Module 3b. Hypothesis testing

Data Science & AI

Sabine De Vreese   Lieven Smits   Bert Van Vreckem

2023–2024

HO
GENT

Contents

HO
GENT

Learning Goals

- Statistical hypothesis testing concepts
- Hypothesis testing procedure
- Apply the *z*- and *t*-test

HO
GENT

Testing procedure

Statistical Hypothesis Testing

Hypothesis Idea that has yet to be proven: statement regarding the numeric value of a population parameter

Hypothesis Test verification of a statement about the values of one or multiple population parameters

Null Hypothesis (H_0) Base hypothesis, we start with assuming it is true

Alternative Hypothesis (H_1, H_a) Conclusion if the null hypothesis is unlikely to be true

HO GENT

Elements of a testing procedure

Test Statistic The value that is calculated from the sample

Region of Acceptance The region of values supporting the null hypothesis

Critical Region / Region of Rejection The region of values rejecting the null hypothesis

Significance Level The probability of rejecting a true null hypothesis H_0

HO GENT

Testing procedure

1. Formulate both hypotheses (H_0 and H_1)
2. Determine the significance level (α)
3. Calculate the test statistic
4. Determine the critical region or the probability value
5. Draw conclusions

HO
GENT

Hypotheses about superheroes

A superhero rescues 3.3 persons a day


敵は階段 駅の杉並戦隊

ベビーカーおろすんジャー

Source: `http://www.cracked.com/quick-fixes/4-people-who-saved-day-while-dressed-as-superheroes/`

Assume that, over a period of 30 days, on average 3.483 people were saved each day ($\overline{x} = 3.483$, $n = 30$)

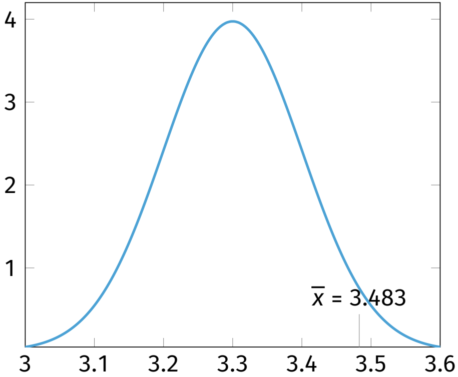
1. Hypothesis: $H_0 : \mu = 3.3$; $H_1 : \mu > 3.3$
2. Significance level: $\alpha = 0.05$
3. Test statistic: $\overline{x} = 3.483$

Standard deviation of the population (assumed to be known): $\sigma = 0.55$

HO GENT

Calculate test statistic

Based on the central limit theorem, we know that:
$M \sim Nor(\mu = 3.3; \sigma = \frac{0.55}{\sqrt{30}} = 0.1)$



Probability Value

HO
GENT

Probability Value

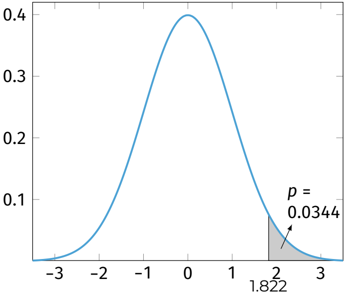
p-value

The p-value is the probability, if the null hypothesis is true, to obtain a value for the test statistic that is at least as extreme as the observed value.

- p-value $< \alpha \Rightarrow$ reject H_0: the discovered value of $\overline{x}$ is too extreme;
- p-value $\geq \alpha \Rightarrow$ do not reject H_0: the discovered value of $\overline{x}$ can still be explained by coincidence.

HO
GENT

Probability Value



$$P(M > 3.483) = P\left(Z > \frac{3.483 - 3.3}{\frac{\sigma}{\sqrt{n}}}\right) = P(Z > 1.822) = 0.0344$$

HO GENT

Critical Region

HO
GENT

Critical Region

Critical region

The critical region is the collection of all values of the test statistic for which we can reject the null hypothesis.

We look for a critical value g for which:

$$P(M > g) = \alpha$$

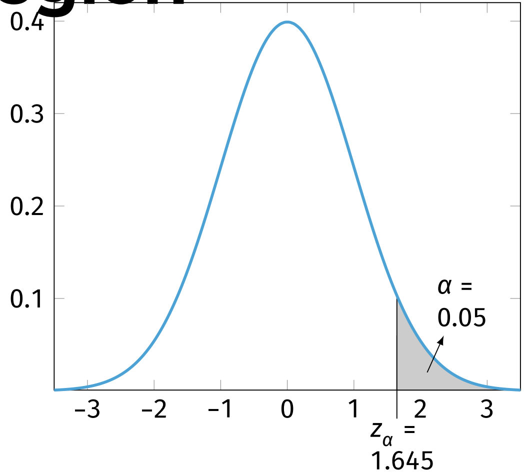
Determine z_α for which:

$$P(Z > z_\alpha) = \alpha \Rightarrow g = \mu + z_\alpha . \frac{\sigma}{\sqrt{n}}$$

- Left of g: region of acceptance (do not reject H_0)
- Right of g: critical region (reject H_0)

HO GENT

Critical Region



significance level α = 0.05 ⇒ critical value z_α=1.645
(z_α =`stats.norm.isf(1 - 0.95)`)

Left-tailed testing

What would you have to change in the equation in order to calculate the correct critical value?

Left-tailed testing

What would you have to change in the equation in order to calculate the correct critical value? Answer:

$$g = \mu - z \times \frac{\sigma}{\sqrt{n}}$$

because

$$P(M < g) = P\left(Z < \frac{g - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = 0.05$$

HO
GENT

Left-tailed testing

Because of symmetry:

$$P\left(Z > -\left(\frac{g-\mu}{\frac{\sigma}{\sqrt{n}}}\right)\right) = 0.05$$

The corresponding *z*-value is 1.645, and therefore:

$$z = \frac{-g+\mu}{\frac{\sigma}{\sqrt{n}}} \Leftrightarrow -g = \frac{\sigma}{\sqrt{n}} z_\alpha - \mu$$

$$\Leftrightarrow g = \mu - z_\alpha \frac{\sigma}{\sqrt{n}}$$

Two-tailed testing

Sometimes it can be necessary to perform a two-tailed test. In this case, two critical values need to be calculated, namely the left and right critical value.

$$g = \mu \pm z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}} \tag{1}$$

HO GENT

Summary

Goal	Test regarding the value of the population mean μ using a sample of n independent values		
Prerequisite	De population has a random distribution, n is sufficiently large		
Test Type	Two-tailed	Left-tailed	Right-tailed
H_0	$\mu = \mu_0$	$\mu = \mu_0$	$\mu = \mu_0$
H_1	$\mu \neq \mu_0$	$\mu < \mu_0$	$\mu > \mu_0$
Critical Region	$\lvert\overline{x}\rvert > g$	$\overline{x} < -g$	$\overline{x} > g$
Test statistic	$z = \frac{\overline{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$		

Table: Summary of Testing Procedures

Requirements for z-test

- The sample needs to be random
- The sample size needs to be sufficiently large ($n \geq 30$)
- The test statistic needs to have a normal distribution
- The standard deviation of the population, σ, is known

Sometimes these assumptions will not hold and in this case we can *not* use the Z-test!

HO
GENT

Examples

Example 1

When drawing a random sample consisting of 50 observations from a population with variance $\sigma^2 = 55$ we obtain as sample mean $\overline{x} = 25$.
We now want to find out if there is a reason to assume that the population mean is smaller than 27.

HO
GENT

Example 1

1. Formulate both hypotheses
   $H_0 : \mu = 27$ en $H_1 : \mu < 27$.
2. Determine significance level α and sample size n
   $\alpha = 0.05$ en $n = 50$
3. Test statistic & value: sample mean $\overline{x} = 25$

HO GENT

Example 1 (cont.)

4a Probability Value

According to the central limit theorem:

$$M \sim Nor(\mu = 27, \frac{\sigma^2}{\sqrt{n}})$$

$$z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{25 - 27}{\sqrt{\frac{55}{50}}} \approx -1.91$$

We therefore have a probability value of **0.0281**.

Using a significance level of 0.05, we can reject H_0.

HO GENT

Example 1 (cont.)

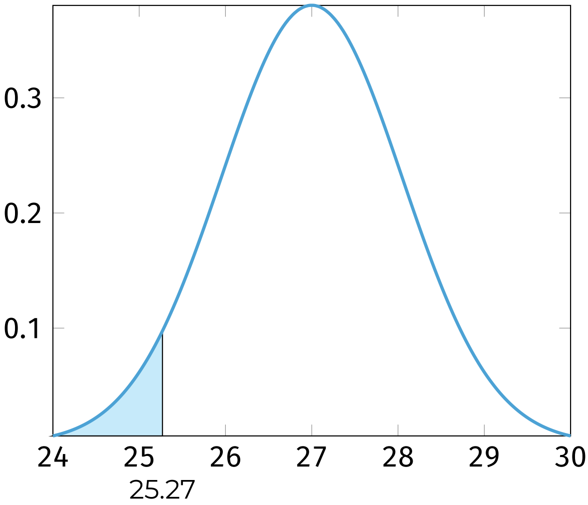
4b Calculate and plot the critical region

$$g = \mu - z \times \frac{\sigma}{\sqrt{n}}$$
$$= 27 - 1.645 \times \sqrt{\frac{\sigma}{n}}$$
$$= 25.27470944$$

We have that $\overline{x} < g$ and therefore we can reject H_0.

HO GENT

Example 1 (cont.)



Example 1 (cont.)

5 Conclusion
We can conclude, based on the sample, that $\mu < 27$ for a significance level of 0.05

HO
GENT

Example 2

In a reseach about the amount of change in the pockets of our superheroes, researchers state that on average a superhero carries €25 of cash. We assume that the standard deviation of the population $\sigma = 7$. For a random sample of size $n = 64$, the average amount of money a superhero carries is $\overline{x} = €23$. For the significance level, $\alpha = 0.05$ is selected.

HO
GENT

Example 2

1. Formulate both hypotheses
   $H_0 : \mu = 25$ en $H_1 : \mu \neq 25$
2. Determine significance level α and sample size n
   $\alpha = 0.05$ en $n = 64$.
3. Test statistic & value: $\overline{x} = 23$

HO GENT

Example 2 (cont.)

4b Calculate the critical region

$$g_1 = \mu - z \times \frac{\sigma}{\sqrt{n}} = 23.28$$

$$g_2 = \mu + z \times \frac{\sigma}{\sqrt{n}} = 26.72$$

We have that $\overline{x}$ is inside the critical region (because 23 < 23.28) so we can reject H_0.

5 Based on this sample we can conclude that superheroes carry on average *less* than 25 €, using a significance level of 5%

HO GENT

Student’s t-test

HO
GENT

Student's t-test

What if the requirements for a z-test are not met? E.g.

- Sample size too small
- Population stdev (σ) unknown

If the variable is normally distributed, you can use the t-test

HO
GENT

The *t*-test

Determine critical value:

$$g = \mu \pm t \times \frac{s}{\sqrt{n}}$$

- t-value is derived from the Student's t-distribution, based on the number of *degrees of freedom*, $n - 1$
- Look for value using the function `t.isf()` in Python
- Apart from this, the procedure is identical to the procedure of the z-test

HO GENT

Errors in Hypothesis Tests

HO
GENT

Errors in Hypothesis Tests

Conclusion	Reality H_0 **True**	 H_1 **True**
H_0 **not rejected**	Correct inference	Type II error (false negative)
H_0 **rejected**	Type I error (false positive)	Correct inference

P(type I error) = α (= significance level)
P(type II error) = β
Calculating β is ***not*** trivial, but if α ↘ then β ↗

HO
GENT