14. Comparing Two Means

Readings - Rosner: 8.1-12
SAS: PROC TTEST, PROC POWER
R: t.test, power

A) Paired samples ~ one-sample test
B) Two independent samples
C) Power and sample size for the two independent sample *t*-test

A) Two Cases: Paired vs. Independent Samples

So far, we have looked at using the sample average $\overline{X}$ to test the hypothesis for a single mean at a fixed value. All of the tests we have discussed have been one-sample tests. We compared the underlying parameters of the population from which the sample was drawn to comparable values from generally large populations whose parameters we assumed to be known.

We'll now look at comparing two means. We've actually already done this for the paired sample case – by taking differences and doing a one-sample analysis of the differences. For independent samples, most of the ideas are familiar.

Designs for Comparative Studies—Two Basic Possibilities:

1. Independent Samples:

Two separate groups of measurements, i.e. two different groups of subjects. The sample sizes need not be the same for the two groups.

2. Paired (matched) Samples:

Two groups of measurements where each value in group 1 has a corresponding measurement in group 2. Measurements are related, e.g. they are taken on the same subject, or on two related subjects – twins, sibs, parent-child, etc.

e.g. Which are paired and which are independent samples?

1. Percent body fat is determined for each of 20 subjects using two methods on each subject—DEXA and UMW

2. Percent body fat is determined for each of 20 subjects using DEXA and for each of 20 other subjects using UMW

3. Baseline measurements are made for each of 50 subjects and one month later the same quantity is measured again on each subject

4. 40 families with at least two obese children are selected and from each family, two obese siblings are selected. In each family, one sibling is randomly chosen to get a new diet drug, the other gets a placebo.

5. Baseline measurements are made on each of 50 subjects. 25 subjects are then given a test drug, the other 25 are given the standard drug and all subjects are later re-tested. What is the effect of each drug? Is the effect of the drugs different?

B) Paired Samples: X_{1i} and X_{2i}, i = 1, …, n

1. We wish to test the hypothesis: H_0: $\Delta = 0$ vs. H_1: $\Delta \neq 0$

2. Compute the differences: $D_i = X_{2i} - X_{1i}$

3. Perform statistical analyses on the single sample of D_i values. Check assumptions as usual.

e.g. Effect of oral contraceptives (OC) on blood pressure

Subject (i)	X_{1i}: SBP no OC	X_{2i}: SBP on OC	$D_i = X_{2i} - X_{1i}$
1	115	128	13
2	112	115	3
3	107	106	-1
4	119	128	9
5	115	122	7
6	138	145	7
7	126	132	6
8	105	109	4
9	104	102	-2
10	115	117	2

$\overline{D} = 4.80$

$s_D = 4.566$

95% CI for μ_D: $\overline{D} \pm t_9\left(s \big/ \sqrt{n} \right) = 4.80 \pm 2.262\left(4.566 \big/ \sqrt{10} \right)$

(1.53 mmHg, 8.07 mmHg)

p-value: H_0: $\mu_D = 0$ vs. H_1: $\mu_D \neq 0$

$$P\left(\frac{|\overline{D} - 0|}{s \big/ \sqrt{n}} > \frac{|4.80 - 0|}{4.566 \big/ \sqrt{10}} \right) = P\left(|t_9| > 3.32 \right) = 0.009$$

Conclusion:

```
/* SAS code for the OC and BP data */
data new;
   input X1 X2;
   D = X2 - X1;
   cards;
       115  128
       112  115
       107  106
       119  128
       115  122
       138  145
       126  132
       105  109
       104  102
       115  117
        ;
  run;

proc univariate data=new;
      var D;
run;
```

```
The UNIVARIATE Procedure
Variable:  D
                            Moments
N                           10    Sum Weights                 10
Mean                       4.8    Sum Observations            48
Std Deviation       4.56557164    Variance            20.8444444
Skewness            0.16147087    Kurtosis            -0.2084615
Uncorrected SS             418    Corrected SS             187.6
Coeff Variation     95.1160759    Std Error Mean      1.44376052

              Basic Statistical Measures
    Location                    Variability
Mean     4.800000     Std Deviation            4.56557
Median   5.000000     Variance                20.84444
Mode     7.000000     Range                   15.00000
                      Interquartile Range      5.00000

           Tests for Location: Mu0=0
Test           -Statistic-    -----p Value------
Student's t    t  3.324651    Pr > |t|    0.0089
Sign           M         3    Pr >= |M|   0.1094
Signed Rank    S        24    Pr >= |S|   0.0117

   Stem Leaf                     #  Boxplot
     12 0                        1     |
     10                                |
      8 0                        1     |
      6 000                      3  +-----+
      4 0                        1  *--+--*
      2 00                       2  +-----+
      0                                |
     -0 0                        1     |
     -2 0                        1     |
        ----+----+----+----+
```

```
                 Normal Probability Plot
  13+                                       *+++++
    |                                    ++++
    |                                ++*+
    |                         * +*+++
   5+                       *+++
    |                 *++*++
    |              ++++
    |          *++++*
  -3+      +++++
      +----+----+----+----+----+----+----+----+----+----+
          -2        -1         0        +1        +2
```

C) Comparing two means: Independent samples

The ideas for hypothesis testing in this context are the same as for the one-sample case. To perform the test correctly, however, we need to obtain the appropriate standard error and degrees of freedom associated with the difference in two means. (Be sure to review Lecture 9 on functions of r.v.)

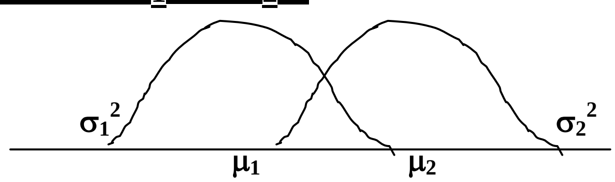
Statistical Model:

Group 1: $X_{1i} \sim N(\mu_1, \sigma_1^2)$ indep i = 1, …, n_1

Group 2: $X_{2i} \sim N(\mu_2, \sigma_2^2)$ indep i = 1, …, n_2

Hypotheses: H_0: $\mu_1 = \mu_2$ versus H_1: $\mu_1 \neq \mu_2$

<u>1) Equal Variance $\sigma_1^2 = \sigma_2^2$:</u>



1. A common assumption to make is that the two groups have equal variances $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

2. If n_1 and n_2 are large (≥ 30) we can proceed even if the data are not normal (by the CLT).

Quantity of interest: $\mu_2 - \mu_1$

Estimate of $\mu_2 - \mu_1$: $\overline{X}_2 - \overline{X}_1$

s.e.$(\overline{X}_2 - \overline{X}_1)$: $\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}} = \sqrt{\sigma^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)} = \sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$

Estimated s.e. $(\overline{X}_2 - \overline{X}_1)$: $s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$ - How do we estimate s?

Distribution of difference? $\left(\overline{X}_2 - \overline{X}_1\right) \sim N\left(\left(\mu_2 - \mu_1\right),\ \sigma^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)\right)$

$$\boxed{t = \frac{estimate - null}{se(estimate)} = \frac{\left(\overline{X}_2 - \overline{X}_1\right) - \left(\mu_2 - \mu_1\right)}{s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}}$$

This quantity has exactly a t-distribution with $n_1 + n_2$ - 2 df. The main new technical aspect is how to estimate the common variance σ^2. In the equation above we use *s* as the estimator for σ. What is *s* in this case?

Recall the one-sample variance estimate: $s^2 = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1}$

A combined or *pooled* variance, s_p^2, could do the same except for using deviations about the separate group means.

$$\text{e.g.}\quad s_p^2 = \frac{\sum_{i=1}^{n_1}\left(X_{1i} - \overline{X}_1\right)^2 + \sum_{i=1}^{n_2}\left(X_{2i} - \overline{X}_2\right)^2}{n_1 + n_2 - 2}$$

The idea is that if s_1^2 and s_2^2 are the sample variances from group 1 and group 2 respectively, then their average could be used as the estimate of the common value of σ^2. However, the sample variance from the larger sample is probably more precise and should be weighted more heavily.

The best pooled estimate of σ^2 is given by a *weighted* average of the two sample variances, where the weights are the number of degrees of freedom in each sample:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

s_p^2 will then have $n_1 - 1$ df from the first sample and $n_2 - 1$ df from the second sample or a total of $(n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$ df.

Two sample t-test for independent samples with *equal* variances:

In a two-sample test, the null hypothesis is H_0: $\mu_1 = \mu_2$ vs. H_1: $\mu_1 \neq \mu_2$ with a significance level α, where the two samples are (approximately) normally distributed, and σ^2 is assumed to be the same for each population. Under these assumptions:

$$t = \frac{\text{estimate} - \text{null}}{\text{se(estimate)}} = \frac{(\overline{X}_2 - \overline{X}_1) - (\mu_2 - \mu_1)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{where } s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

If $t > t_{n1+n2-2,\, 1-\alpha/2}$ or $t < -\,t_{n1+n2-2,\, 1-\alpha/2}$ then we reject H_0.

$$p-value = P\left(|T| > \frac{(\bar{X}_2 - \bar{X}_1) - (\mu_2 - \mu_1)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \,\middle|\, H_0 : \mu_1 = \mu_2 \right)$$

If $t \le 0$, p = 2 x (area to the left of *t* under a $t_{n1+n2-2}$ distribution).

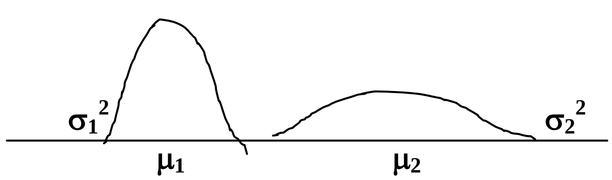
If $t > 0$, p = 2 x (area to the right of *t* under a $t_{n1+n2-2}$ distribution).

Confidence interval for the underlying mean difference (μ1 - μ2) between two groups assuming equal variances:

$$\text{CI for } \mu_2 - \mu_1: \ \left(\bar{X}_2 - \bar{X}_1\right) \pm t_{n1-1+n2-2,1-\alpha/2}\left(s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

2) Unequal Variances $\sigma_1^2 \neq \sigma_2^2$: (the Behrens-Fisher problem)

The assumption of equal variances may not be valid. The approach to developing a valid test allows the two groups to have unequal variances, but assumes they are both normal. We proceed as usual to construct a CI and hypothesis test. The distribution involved is not quite a *t* but is approximated by a *t* with adjusted df (usually using Satterthwaite's method).



Quantity of interest: $\mu_2 - \mu_1$

Estimate of $\mu_2 - \mu_1$: $\bar{X}_2 - \bar{X}_1$

s.e. $(\bar{X}_2 - \bar{X}_1)$: $\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$

Estimated SE $(\bar{X}_2 - \bar{X}_1)$: $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

Distribution of difference? $(\bar{X}_2 - \bar{X}_1) \sim N\left((\mu_2 - \mu_1), \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)\right)$

$$\boxed{t = \frac{\text{estimate} - \text{null}}{\text{se(estimate)}} = \frac{(\bar{X}_2 - \bar{X}_1) - (\mu_2 - \mu_1)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}}$$

This *does not* have exactly a *t*-distribution.

Satterthwaite's Method: (see Rosner text, section 8.7)
Approximate degrees of freedom d' :

$$d' = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$

, and round down to the nearest integer, d''

$$t = \frac{\text{estimate} - \text{null}}{\text{se(estimate)}} = \frac{(\bar{X}_2 - \bar{X}_1) - (\mu_2 - \mu_1)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

If $t > t_{d'', 1-\alpha/2}$ or $t < -t_{d'', 1-\alpha/2}$ then H_0 is rejected

$$p\text{-}value = P\left(|T| > \frac{(\bar{X}_2 - \bar{X}_1) - (\mu_2 - \mu_1)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \,\middle|\, H_0 : \mu_1 = \mu_2 \right)$$

If $t \le 0$, p = 2 x (area to the left of *t* under a $t_{d''}$ distribution).

If $t > 0$, p = 2 x (area to the right of *t* under a $t_{d''}$ distribution).

Confidence interval for the underlying mean difference (μ_1 - μ_2) between two groups assuming unequal variances:

$$\text{CI for } \mu_2 - \mu_1: \ \bar{x}_2 - \bar{x}_1 \pm t_{d'',1-\alpha/2}\left(\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)$$

<u>Testing for equality of variances:</u>
Controversy exists over whether to test the hypothesis H_0: $\sigma_1^2 = \sigma_2^2$ to determine which form of the two-sample test should be used. Rosner seems to advocate for using this test. ***I do not.***

The authors below suggest that the unequal variances version should *always* be used when the *ratio of the two variances is unknown – which is most often the case*:
Moser, B.K. and Stevens, G.R. (1992). Homogeneity of variance in the two-sample means test. *American Statistician*, 46:19-21. (posted to the course website under Files -> Papers)

Example: comparative study of two adult homes in western Virginia. Does mean age differ across the two homes?

$$H_0: \mu_{1age} = \mu_{2age} \text{ vs. } H_1: \mu_{1age} \neq \mu_{2age}$$

```
filename home 'c:\annab\adulthome.dat';

data one;
infile home;
input Home Gender $ Diagnos $ Age Destin;

proc ttest data=one;
class home;
var age;
run;
```

The TTEST Procedure

Variable: Age

Home	N	Mean	Std Dev	Std Err	Minimum	Maximum
1	39	44.6154	18.0899	2.8967	18.0000	89.0000
2	24	66.4167	13.9780	2.8532	36.0000	92.0000
Diff (1-2)		-21.8013	16.6591	4.3220		

Home	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		44.6154	38.7513	50.4794	18.0899	14.7839	23.3138
2		66.4167	60.5143	72.3191	13.9780	10.8639	19.6078
Diff (1-2)	Pooled	-21.8013	-30.4436	-13.1589	16.6591	14.1563	20.2453
Diff (1-2)	**Satterthwaite**	**-21.8013**	**-29.9410**	**-13.6616**			

Method	Variances	DF	t Value	Pr > |t|	Using probt()
Pooled	Equal	61	-5.04	<.0001	4.4332291E-6
Satterthwaite	**Unequal**	**57.727**	**-5.36**	**<.0001**	1.5198918E-6

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	38	23	1.67	0.1924

Conclusion:

C) Two independent samples: Power, sample size and detectable difference

The concepts are identical to the one-sample case. The formulas below are based on the Z-test assuming known, unequal variances and unequal sample sizes. Often equal sample sizes are desired (k=1) but the formulas below are more general.

To find power we must specify:

1) Difference to be detected: $\Delta = |\mu_2 - \mu_1|$
2) Significance level α
3) Sample sizes n_1 and n_2
4) Variances σ_1^2 and σ_2^2

$$\text{Power} = \Phi\left[-Z_{1-\alpha/2} + \frac{\sqrt{n_1}\Delta}{\sqrt{\sigma_1^2 + \sigma_2^2/k}}\right] = \Phi\left[-Z_{1-\alpha/2} + \frac{\sqrt{n_2}\Delta}{\sqrt{k\sigma_1^2 + \sigma_2^2}}\right], \textit{where } k = \frac{n_2}{n_1}.$$

To find sample sizes we must specify:

1) Difference to be detected: $\Delta = |\mu_2 - \mu_1|$
2) Significance level α
3) Power = 1 - β
4) Variances σ_1^2 and σ_2^2

$$n_1 = \frac{\left(\sigma_1^2 + \sigma_2^2/k\right)\left(Z_{1-\alpha/2} + Z_{1-\beta}\right)^2}{\Delta^2} \quad \text{and} \quad n_2 = \frac{\left(k\sigma_1^2 + \sigma_2^2\right)\left(Z_{1-\alpha/2} + Z_{1-\beta}\right)^2}{\Delta^2}$$

To find the detectable difference we must specify:

1) Significance level α
2) Power = 1 - β
3) Variances σ_1^2 and σ_2^2
4) n_1 and n_2 <u>or</u> n_1 (or n_2) and k

$$\Delta = \sqrt{\frac{\left(\sigma_1^2 + \sigma_2^2/k\right)\left(Z_{1-\alpha/2} + Z_{1-\beta}\right)^2}{n_1}} \quad or \quad \Delta = \sqrt{\frac{\left(k\sigma_1^2 + \sigma_2^2\right)\left(Z_{1-\alpha/2} + Z_{1-\beta}\right)^2}{n_2}}$$

We can use SAS PROC POWER for this as well, … sort of.

Example: from the adult home study – smallest detectable difference with $n_1 = 39$, $n_2 = 24$, $\mu_{1age} = 45$ yrs, $\sigma_{1age} = 18$ yrs, $\sigma_{2age} = 14$ yrs - Consider known vs. unknown variances

<u>SAS Code using PROC POWER:</u>

Example 1: $\mu 1 = 45$, $\sigma_1 = 18$, $\sigma_2 = 14$, $n_1 = 39$, $n_2 = 24$; $\alpha = 0.05$ (two-sided); Power = 0.80, 0.90. Find detectable difference. Assume σ_1 and σ_2 known.

……….. SAS won't do this.

Example 2: $\mu 1 = 45$, $\sigma_1 = 18$, $\sigma_2 = 14$, $n_1 = 39$, $n_2 = 24$; $\alpha = 0.01$, 0.05, 0.10 (two-sided); Power = 0.80, 0.90, 0.95. Find detectable difference. Assume σ_1 and σ_2 *unknown*.

………. SAS sort of does this; in order to solve for detectable difference the s.d. must be equal. ☹

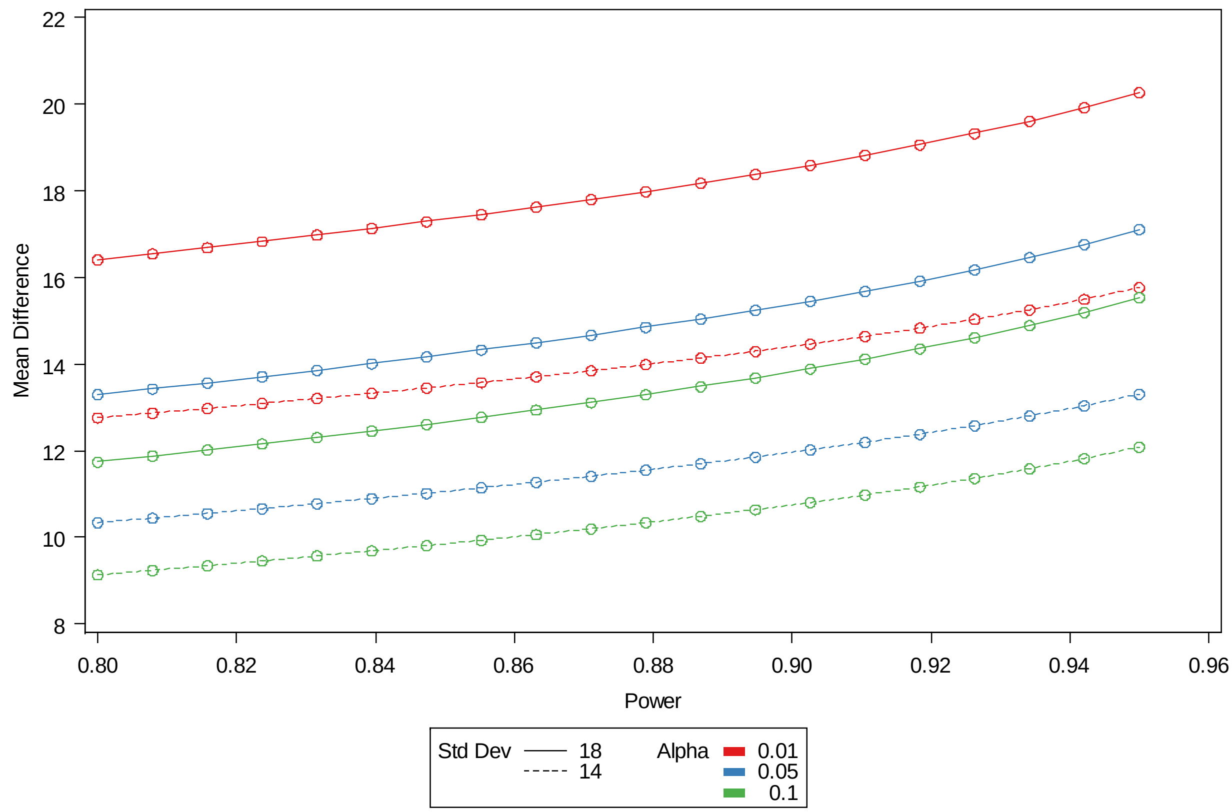
```
PROC POWER;
 TWOSAMPLEMEANS TEST=DIFF ALPHA=0.01 0.05 0.10 MEANDIFF = .
      GROUPNS = 39 | 24  STDDEV = 18 14 POWER= 0.8 0.9 0.95;
 PLOT INTERPOL = join X = power VARY(color by alpha) ;
 TITLE1 'Detectable Difference of Means, Unequal Sample Sizes,
Equal Variances';
 TITLE2 'Two-Sample t-test with Pooled Variances';

RUN;
```

```
Detectable Difference of Means, Unequal Sample Size, Unequal Sigma,
Two-Sample t-test with Pooled Variance

The POWER Procedure
Two-sample t Test for Mean Difference

      Fixed Scenario Elements

Distribution                 Normal
Method                        Exact
Group 1 Sample Size              39
Group 2 Sample Size              24
Number of Sides                   2
Null Difference                   0

             Computed Mean Diff
```

Index	Alpha	Std Dev	Power	Mean Diff
1	0.01	18	0.800	16.41
2	0.01	18	0.900	18.52
3	0.01	18	0.950	20.27
4	0.01	14	0.800	12.76
5	0.01	14	0.900	14.41
6	0.01	14	0.950	15.76
7	0.05	18	0.800	13.29
8	0.05	18	0.900	15.38
9	0.05	18	0.950	17.11
10	0.05	14	0.800	10.34
11	0.05	14	0.900	11.96
12	0.05	14	0.950	13.30
13	0.10	18	0.800	11.74
14	0.10	18	0.900	13.82
15	0.10	18	0.950	15.54
16	0.10	14	0.800	9.13
17	0.10	14	0.900	10.75
18	0.10	14	0.950	12.08

Detectable Difference of Means, Unequal Sample Sizes, Equal Variances

Two-Sample t-test with Pooled Variances



Example 3: $\mu 2-\mu 1$ = 10, 15, 20, σ_1 = 18, σ_2 = 14, n_1 = 39, n_2 = 24; α = 0.01, 0.05, 0.10 (two-sided); Power = 0.80, 0.90, 0.95. Find N = n1 + n2; k =1. Assume σ_1 and σ_2 *unknown and unequal*.

```
PROC POWER;
 TWOSAMPLEMEANS TEST=DIFF_SATT ALPHA=0.01 0.05 0.10 MEANDIFF = 10 15 20
      GROUPWEIGHTS = (1 1) NTOTAL=. GROUPSTDDEVS = 18 | 14 POWER= 0.8 0.9
0.95;
 PLOT INTERPOL = join X = effect VARY(color by alpha) ;
 TITLE1 'Sample Sizes for Two-sample Means, Unequal Variances';
 TITLE2 'Satterthwaite Two-Sample t-test';

RUN;
```

Sample Sizes for Two-sample Means, Unequal Variances

Satterthwaite Two-Sample t-test

The POWER Procedure

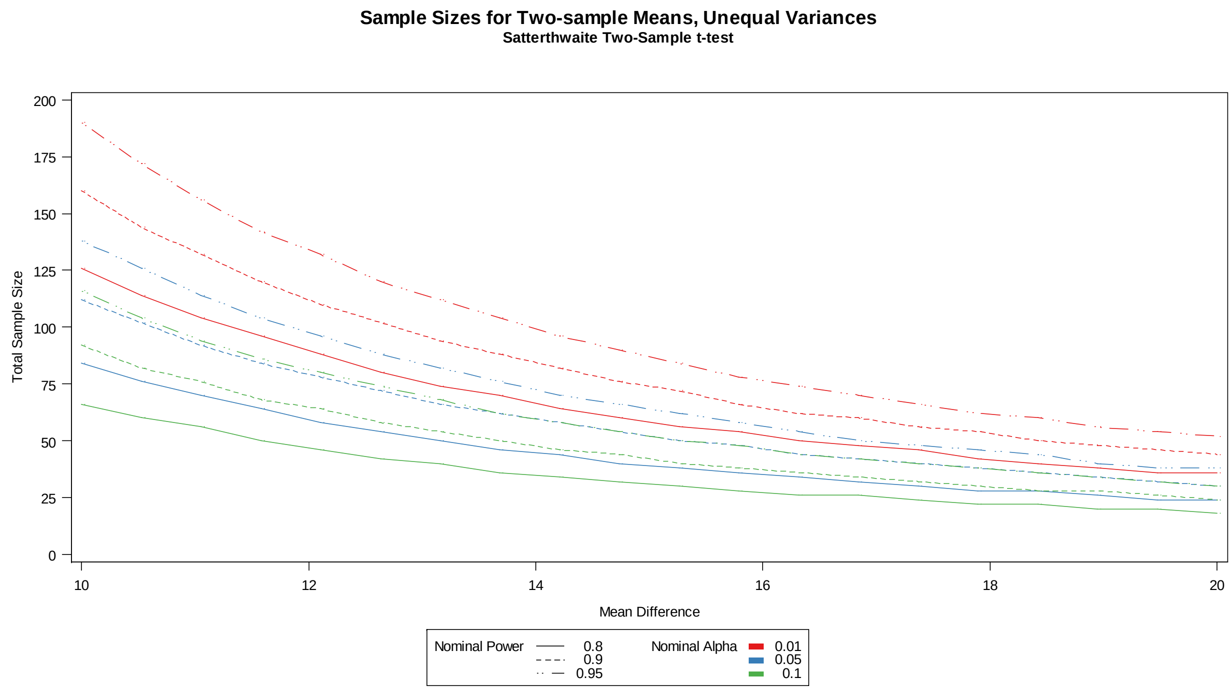
Two-Sample t Test for Mean Difference with Unequal Variances

Fixed Scenario Elements

Distribution	Normal
Method	Exact
Group 1 Standard Deviation	18
Group 2 Standard Deviation	14
Group 1 Weight	1
Group 2 Weight	1
Number of Sides	2
Null Difference	0

Computed N Total

Index	Nominal Alpha	Mean Diff	Nominal Power	Actual Alpha	Actual Power	N Total
1	0.01	10	0.80	0.00999	0.804	126
2	0.01	10	0.90	0.00999	0.904	160
3	0.01	10	0.95	0.01000	0.951	190
4	0.01	15	0.80	0.00995	0.803	58
5	0.01	15	0.90	0.00997	0.907	74
6	0.01	15	0.95	0.00998	0.950	86
7	0.01	20	0.80	0.00985	0.827	36
8	0.01	20	0.90	0.00990	0.913	44
9	0.01	20	0.95	0.00993	0.958	52
10	0.05	10	0.80	0.04994	0.801	84
11	0.05	10	0.90	0.04997	0.902	112
12	0.05	10	0.95	0.04998	0.951	138
13	0.05	15	0.80	0.04973	0.816	40
14	0.05	15	0.90	0.04985	0.907	52
15	0.05	15	0.95	0.04990	0.955	64
16	0.05	20	0.80	0.04923	0.823	24
17	0.05	20	0.90	0.04952	0.904	30
18	0.05	20	0.95	0.04970	0.960	38
19	0.10	10	0.80	0.09989	0.801	66
20	0.10	10	0.90	0.09994	0.904	92
21	0.10	10	0.95	0.09996	0.953	116
22	0.10	15	0.80	0.09949	0.821	32
23	0.10	15	0.90	0.09971	0.905	42
24	0.10	15	0.95	0.09981	0.951	52
25	0.10	20	0.80	0.09828	0.804	18
26	0.10	20	0.90	0.09907	0.901	24
27	0.10	20	0.95	0.09942	0.951	30



Example 4: $\mu2-\mu1 = 10, 15, 20$, $\sigma_1 = 18$, $\sigma_2 = 14$, N = n1 + n2 = 50, 70, 90, 1001; k =1; α = 0.01, 0.05, 0.10 (two-sided). Find Power. Assume σ_1 and σ_2 *unknown and unequal*.

```
PROC POWER;
 TWOSAMPLEMEANS TEST=DIFF_SATT ALPHA=0.01 0.05 0.10 MEANDIFF = 10 15 20
      GROUPWEIGHTS = (1 1) NTOTAL= 50 70 90 110 GROUPSTDDEVS = 18 | 14 POWER=
.;
 PLOT INTERPOL = join X = effect VARY(color by alpha) ;
 TITLE1 'Power for Two-sample Means, Unequal Variances';
 TITLE2 'Satterthwaite Two-Sample t-test';

RUN;
```

Power for Two-sample Means, Unequal Variances

Satterthwaite Two-Sample t-test

The POWER Procedure

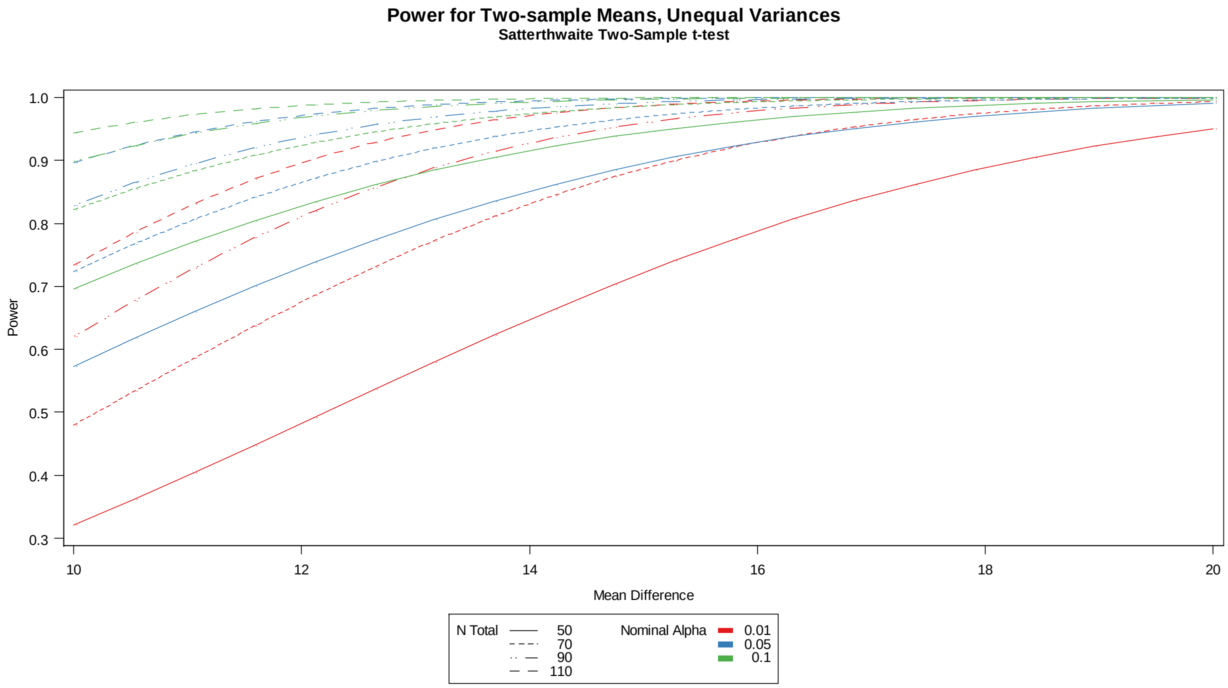
Two-Sample t Test for Mean Difference with Unequal Variances

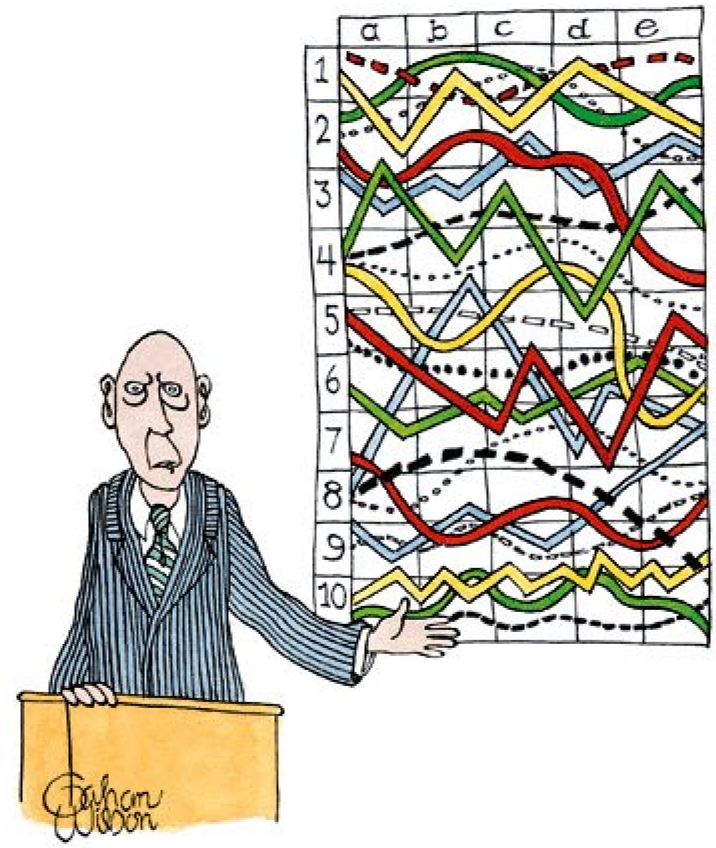
Fixed Scenario Elements

Distribution	Normal
Method	Exact
Group 1 Standard Deviation	18
Group 2 Standard Deviation	14
Group 1 Weight	1
Group 2 Weight	1
Number of Sides	2
Null Difference	0

Computed Power

Index	Nominal Alpha	Mean Diff	N Total	Actual Alpha	Power
1	0.01	10	50	0.00993	0.320
2	0.01	10	70	0.00996	0.480
3	0.01	10	90	0.00998	0.620
4	0.01	10	110	0.00999	0.733
5	0.01	15	50	0.00993	0.722
6	0.01	15	70	0.00996	0.887
7	0.01	15	90	0.00998	0.960
8	0.01	15	110	0.00999	0.987
9	0.01	20	50	0.00993	0.950
10	0.01	20	70	0.00996	0.993
11	0.01	20	90	0.00998	>.999
12	0.01	20	110	0.00999	>.999
13	0.05	10	50	0.04983	0.573
14	0.05	10	70	0.04992	0.724
15	0.05	10	90	0.04995	0.828
16	0.05	10	110	0.04997	0.896
17	0.05	15	50	0.04983	0.896
18	0.05	15	70	0.04992	0.969
19	0.05	15	90	0.04995	0.992
20	0.05	15	110	0.04997	0.998
21	0.05	20	50	0.04983	0.990
22	0.05	20	70	0.04992	>.999
23	0.05	20	90	0.04995	>.999
24	0.05	20	110	0.04997	>.999
25	0.10	10	50	0.09980	0.696
26	0.10	10	70	0.09990	0.822
27	0.10	10	90	0.09994	0.898
28	0.10	10	110	0.09996	0.944
29	0.10	15	50	0.09980	0.944
30	0.10	15	70	0.09990	0.986
31	0.10	15	90	0.09994	0.997
32	0.10	15	110	0.09996	>.999
33	0.10	20	50	0.09980	0.996
34	0.10	20	70	0.09990	>.999
35	0.10	20	90	0.09994	>.999
36	0.10	20	110	0.09996	>.999


Power for Two-sample Means, Unequal Variances
Satterthwaite Two-Sample t-test
Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
10
12
14
16
18
20
Mean Difference
N Total
50
70
90
110
Nominal Alpha
0.01
0.05
0.1



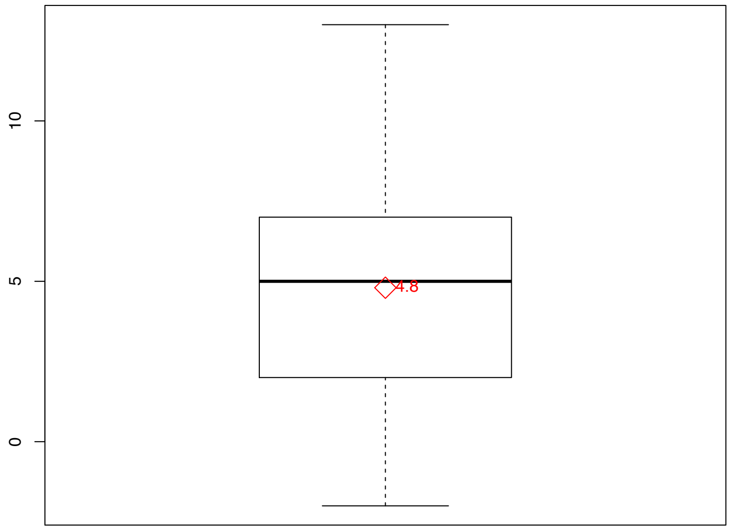
"I'll pause for a moment so you can let this information sink in."

These examples follow the SAS examples in Unit 17.

B. Paired Samples.

```
library(psych)
library(pastecs)
oc<-read.table(header=T, con <- textConnection('
id x1 x2 di
1  115  128 13
2 112 115 3
3 107 106 -1
4 119 128 9
5 115 122 7
6 138 145 7
7 126 132 6
8 105 109 4
9 104 102 -2
10 115 117 2
')
)
```

```
summary(oc$di)

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -2.00    2.25    5.00    4.80    7.00   13.00

# alternative summaries using the pastecs packages
stat.desc(oc$di)

     nbr.val     nbr.null       nbr.na          min          max
    10.00000      0.00000      0.00000     -2.00000     13.00000
       range          sum       median         mean      SE.mean
    15.00000     48.00000      5.00000      4.80000      1.44376
CI.mean.0.95          var      std.dev     coef.var
     3.26601     20.84444      4.56557      0.95116

# verify the data. This step ensures that the difference you are given is
# equal to the calculated difference. If there was a FALSE value in this
# output, you would double-check your data.
oc$d2 <- oc$x2 - oc$x1
oc$di == oc$d2

 [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE

boxplot(oc$di)
points(mean(oc$di), col = "red", pch = 5, cex = 2)
text(1, mean(oc$di), mean(oc$di), col = "red", pos = 4)
```




```
t.test(oc$di)


One Sample t-test

data:  oc$di
t = 3.3247, df = 9, p-value = 0.008874
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 1.534 8.066
sample estimates:
mean of x
      4.8

# You don't need to compute the differences, you can select the option
# paired=TRUE
t.test(oc$x1, oc$x2, data = oc, paired = T)


Paired t-test

data:  oc$x1 and oc$x2
t = -3.3247, df = 9, p-value = 0.008874
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
```

```
 -8.066 -1.534
sample estimates:
mean of the differences
                   -4.8
```

2. Unequal Variance (Satterthwaite Approximation)

```
adulthome <- read.table("~/Dropbox/6611METHODS/6611/AdultHome.dat", quote = "\"")
names(adulthome) <- c("home", "gender", "diagnos", "age", "destin")
adulthome$home <- factor(adulthome$home)
t.test(age ~ home, adulthome, var.equal = T)


Two Sample t-test

data:  age by home
t = -5.0443, df = 61, p-value = 4.364e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -30.444 -13.159
sample estimates:
mean in group 1 mean in group 2
         44.615          66.417

t.test(age ~ home, adulthome, var.equal = F)


Welch Two Sample t-test

data:  age by home
t = -5.3619, df = 57.727, p-value = 1.509e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -29.941 -13.662
sample estimates:
mean in group 1 mean in group 2
         44.615          66.417
```

C. Power with two independent samples (using package `pwr`)

```
library(pwr)
alpha <- rep(c(0.01, 0.05, 0.1), 3)
power <- c(rep(0.8, 3), rep(0.9, 3), rep(0.95, 3))
n1 <- 39
n2 <- 24
sd <- c(18, 14)
diffs <- matrix(rep(NA, 18), ncol = 2)
for (d in 1:9) {
    p <- pwr.t2n.test(n1 = n1, n2 = n2, sig.level = alpha[d], power = power[d])
```

```
    diffs[d, ] <- p$d * sd
}
diffs <- as.data.frame(diffs)
diffs[, 3] <- alpha
diffs[, 4] <- power
names(diffs) <- c("sd18", "sd14", "alpha", "power")
library(reshape2)
library(ggplot2)
diffs.long <- melt(diffs, measure.vars = 1:2)
names(diffs.long) <- c("alpha", "power", "sd", "differences")
```

```
qplot(power,differences,data=diffs.long,colour=factor(alpha),linetype=sd,shape=sd,geom="line")
+geom_point()+theme_bw()
+ggtitle("Detectable Difference of Means, Unequal Sample Sizes, Equal Variances")
+xlab("Power")+ylab("Mean Difference")
```

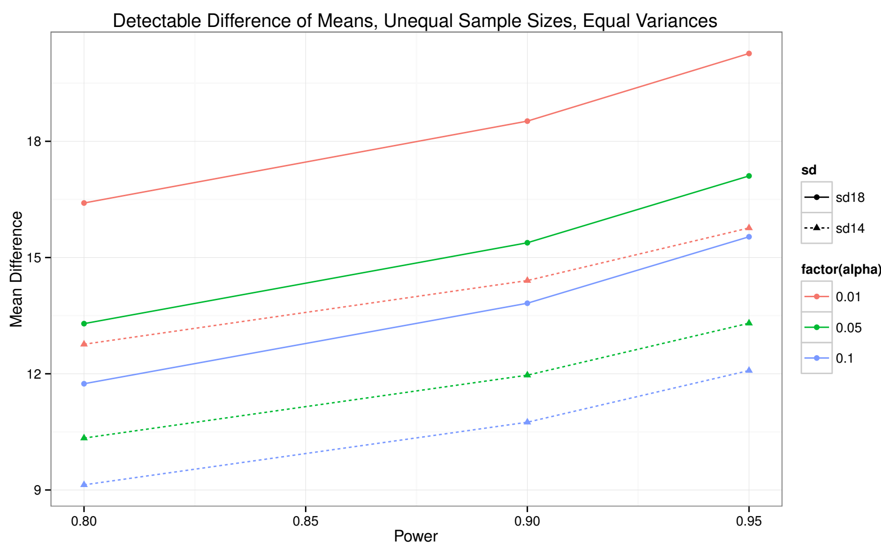


Figure 1: Detectable Difference of Means, Unequal Sample Sizes, Equal Variances

Appendix: Code

```
# unit 17

## ---- ex0 ----
library(pastecs)
oc<-read.table(header=T, con <- textConnection('
id x1 x2 di
1  115  128     13
2       112     115     3
3       107     106     -1
4       119     128     9
5       115     122     7
6       138     145     7
7       126     132     6
8       105     109     4
9       104     102     -2
10      115     117     2
')
)
## ---- ex1 ----
summary(oc$di)
#alternative summaries using the pastecs packages
stat.desc(oc$di)
# verify the data. This step ensures that the difference you are given is equal to
     the calculated difference. If there was a FALSE value in this output, you
    would double-check your data.
oc$d2<-oc$x2-oc$x1
oc$di==oc$d2

boxplot(oc$di)
points(mean(oc$di),col="red",pch=5,cex=2)
text(1,mean(oc$di),mean(oc$di),col="red",pos=4)
t.test(oc$di)
#You don't need to compute the differences, you can select the option paired=TRUE
t.test(oc$x1,oc$x2,data=oc,paired=T)

## ---- ex2 ----
adulthome<-read.table("~/Dropbox/6611METHODS/6611/AdultHome.dat", quote="\"")
names(adulthome)<-c("home","gender","diagnos","age","destin")
adulthome$home<-factor(adulthome$home)
t.test(age~home,adulthome,var.equal=T)
t.test(age~home,adulthome,var.equal=F)
## ---- ex3 ----
library(pwr)
alpha<-rep(c(0.01,0.05,0.10),3)
power<-c(rep(0.8,3),rep(0.9,3),rep(0.95,3))
n1<-39
n2<-24
sd<-c(18,14)
diffs<-matrix(rep(NA,18),ncol=2)
for (d in 1:9){
  p<-pwr.t2n.test(n1=n1,n2=n2,sig.level=alpha[d],power=power[d])
  diffs[d,]<-p$d*sd
}
diffs<-as.data.frame(diffs)
diffs[,3]<-alpha
diffs[,4]<-power
names(diffs)<-c("sd18","sd14","alpha","power")
library(reshape2)
library(ggplot2)
```

```
diffs.long<-melt(diffs,measure.vars=1:2)
names(diffs.long)<-c("alpha","power","sd","differences")
## ---- ex3b ----
qplot(power,differences,data=diffs.long,colour=factor(alpha),linetype=sd,shape=sd,
    geom="line")+geom_point()+theme_bw()+ggtitle("Detectable Difference of Means,
    Unequal Sample Sizes, Equal Variances")+xlab("Power")+ylab("Mean Difference")
```