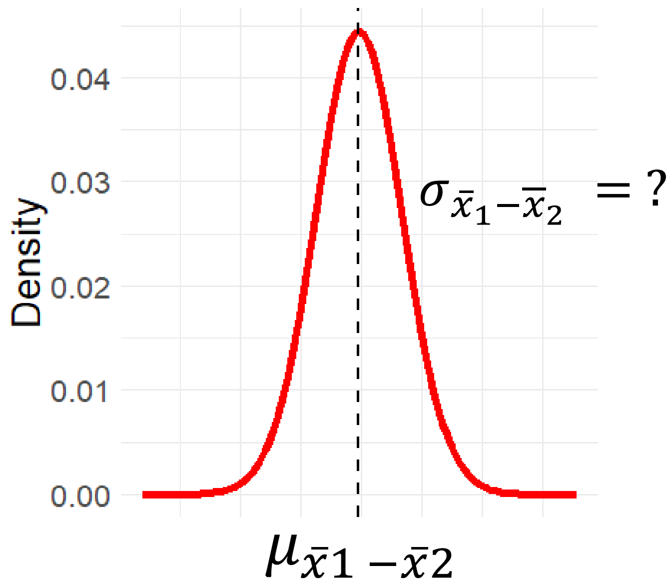
$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{x}1 - \bar{x}2}}$$

$$t = \frac{5.48}{2.59}$$

$$t = 2.12$$

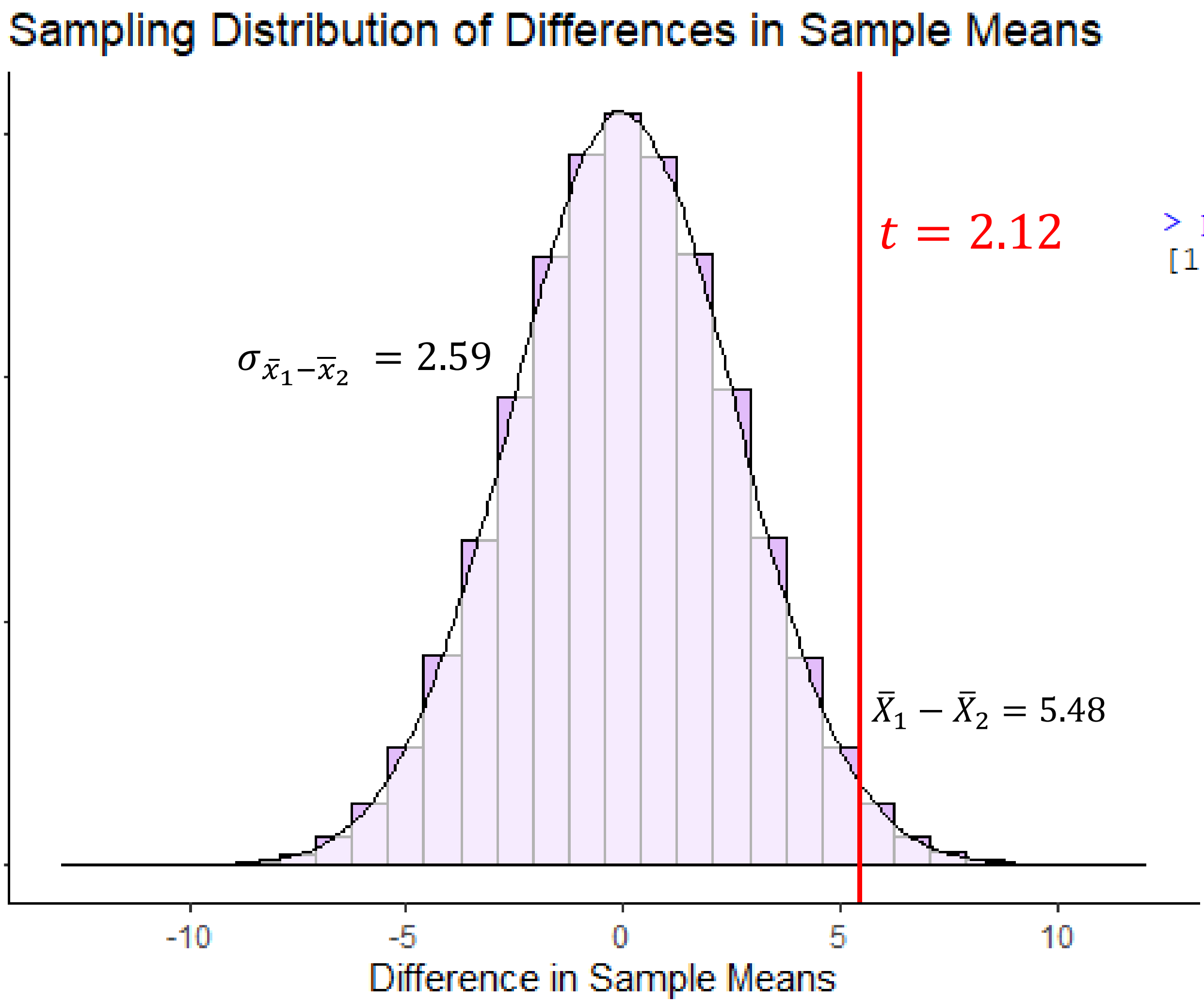
Where did we get our value of $\sigma_{\bar{x}1-\bar{x}2}$ from ?

$$\hat{\sigma}_p = \sqrt{\frac{w_1 {s_1}^2 + w_2 {s_2}^2}{w_1 + w_2}} \qquad w_1 = n_1 - 1 \qquad w_2 = n_2 - 1$$



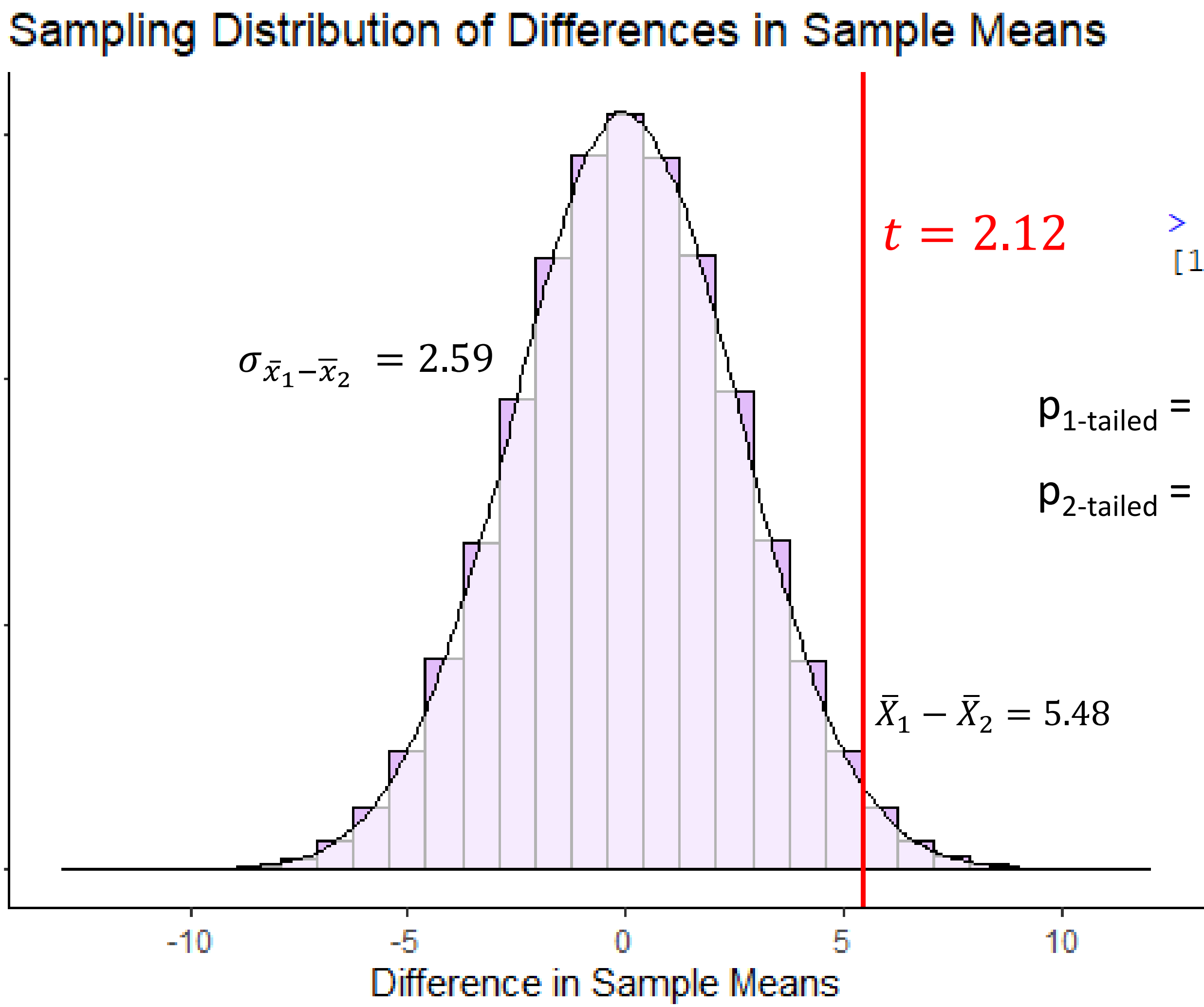
$$\hat{\sigma}_p = \sqrt{\frac{\sum(x_{ik} - \bar{x}_k)^2}{n-2}}$$

$$\sigma_{\bar{x}_1-\bar{x}_2} = \hat{\sigma}_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$




```
> pt(2.12, df = 31)
[1] 0.9789433
```

$$\mu_{\bar{x}1-\bar{x}2} = 0$$



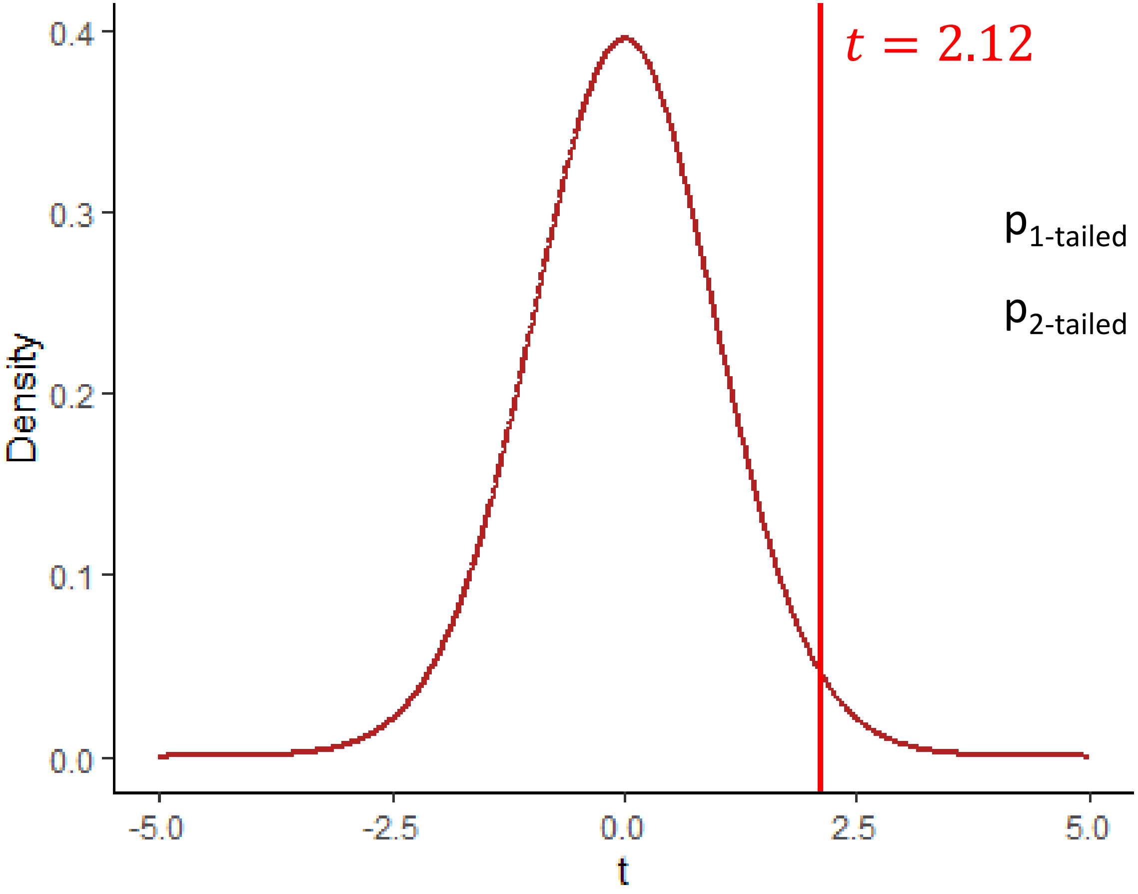

```
> pt(2.12, df = 31)
[1] 0.9789433
```

$p_{\text{1-tailed}} = 1 - 0.979 = 0.021$

$p_{\text{2-tailed}} = 0.021 * 2 = 0.042$

$$\mu_{\bar{x}1 - \bar{x}2} = 0$$

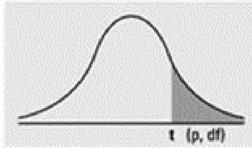
t distribution with d.f.=31




```
> pt(2.12, df = 31)
[1] 0.9789433
```

$p_{1\text{-tailed}} = 1 - 0.979 = 0.021$

$p_{2\text{-tailed}} = 0.021 * 2 = 0.042$

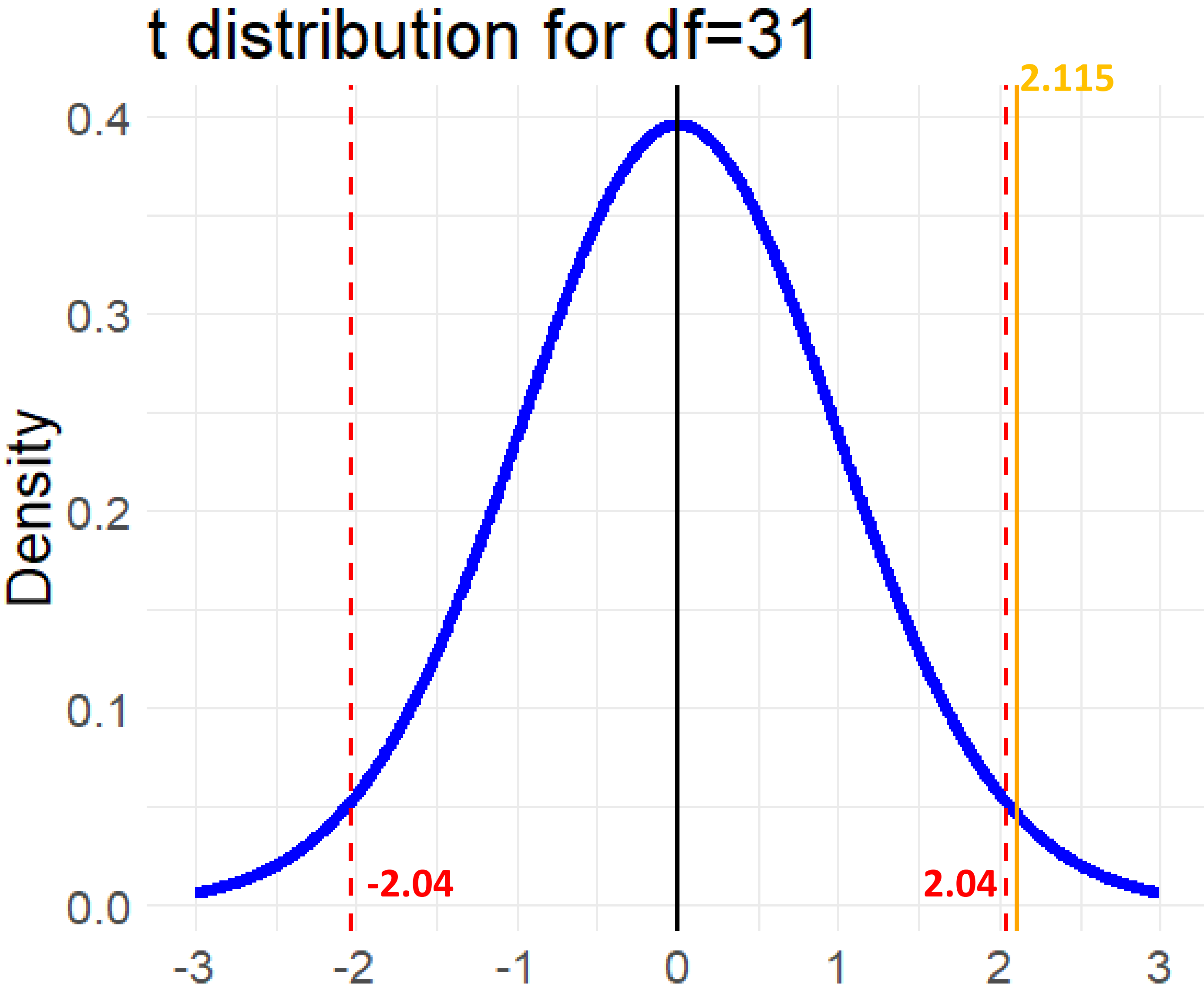


df/p	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.54070	5.84091	12.9240
4	0.270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.943180	2.44691	3.14267	3.70743	5.9588
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079
8	0.261921	0.706387	1.396815	1.859548	2.30600	2.89646	3.35539	5.0413
9	0.260955	0.702722	1.383029	1.833113	2.26216	2.82144	3.24984	4.7809
10	0.260185	0.699812	1.372184	1.812461	2.22814	2.76377	3.16927	4.5869
11	0.259556	0.697445	1.363430	1.795885	2.20099	2.71808	3.10581	4.4370
12	0.259033	0.695483	1.356217	1.782288	2.17881	2.68100	3.05454	43178
13	0.258591	0.693829	1.350171	1.770933	2.16037	2.65031	3.01228	4.2208
14	0.258213	0.692417	1.345030	1.761310	2.14479	2.62449	2.97684	4.1405
15	0.257885	0.691197	1.340606	1.753050	2.13145	2.60248	2.94671	4.0728
16	0.257599	0.690132	1.336757	1.745884	2.11991	2.58349	2.92078	4.0150
17	0.257347	0.689195	1.333379	1.739607	2.10982	2.56693	2.89823	3.9651
18	0.257123	0.688364	1.330391	1.734064	2.10092	2.55238	2.87844	3.9216
19	0.256923	0.687621	1.327728	1.729133	2.09302	2.53948	2.86093	3.8834
20	0.256743	0.686954	1.325341	1.724718	2.08596	2.52798	2.84534	3.8495
21	0.256580	0.686352	1.323188	1.720743	2.07961	2.51765	2.83136	3.8193
22	0.256432	0.685805	1.321237	1.717144	2.07387	2.50832	2.81876	3.7921
23	0.256297	0.685306	1.319460	1.713872	2.06866	2.49987	2.80734	3.7676
24	0.256173	0.684850	1.317836	1.710882	2.06390	2.49216	2.79694	3.7454
25	0.256060	0.684430	1.316345	1.708141	2.05954	2.48511	2.78744	3.7251
26	0.255955	0.684043	1.314972	1.705618	2.05553	2.47863	2.77871	3.7066
27	0.255858	0.683685	1.313703	1.703288	2.05183	2.47266	2.77068	3.6896
28	0.255768	0.683353	1.312527	1.701131	2.04841	2.46714	2.76326	3.6739
29	0.255684	0.683044	1.311434	1.699127	2.04523	2.46202	2.75639	3.6594
30	0.255605	0.682756	1.310415	1.697261	2.04227	2.45726	2.75000	3.6460
z	0.253347	0.674490	1.281552	1.644854	1.95996	2.32635	2.57583	3.2905
CI	———	———	80%	90%	95%	98%	99%	99.9%

Alternatively, we could calculate so-called "critical values" of t

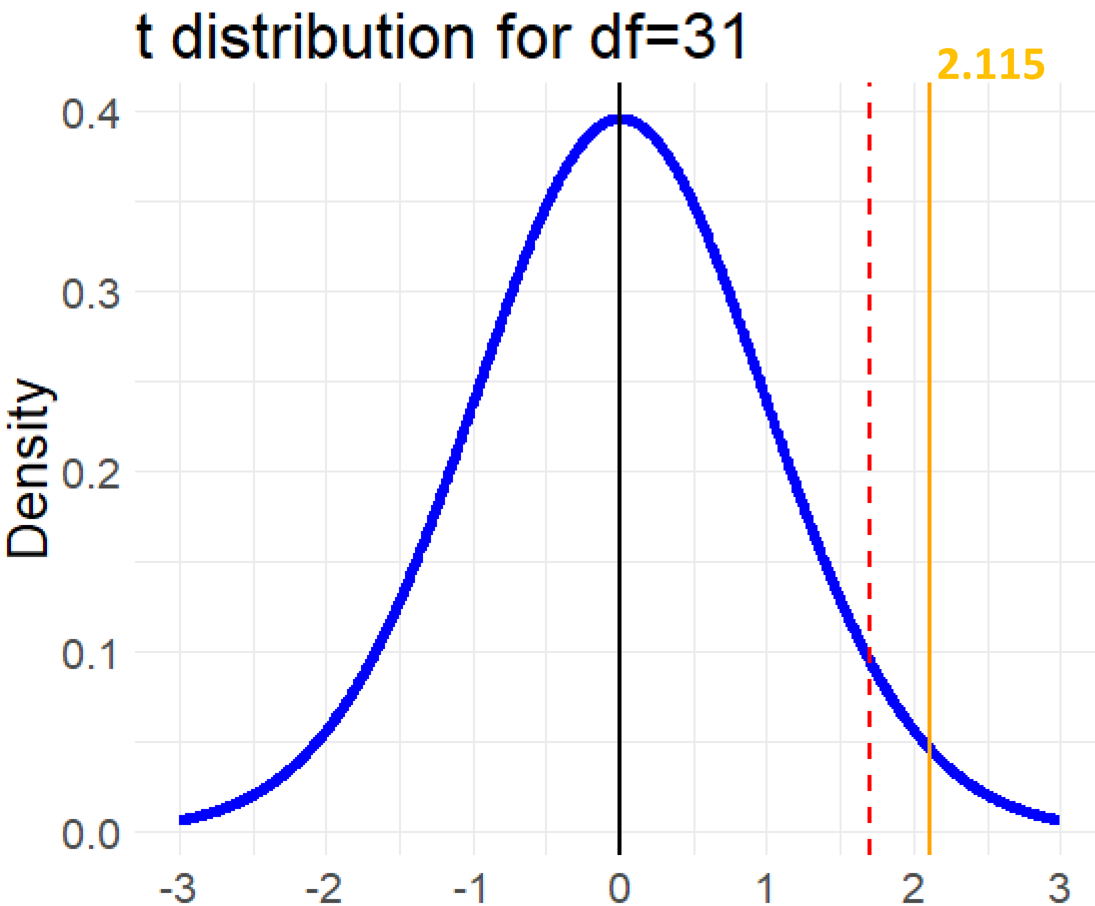
e.g. for a 2-tailed test:

```
> qt(c(0.025,0.975),df=31)
[1] -2.039513  2.039513
```



t distribution for df=31
Density
0.4
0.3
0.2
0.1
0.0
-3
-2
-1
0
1
2
3
-2.04
2.04
2.115

What if we had decided upon a one-tailed t-test ?




```
> qt(c(0.95),df=31)
[1] 1.695519
```

e.g. if we'd predicted that Anastasia would have students that perform better than Bernadette

Let’s look at this in R

```
> anastasia
 [1] 65 74 73 83 76 65 86 70 80 55 78 78 90 77 68
> bernadette
 [1] 72 66 71 66 76 69 79 73 62 69 68 60 73 68 67 74 56 74
```

```
> t.test(anastasia,bernadette, var.equal = T)

        Two Sample t-test

data:  anastasia and bernadette
t = 2.1154, df = 31, p-value = 0.04253
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.1965873 10.7589683
sample estimates:
mean of x mean of y
 74.53333  69.05556
```

The mean grade in Anastasia's class was 74.5 (std dev = 9.0), whereas the mean in Bernadette's class was 69.1 (std dev = 5.8). A Student's independent samples t-test showed that this 5.4 difference was significantly different from 0 ($t(31)=2.1$, $p<.05$, $CI_{95}=[0.2,10.8]$, $d=.74$), suggesting that a genuine difference in learning outcomes has occurred.

Remember there's another version of this test…

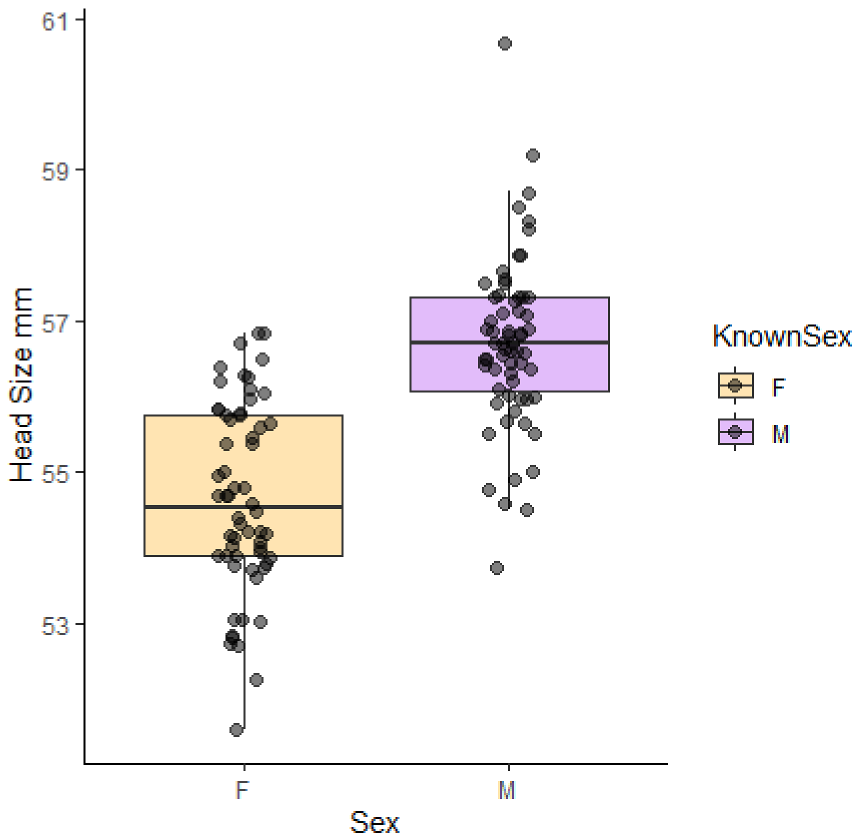
```
> anastasia
 [1] 65 74 73 83 76 65 86 70 80 55 78 78 90 77 68
> bernadette
 [1] 72 66 71 66 76 69 79 73 62 69 68 60 73 68 67 74 56 74
```

```
> t.test(anastasia,bernadette)

        Welch Two Sample t-test

data:  anastasia and bernadette
t = 2.0342, df = 23.025, p-value = 0.05361
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.09249349 11.04804904
sample estimates:
mean of x mean of y
 74.53333  69.05556
```

Let's briefly look at a second Student's t-test example …

(Head size in BlueJays)

```
> headf
 [1] 53.77 52.25 52.83 54.94 54.69 55.75 53.05 54.81 54.09 56.82 55.39 54.21 56.37 53.70 55.83 55.47
[17] 54.49 56.04 54.13 56.50 54.31 53.04 55.58 56.84 53.74 55.74 55.37 52.80 54.20 53.90 53.90 54.00
[33] 55.00 54.40 53.60 56.20 56.70 55.76 55.70 52.74 55.97 53.88 54.04 53.03 55.65 54.70 51.60 54.69
[49] 54.58 53.80 54.15 56.25 54.80 56.09 55.82 53.95 56.28 53.90 54.19 52.70
> headm
 [1] 56.58 56.36 57.32 57.32 57.12 60.67 56.54 56.48 57.34 58.70 57.07 55.79 58.31 54.51 58.21 56.80
[17] 56.88 55.96 55.96 56.59 57.50 56.01 53.74 56.80 57.86 56.82 56.35 56.40 57.00 56.86 57.64 57.86
[33] 55.02 57.30 56.90 57.30 56.70 56.30 55.50 56.10 54.90 59.20 56.70 55.50 56.60 57.50 56.20 56.48
[49] 56.45 54.76 56.44 58.50 55.64 57.56 56.00 55.91 57.10 56.85 56.71 57.25 54.58 56.61 55.68
```



$n_F = 60$ $\quad n_M = 63$

$$\bar{x}_F = 54.65$$

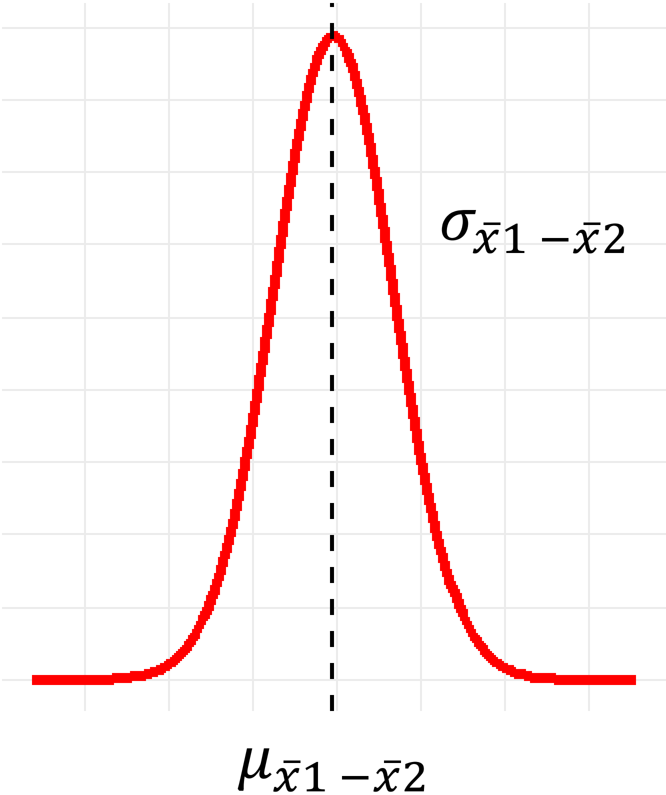
$$s_F = 1.24$$

$$\bar{x}_M = 56.69$$

$$s_M = 1.14$$

$$\bar{X}_1 - \bar{X}_2 = 2.05$$

What does the sampling distribution of differences in sample means look like?



df =

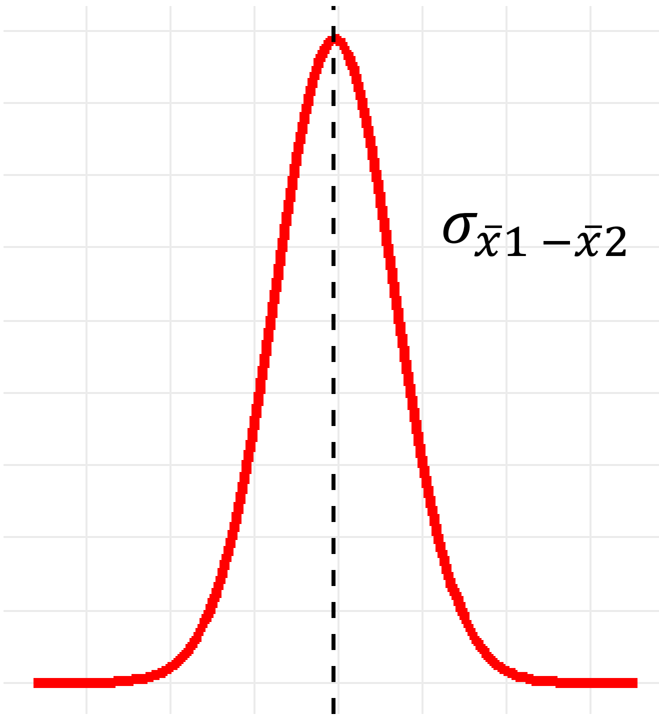
$\mu_{\bar{x}1-\bar{x}2} =$

$\sigma_{\bar{x}1-\bar{x}2} =$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{x}1-\bar{x}2}}$$

$$\bar{X}_1 - \bar{X}_2 = 2.05$$

How do we calculate $\sigma_{\bar{x}1 - \bar{x}2}$?



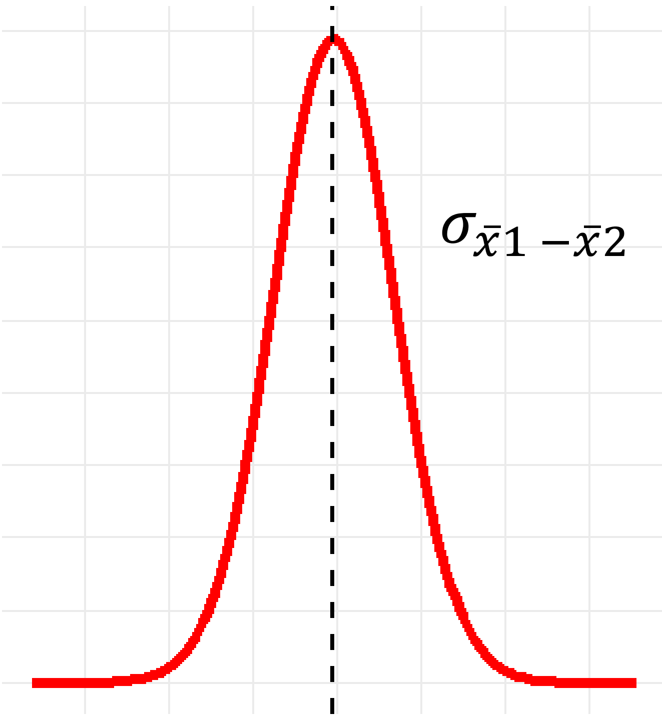
$$\sigma_{\bar{x}_1 - \bar{x}_2} = \hat{\sigma}_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Two options for calculating $\hat{\sigma}_p$ …

$$\hat{\sigma}_p = \sqrt{\frac{w_1 {s_1}^2 + w_2 {s_2}^2}{w_1 + w_2}}$$

$$\hat{\sigma}_p = \sqrt{\frac{\sum (x_{ik} - \bar{x}_k)^2}{n - 2}}$$

How do we calculate $\sigma_{\bar{x}1-\bar{x}2}$?



$$\sigma_{\bar{x}_1-\bar{x}_2} = \hat{\sigma}_p \times \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$\sigma_{\bar{x}_1-\bar{x}_2} = 1.19 \times \sqrt{\frac{1}{63}+\frac{1}{60}}$$

$$\sigma_{\bar{x}_1-\bar{x}_2} = 0.215$$

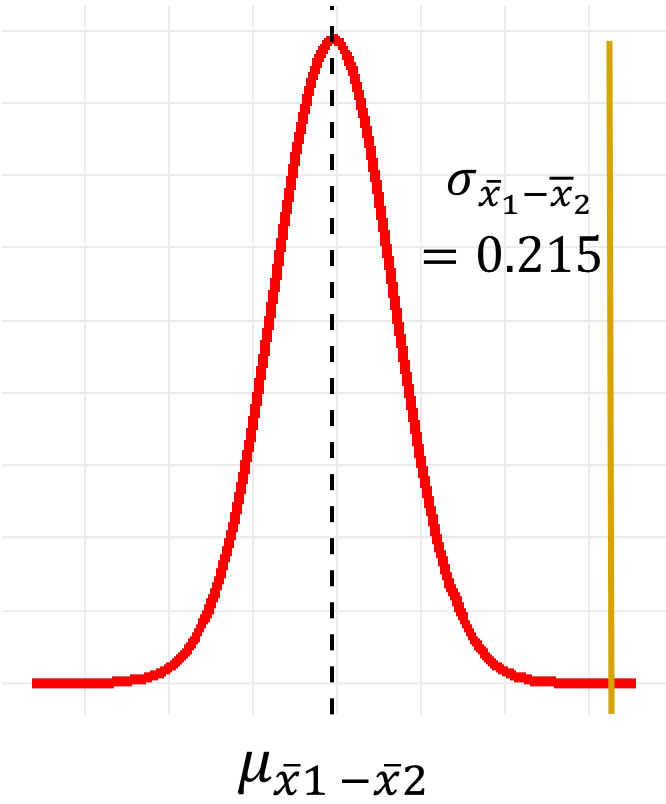
$\hat{\sigma}_p$ = 1.19

$s_F = 1.24$

$s_M = 1.14$

$\bar{X}_1 - \bar{X}_2 = 2.05$

Calculating *t* – how unusual was our one sample difference in means?



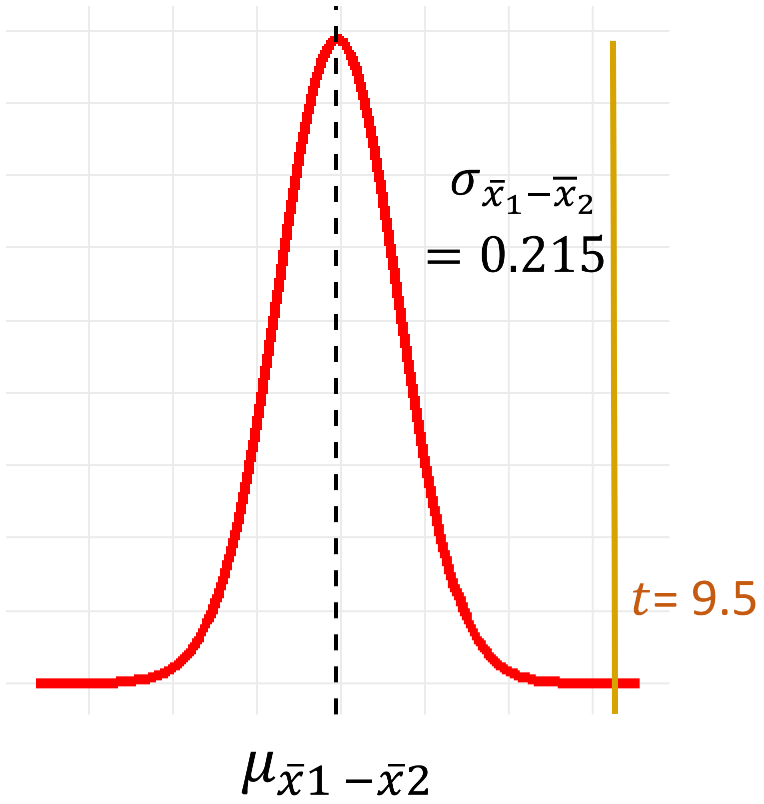
$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{x}1 - \bar{x}2}}$$

$$\bar{X}_1 - \bar{X}_2 = 2.05$$

$$t = \frac{2.05}{0.215} = 9.5$$

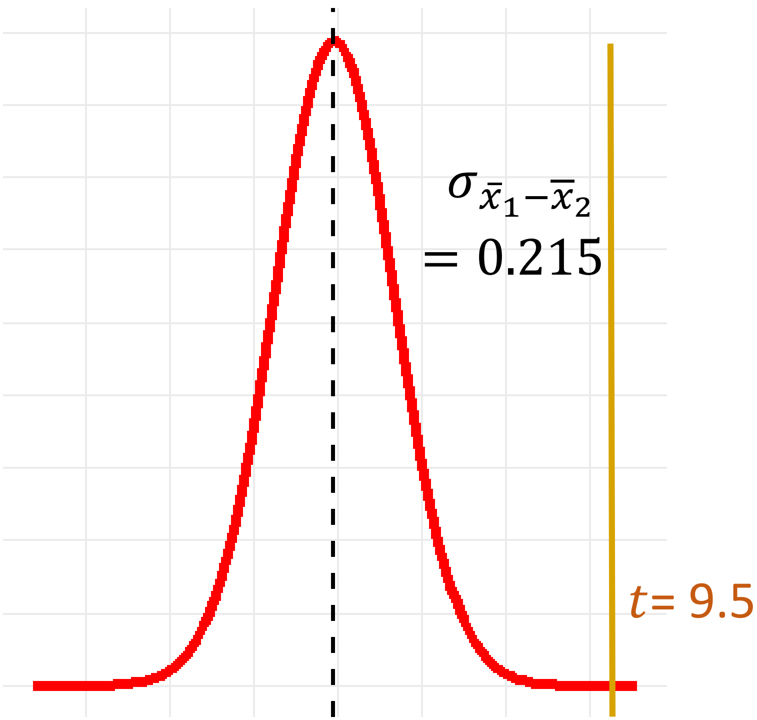
```
> 1 - (pt(9.520152, df=121))
[1] 1.110223e-16
```

If we cared about “critical values” we could calculate them for 2-tailed and 1-tailed tests

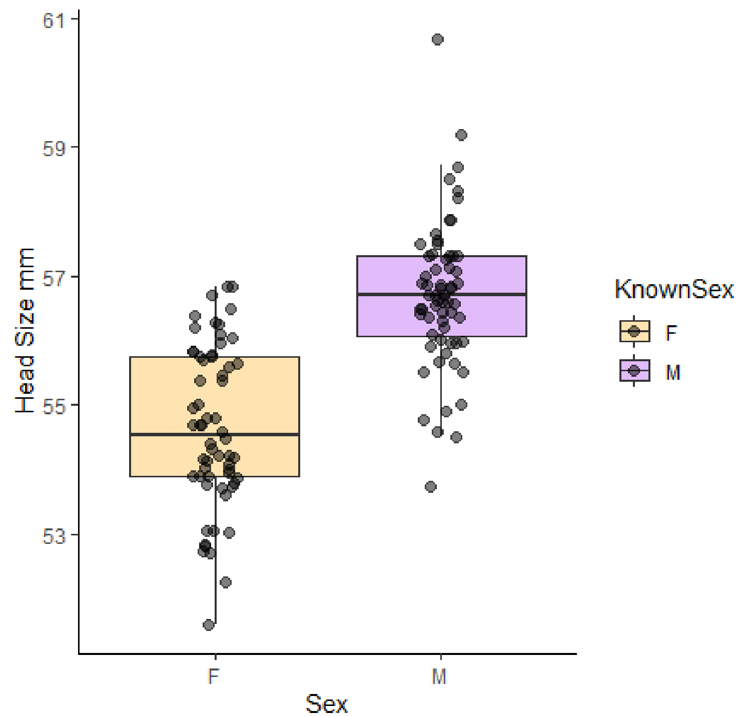



```
> qt(.975, df = 121)
[1] 1.979764
> qt(.95, df = 121)
[1] 1.657544
```

```
> t.test(headm, headf, var.equal = T)

        Two Sample t-test

data:  headm and headf
t = 9.5202, df = 121, p-value = 2.238e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.620977 2.472166
sample estimates:
mean of x mean of y
 56.69190  54.64533
```



$\mu_{\bar{x}1-\bar{x}2}$



$\bar{X}_1 - \bar{X}_2 = 2.05$

Remember the Assumptions of the independent t-test

Normality
Assumes that the true population distribution is normal

Independence
Observations are not correlated with each other (are independent of each other)

Homogeneity of Variance (homoscedasticity)
Assumes population variances/standard deviations are equal between groups (unless do Welch’s version)

Effect Sizes

Just because you observe a "significant" difference in means between two groups doesn't mean that it's interesting or relevant….

i.e. being 'significantly different' doesn't tell you how ***BIG*** *the difference is – i.e. how* ***LARGE*** *the* ***effect size*** *is.*

Cohen’s d - Measure of Effect Size

$$\delta = \frac{\bar{X}_1 - \bar{X}_2}{\hat{\sigma}_p}$$

e.g. for Student’s T-test

d-value	rough interpretation
about 0.2	small effect
about 0.5	moderate effect
about 0.8	large effect

There are several versions of this measure – this is the one we’ll use

$\hat{\sigma}_p = 1.19$

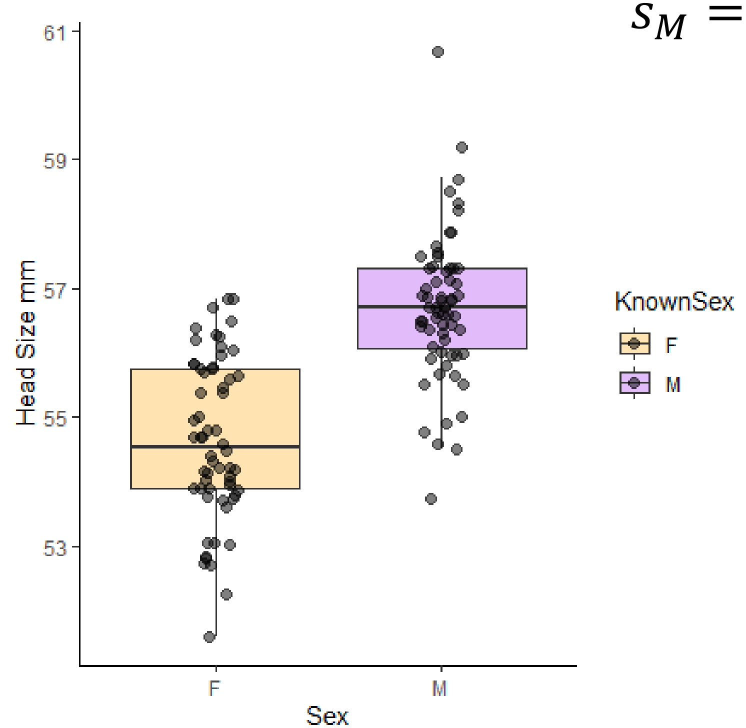
$s_F = 1.24$

$s_M = 1.14$

$$\delta = \frac{\bar{X}_1 - \bar{X}_2}{\hat{\sigma}_p}$$

$$\delta = \frac{2.05}{1.19}$$

$$\delta = 1.72$$



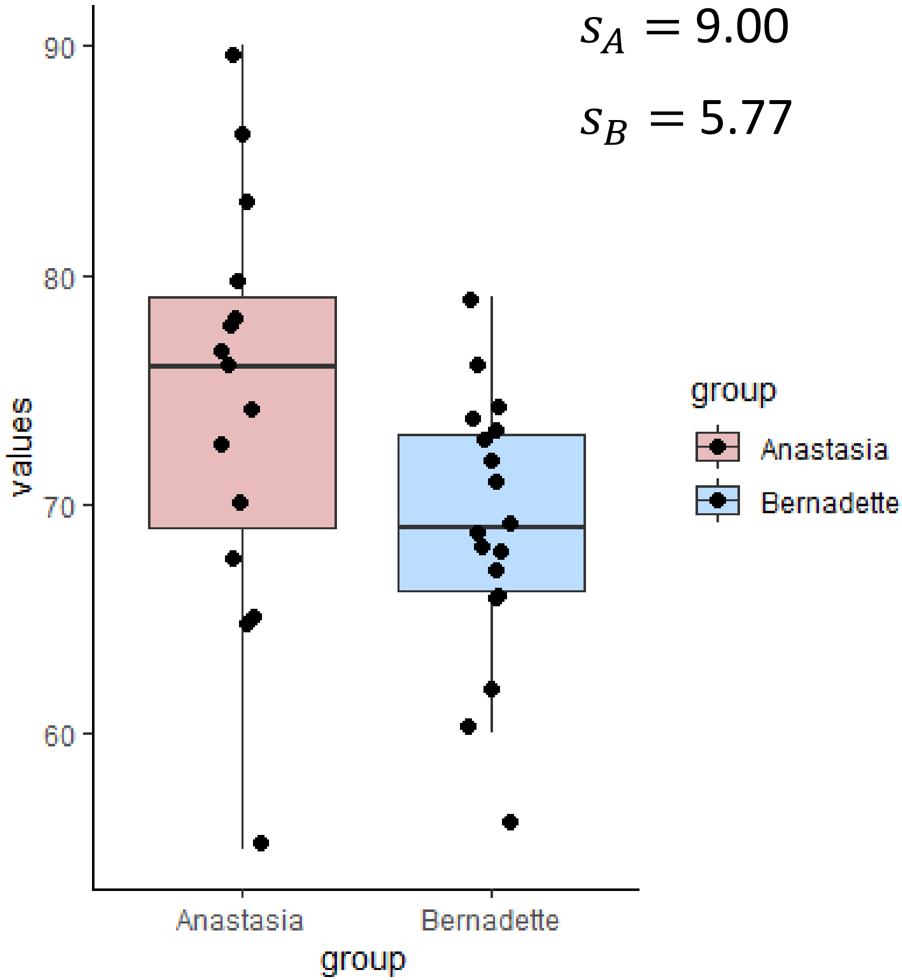
t(121) = 9.5, p<0.001

$$\bar{X}_1 - \bar{X}_2 = 2.05$$

$\hat{\sigma}_p = 7.41$

$s_A = 9.00$

$s_B = 5.77$



t(31) = 2.12, p=0.042

$\bar{X}_1 - \bar{X}_2 = 5.48$

$$\delta = \frac{\bar{X}_1 - \bar{X}_2}{\hat{\sigma}_p}$$

$$\delta = \frac{5.48}{7.41}$$

$$\delta = 0.74$$

Paired Data

id	grade_test1	grade_test2
student1	42.9	44.6
student2	51.8	54
student3	71.7	72.3
student4	51.6	53.4
student5	63.5	63.8
student6	58	59.3
student7	59.8	60.8
student8	50.8	51.6
student9	62.5	64.3
student10	61.9	63.2
student11	50.4	51.8
student12	52.6	52.2
student13	63	63
student14	58.3	60.5
student15	53.3	57.1
student16	58.7	60.1
student17	50.1	51.7
student18	64.2	65.6
student19	57.4	58.3
student20	57.1	60.1

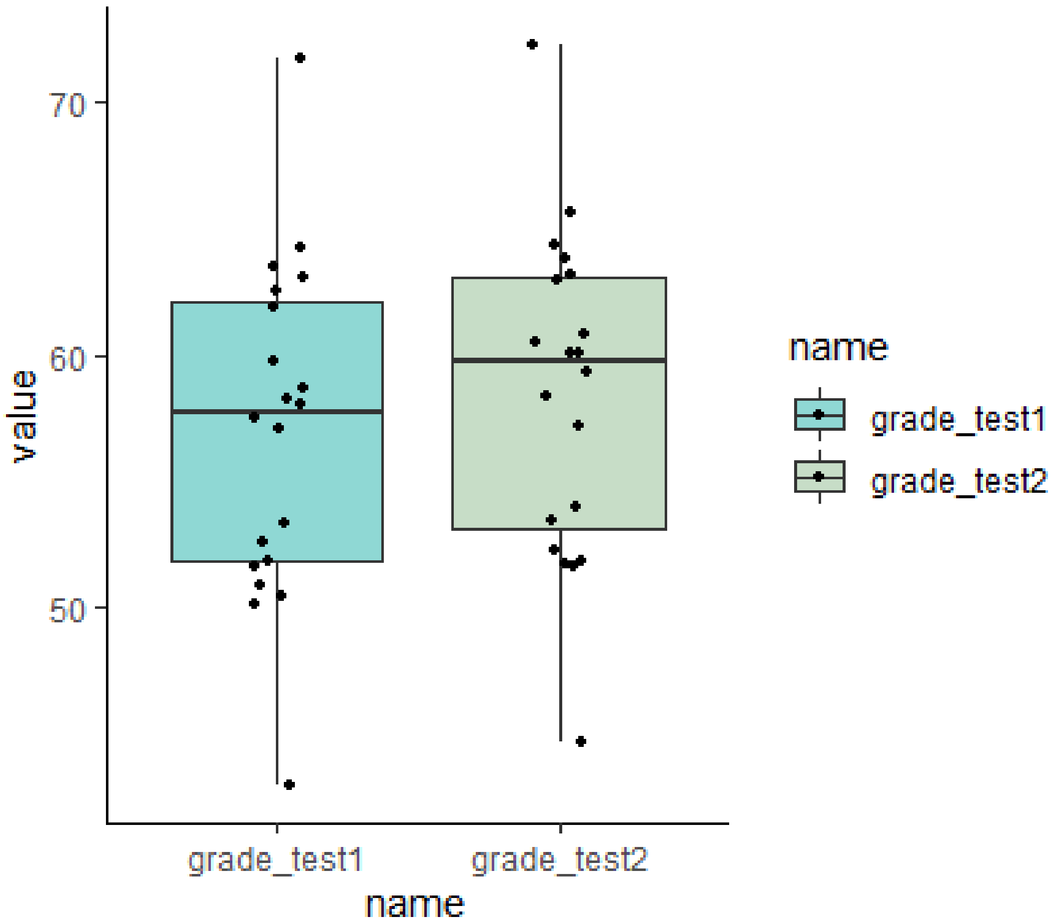
$$n_1 = 20$$
$$\bar{x}_1 = 56.98 \quad s_1 = 6.62$$

$$n_2 = 20$$
$$\bar{x}_2 = 58.38 \quad s_2 = 6.41$$

Did students improve their grades scores on test 2 compared to test 1 ?

Boxplots of Data

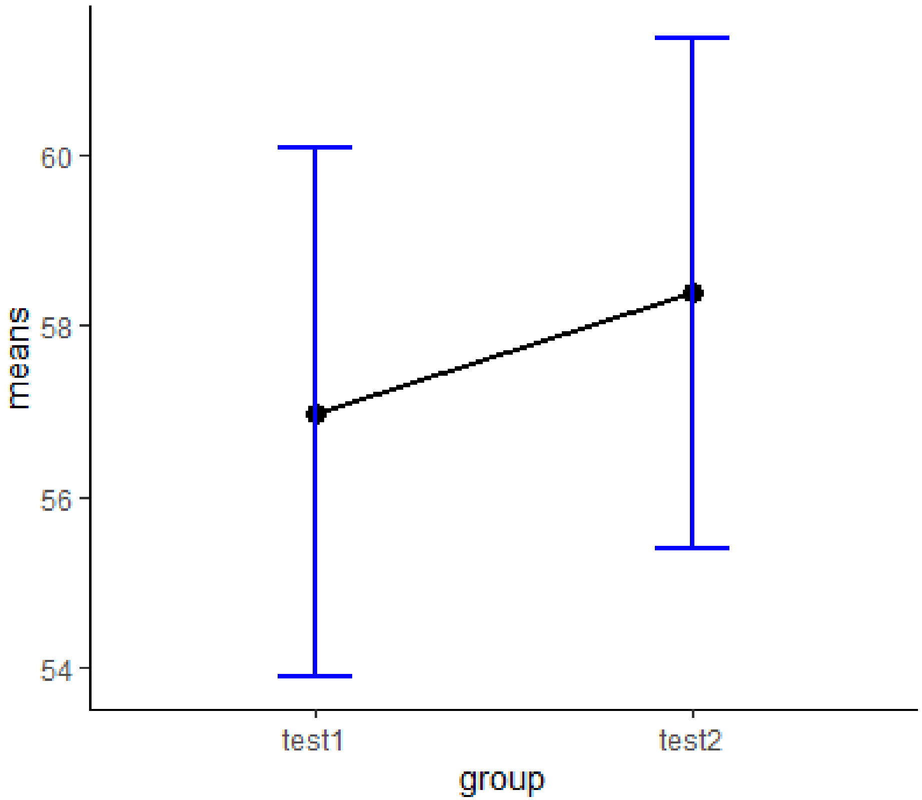
Boxplots are not necessarily the best way of plotting paired data



$$n_1 = 20$$
$$\bar{x}_1 = 56.98 \quad s_1 = 6.62$$

$$n_2 = 20$$
$$\bar{x}_2 = 58.38 \quad s_2 = 6.41$$

95% Confidence Intervals of Data

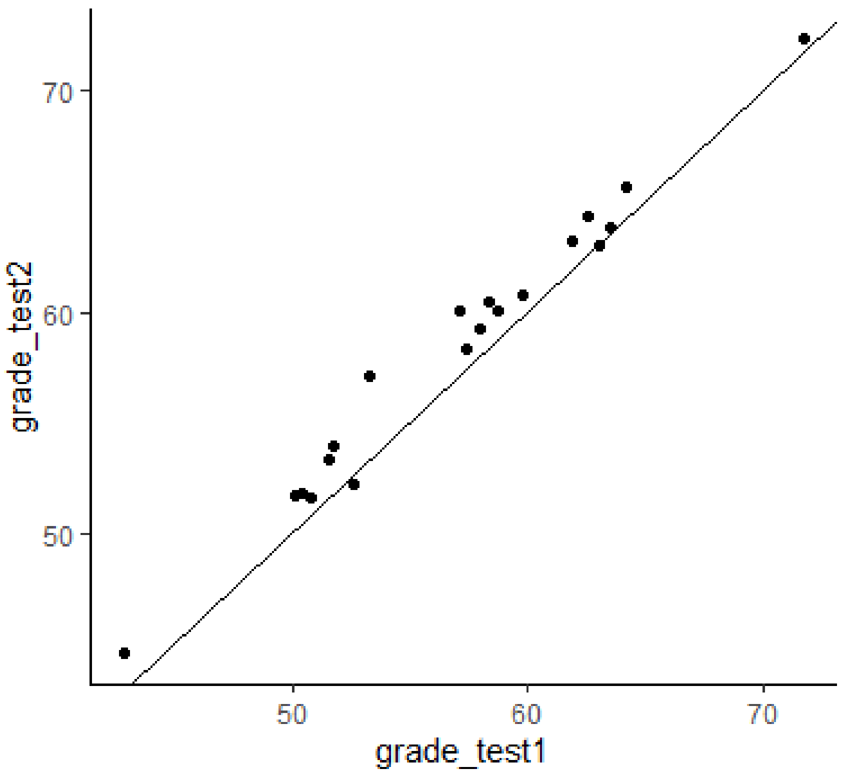


95% Confidence Intervals of each group also doesn't quite convey if there is a strong difference or not

$$n_1 = 20$$
$$\bar{x}_1 = 56.98 \quad s_1 = 6.62$$

$$n_2 = 20$$
$$\bar{x}_2 = 58.38 \quad s_2 = 6.41$$

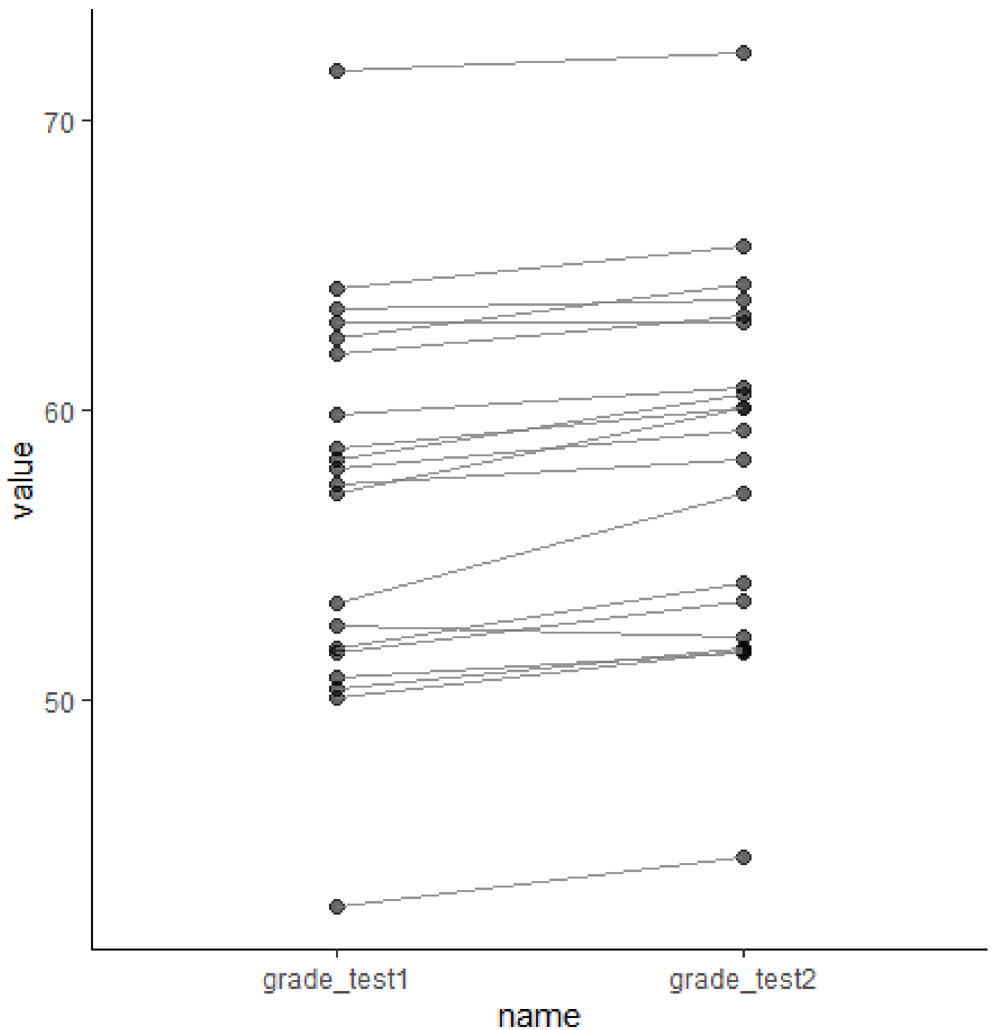
Scatterplot of Data



The diagonal line represents equal scores on test1 and test2

Now it’s clear that the vast majority of students are doing better on test2 compared to test1

Slope Graph of Data



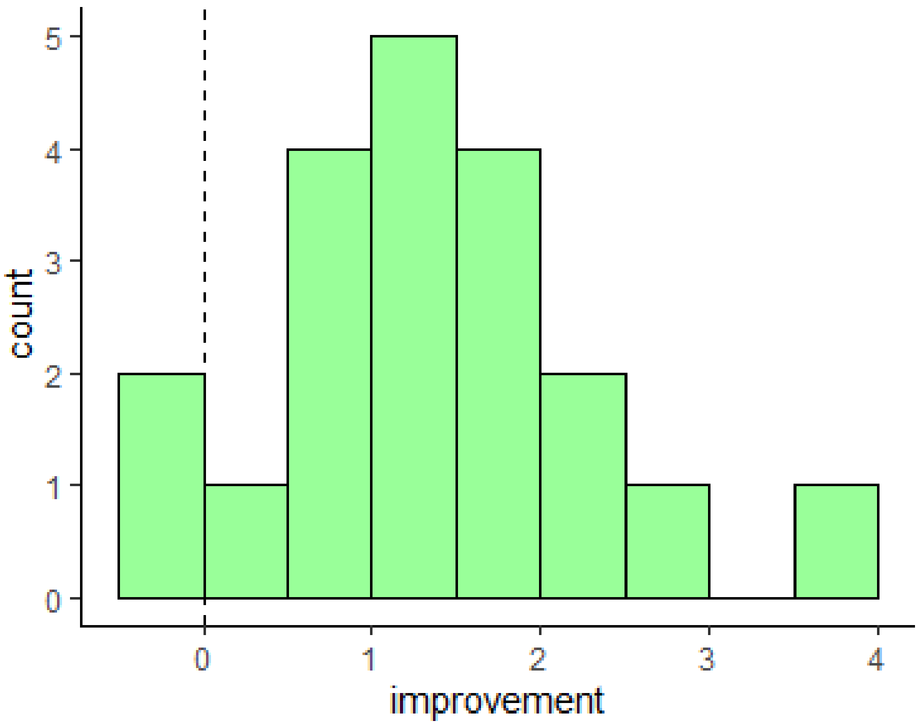
Slope graphs connect the scores of each individual in the two groups

This graph shows moderate increases on average in test2 compared to test1

id	grade_test1	grade_test2	improvement
student1	42.9	44.6	1.7
student2	51.8	54	2.2
student3	71.7	72.3	0.6
student4	51.6	53.4	1.8
student5	63.5	63.8	0.3
student6	58	59.3	1.3
student7	59.8	60.8	1.0
student8	50.8	51.6	0.8
student9	62.5	64.3	1.8
student10	61.9	63.2	1.3
student11	50.4	51.8	1.4
student12	52.6	52.2	-0.4
student13	63	63	0.0
student14	58.3	60.5	2.2
student15	53.3	57.1	3.8
student16	58.7	60.1	1.4
student17	50.1	51.7	1.6
student18	64.2	65.6	1.4
student19	57.4	58.3	0.9
student20	57.1	60.1	3.0

We can measure the difference between test2 and test1

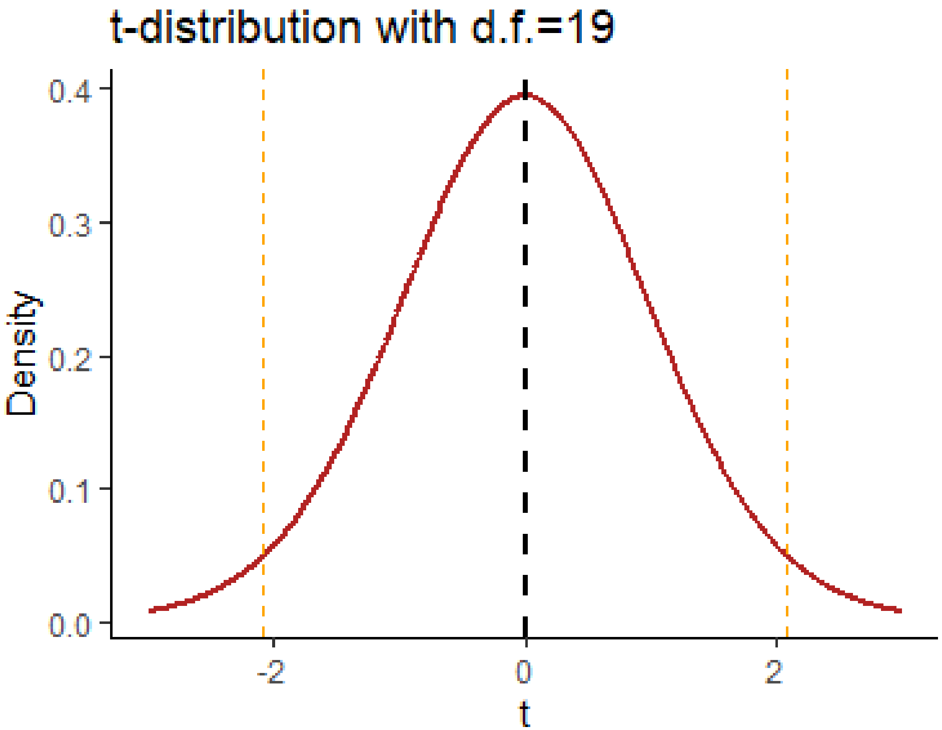
Histogram of Difference Scores



Most difference scores are above 0, indicating that the majority of students showed improvement

$\bar{x}_d = 1.41 \quad s_d = 0.97$

Confidence interval for one-sample



$\bar{x}_d = 1.41$

$$CI = \bar{X} \pm t^*(\frac{s}{\sqrt{n}})$$

$$CI = 1.41 \pm 2.09^*(\frac{0.97}{\sqrt{20}})$$

$$CI = 1.41[0.95 , 1.86]$$

```
> qt(.975, df=19)
[1] 2.093024
```

Paired Samples T-test

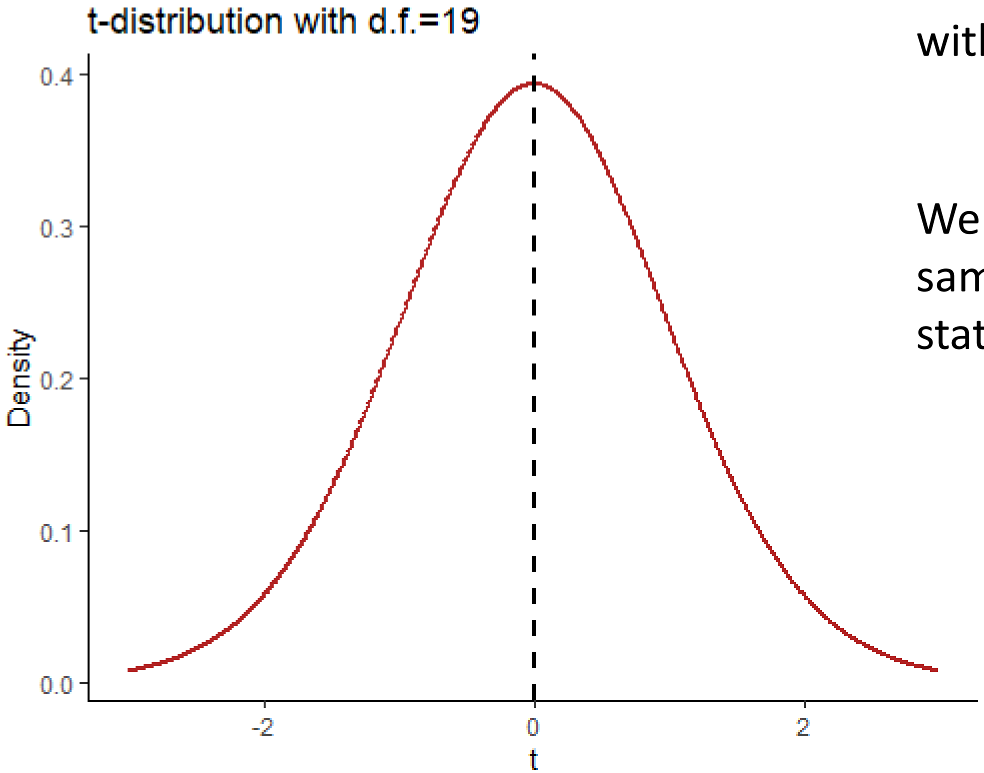
id	grade_test1	grade_test2	improvement
student1	42.9	44.6	1.7
student2	51.8	54	2.2
student3	71.7	72.3	0.6
student4	51.6	53.4	1.8
student5	63.5	63.8	0.3
student6	58	59.3	1.3
student7	59.8	60.8	1.0
student8	50.8	51.6	0.8
student9	62.5	64.3	1.8
student10	61.9	63.2	1.3
student11	50.4	51.8	1.4
student12	52.6	52.2	-0.4
student13	63	63	0.0
student14	58.3	60.5	2.2
student15	53.3	57.1	3.8
student16	58.7	60.1	1.4
student17	50.1	51.7	1.6
student18	64.2	65.6	1.4
student19	57.4	58.3	0.9
student20	57.1	60.1	3.0

$\bar{x}_d = 1.41 \quad s_d = 0.97$

$$H_0:\ \mu_d = 0$$

$$H_1:\ \mu_d \neq 0$$

We can assume that our one sample mean of difference scores, comes from a sampling distribution of sample means



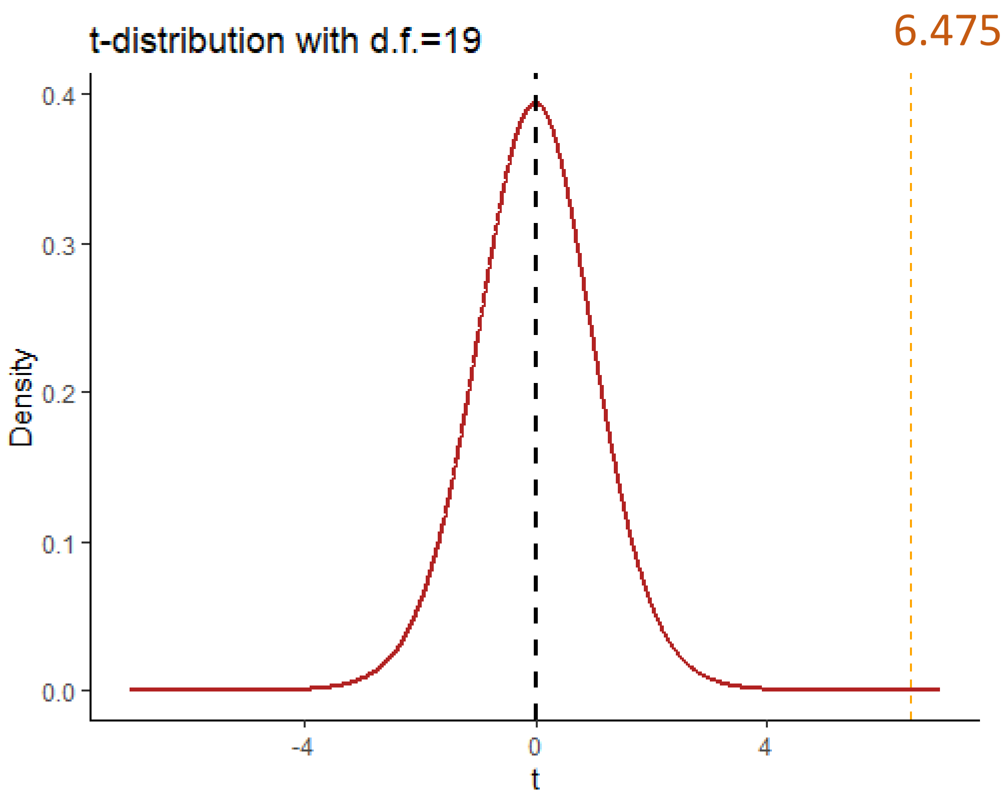
This sampling distribution is a t-distribution with d.f. n-1 = 19

We calculate how surprising our observed sample mean difference was using a t-statistic:

$$t = \frac{\bar{x}_d}{\sigma_{\bar{x}_d}}$$

$$t = \frac{\bar{x}_d}{s_d/\sqrt{n}}$$

The null hypothesis assumes that $\mu_{\bar{x}_d} = 0$



$$t = \frac{\bar{x}_d}{s_d/\sqrt{n}}$$

$$t = \frac{1.41}{0.97/\sqrt{20}}$$

$$t = 6.475$$

```
> pt(6.475, df=19)
[1] 0.9999983
```

p = 0.000017 (1-tailed test)

p = 0.000034 (2-tailed test)

Paired t-test in R (two ways ….)

```
> t.test(chico$improvement, mu=0)

        One Sample t-test

data:  chico$improvement
t = 6.4754, df = 19, p-value = 3.321e-06
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.9508686 1.8591314
sample estimates:
mean of x
    1.405

>
> t.test(chico$grade_test2, chico$grade_test1, paired = T)

        Paired t-test

data:  chico$grade_test2 and chico$grade_test1
t = 6.4754, df = 19, p-value = 3.321e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.9508686 1.8591314
sample estimates:
mean of the differences
                  1.405
```

Effect Sizes for Paired t-tests

$$\delta = \frac{\bar{x}_d}{s_d}$$

$$\delta = \frac{1.41}{0.97} = 1.45$$