Statistics MM5: Hypothesis testing 2

Lecturer: Israel Leyva-Mayorga

email: ilm@es.aau.dk

AALBORG UNIVERSITY
DENMARK

Connectivity

Schedule

1. Introduction to statistics
2. Parameter estimation
3. Confidence intervals
4. Hypothesis testing 1
5. **Hypothesis testing 2**
6. Regression
7. Workshop: wrap-up and exam problems

Outline

Recap on hypothesis testing

- Tests for the mean with known variance
- Type II error probabilities

Tests with one normally distributed population and unknown variance

- Two-sided test for the mean
- One-sided test for the mean

Tests for the difference of mean of two normal populations

Tests with Bernoulli RVs

Recap on hypothesis testing

Types of tests based on the populations

Parameter testing with 1 population

There is some idea about the value of a parameter

Is that idea correct?

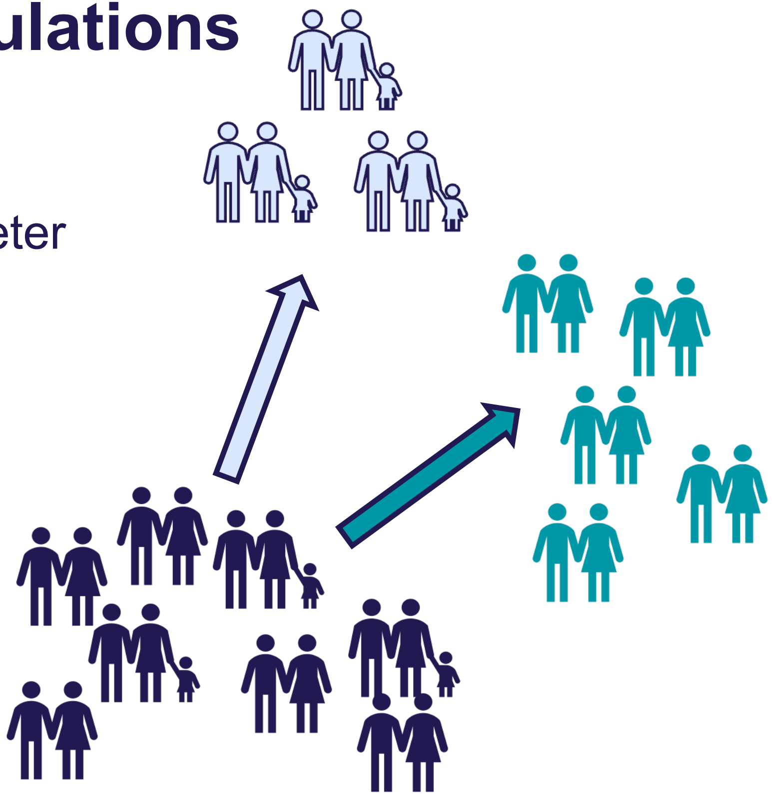
Example: Is it true that the average age is 20?

Compare 2 populations with each other

No parameter known a priori

- Begin with different populations
- One initial population divided into 2

Can we find differences between populations?



How do we conduct the tests?

We begin with two opposing hypotheses:

H_0: The null hypothesis, the one assumed to be true

H_1: The alternative hypothesis, which contradicts H_0

We try to find evidence to support H_0

If we cannot, then we say we can **reject** H_0

Accepting a hypothesis does not mean it is true, but that the data support it

Types of tests with a single population

One-sided tests

The null hypothesis is that the true value lies in an interval with one finite limit

$\mathbf{H_0}$: $\theta \in \Theta$ where $\Theta = (-\infty, b]$ or $\Theta = [a, \infty)$

The same as $\mathbf{H_0}$: $\theta \leq b$ or $\theta \geq a$

Two-sided tests

The null hypothesis is that the true value lies in an interval with finite limits

$\mathbf{H_0}$: $\theta \in \Theta = [a, b]$

The same as $\mathbf{H_0}$: $a \leq \theta \leq b$

There is a special case: If Θ has a single element $\theta = a = b$

Procedure for testing

1. Choose a parameter for testing θ
2. Formulate the null hypothesis H_0 and the alternative hypothesis H_1 about θ
3. Design the test and define the rejection region $R = \{x: T(x) > c\}$
4. Collect a sample $X_1, X_2, X_3, \dots, X_n$ of i.i.d. RVs
5. If the observation regarding θ is close enough to the value(s) in H_0

   H_0 cannot be rejected
6. Else:

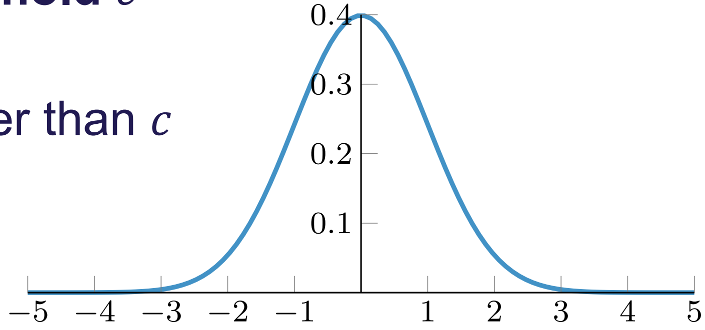
   Reject H_0 and accept H_1

Two ways of testing

1. Comparing the test statistic T with the threshold c

These values are points in the x-axis

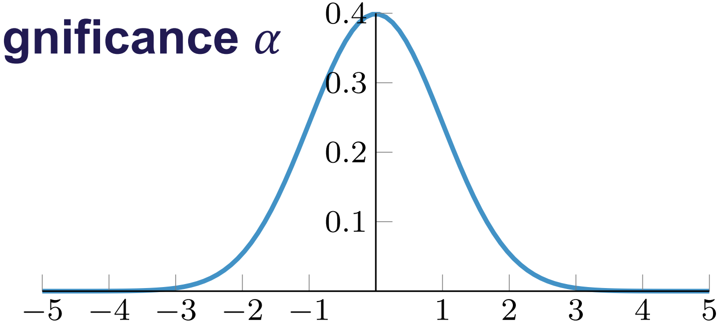
Reject the null hypothesis if T is farther from center than c



2. Comparing the p-value v with the level of significance α

These are the areas under the curve

Reject the null hypothesis if $v < \alpha$



Test outcomes and errors

Errors can occur even when we follow a proven methodology

If a test statistic is in the rejection region R, **H_0** is rejected

	Accept	Reject
H_0 is True	**All good!**	**Type I error**
H_0 is False	**Type II error**	**All good!**

Significance level α

The probability of Type I error should not exceed α

$$P(\text{Type I error}) = \alpha$$

We don't want to reject **H_0** when it's true

Tests for the mean:
Normal population with known variance

Two-sided test for μ with normal population and known σ

The null hypothesis H$_0$: $\mu = \mu_0$, so **H$_1$:** $\mu \neq \mu_0$

Test statistic: $T = |\hat{\mu}_n - \mu_0|\sqrt{n}/\sigma$

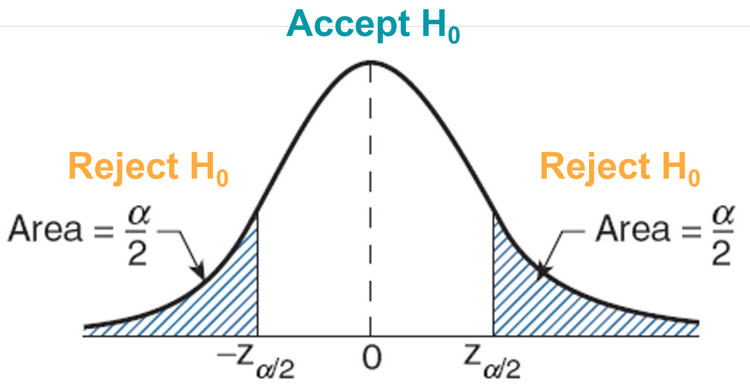
Rejection region: $R = \{X_1, X_2, \dots, X_n : T > c\}$

The total area of the region outside the CI is α

$$C_{1-\alpha} = \left(\hat{\mu}_n - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \hat{\mu}_n + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

So $c = Z_{\alpha/2}$

Reject H$_0$ if $\frac{|\hat{\mu}_n - \mu_0|\sqrt{n}}{\sigma} > Z_{\alpha/2}$

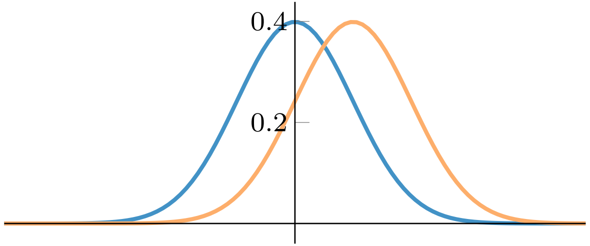


Type II error for two-sided test

The probability of Type II error is the probability of not rejecting H_0 given its False

$$P(\text{Type II error}) = P(T < c \mid \text{H}_0 \text{ is False}) = \beta$$

For our case, H_0 is false if $\mu \neq \mu_0$ and we can calculate



$$\beta(\mu) = P\left(\frac{|\hat{\mu}_n - \mu_0|\sqrt{n}}{\sigma} < Z_{\alpha/2} \,\middle|\, \text{H}_0 \text{ is False} \right)$$

$$= P\left(\mu_0 - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \hat{\mu}_n \leq \mu_0 + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \,\middle|\, \text{H}_0 \text{ is False} \right)$$

$$= \Phi\left(\frac{(\mu_0 - \mu)\sqrt{n}}{\sigma} + Z_{\alpha/2} \right) - \Phi\left(\frac{(\mu_0 - \mu)\sqrt{n}}{\sigma} - Z_{\alpha/2} \right)$$

One-sided test for μ with normal population and known σ

The null hypothesis H$_0$: $\mu \leq \mu_0$, so **H$_1$:** $\mu > \mu_0$

Test statistic: $T = (\hat{\mu}_n - \mu_0)\sqrt{n}/\sigma$

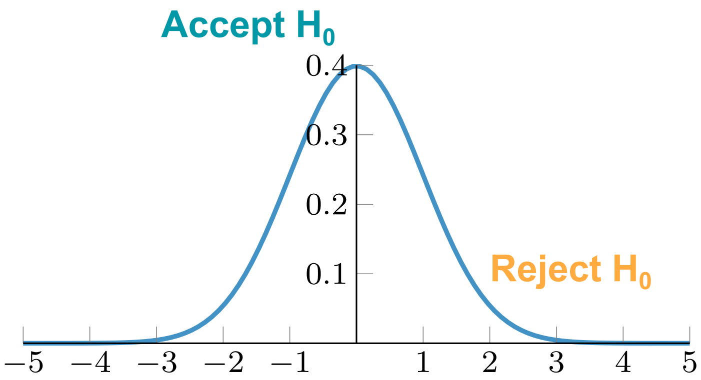
Rejection region: $R = \{X_1, X_2, \dots, X_n : T > c\}$

The total area of the region outside the CI is α

$$C_{1-\alpha} = \left(-\infty, \hat{\mu}_n + Z_\alpha \frac{\sigma}{\sqrt{n}}\right)$$

So $c = Z_\alpha$

Reject H$_0$ if $\frac{(\hat{\mu}_n - \mu_0)\sqrt{n}}{\sigma} > Z_\alpha$



Type II error for one-sided test

The probability of Type II error is the probability of not rejecting H_0 given its False

$$P(\text{Type II error}) = P(T < c \mid \mathrm{H}_0 \text{ is False}) = \beta$$

For our case, H_0 is false if $\mu > \mu_0$ and we can calculate

$$\beta(\mu) = P\left(\hat{\mu}_n - \mu_0 > Z_\alpha \frac{\sigma}{\sqrt{n}} \,\middle|\, \mathrm{H}_0 \text{ is False}\right)$$

$$= \Phi\left(\frac{(\mu_0 - \mu)\sqrt{n}}{\sigma} + Z_\alpha\right)$$

One-sided test for μ with normal population and known σ

The null hypothesis is H$_0$: $\mu \geq \mu_0$, so **H$_1$:** $\mu < \mu_0$ — **Changed!**

Test statistic: $T = (\hat{\mu}_n - \mu_0)\sqrt{n}/\sigma$ — **Same as before**

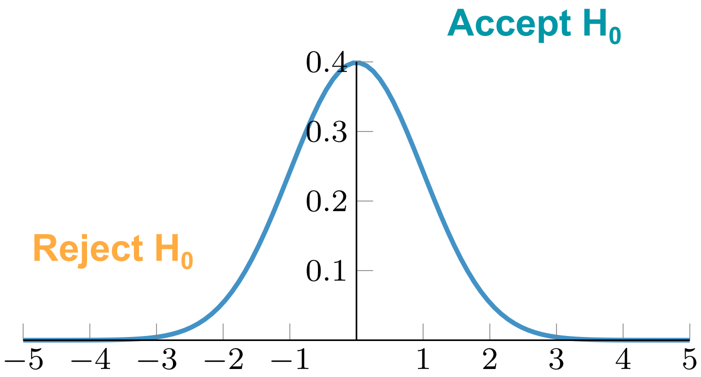
Rejection region: $R = \{X_1, X_2, \dots, X_n : T < c\}$ — **Changed!**

The total area of the region outside the CI is α

$$C_{1-\alpha} = \left(\hat{\mu}_n - Z_\alpha \frac{\sigma}{\sqrt{n}}, \infty\right)$$

So $c = -Z_\alpha$

Reject H$_0$ if $\frac{(\hat{\mu}_n - \mu_0)\sqrt{n}}{\sigma} < -Z_\alpha$



Type II error for one-sided test

The probability of Type II error is the probability of not rejecting H_0 given its False

$$P(\text{Type II error}) = P(T < c \mid \text{H}_0 \text{ is False}) = \beta$$

For our case, H_0 is false if $\mu > \mu_0$ and we can calculate

$$\beta(\mu) = P\left(\hat{\mu}_n - \mu_0 < -Z_\alpha \frac{\sigma}{\sqrt{n}} \middle| \text{H}_0 \text{ is False}\right)$$

$$= \Phi\left(\frac{(\mu_0 - \mu)\sqrt{n}}{\sigma} - Z_{\alpha/2}\right)$$

Summary table

Testing for the mean of one normal population with known variance σ

H_0	H_1	Test statistic T	Rejection region	P-value
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert\hat{\mu}_n - \mu_0\rvert\sqrt{n}/\sigma$	$T > Z_{\alpha/2}$	$2(1 - \Phi(T))$
$\mu \leq \mu_0$	$\mu > \mu_0$	$(\hat{\mu}_n - \mu_0)\sqrt{n}/\sigma$	$T > Z_{\alpha}$	$(1 - \Phi(T))$
$\mu \geq \mu_0$	$\mu < \mu_0$	$(\hat{\mu}_n - \mu_0)\sqrt{n}/\sigma$	$T < -Z_{\alpha}$	$\Phi(T)$

Example

It is known that if a signal of value μ is sent from location A, then the received value at location B is normally distributed with mean μ and $\sigma = 2$.

This means that Gaussian noise that is added to the signal is a RV with distribution $N(0,4)$.

The signal value is sent 5 times and the sample average at location B is $\bar{X}_5 = 9.5$.

Can we reject H_0 using a 5% level of significance for the following tests?

a) We define H_0: $\mu \leq 8$

b) We define H_0: $\mu \geq 8$

We have $\hat{\mu}_5 = \bar{X}_5 = 9.5$ and $Z_\alpha = 1.645$

Example

H$_0$: $\mu \leq 8$

Our test statistic is

$$T = \frac{(\hat{\mu}_n - \mu_0)\sqrt{n}}{\sigma} = 1.677$$

Rejection region is

$$R = \{x : T(x) > Z_\alpha\}$$

$$T = 1.677 > 1.645$$

We reject H$_0$

H$_0$: $\mu \geq 8$

Our test statistic is

$$T = \frac{(\hat{\mu}_n - \mu_0)\sqrt{n}}{\sigma} = 1.677$$

Rejection region is

$$R = \{x : T(x) < -Z_\alpha\}$$

$$T = 1.677 > -1.645$$

We cannot reject H$_0$

Tests for the mean:
Normal population with unknown variance

The t-test

In the previous examples, we knew the variance σ^2

We can use the standard normal distribution $N(0,1)$ and its quantiles Z_α and $Z_{\alpha/2}$

If σ^2 is unknown, it must be estimated we cannot use $N(0,1)$ anymore

Now, we estimate σ^2 using our unbiased estimator for the sample variance

$$\hat{\sigma}_n^2 = S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

And define the **test statistic** based on S_n^2

T-test:

Since we are not using σ^2, we now have to use the **t-distribution** for testing

Two-sided test for μ with normal population and unknown σ

The null hypothesis H₀: $\mu = \mu_0$, so **H₁:** $\mu \neq \mu_0$

Test statistic: $T = |\hat{\mu}_n - \mu_0|\sqrt{n}/S_n$

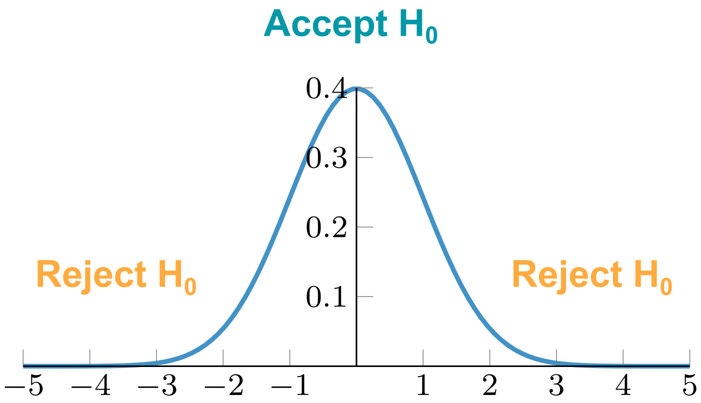
Rejection region: $R = \{X_1, X_2, \ldots, X_n : T > c\}$

The total area of the region outside the CI is α

$$C_{1-\alpha} = \left(\hat{\mu}_n - T_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}, \hat{\mu}_n + T_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}\right)$$

So $c = T_{n-1,\alpha/2}$

Reject H_0 if $\frac{|\hat{\mu}_n - \mu_0|\sqrt{n}}{S_n} > T_{n-1,\alpha/2}$



Example

A group of 50 patients with high cholesterol levels were given a new drug. The reduction in levels of cholesterol were recorded.

The reduction of cholesterol levels had a sample mean of 14.8 with sample standard deviation of 6.4. Was the drug effective with 5% significance?

1. Formulate the null and alternative hypothesis

2. Calculate the test statistic

3. Can you reject the null hypothesis?

Example

A group of 50 patients with high cholesterol levels were given a new drug. The reduction in levels of cholesterol were recorded.

The reduction of cholesterol levels had a sample mean of 14.8 with sample standard deviation of 6.4. Was the drug effective with 5% significance?

1. Formulate the null and alternative hypothesis

$$H_0\colon \mu = 0 \text{ and } H_1\colon \mu \neq 0$$

2. Calculate the test statistic

$$T = |\hat{\mu}_n - \mu_0|\sqrt{n}/S_n = |14.8 - 0|\sqrt{50}/6.4 =$$

3. Can you reject the null hypothesis?

Reject if $T > T_{n-1,\alpha/2}$. Since $T = 16.35 > 2.009$, **we reject H_0**

One-sided t-test for μ with normal population and unknown σ

The null hypothesis H$_0$: $\mu \leq \mu_0$
H$_1$: $\mu > \mu_0$
Test statistic: $T = (\hat{\mu}_n - \mu_0)\sqrt{n}/S_n$
Rejection region:

$$R = \{X_1, X_2, \dots, X_n : T > T_{n-1,\alpha}\}$$

CI is $C_{1-\alpha} = \left(-\infty, \hat{\mu}_n + T_{n-1,\alpha}\frac{S_n}{\sqrt{n}}\right)$

Reject H$_0$ if $\frac{(\hat{\mu}_n - \mu_0)\sqrt{n}}{S_n} > T_{n-1,\alpha}$

The null hypothesis H$_0$: $\mu \geq \mu_0$
H$_1$: $\mu < \mu_0$
Test statistic: $T = (\hat{\mu}_n - \mu_0)\sqrt{n}/S_n$
Rejection region:

$$R = \{X_1, X_2, \dots, X_n : T < -T_{n-1,\alpha}\}$$

CI is $C_{1-\alpha} = \left(\hat{\mu}_n - T_{n-1,\alpha}\frac{S_n}{\sqrt{n}}, \infty\right)$

Reject H$_0$ if $\frac{(\hat{\mu}_n - \mu_0)\sqrt{n}}{S_n} < -T_{n-1,\alpha}$

Summary table

Testing for the mean of one normal population with **unknown** variance

H_0	H_1	Test statistic T	Rejection region	P-value
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert\hat{\mu}_n - \mu_0\rvert\sqrt{n}/S_n$	$T > T_{n-1,\alpha/2}$	$2(1 - P(T_{n-1} \leq T))$
$\mu \leq \mu_0$	$\mu > \mu_0$	$(\hat{\mu}_n - \mu_0)\sqrt{n}/S_n$	$T > T_{n-1,\alpha}$	$(1 - P(T_{n-1} \leq T))$
$\mu \geq \mu_0$	$\mu < \mu_0$	$(\hat{\mu}_n - \mu_0)\sqrt{n}/S_n$	$T < -T_{n-1,\alpha}$	$P(T_{n-1} \leq T)$

To calculate $P(T_{n-1} \leq T)$ where T_{n-1} : RV with t-distribution and parameter n-1

Matlab: tcdf(T,n-1) **Python:** X.cdf(T) where X is the appropriate RV

Tests with two normal populations: Testing the equality of means

Two-sided test with known variance

We want to test $\bar{X} - \bar{Y}$ where both RVs have normal distribution

Null hypothesis is H$_0$: $\mu_X = \mu_Y$ and so $\mu_X - \mu_Y = 0$

We calculate $\mu_{\bar{X}-\bar{Y}} = \mu_X - \mu_Y$ $\mathrm{var}(\bar{X} - \bar{Y}) = \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}$

Test statistic: $T = \frac{|\hat{\mu}_X - \hat{\mu}_Y|}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}}$

Rejection region: $R = \{X_1, X_2, \dots, X_n : T > Z_{\alpha/2}\}$

Reject H$_0$ if $\frac{|\hat{\mu}_X - \hat{\mu}_Y|}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} > Z_{\alpha/2}$

Two-sided test with unknown variance

Null hypothesis is H$_0$: $\mu_X = \mu_Y$

We calculate $\mu_{\bar{X}-\bar{Y}} = \mu_X - \mu_Y$

Non-biased estimator for the variance is $S^2_{\bar{X}-\bar{Y}} = \frac{\sum_{i=1}^{n_X}(X_i-\bar{X})^2 + \sum_{j=1}^{n_Y}(Y_i-\bar{Y})^2}{n_X+n_Y-2}$

Test statistic: $T = \frac{|\hat{\mu}_X-\hat{\mu}_Y|}{S_{\bar{X}-\bar{Y}}\sqrt{\frac{1}{n_X}+\frac{1}{n_Y}}}$

Rejection region: $R = \left\{X_1, X_2, \dots, X_n : T > T_{n_X+n_Y-2,\alpha/2}\right\}$

Reject H$_0$ if $\frac{|\hat{\mu}_X-\hat{\mu}_Y|}{S_{\bar{X}-\bar{Y}}\sqrt{\frac{1}{n_X}+\frac{1}{n_Y}}} > T_{n_X+n_Y-2,\alpha/2}$

Example

Twenty two volunteers at a cold research institute caught a cold.

A random selection of 10 of these were given tablets with vitamin C

Others were given a placebo and the time that the cold lasted was recorded

Did the vitamin C have an effect on the mean length of the cold?

At what level of significance?

Treated with vitamin C		Given Placebo	
5.5	6.0	6.5	7.5
6.0	7.5	6.0	6.5
7.0	5.5	8.5	7.5
6.0	7.0	7.0	6.0
7.5	6.5	6.5	8.5
		8.0	7.0

Example

We calculate

$$\hat{\mu}_X = 6.450 \text{ and } \hat{\mu}_Y = 7.125$$

$$S^2_{\bar{X}-\bar{Y}} = 0.689$$

$$T = \frac{|\hat{\mu}_X - \hat{\mu}_Y|}{S_{\bar{X}-\bar{Y}}\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}} = 1.899$$

Since $T = 1.899 < T_{n_X+n_Y-2,\alpha/2} = 2.0859$, we cannot reject H_0 with $\alpha = 0.05$

The p-value is $v = 0.0721$, so we cannot reject H_0 for all $\alpha < 0.0721$

Hypothesis testing for Bernoulli population

Bernoulli RVs

Bernoulli RVs have only two possible outcomes: 1 or 0

$$P(X_i = 1) = p \text{ and } P(X_i = 0) = 1 - p$$

The sum of $X_1, X_2, \ldots, X_n \sim \text{Bernoulli}(p)$ RVs is a binomial RV with parameters n, p:

$$\sum_{i=1}^{n} X_i \sim B(n, p)$$

where $P(X = k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$ and $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

We also know that the MLE for p is

$$\hat{p}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

One-side tests with p-value for Bernoulli RVs

We formulate a null hypothesis H_0: $p \le p_0$

We begin with an observation where k out of the n outcomes are 1, so $X = k$

Recall: The p-value is the **probability** (under H_0) of observing a value of the test statistic the same as or more extreme than what was actually observed

Hence, we calculate the p-value in this case as

$$v = P(X \ge k; n, p_0) = \sum_{i=k}^{n} \binom{n}{i} p_0^i (1 - p_0)^{n-i}$$

Reject H_0 if $v < \alpha$

Example

A chip manufacturer claims that no more than 2% of its chips are defective.

A company has purchased 300 of these chips.

If 10 of these chips are defective, can the claim be rejected with 5% significance?

Solution

H_0: $p \leq p_0 = 0.02$, $n = 300$, and $k = 10$

We calculate the p-value as

$$v = \sum_{i=10}^{300} \binom{300}{i} (0.02)^i (1 - 0.02)^{300-i} = 0.0818 > 0.05$$

Since $v = 0.0818 > 0.05$, **we cannot reject $\mathbf{H_0}$**

Solution with normal approximation

Central Limit Theorem:

For $X_1, X_2, \ldots, X_n$ be i.i.d. RVs with mean μ and variance σ^2 if n is large:

$$\sum_{i=1}^{n} X_i \approx N(n\mu, n\sigma^2)$$

Hence, if n is large and the RVs are Bernoulli, we can use the z-test with

$$T = \frac{k - np_0}{\sqrt{np_0(1 - p_0)}}$$

For the example $T = 1.1605 < Z_{0.05} = 1.6448$, so **we cannot reject H_0**

Finally, the coin toss example from the first lecture

We start with the null hypothesis H_0: $p = 1/2$

Toss the coin a large number of times n and count the number of heads k

Generate an estimate for p as

$$\hat{p}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{k}{n}$$

We define our **test statistic** as $T = |\hat{p}_n - p_0|/\sqrt{p_0(1-p_0)/n}$ where $p_0 = 1/2$

If $T \leq Z_{\alpha/2}$, **we cannot reject** H_0

Rejection region: $R = \{\hat{p}_n : T > Z_{\alpha/2}\}$

It is concluded with α significance that the coin is not fair if $|\hat{p}_n - p_0| > Z_{\alpha/2}/2\sqrt{n}$

Summary

Summary

The type of test to use depends on:

Number of populations involved, types of hypothesis, and assumptions

Z-test for normal population and the variance is **known**

T-test for normal population and the variance is **unknown**

We need to adapt the estimator for the variance when we have two populations

Hypothesis testing for Bernoulli distributions relies on the p-value

But, if the sample size is large, we can use the **Central Limit Theorem**