Chapter 4.5 Tests of the Equality of two parameters

Grethe Hystad

December 2, 2012

Let $X = (X_1, ..., X_n)$ and $Y = (Y_1, ..., Y_m)$ be two independent random samples, where X is $N(\mu_X, \sigma_X^2)$ and Y is $N(\mu_Y, \sigma_Y^2)$. We want to test whether

- the two means are equal (for equal or unequal variances)
- the two variances are equal.

Test of difference in mean, equal variance

- We assume first that the variances are equal and test their mean.
- We test $H_0 : \mu_X = \mu_Y$ against H_1: $\mu_X > \mu_Y$ or $\mu_X < \mu_Y$ or $\mu_X \neq \mu_Y$.
- Recall that $E(\overline{X} - \overline{Y}) = \mu_X - \mu_Y$ and $Var(\overline{X} - \overline{Y}) = \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}$ and
- $$T = \frac{\overline{X} - \overline{Y}}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}},$$

  where $S_p = \sqrt{\frac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2}}$
- T has a t-distribution with $r = n + m - 2$ degrees of freedom when H_0 is true and the variances are approximately equal.

Test of hypothesis

One sided composite hypothesis for equality of two means:

$$H_0 : \mu_X = \mu_Y \quad \textit{against} \quad H_1 : \mu_X > \mu_Y.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$t \geq t_\alpha(n + m - 2)$$

or

$$\overline{x} - \overline{y} \geq t_\alpha(n + m - 2)s_p\sqrt{\frac{1}{n} + \frac{1}{m}}$$

Test of hypothesis

One sided composite hypothesis for equality of two means:

$$H_0 : \mu_X = \mu_Y \quad \textit{against} \quad H_1 : \mu_X < \mu_Y.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$t \leq -t_\alpha(n+m-2)$$

or

$$\overline{x} - \overline{y} \leq -t_\alpha(n+m-2)s_p\sqrt{\frac{1}{n}+\frac{1}{m}}$$

Test of hypothesis

Two sided composite hypothesis for equality of two means:

$$H_0 : \mu_X = \mu_Y \quad \textit{against} \quad H_1 : \mu_X \neq \mu_Y.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$|t| \geq t_{\alpha/2}(n + m - 2)$$

or

$$|\overline{x} - \overline{y}| \geq t_{\alpha/2}(n + m - 2)s_p\sqrt{\frac{1}{n} + \frac{1}{m}}$$

Test of hypothesis, variance is known

If the variances are **known**, we would use the statistics,

$$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

to test $H_0 : \mu_X = \mu_Y$. If the null hypothesis, H_0, is true, Z would be $N(0,1)$ or approximately normal if the underlying distributions are not normal.

If the variances are **unknown**, but the sample sizes large enough, we would replace σ_X^2 by S_X^2 and σ_Y^2 by S_Y^2 and

$$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}}$$

would be approximately $N(0,1)$.

Women in the Labor force

The dataset on the next page contains the labor force participation rate (LFPR) of women in 19 cities in the United States in 1968 and 1972. The data help to measure the growing presence of women in the labor force over this period.
Source: DASL at http://lib.stat.cmu.edu/DASL/DataArchive.html

Did the labor force participation rate of women increase from 1968 to 1972 at the 5% significance level?

Women in Labor Force

```
> z<-read.delim("C:/Documents and Settings/ghystad/Desktop/Math 361/labor.txt"
> z
            City X1972 X1968
1           N.Y.  0.45  0.42
2           L.A.  0.50  0.50
3        Chicago  0.52  0.52
4   Philadelphia  0.45  0.45
5        Detroit  0.46  0.43
6  San Francisco  0.55  0.55
7         Boston  0.60  0.45
8          Pitt.  0.49  0.34
9      St. Louis  0.35  0.45
10   Connecticut  0.55  0.54
11   Wash., D.C.  0.52  0.42
12         Cinn.  0.53  0.51
13     Baltimore  0.57  0.49
14        Newark  0.53  0.54
15 Minn/St. Paul  0.59  0.50
16       Buffalo  0.64  0.58
17       Houston  0.50  0.49
18     Patterson  0.57  0.56
19        Dallas  0.64  0.63
```

Women in Labor Force

```
> x=z[,2]
> y=z[,3]
> x
 [1] 0.45 0.50 0.52 0.45 0.46 0.55 0.60 0.49 0.35 0.55 0.52 0.53 0.57 0.53 0.59
[16] 0.64 0.50 0.57 0.64
> y
 [1] 0.42 0.50 0.52 0.45 0.43 0.55 0.45 0.34 0.45 0.54 0.42 0.51 0.49 0.54 0.50
[16] 0.58 0.49 0.56 0.63
```

Women in the Labor force

Solution

- *1. method. We start with a pooled t-test since the United States did not change much from* 1968 *to* 1972 *so we can assume the variances are the same.*

Women in the Labor force

Solution

- *1. method. We start with a pooled t-test since the United States did not change much from* 1968 *to* 1972 *so we can assume the variances are the same.*
- *It can be checked that the data appears to be approximately normal with no skewness or outliers. We test,*

Women in the Labor force

Solution

- *1. method. We start with a pooled t-test since the United States did not change much from* 1968 *to* 1972 *so we can assume the variances are the same.*
- *It can be checked that the data appears to be approximately normal with no skewness or outliers. We test,*
- $$H_0 : \mu_X = \mu_Y \quad \text{against} \quad H_1 : \mu_X > \mu_Y .$$

Women in the Labor force

Solution

- *1. method. We start with a pooled t-test since the United States did not change much from* 1968 *to* 1972 *so we can assume the variances are the same.*
- *It can be checked that the data appears to be approximately normal with no skewness or outliers. We test,*
- $$H_0 : \mu_X = \mu_Y \quad \textit{against} \quad H_1 : \mu_X > \mu_Y.$$
- *We have* $\overline{x} = 0.5268$, $\overline{y} = 0.4932$ *and* $s_x^2 = 0.005012$, $s_y^2 = 0.004623$ *with* $n = m = 19$.

Solution continue, Women in the Labor force

Solution

- *This gives* $s_p = \sqrt{\frac{(n-1)s_x^2+(m-1)s_y^2}{n+m-2}} = 0.06941$ *and the t-distribution,*

Solution continue, Women in the Labor force

Solution

- *This gives* $s_p = \sqrt{\frac{(n-1)s_x^2+(m-1)s_y^2}{n+m-2}} = 0.06941$ *and the t-distribution,*
- $t = \frac{\overline{x}-\overline{y}}{s_p\sqrt{\frac{1}{n}+\frac{1}{m}}} = \frac{0.5268-0.4932}{0.06941(\sqrt{2\frac{1}{19}})} = 1.4959$ *has* $19+19-2=36$ *degrees of freedom.*

Solution continue, Women in the Labor force

Solution

- *This gives* $s_p = \sqrt{\frac{(n-1)s_x^2+(m-1)s_y^2}{n+m-2}} = 0.06941$ *and the t-distribution,*
- $t = \frac{\overline{x}-\overline{y}}{s_p\sqrt{\frac{1}{n}+\frac{1}{m}}} = \frac{0.5268-0.4932}{0.06941(\sqrt{2\frac{1}{19}})} = 1.4959$ *has* $19+19-2 = 36$ *degrees of freedom.*
- *Since* $t = 1.4959 < t_{0.05}(36) = 1.6883$*, we accept* H_0*. Thus there was not a significant increase at the* 5% *level in the labor force participation rate of women from 1968 to 1972. Notice that the test just tells you that there was not strong enough evidence against,* H_0*, i.e not strong enough evidence against an increase in LFPR.*

Women in the Labor force

Solution

- *2. method. The data are naturally paired because the measurements were made in the same cities for each of the two years. It is better to compare each city in 1972 to its own value in 1968. Thus, we will use a Matched-pairs test.*

Women in the Labor force

Solution

- *2. method. The data are naturally paired because the measurements were made in the same cities for each of the two years. It is better to compare each city in 1972 to its own value in 1968. Thus, we will use a Matched-pairs test.*
- *Define* $W = X - Y$.

Women in the Labor force

Solution

- *2. method. The data are naturally paired because the measurements were made in the same cities for each of the two years. It is better to compare each city in 1972 to its own value in 1968. Thus, we will use a Matched-pairs test.*
- *Define* $W = X - Y$.
- *We have* $\overline{w} = \overline{x} - \overline{y} = 0.5268 - 0.4932 = 0.03368$ *and*

Women in the Labor force

Solution

- *2. method. The data are naturally paired because the measurements were made in the same cities for each of the two years. It is better to compare each city in 1972 to its own value in 1968. Thus, we will use a Matched-pairs test.*
- *Define* $W = X - Y$.
- *We have* $\overline{w} = \overline{x} - \overline{y} = 0.5268 - 0.4932 = 0.03368$ *and*
- $s_w = 0.05974$.

Women in the Labor force

Solution

- *2. method. The data are naturally paired because the measurements were made in the same cities for each of the two years. It is better to compare each city in 1972 to its own value in 1968. Thus, we will use a Matched-pairs test.*
- *Define* $W = X - Y$.
- *We have* $\overline{w} = \overline{x} - \overline{y} = 0.5268 - 0.4932 = 0.03368$ *and*
- $s_w = 0.05974$.
- *The t-distribution,* $T = \frac{\overline{W}-0}{\frac{s_W}{\sqrt{19}}}$, *has* 18 *degrees of freedom.*

Solution continue, Women in the Labor force

Solution

- *Then*
  $t = \frac{\overline{w}-0}{\frac{s_w}{\sqrt{n}}} = \frac{0.03368}{0.05974/\sqrt{19}} = 2.4577 > t_{0.05}(18) = 1.734064$

Solution continue, Women in the Labor force

Solution

- *Then*
  $t = \frac{\overline{w}-0}{\frac{s_W}{\sqrt{n}}} = \frac{0.03368}{0.05974/\sqrt{19}} = 2.4577 > t_{0.05}(18) = 1.734064$
- *Thus, we reject H_0 at the 5% significance level. Thus, there was a significant increase at the 5% level in the labor force participation rate of women from 1968 to 1972.*

Solution continue, Women in the Labor force

Solution

- *Then*
  $t = \frac{\overline{w}-0}{\frac{s_w}{\sqrt{n}}} = \frac{0.03368}{0.05974/\sqrt{19}} = 2.4577 > t_{0.05}(18) = 1.734064$
- *Thus, we reject H_0 at the 5% significance level. Thus, there was a significant increase at the 5% level in the labor force participation rate of women from 1968 to 1972.*
- *The data offer a good example of how a matched pairs t-test can be more effective when it is appropriate.*

Women in the Labor force

```
> t.test(x,y,alternative = c("greater"),var.equal=TRUE)

        Two Sample t-test

data:  x and y
t = 1.4959, df = 36, p-value = 0.0717
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.004333581          Inf
sample estimates:
mean of x mean of y
0.5268421 0.4931579
```

Notice that we apply the command: "var.equal=TRUE" when the population variances are the same for the two populations under test.

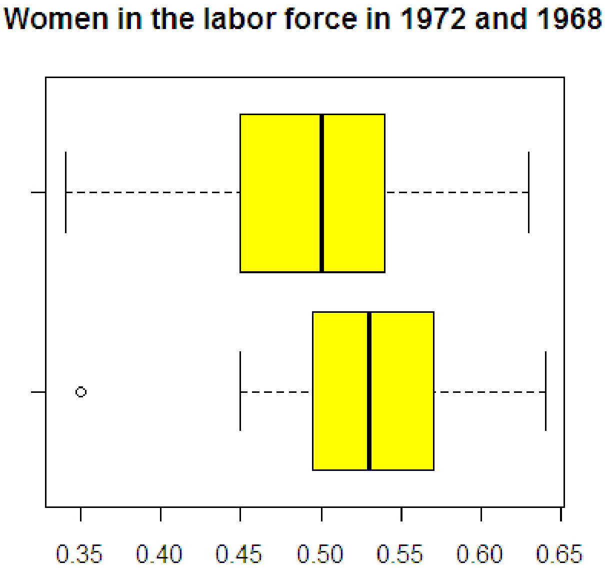
Women in the Labor force

```
> x
 [1] 0.45 0.50 0.52 0.45 0.46 0.55 0.60 0.49 0.35 0.55 0.52 0.53 0.57 0.53 0.59
[16] 0.64 0.50 0.57 0.64
> y
 [1] 0.42 0.50 0.52 0.45 0.43 0.55 0.45 0.34 0.45 0.54 0.42 0.51 0.49 0.54 0.50
[16] 0.58 0.49 0.56 0.63
> W=x-y
> W
 [1]  0.03  0.00  0.00  0.00  0.03  0.00  0.15  0.15 -0.10  0.01  0.10  0.02
[13]  0.08 -0.01  0.09  0.06  0.01  0.01  0.01
> t.test(x,y, alternative = c("greater"),mu = 0, paired = TRUE)

        Paired t-test

data:  x and y
t = 2.4577, df = 18, p-value = 0.01218
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.009917895         Inf
sample estimates:
mean of the differences
             0.03368421
```

Notice that we apply the command: "paired=TRUE" when we are using a matched pairs test.


Women in the labor force in 1972 and 1968
0.35
0.40
0.45
0.50
0.55
0.60
0.65

Test of hypothesis

- With the previous normal and independent assumptions, we test $H_0 : \sigma_X^2 = \sigma_Y^2$ using the statistics

$$F = \frac{\frac{(n-1)S_X^2}{\sigma_X^2(n-1)}}{\frac{(m-1)S_Y^2}{\sigma_Y^2(m-1)}} = \frac{S_X^2}{S_Y^2}$$

  which has an F-distribution with $r_1 = n-1$ and $r_2 = m-1$ degrees of freedom provided H_0 is true.
- It can be shown that $\frac{1}{F}$ has an F-distribution with $r_1 = m-1$ and $r_2 = n-1$ degrees of freedom.

Test of hypothesis

One sided composite test of hypothesis of the equality of variances:

$$H_0 : \sigma_X^2 = \sigma_Y^2 \quad \textit{against} \quad H_1 : \sigma_X^2 > \sigma_Y^2.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$\frac{s_x^2}{s_y^2} \geq F_\alpha(n-1, m-1)$$

Test of hypothesis

One sided composite test of hypothesis of the equality of variances:

$$H_0 : \sigma_X^2 = \sigma_Y^2 \quad \textit{against} \quad H_1 : \sigma_X^2 < \sigma_Y^2.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$\frac{s_y^2}{s_x^2} \geq F_\alpha(m-1, n-1)$$

Test of hypothesis

Two sided composite test of hypothesis of the equality of variances:

$$H_0 : \sigma_X^2 = \sigma_Y^2 \quad \textit{against} \quad H_1 : \sigma_X^2 \neq \sigma_Y^2.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$\frac{s_x^2}{s_y^2} \geq F_{\alpha/2}(n-1, m-1)$$

or

$$\frac{s_y^2}{s_x^2} \geq F_{\alpha/2}(m-1, n-1).$$

Example

Test the equality of the variance under the pooled t-test used in the previous problem. (The labor force participation rate of women in the United States in 1968 and 1972.)

Solution

- *We used the pooled t-test under the assumption that the variances of the two populations were the same.*
- *Test of variance:*

$$H_0 : \sigma_X^2 = \sigma_Y^2 \quad against \quad H_1 : \sigma_X^2 \neq \sigma_Y^2$$

- $\frac{s_x^2}{s_y^2} = \frac{0.005012}{0.004623} = 1.0841 < 2.52655 = F_{0.025}(19, 19)$.
- *Thus, we accept* H_0. *Thus the difference in the variances are not significantly different at the 5% level.*

In R,

$> qf(0.975, 19, 19)$

[1] 2.52655

so $F_{0.025}(19, 19) = 2.52655$.

In general:

$qf((1-\alpha)/2, df1, df2)$

- Let $X = (X_1, ..., X_n)$ and $Y = (Y_1, ..., Y_m)$ be two independent random samples, where X is $N(\mu_X, \sigma_X^2)$ and Y is $N(\mu_Y, \sigma_Y^2)$. We want to test whether the two means are equal when the variance are unequal.
- Recall that $E(\overline{X} - \overline{Y}) = \mu_X - \mu_Y$ and $Var(\overline{X} - \overline{Y}) = \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}$ and
- $$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}}.$$
- T does not have a t-distribution but can be approximated by a Student's t-distribution with $\lfloor v \rfloor$ degrees of freedom, where

$$v = \frac{\left(\frac{S_x^2}{n} + \frac{S_y^2}{m}\right)^2}{\frac{1}{n-1}\left(\frac{S_x^2}{n}\right)^2 + \frac{1}{m-1}\left(\frac{S_y^2}{m}\right)^2}.$$

Test of hypothesis, two-sample Welch test

One sided composite test of hypothesis for equality of two means with unequal variances:

$$H_0 : \mu_X = \mu_Y \quad \textit{against} \quad H_1 : \mu_X > \mu_Y.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$t \geq t_\alpha(\lfloor v \rfloor)$$

or

$$\overline{x} - \overline{y} \geq t_\alpha(\lfloor v \rfloor)\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}.$$

Test of hypothesis, two-sample Welch test

One sided composite test of hypothesis for equality of two means with unequal variances:

$$H_0 : \mu_X = \mu_Y \quad \textit{against} \quad H_1 : \mu_X < \mu_Y.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$t \leq -t_\alpha(\lfloor v \rfloor)$$

or

$$\bar{x} - \bar{y} \leq -t_\alpha(\lfloor v \rfloor)\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}.$$

Test of hypothesis, two-sample Welch test

Two sided composite test of hypothesis for equality of two means with unequal variances:

$$H_0 : \mu_X = \mu_Y \quad against \quad H_1 : \mu_X \neq \mu_Y.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$|t| \geq t_{\alpha/2}(\lfloor v \rfloor)$$

or

$$|\overline{x} - \overline{y}| \geq t_{\alpha/2}(\lfloor v \rfloor)\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}.$$

Speed of light

In 1879 and 1882 Michelson determined the speed of light. He had $n = 100$ observations in 1879 and $n=23$ observations in 1882 (suitably coded (km/sec, with 299000 subtracted)).
Source: DASL at http://lib.stat.cmu.edu/DASL/DataArchive.html
The data is provided on the next page.
Is there a difference in the mean of the speed of light of the 1879 observations and the 1882 observations at the 5% confidence level?

Speed of light

```
> data<-read.delim("C:/Documents and Settings/ghystad/Desktop/Math 361/differenceVel.txt")
> VEL1=data[,1]
> VEL1
  [1]  850  740  900 1070  930  850  950  980  980  880 1000  980  930  650  760
 [16]  810 1000 1000  960  960  960  940  960  940  880  800  850  880  900  840
 [31]  830  790  810  880  880  830  800  790  760  800  880  880  880  860  720
 [46]  720  620  860  970  950  880  910  850  870  840  840  850  840  840  840
 [61]  890  810  810  820  800  770  760  740  750  760  910  920  890  860  880
 [76]  720  840  850  850  780  890  840  780  810  760  810  790  810  820  850
 [91]  870  870  810  740  810  940  950  800  810  870
> VEL2=data[1:23,2]
> VEL2
 [1]  883  816  778  796  682  711  611  599 1051  781  578  796  774  820  772
[16]  696  573  748  748  797  851  809  723
```

Speed of light

Solution

- *It can be checked that the data appears to be approximately normal. Since we don't have any information about the variances, we apply the two-sided two-sample Welch test.*
- *Let* X=*vel1 and* Y = *vel2. We test*

$$H_0 : \mu_X = \mu_Y \quad \text{against} \quad H_1 : \mu_X \neq \mu_Y.$$

- *We have* $\overline{x} = 852.4$, $\overline{y} = 756.2174$, $s_x = 79.01055$ *and* $s_y = 107.1146$.
- *We have* $t = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} \approx 4.0598$.

Speed of light

Solution

- *We have*

$$v = \frac{\left(\frac{S_x^2}{n} + \frac{S_y^2}{m}\right)^2}{\frac{1}{n-1}\left(\frac{S_x^2}{n}\right)^2 + \frac{1}{m-1}\left(\frac{S_y^2}{m}\right)^2} = 27.754.$$

- *Since* $t = 4.0598 > t_{0.025}(\lfloor 27.754 \rfloor) = 2.0518$, *we reject* H_0*. Thus there is a difference in the mean of the speed of light of the 1879 observations and the 1882 observations at the* 5% *confidence level.*

Speed of light

```
>  t.test(VEL1,VEL2,alternative = c("two.sided"))

        Welch Two Sample t-test

data:  VEL1 and VEL2
t = 4.0598, df = 27.754, p-value = 0.0003625
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  47.63387 144.73135
sample estimates:
mean of x mean of y 
 852.4000  756.2174 
```

Speed of light, 99% confidence interval

```
>   t.test(VEL1,VEL2,conf.level = 0.99)

        Welch Two Sample t-test

data:  VEL1 and VEL2
t = 4.0598, df = 27.754, p-value = 0.0003625
alternative hypothesis: true difference in means is not equal to 0
99 percent confidence interval:
  30.67544 161.68977
sample estimates:
mean of x mean of y
 852.4000  756.2174
```

Test of hypothesis

- Let Y_1 and Y_2 be binomial random variables, $b(n_1, p_1)$ and $b(n_2, p_2)$ respectively, where p_1 and p_2 are unknown.

Test of hypothesis

- Let Y_1 and Y_2 be binomial random variables, $b(n_1, p_1)$ and $b(n_2, p_2)$ respectively, where p_1 and p_2 are unknown.
- We know from the Central limit theorem that $\hat{p}_1 = \frac{Y_1}{n_1}$ and $\hat{p}_2 = \frac{Y_2}{n_2}$ have approximate normal distributions so that

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}}$$

is approximate $N(0, 1)$.

Test of hypothesis

- Let Y_1 and Y_2 be binomial random variables, $b(n_1, p_1)$ and $b(n_2, p_2)$ respectively, where p_1 and p_2 are unknown.
- We know from the Central limit theorem that $\hat{p}_1 = \frac{Y_1}{n_1}$ and $\hat{p}_2 = \frac{Y_2}{n_2}$ have approximate normal distributions so that

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}}$$

  is approximate $N(0, 1)$.
- In testing $H_0 : p_1 = p_2$ against H_1: $(p_1 > p_1$ or $p_1 < p_1$ or $p_1 \neq p_1)$, we use the test-statistics

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}}$$

Test of hypothesis

One sided composite test of hypothesis for two proportions:

$$H_0 : p_1 = p_2 \quad \textit{against} \quad H_1 : p_1 > p_1.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1 - \hat{p}_1)/n_1 + \hat{p}_2(1 - \hat{p}_2)/n_2}} \geq z_\alpha.$$

Test of hypothesis

One sided composite test of hypothesis for two proportions:

$$H_0 : p_1 = p_2 \quad \textit{against} \quad H_1 : p_1 < p_1.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} \leq -z_\alpha.$$

Test of hypothesis

Two sided composite test of hypothesis for two proportions:

$$H_0 : p_1 = p_2 \quad \textit{against} \quad H_1 : p_1 \neq p_1.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$|z| = \frac{|\hat{p}_1 - \hat{p}_2|}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} \geq z_{\alpha/2}.$$

Test of hypothesis

Note that in testing $H_0 : p_1 = p_2$ some statisticians replace both $\hat{p}_1$ and $\hat{p}_2$ in the denominator of Z by an estimate of the common $p_1 = p_2$, namely

$$\hat{p} = \frac{Y_1 + Y_2}{n_1 + n_2}.$$

The test statistics is then

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}.$$

Test of hypothesis

- The estimate of the standard deviation of $\hat{p}_1 - \hat{p}_2$ using the common estimate $\hat{p}$ is better if H_0 is true.
- The estimate of the standard deviation of $\hat{p}_1 - \hat{p}_2$ using the individual $\hat{p}_1$ and $\hat{p}_2$ in the denominator is better when H_0 is clearly false.
- However, the numerical results are about the same.
- The estimate of the standard deviation is called the standard error of $\hat{p}_1 - \hat{p}_2$.

Test of hypothesis

The Trial Urban District Assessment (TUDA) is a study sponsored by the government of student achievement in large urban school district. The math test-score is on a scale from 0 to 500. A "basic" math level is a score of 262, a "proficient" level is a score of 299 and a "advanced" level is a score of 333. In 2007, 715 of a random sample of 1100 *eighth*-graders from Houston performed at or above the basic level. In 2011, 864 of a random sample of 1200 *eighth*-graders from Houston performed at or above the basic level. (The study reports the proportions). Source: TUDA results for 2007 and 2011 from the National Center for Education Statistics, at nces.ed.gov/nationsreportcard

Is there an increase in the proportions of students who performed at or above the basic level from 2007 to 2011 at the 5% significance level?

Solution

Let Y_1 be the number of eighth-graders from Houston that performed at or above the basic level in 2011 and let Y_2 be the number of eighth-graders from Houston that performed at or above the basic level in 2007. Let p_1 and p_2 be the proportions of eighth-graders that performed at or above the basic level in 2011 and 2007 respectively. We test

$$H_0 : p_1 = p_2 \quad \textit{against} \quad H_1 : p_1 > p_1.$$

We have $\hat{p}_1 = \frac{Y_1}{n_1} = \frac{864}{1200} = 0.72$ *and* $\hat{p}_2 = \frac{Y_2}{n_2} = \frac{715}{1100} = 0.65$*. Then*
$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} =$
$\frac{0.72-0.65}{\sqrt{0.72(1-0.72)/1200+0.65(1-0.65)/1100}} \approx 3.6157 > 1.6449 = z_{0.05}$*. We reject H_0. Thus, there is an increase from 2007 to 2011 in the proportions of eighth-graders who performed at or above the basic level at the 5% significance level.*

```
> prop.test(c(864,715),c(1200,1100),alternative=c("greater"))

        2-sample test for equality of proportions with continuity correction

data:  c(864, 715) out of c(1200, 1100)
X-squared = 12.7439, df = 1, p-value = 0.0001786
alternative hypothesis: greater
95 percent confidence interval:
 0.03728406 1.00000000
sample estimates:
prop 1 prop 2
  0.72   0.65
```