# Math 4501 - Probability and Statistics II

8.3 - Hypothesis testing: variances

# Overview

We will study hypothesis tests and confidence intervals for:

- the variance or standard deviation of a normal distribution
- the ratio of the variances or standard deviations of two independent normal distributions

# Tests of hypotheses for one variance

Let $X$ be a normally distributed random variable with mean $\mu$ and variance $\sigma^2$.

both unknown

We are interested in testing the null hypothesis

$$H_0 : \sigma^2 = \sigma_0^2$$

fixed value

against one of the following three composite alternative hypotheses:

(i) $H_1 : \sigma^2 > \sigma_0^2$ ✓

(ii) $H_1 : \sigma^2 < \sigma_0^2$ ✓

(iii) $H_1 : \sigma^2 \neq \sigma_0^2$ ✓

One random sample $X_1, \ldots, X_n$ is taken from the distribution of $X$.

The sample variance $s^2$ is found:

- A decision concerning rejection of $H_0$, is related with how close the sample variance $s^2$ is to $\sigma_0^2$.

Under the null hypothesis $H_0 : \sigma^2 = \sigma_0^2$, the test statistic

$$\boxed{Q = \frac{(n-1)S^2}{\sigma_0^2} \quad \text{is } \chi^2(n-1)}$$

From Chp. 5

Let $s^2$ be the observed value of the sample variance, and let

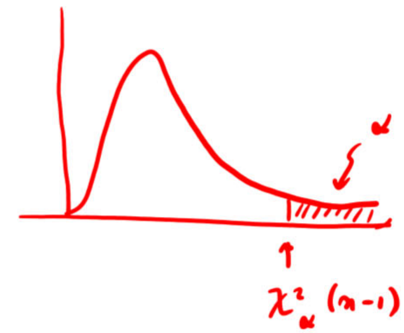
$$\boxed{q = \frac{(n-1)s^2}{\sigma_0^2}}$$

be the corresponding value of the test statistic.

$q = \frac{(n-1)s^2}{\sigma^2}$

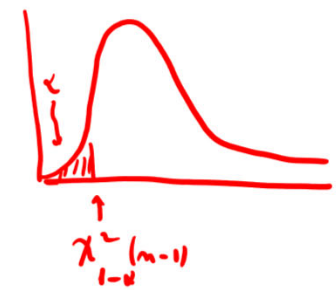
(i) When testing $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 > \sigma_0^2$, we reject $H_0$ if

$$q \geq \chi^2_\alpha(n-1) \quad \text{or, equivalently,} \quad s^2 \geq \sigma_0^2 \frac{\chi^2_\alpha(n-1)}{n-1}$$




(ii) When testing $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 < \sigma_0^2$, we reject $H_0$ if

$$q \leq \chi^2_{1-\alpha}(n-1) \quad \text{or, equivalently,} \quad s^2 \leq \sigma_0^2 \frac{\chi^2_{1-\alpha}(n-1)}{n-1}$$

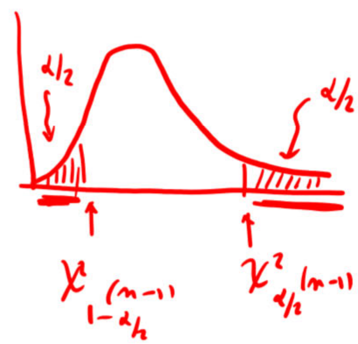



(iii) When testing $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 \neq \sigma_0^2$, we reject $H_0$ if

$$q \leq \chi^2_{1-\alpha/2}(n-1) \quad \text{or} \quad q \geq \chi^2_{\alpha/2}(n-1)$$

or, equivalently,

$$s^2 \leq \sigma_0^2 \frac{\chi^2_{1-\alpha/2}(n-1)}{n-1} \quad \text{or} \quad s^2 \geq \sigma_0^2 \frac{\chi^2_{\alpha/2}(n-1)}{n-1}$$



## Example

Suppose that a well-established manufacturing process produces pills for which the standard deviation of the amount of active ingredient is 0.6 micrograms. Suppose further that a pharmaceutical company has developed a new process for producing the pills that it thinks may reduce this standard deviation.

The company wants to test the null hypothesis $H_0 : \sigma = 0.6$ against the alternative hypothesis $H_1 : \sigma < 0.6$.

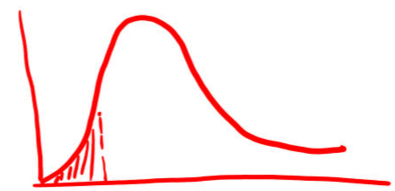
A random sample of $n = 23$ pills is taken from the production line using the new process, yielding the sample standard deviation $s = 0.42$

Under the assumption that the amount of active ingredient in each pill is normally distributed, test the null hypothesis $H_0 : \sigma = 0.6$ against the alternative hypothesis $H_1 : \sigma < 0.6$ at significance level $\alpha = 0.05$

To test $H_0 : \sigma = 0.6$ against $H_1 : \sigma < 0.6$, we compute the test statistic $\quad Q = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2(n-1)$

$$q = \frac{(\overset{23}{n}-1)s^2}{\sigma_0^2} = \frac{(22)(0.\overset{s}{4}2)^2}{0.6^2} = 10.78. \quad < \chi^2_{0.95}(22)$$

$\chi^2_{1-\alpha}(n-1)$



$\chi^2_{1-\alpha}(n-1)$

Comparing this to $\boxed{\chi^2_{0.95}(22) = 12.34}$ ← to be, we reject $H_0$.

In words, it can be concluded, at the 0.05 level of significance, that the new production process reduces the variability in the amount of active ingredient in the pills.

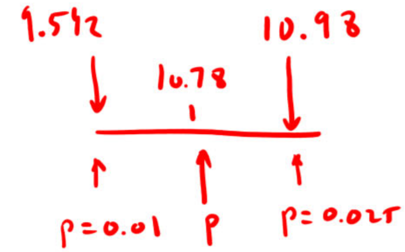
We could also find $p$-values for tests about a variance. In this example:

under $H_0$: $\sigma = \sigma_0 \Rightarrow Q = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2(n-1)$

$$p\text{-value} = P(Q \le 10.78) \; ,$$

where $\underbrace{Q \sim \chi^2(22)}_{}$ under $H_0$. Using the tables for the chi-square distribution, we see that

$$P(Q \le 10.98) = 0.025 \quad \text{and} \quad P(Q \le 9.542) = 0.01$$




and so

$$0.01 < p\text{-value} < 0.025 \; .$$

$\boxed{p < \alpha = 0.05 \Downarrow \text{Rej } H_0}$

The null hypothesis would be rejected at the 0.025 level but not at the 0.01 level of significance.

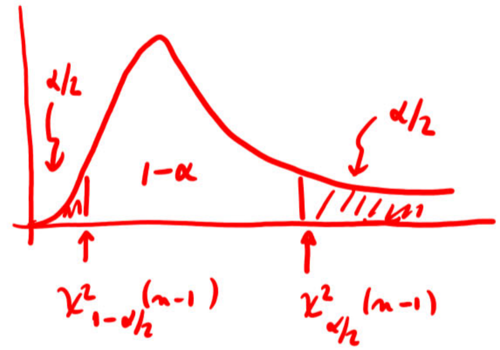
**Note**: If the alternative hypothesis had been two-sided, the $p$-value would have been $2P(Q \le 10.78)$, which lies between 0.02 and 0.05.

## Confidence intervals for one variance (related with content of Chp. 7)

Let $X_1, X_2, \ldots X_n$ be a random sample from a normal distribution (with unknown mean $\mu$ and variance $\sigma^2$).

Since

$$\boxed{Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \ ,}$$




for any $\alpha \in (0,1)$, we have

$$P\left(\chi^2_{1-\alpha/2}(n-1) \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{\alpha/2}(n-1)\right) = \underbrace{1-\alpha}_{\text{confidence level}}.$$

take reciprocals

Solving the inequalities inside the parentheses for $\sigma^2$ yields

$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right) = 1-\alpha.$$

Once the sample variance is $s^2$ is observed, a $100(1-\alpha)\%$ two-sided confidence interval for $\sigma^2$ is given by

$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}\right] .$$

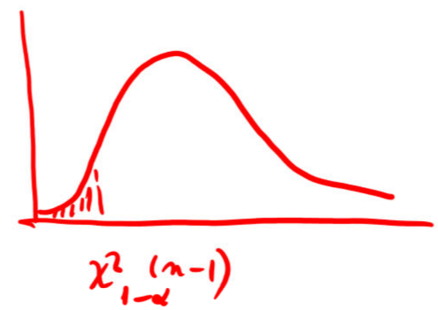
One-sided confidence intervals can be obtained similarly:

- $100(1-\alpha)\%$ lower bound for $\sigma^2$:

$$\frac{(n-1)s^2}{\chi^2_{\alpha}(n-1)}$$


- $100(1-\alpha)\%$ upper bound for $\sigma^2$:

$$\frac{(n-1)s^2}{\chi^2_{1-\alpha}(n-1)}$$



## Example (Continued) $\rightsquigarrow$ data from new manufacturing process

Construct a 95% two-sided confidence interval for the variance of a normal distribution given a sample of size $n = 23$ and sample standard deviation $s = 0.42$.

Since $\alpha = 0.05$, $n = 23$, and $s = 0.42$, the desired interval is:

$$\sigma^2 \in \left[\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}\right] = \left[\frac{(22)(0.42)^2}{33.92}, \frac{(22)(0.42)^2}{12.34}\right]$$

$$= [0.1144, 0.3145],$$

where we got

$$\chi^2_{0.025}(22) = 33.92 \quad \text{and} \quad \chi^2_{0.975}(22) = 12.34$$

from the chi-square distribution tables.

# Tests of hypotheses for equality of variances

Let $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ be independent.

We are interested in testing the null hypothesis

$$\boxed{H_0 : \frac{\sigma_X^2}{\sigma_Y^2} = 1} \quad \text{or, equivalently,} \quad \sigma_X^2 = \sigma_Y^2$$

against one of the following three composite alternative hypotheses:

(i) $H_1 : \frac{\sigma_X^2}{\sigma_Y^2} > 1$ $\rightsquigarrow$ $\sigma_x^2 > \sigma_y^2$

(ii) $H_1 : \frac{\sigma_X^2}{\sigma_Y^2} < 1$ $\longrightarrow$ $\sigma_x^2 < \sigma_y^2$

(iii) $H_1 : \frac{\sigma_X^2}{\sigma_Y^2} \neq 1$ $\longrightarrow$ $\sigma_x^2 \neq \sigma_y^2$

Extract independent random samples:

- $X_1, \ldots, X_{n_X}$ from $N(\mu_X, \sigma_X^2)$
- $Y_1, \ldots, Y_{n_Y}$ from $N(\mu_Y, \sigma_Y^2)$

possibly different sample sizes!

Sample variances $s_X^2$ and $s_Y^2$ are determined.

Observe that

$$\boxed{Q_X = \frac{(n_X - 1)S_X^2}{\sigma_X^2} \sim \chi^2(n_X - 1)} \quad \text{and} \quad \boxed{Q_Y = \frac{(n_Y - 1)S_Y^2}{\sigma_Y^2} \sim \chi^2(n_Y - 1)}$$

Chap. 5

are independent

or $\sigma_x^2/\sigma_y^2 = 1$

Under the null hypothesis $H_0 : \sigma_X^2 = \sigma_Y^2$, the test statistic

$$F = \frac{Q_X/(n_X - 1)}{Q_Y/(n_Y - 1)} = \frac{\frac{\cancel{(n_X-1)}S_X^2}{\cancel{\sigma_X^2}\cancel{(n_X-1)}}}{\frac{\cancel{(n_Y-1)}S_Y^2}{\cancel{\sigma_Y^2}\cancel{(n_Y-1)}}} = \frac{S_X^2}{S_Y^2} \text{ is } F(n_X - 1, n_Y - 1)$$

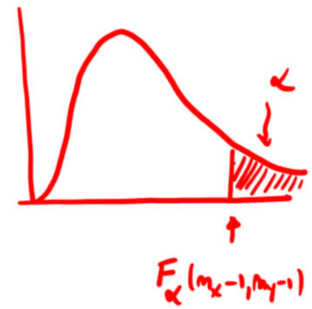
Let $s_X^2$ and $s_Y^2$ be the observed values of the two sample variances, and let

$$\boxed{f = \frac{s_X^2}{s_Y^2}}$$

be the corresponding value of the test statistic.

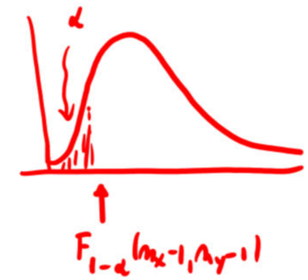
(i) When testing $H_0 : \sigma_X^2 = \sigma_Y^2$ against $H_1 : \sigma_X^2 > \sigma_Y^2$, we reject $H_0$ if

$$f \geq F_{\alpha}(n_X - 1, n_Y - 1) \qquad \text{or, equivalently,} \qquad \frac{s_X^2}{s_Y^2} \geq F_{\alpha}(n_X - 1, n_Y - 1)$$




(ii) When testing $H_0 : \sigma_X^2 = \sigma_Y^2$ against $H_1 : \sigma_X^2 < \sigma_Y^2$, we reject $H_0$ if

$$f \leq F_{1-\alpha}(n_X - 1, n_Y - 1) \qquad \text{or, equivalently,} \qquad \frac{s_Y^2}{s_X^2} \geq F_{\alpha}(n_Y - 1, n_X - 1)$$

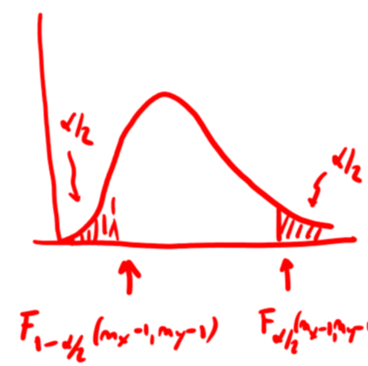



(iii) When testing $H_0 : \sigma_X^2 = \sigma_Y^2$ against $H_1 : \sigma_X^2 \neq \sigma_Y^2$, we reject $H_0$ if

$$f \leq F_{1-\alpha/2}(n_X - 1, n_Y - 1) \quad \text{or} \quad f \geq F_{\alpha/2}(n_X - 1, n_Y - 1)$$

or, equivalently,

$$\frac{s_Y^2}{s_X^2} \geq F_{\alpha/2}(n_Y - 1, n_X - 1) \quad \text{or} \quad \frac{s_X^2}{s_Y^2} \geq F_{\alpha/2}(n_X - 1, n_Y - 1)$$



$2^{nd}$ case: $H_0: \sigma_x^2 = \sigma_y^2$ vs $H_1: \sigma_x^2 < \sigma_y^2$

Rej $H_0$ if $f \le F_{1-\alpha}(n_x - 1, n_y - 1)$

$$\Leftrightarrow \frac{S_x^2}{S_y^2} \le F_{1-\alpha}(n_x-1, n_y-1)$$

take reciprocals

$$\Leftrightarrow \frac{S_y^2}{S_x^2} \geq \frac{1}{F_{1-\alpha}(n_x-1, n_y-1)} = F_\alpha(n_y-1, n_x-1)$$

Recall that if $U \sim \chi^2(m)$ and $V \sim \chi^2(n)$

$\Rightarrow F = \dfrac{U/m}{V/n} \sim F(m,n)$

$\Rightarrow \dfrac{1}{F} = \dfrac{V/n}{U/m} \sim F(n,m)$

## Example

Given $n_X = 11$ observations of $X \sim N\left(\mu_X, \sigma_X^2\right)$ and $n_Y = 13$ observations of $Y \sim N\left(\mu_Y, \sigma_Y^2\right)$ yielding the sample variances $s_X^2 = 0.24$ and $s_Y^2 = 0.35$, test the null hypothesis $H_0 : \sigma_X^2 = \sigma_Y^2$ against a $\underbrace{\text{two-sided alternative hypothesis}}_{H_1:\ \sigma_X^2 \neq \sigma_Y^2}$ at $\alpha = 0.05$ significance level.

At an $\alpha = 0.05$ significance level, $H_0$ is rejected if

$$f = \frac{s_X^2}{s_Y^2} \geq F_{0.025}(10, 12) = 3.37$$

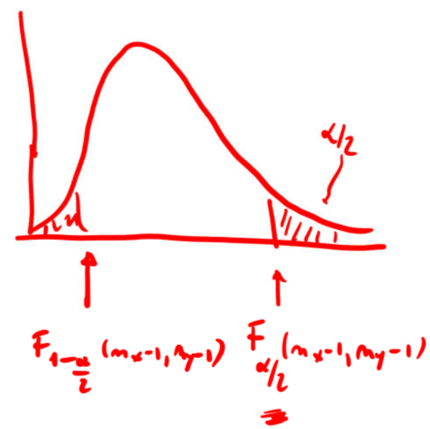
$\alpha/2$   $n_x - 1$   $n_y - 1$

or $f \leq F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1)$

$\frac{1}{f} \geq F_{\alpha/2}(n_y - 1, n_x - 1)$

$$\frac{1}{f} = \frac{s_Y^2}{s_X^2} \geq F_{0.025}(12, 10) = 3.62.$$

$n_y - 1$   $n_x - 1$




Using the data provided, we have

$$\frac{s_X^2}{s_Y^2} = \frac{0.24}{0.35} = 0.686 \quad \text{and} \quad \frac{s_Y^2}{s_X^2} = \frac{0.35}{0.24} = 1.458\ ,$$

so we do not reject $H_0$.

# Confidence intervals for the ratio of two variances

$\sigma_x^2/\sigma_y^2$

Consider the following independent random samples:

- $X_1, X_2, \ldots X_{n_X}$ from a $N(\mu_X, \sigma_X^2)$ distribution (with unknown $\mu_X$ and $\sigma_X^2$)
- $Y_1, Y_2, \ldots Y_{n_X}$ from a $N(\mu_Y, \sigma_Y^2)$ distribution (with unknown $\mu_Y$ and $\sigma_Y^2$)

We have that

$$Q_X = \frac{(n_X - 1)S_X^2}{\sigma_X^2} \sim \chi^2(n_X - 1) \quad \text{and} \quad Q_Y = \frac{(n_Y - 1)S_Y^2}{\sigma_Y^2} \sim \chi^2(n_Y - 1)$$

are independent and

$$\frac{Q_Y/(n_Y - 1)}{Q_X/(n_X - 1)} = \frac{\frac{\cancel{(n_Y-1)}S_Y^2}{\sigma_Y^2\cancel{(n_Y-1)}}}{\frac{\cancel{(n_X-1)}S_X^2}{\sigma_X^2\cancel{(n_X-1)}}} = \frac{S_Y^2/\sigma_Y^2}{S_X^2/\sigma_X^2} \text{ is } F(\underbrace{n_Y - 1}_{\text{numerator}}, \underbrace{n_X - 1}_{\text{denominator}})\ .$$

$$\frac{\sigma_x^2}{\sigma_y^2} \cdot \frac{s_y^2}{s_x^2}$$

Thus, for any $\alpha \in (0, 1)$, we have

$$P\left(F_{1-\alpha/2}(n_Y - 1, n_X - 1) \le \frac{S_Y^2/\sigma_Y^2}{S_X^2/\sigma_X^2} \le F_{\alpha/2}(n_Y - 1, n_X - 1)\right) = 1 - \alpha.$$

Solving the inequalities inside the parentheses for $\sigma_X^2/\sigma_Y^2$ yields

$$P\left[F_{1-\alpha/2}\left(n_Y-1,n_X-1\right)\frac{S_X^2}{S_Y^2}\le\frac{\sigma_X^2}{\sigma_Y^2}\le F_{\alpha/2}\left(n_Y-1,n_X-1\right)\frac{S_X^2}{S_Y^2}\right]=1-\alpha.$$

Once the sample variances $s_X^2$ and $s_Y^2$ are observed, a $100(1-\alpha)\%$ two-sided confidence interval for $\sigma_X^2/\sigma_Y^2$ is given by

$$\left[F_{1-\alpha/2}\left(n_Y-1,n_X-1\right)\frac{s_X^2}{s_Y^2},F_{\alpha/2}\left(n_Y-1,n_X-1\right)\frac{s_X^2}{s_Y^2}\right].$$

One-sided confidence intervals can be obtained similarly:

- $100(1-\alpha)\%$ upper bound for $\sigma_X^2/\sigma_Y^2$:

$$F_{\alpha}\left(n_Y-1,n_X-1\right)\frac{s_X^2}{s_Y^2}$$

- $100(1-\alpha)\%$ lower bound for $\sigma_X^2/\sigma_Y^2$:

$$F_{\underline{1-\alpha}}\left(n_Y-1,n_X-1\right)\frac{s_X^2}{s_Y^2}$$

**EXAMPLE**: Same data as in previous example

Find a $\underbrace{95\% \text{ confidence}}_{\alpha = 0.05}$ interval for $\dfrac{\sigma_x^2}{\sigma_y^2}$

Recall data:

$n_x = 11$ , $s_x^2 = 0.24$

$n_y = 13$ , $s_y^2 = 0.35$

Confidence interval for $\dfrac{\sigma_x^2}{\sigma_y^2}$ is

$$\left[ F_{1-\alpha/2}(n_y-1, n_x-1) \cdot \frac{S_x^2}{S_y^2} ,\ F_{\alpha/2}(n_y-1, n_x-1) \frac{S_x^2}{S_y^2} \right]$$

$$\left[ \underbrace{F_{0.975}(12,10)}_{\Downarrow} \frac{0.24}{0.35} ,\ \underbrace{F_{0.025}(12,10)}_{\text{table}} \frac{0.24}{0.35} \right]$$

$$\frac{1}{F_{0.025}(10,12)} \cdot \frac{0.24}{0.35}$$

table!

$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$

$\Rightarrow 1 - \frac{\alpha}{2} = 0.975$

# Math 4501 - Probability and Statistics II

8.4 - Hypothesis testing: proportions $\Bigg\}$ similar to 8.1, 8.2 and also 7.3

# Overview

We will study approximate hypothesis tests for:

- proportions
- equality of proportions

# Hypothesis testing for one proportion: setup

Let $Y$ be a random variable representing the number of observed successes in $n$ independent Bernoulli trials with some unknown probability of success $p$.

$\hat{p} = \frac{Y}{n}$

point estimate for probability of success

We are interested in testing the null hypothesis

$$H_0 : p = p_0$$

against one of the following three composite alternative hypothesis:

$$H_1 : p > p_0 \qquad \text{or} \qquad H_1 : p < p_0 \qquad \text{or} \qquad H_1 : p \neq p_0$$

Let $y$ denote the number of observed successes on such sequence of trials.

The point estimate $y/n$ for the unknown proportion $p$ is determined.

- A decision concerning rejection of $H_0$, is related with how close $y/n$ is to $p_0$.

# Hypothesis testing for one proportion

Let $Y$ be a random variable representing the number of observed successes in $n$ independent Bernoulli trials with some unknown probability of success $p$.

- Recall that $Y$ is $b(n, p)$

Note that, under the null hypothesis $H_0 : p = p_0$, for $n$ sufficiently large

$$\frac{Y}{n} \text{ is approximately } N\left(p_0, \frac{p_0(1-p_0)}{n}\right) \quad \leftarrow \text{ under } H_0$$

CLT

Thus, under the null hypothesis $H_0 : p = p_0$, for $n$ sufficiently large, the test statistic

$$Z = \frac{Y/n - p_0}{\sqrt{p_0\,(1-p_0)\,/n}} \quad \text{is approximately } N(0,1)$$

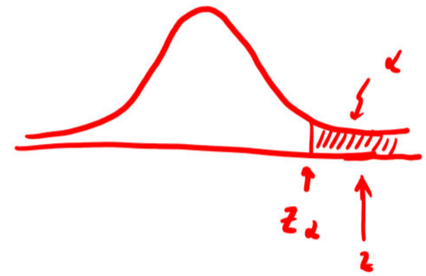
Let $y$ be the observed number of successes in $n$ independent Bernoulli trials, and let

$$z = \frac{y/n - p_0}{\sqrt{p_0\,(1-p_0)\,/n}}$$

be the corresponding value of the test statistic.

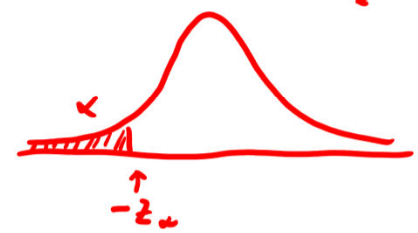
(i) When testing $H_0 : p = p_0$ against $H_1 : p > p_0$, we reject $H_0$ if

$$z \geq z_\alpha \qquad \text{or, equivalently,} \qquad \frac{y}{n} \geq p_0 + z_\alpha \sqrt{\frac{p_0(1-p_0)}{n}}$$




(ii) When testing $H_0 : p = p_0$ against $H_1 : p < p_0$, we reject $H_0$ if

$$z \leq -z_\alpha \qquad \text{or, equivalently,} \qquad \frac{y}{n} \leq p_0 - z_\alpha \sqrt{\frac{p_0(1-p_0)}{n}}$$




(iii) When testing $H_0 : p = p_0$ against $H_1 : p \neq p_0$, we reject $H_0$ if

$$|z| \geq z_{\alpha/2} \qquad \text{or, equivalently,} \qquad \left|\frac{y}{n} - p_0\right| \geq z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}}$$

↳ $z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$

## Example

Let $Y$ be the random variable representing the number of successes in $n$ Bernoulli trials with probability of success $p$.

Test the hypotheses

$$\begin{cases} H_0 : p = 1/6 \\ H_1 : p > 1/6 \end{cases}$$

at significance level $\alpha = 0.05$ knowing that on $n = 8000$ trials, $y = 1389$ successes were observed.

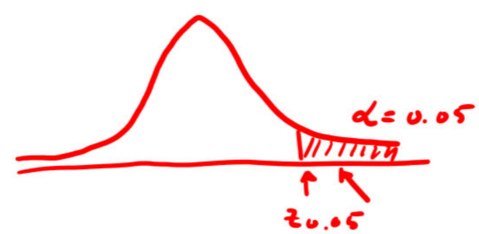
The test statistic is

Under $H_0$

$$\boxed{Z = \frac{Y/n - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{Y/8000 - 1/6}{\sqrt{(1/6)(5/6)/8000}}} \longleftarrow \text{approx. } N(0,1)$$

Using a significance level of $\alpha = 0.05$, the critical region is

$$z \geq z_{0.05} = 1.645$$



Since $y = 1389$ successes were observed, the calculated value of the test statistic is

$$z = \frac{1389/8000 - 1/6}{\sqrt{(1/6)(5/6)/8000}} = 1.67$$

since

$$z = 1.67 > 1.645 = z_{0.05}$$

the null hypothesis $H_0$ is rejected.

These results indicate that the probability of success in such Bernoulli trials is greater than $1/6$.

# Hypothesis testing for two proportions: setup

Let $Y_1$ and $Y_2$ represent, respectively, the numbers of observed successes in $n_1$ and $n_2$ independent Bernoulli trials with probabilities of success $p_1$ and $p_2$.

We are interested in testing the null hypothesis

$$\boxed{H_0 : p_1 = p_2} \quad p_1 - p_2 = 0$$

against one of the following three composite alternative hypothesis:

$$H_1 : p_1 > p_2 \qquad \text{or} \qquad H_1 : p_1 < p_2 \qquad \text{or} \qquad H_1 : p_1 \neq p_2$$

Let $y_1$ and $y_2$ denote the number of observed successes on such sequence of trials.

Point estimates $y_1/n_1$ and $y_2/n_2$ are determined for the respective unknown proportions $p_1$ and $p_2$.

- A decision concerning rejection of $H_0$ is related with how close the difference $y_1/n_1 - y_2/n_2$ is to 0.

# Hypothesis testing for two proportions

Let $Y_1$ and $Y_2$ be random variables representing, respectively, the numbers of observed successes in $n_1$ and $n_2$ independent Bernoulli trials with probabilities of success $p_1$ and $p_2$.

- Recall that $Y_1$ is $b(n_1, p_1)$ and $Y_2$ is $b(n_2, p_2)$

For $n_1$ and $n_2$ large enough:

CLT

- $\widehat{p}_1 = Y_1/n_1$ is approximately $N\left(p_1, p_1(1-p_1)/n_1\right)$
- $\widehat{p}_2 = Y_2/n_2$ is approximately $N\left(p_2, p_2(1-p_2)/n_2\right)$

CLT + large sample

As a consequence, for $n_1$ and $n_2$ large enough

Standardize

$$\overbrace{\widehat{p}_1 - \widehat{p}_2}^{Y_1/n_1 - Y_2/n_2} \quad \text{is approximately } N\left[\underbrace{p_1 - p_2}_{E[\hat{p}_1 - \hat{p}_2]}, \underbrace{p_1\left(1-p_1\right)/n_1 + p_2\left(1-p_2\right)/n_2}_{Var(\hat{p}_1 - \hat{p}_2)}\right]$$

and so

$$Z = \frac{\widehat{p}_1 - \widehat{p}_2 - \overbrace{\left(p_1 - p_2\right)}}{\underbrace{\sqrt{p_1\left(1-p_1\right)/n_1 + p_2\left(1-p_2\right)/n_2}}} \quad \text{is approximately } N(0,1)$$

Under the null hypothesis $H_0 : p_1 = p_2$, let $p = p_1 = p_2$ be the common value of the proportions under $H_0$.

Estimate $p$ using (the pooled estimate)

total # of successes

$p_1$ and $p_2$ under $H_0$

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

combined the two samples to obtain a more precise estimate for $p_1$ and $p_2$ under $H_0 : p_1 = p_2$

total number of trial

Replacing $p_1$ and $p_2$ in the denominator of

$$Z = \frac{\widehat{p}_1 - \widehat{p}_2 - \overbrace{(p_1 - p_2)}^{=0}}{\sqrt{p_1\,(1 - p_1)\,/\,n_1 + p_2\,(1 - p_2)\,/\,n_2}}$$

with this pooled estimate to obtain, under the null hypothesis $H_0 : p_1 = p_2$ and for $n_1$ and $n_2$ sufficiently large, that the test statistic

$$Z = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1 - \widehat{p})\,(1/n_1 + 1/n_2)}} \quad \text{is approximately } N\,(0,1)$$

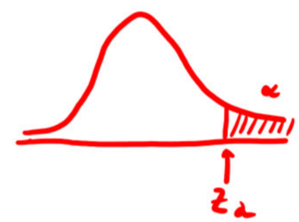
Let $y_1$ and $y_2$ be the observed number of successes in the respective independent Bernoulli trials, set $\widehat{p}_1 = y_1/n_1$, $\widehat{p}_2 = y_2/n_2$ and $\widehat{p} = (y_1 + y_2)/(n_1 + n_2)$, and let

$$z = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1-\widehat{p})\,(1/n_1 + 1/n_2)}}$$

be the corresponding value of the test statistic.

(i) When testing $H_0 : p_1 = p_2$ against $H_1 : p_1 > p_2$, we reject $H_0$ if

$$z \geq z_\alpha \quad \text{or, equivalently,} \quad \widehat{p}_1 - \widehat{p}_2 \geq z_\alpha \sqrt{\widehat{p}(1-\widehat{p})\,(1/n_1 + 1/n_2)}$$

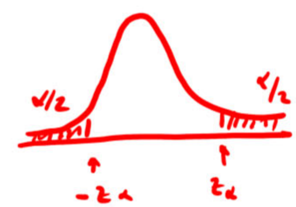



(ii) When testing $H_0 : p_1 = p_2$ against $H_1 : p_1 < p_2$, we reject $H_0$ if

$$z \leq -z_\alpha \quad \text{or, equivalently,} \quad \widehat{p}_1 - \widehat{p}_2 \leq -z_\alpha \sqrt{\widehat{p}(1-\widehat{p})\,(1/n_1 + 1/n_2)}$$


(iii) When testing $H_0 : p_1 = p_2$ against $H_1 : p_1 \neq p_2$, we reject $H_0$ if

$$|z| \geq z_{\alpha/2} \quad \text{or, equivalently,} \quad |\widehat{p}_1 - \widehat{p}_2| \geq z_{\alpha/2} \sqrt{\widehat{p}(1-\widehat{p})\,(1/n_1 + 1/n_2)}$$




$z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$

## Example

Let $Y_1$ and $Y_2$ be random variables representing, respectively, the numbers of observed successes in $n_1 = 900$ and $n_2 = 700$ independent Bernoulli trials with probabilities of success $p_1$ and $p_2$.

Test the hypotheses

$$\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 > p_2 \end{cases}$$

at significance level $\alpha = 0.05$ knowing that the respective number of successes observed were $y_1 = 135$ and $y_2 = 77$

pooled estimator for $p_1$ and $p_2$ under $H_0: p_1 = p_2$

Note that

$$\boxed{\widehat{p}_1 = \frac{y_1}{n_1} = \frac{135}{900}} \quad \boxed{\widehat{p}_2 = \frac{y_2}{n_2} = \frac{77}{700}} \quad \boxed{\widehat{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{212}{1600}}\ .$$

The value of the test statistic is then

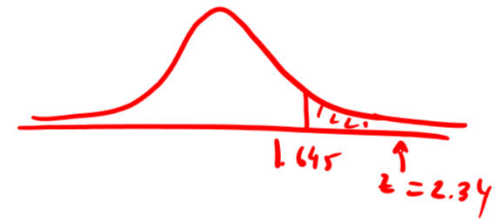
$$z = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1 - \widehat{p})\,(1/n_1 + 1/n_2)}} \approx 2.34$$

Since

$$z = 2.34 > 1.645 = z_{0.05}$$

table

the null hypothesis $H_0$ is rejected.




These results indicate that the two probabilities of success (true proportions) are not equal.

# Math 4501 - Probability and Statistics II

8.6 - Hypothesis testing: $\underbrace{\text{power of a statistical test}}$

Builds on 8.1
AND
also similar to 7.4

and how to pick the sample size to perform a test

## Type I error and significance level (review)

Type I error for a statistical test with critical region $C$:

- to reject $H_0$ when $H_0$ is true
- occurs if $(x_1, x_2, \ldots, x_n) \in C$ even if $H_0$ is true

The *significance level* of a statistical test is

$$\left.\begin{aligned} \alpha &= P(\text{Type I error}) \\ &= P(\text{reject } H_0 | H_0 \text{ true}) \end{aligned}\right\} \text{we want } \alpha \text{ small!}$$

$$= P((X_1, \ldots, X_n) \in C | H_0 \text{ true}) \; .$$

# Type II error and power of a test

Type II error for a statistical test with critical region $C$:

- to not reject $H_0$ when $H_1$ is true
- occurs if $(x_1, x_2, \ldots, x_n) \notin C$ even if $H_1$ is true

The probability of occurrence of an error of type II is denoted as $\beta$:

$$\begin{aligned}\beta &= P(\text{Type II error}) \\ &= P(\text{do not reject } H_0 | H_1 \text{ true}) \\ &= P((X_1, \ldots, X_n) \notin C | H_1 \text{ true}) .\end{aligned}$$

power $= 1-\beta = P(\text{Rej } H_0 \mid H_1 \text{ true})$ ← by definition

**Definition**

The *power* of a test is the quantity $1 - \beta$, the probability of correctly rejecting the null hypothesis $H_0$ when the alternative hypothesis $H_1$ is true.

**Notes**:

- The power of the test is a function of the parameter we are testing, ranging over all possible values in the corresponding parameter space.
- The sample size can be selected to create a test with some desired significance and power (at a given parameter value).

$1-\beta = P(\text{Rej } H_0 \mid \underbrace{H_1 \text{ true}})$

function of unknown parameter

composite hypothesis ← it has multiple values for the unknown parameter

## Example

Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from the $N(\mu, 100)$ distribution.

We wish to test

$$H_0 : \mu = 60 \qquad \text{against} \qquad H_1 : \mu > 60$$

using a test of the form

$$\boxed{\textit{Reject } H_0 \textit{ if and only if } \bar{X} > c}$$

??

Determine the value of $c$ and the sample size $n$ so that $\alpha = 0.025$ and, when $\mu = 65$, $\beta = 0.05$ .

to be determined

c specifies critical region

$\alpha = P(\text{Rej } H_0 \mid \underbrace{\mu = 60}_{H_0 \text{ true}})$

$\beta = 0.05 \implies \underbrace{1 - \beta = 0.95 = P(\text{Rej } H_0 \mid \mu = 65)}_{\text{power of test when } \mu = 65}$

→ two equations for c and n

$0.025 = \alpha = P(\text{Rej } H_0 \mid H_0 \text{ true})$

$X_1, \dots, X_n \sim N(\mu, 100)$

$= P(\bar{X} > c \mid \mu = 60)$

$\Rightarrow \bar{X}_n \sim N\left(\mu, \frac{100}{n}\right)$

$Z$

$= P\left( \frac{\bar{X} - 60}{10/\sqrt{n}} > \frac{c-60}{10/\sqrt{n}} \right)$

$\Rightarrow Z = \frac{\bar{X}_n - \mu}{10/\sqrt{n}} \sim N(0,1)$

$= 1 - P\left( Z < \frac{c-60}{10/\sqrt{n}} \right) = 1 - \phi\left( \frac{c-60}{10/\sqrt{n}} \right)$

$\Rightarrow \phi\left( \frac{c-60}{10/\sqrt{n}} \right) = 1 - 0.025 = 0.975 \underset{\text{table}}{\Longrightarrow} \frac{c-60}{10/\sqrt{n}} = 1.96$

Use info on power:

$$0.95 = 1 - \beta = P(\text{Rej } H_0 \mid \mu = 65)$$


$$= P(\bar{X} > c \mid \mu = 65)$$

$Z \sim N(0,1)$

$$= P\left( \frac{\bar{X} - 65}{10/\sqrt{n}} > \frac{c - 65}{10/\sqrt{n}} \right)$$

$$= 1 - P\left( Z \leq \frac{c-65}{10/\sqrt{n}} \right) = 1 - \phi\left( \frac{c-65}{10/\sqrt{n}} \right)$$

$$\Rightarrow \phi\left( \frac{c-65}{10/\sqrt{n}} \right) = 0.05 \underset{\text{table}}{\Longrightarrow} \frac{c-65}{10/\sqrt{n}} = -1.645$$

CONCLUSION:

$$\frac{c-60}{10/\sqrt{n}} = 1.96$$

system of two eqs in two unknowns.

$$\frac{c-65}{10/\sqrt{n}} = -1.645$$

Solve for $\sqrt{n}$ first $\Longrightarrow$ $\sqrt{n} = 7.21$ $\Longrightarrow$ $n = \underline{\underline{51.98}}$

$n = 52$ (so that $n$ is integer)

Plug back in $n=52$ into one of the equations to get $c \approx 62.718$