Lecture Notes: Chapter 6

20/21

6 Hypothesis Testing

Goal: Show by an experiment that some measure has an effect.

The statement about the effect has the following form:

- the mean of a new distribution is different ($\neq$, $<$, $>$) from a standard mean μ_0
- the difference between two means μ_1, μ_2 (of the distribution with and without the measure is different from 0 ($\neq$, $<$, $>$)

The default assumption is:

The new measure has no effect (null hypothesis)

We only believe that the measure has an effect if experiment data would have a very low probability (5%, 1%) in case the null hypothesis were to hold.

Technically, the null hypothesis has the form (if it is about the mean):

- $\mu = \mu_0$ $(\mu \leq \mu_0, \mu \geq \mu_0)$, or

- $\mu_1 - \mu_2 = 0$, $(\mu_1 \leq \mu_2, \mu_1 \geq \mu_2)$.

If the null hypothesis lets us conclude that the outcome of our experiments is unlikely, we adopt the alternative hypothesis:

- $\mu \neq \mu_0$ $(\mu > \mu_0, \mu < \mu_0)$,

- $\mu_1 - \mu_2 \neq 0$, $(\mu_1 > \mu_2, \mu_1 < \mu_2)$.

Concretely,

- we fix a low probability level (called "significance level") α, e.g. $\alpha = 5\%$, $\alpha = 1\%$ etc.

- we take an independent sample of a random variable X of size n and with average $\overline{x}$ *)

Suppose, our null hypothesis is:

- The mean of the distribution we measure is μ_0, symbolically

$$H_0 : \quad \mu = \mu_0$$

- Our measurement of X has average $\overline{x}$, which is different from μ by $|\mu_0 - \overline{x}|$.

---

*) We distinguish now between the outcome of a specific measurement, which leads to a number $\overline{x} \in \mathbb{R}$, and the approach of taking measurements and averages in general, modeled by the RVs X_i and $\overline{X}$.

So, $H_0 : \mu = \mu_0$ and found $|\mu_0 - \bar{x}|$

We accept H_0 if the probability to see a difference of size $|\mu_0 - x_0|$ is $\geq \alpha$, under the condition that the mean of the current distribution μ equals μ_0:

$$P\left[\, |\bar{X} - \mu_0| \geq |\bar{x} - \mu_0| \;\middle|\; \mu = \mu_0 \right] \geq \alpha .$$

We reject H_0 and accept the alternative hypothesis

$$H_a : \mu \neq \mu_0$$

if this probability is $< \alpha$:

$$P\left[\, |\bar{X} - \mu_0| \geq |\bar{x} - \mu_0| \;\middle|\; \mu = \mu_0 \right] < \alpha .$$

If the RV $\mathcal{X}$ is normally distributed, and we know the variance σ^2, we can effectively perform the test of H_0:

$$P[\,|\mathcal{X}-\mu_0| \geq |\bar{x}-\mu_0| \mid \mu=\mu_0\,]$$

normalization to $\mathcal{N}(0,1)$

$$= P\left[\sqrt{n}\,\frac{|\bar{\mathcal{X}}-\mu_0|}{\sigma} \geq \sqrt{n}\,\frac{|\bar{x}-\mu_0|}{\sigma} \mid \mu=\mu_0\right]$$

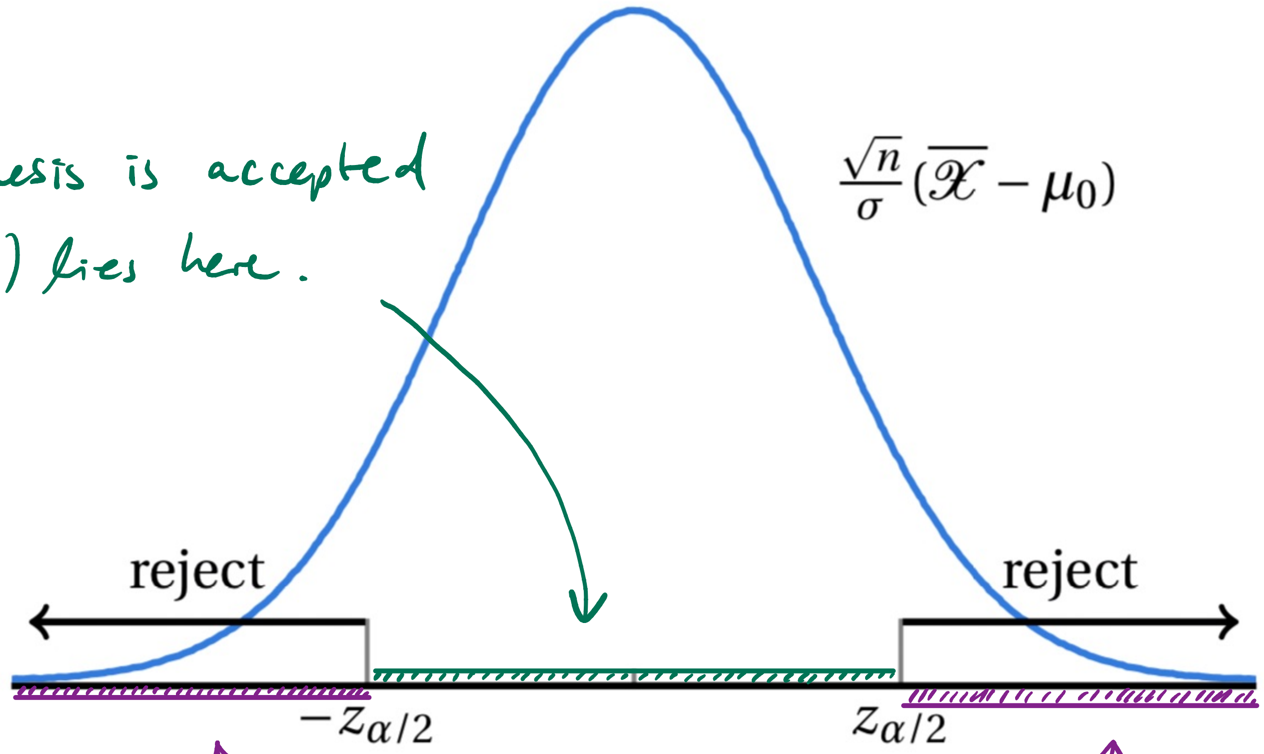
$$= P\left[\,|Z| \geq \sqrt{n}\,\frac{|\bar{x}-\mu_0|}{\sigma} \mid \mu=\mu_0\right] \geq \alpha$$

This inequality holds iff

$$\sqrt{n}\,\frac{|\bar{x}-\mu_0|}{\sigma} \leq z_{\alpha/2}$$

We call $v = \sqrt{n}\,\frac{|\bar{x}-\mu_0|}{\sigma}$ the test statistic for this test

Null hypothesis is accepted if $\frac{\sqrt{n}}{\sigma}(\bar{X}-\mu_0)$ lies here.



Null hypothesis is rejected if $\frac{\sqrt{n}}{\sigma}(\bar{X}-\mu_0)$ lies here.

Example 76: Suppose the normal RV X has variance $\sigma^2 = 4$.

We want to test wheter X has mean $\mu = 8$ (i.e. $\mu_0 = 8$).

We take a sample of size $n = 9$ and obtain $\bar{X} = 9.2$

The required significance level is $\alpha = 0.05$ $(= 5\%)$.

The corresponding z-value is

$$z_{0.025} = 1.98.$$

Example 76: Suppose the normal RV X has variance $\sigma^2 = 4$.

We want to test wheter X has mean $\mu = 8$ (i.e. $\mu_0 = 8$).

We take a sample of size $n = 9$ and obtain $\bar{X} = 9.2$

The required significance level is $\alpha = 0.05$ ($= 5\%$).

The corresponding z-value is

$$z_{0.025} = 1.98.$$

We compute the test statistic

$$v = \sqrt{4}\,\frac{|\bar{X} - \mu_0|}{\sigma} = 3\,\frac{9.2 - 8}{2} = \frac{3}{2} \cdot 1.2 = 1.8.$$

Since $v < z_{0.025}$, we accept the null hypothesis.

p-Values

Suppose we have the test statistic $\sqrt{n}\,\frac{|\bar{X}-\mu_0|}{\sigma}$.

We can compute / look up the probability of a result at least as extreme as our observation as follows:

$$p_v = P[\,|Z| \geq v\,] = 2 \cdot P[\,Z \geq v\,]$$

$$= 2 \cdot (1 - P[\,Z \leq v\,]) = 2\,(1 - \Phi(v))$$

We call this probability the p-value of our experiment.

We accept H_0 if $p_v \geq \alpha$, other we reject H_0 and accept the alternative hypothesis H_a.

Example 77: In Example 76, we have computed the test statistic as

$$v = \frac{\sqrt{n}}{\sigma} |\bar{x} - \mu_0| = 1.8.$$

Then the p-value is

$$p_v = P[\,|Z| > v\,]$$

Example 77: In Example 76, we have computed the test statistic as

$$v = \frac{\sqrt{n}}{\sigma} |\bar{x} - \mu_0| = 1.8.$$

Then the p-value is

$$p_v = P[|Z| > v]$$

$$= P[|Z| > 1.8] = 2 P[Z > 1.8]$$

$$= 2 \cdot 0.036 = 0.072$$

This means, the null hypothesis will be rejected for any significance level

$$\alpha > p_v = 0.072 = 7.2\%$$

Two-sided Tests

What we have described right now is a two-sided test:

We reject H_0 if the probability of $\overline{X}$ being away from μ_0 at least as far as $\bar{x}$ is less than α.

Here, far away includes far to the right and far to the left.

Note that the condition for acceptance is equivalent to

$$\bar{x} \in \left(\mu_0 - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} , \mu_0 + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right)$$

The test is called two-sided because it imposes a lower and an upper bound on $\bar{x}$.

Test imposing only an upper or only a lower bound on $\bar{x}$ are called one-sided.

6.1 One-sided Tests

If we want to check whether a new method leads to greater values than the previous (with mean μ_0), our null hypothesis should be

$$H_0: \mu \leq \mu_0$$

that is, the mean of the new method is not greater than the one of the old method.

Given α and the observed average $\bar{x}$, we reject H_0 if the probability of seeing an average $\geq \bar{x}$ is less than α, that is:

$$P[\bar{X} \geq \bar{x} \mid \mu = \mu_0] < \alpha.$$

Test Statistic of a One-Sided Test

How can we check that

$$P[\bar{\mathcal{X}} \geq \bar{x} \mid \mu = \mu_0] < \alpha.$$

By normalization, we obtain

$$P[\bar{\mathcal{X}} \geq \bar{x} \mid \mu = \mu_0] = P\left[\sqrt{n}\,\frac{\bar{\mathcal{X}} - \mu_0}{\sigma} \geq \sqrt{n}\,\frac{\bar{x} - \mu_0}{\sigma} \;\middle|\; \mu = \mu_0\right]$$

$$P\left[Z \geq \sqrt{n}\,\frac{\bar{x} - \mu_0}{\sigma} \;\middle|\; \mu = \mu_0\right]$$

The value

$$v = \sqrt{n}\,\frac{\bar{x} - \mu_0}{\sigma}$$

is the test statistic of this one-sided test.

p-Value of a One-Sided Test

The p-value corresponding to v, that is, the probability of $\bar{X}$ being at least as extreme as $\bar{x}$, is

$$p_v = P[Z \geq v] = 1 - \Phi(v).$$

Again, the null hypothesis is accepted if

$$p_v \geq \alpha$$

which is equivalent to

$$v \leq z_\alpha$$

and rejected if

$$p_v < \alpha,$$

which is equivalent to $v > z_\alpha$.

If the null hypothesis is

$$H_0 : \mu \geq \mu_0$$

and we have $\bar{x}$ and α as before, then we reject H_0 if

$$P[\bar{X} \leq \bar{x} \mid \mu = \mu_0] < \alpha,$$

which is equivalent to

$$P\left[\sqrt{n}\frac{\bar{X}-\mu_0}{\sigma} \leq \sqrt{n}\frac{\bar{x}-\mu_0}{\sigma} \mid \mu=\mu_0\right]$$

$$= P\left[\sqrt{n}\frac{\bar{X}-\mu_0}{\sigma} \leq \sqrt{n}\frac{\bar{x}-\mu_0}{\sigma} \mid \mu=\mu_0\right] < \alpha.$$

Again, $v = \sqrt{n}\frac{\bar{x}-\mu_0}{\sigma}$ is the test statistic. The corresponding p-value is

$$p_v = P[Z \leq v] = \Phi(v).$$

We accept H_0 if $p_v \geq \alpha$ (i.e., $v \geq -z_\alpha$) and accept H_a otherwise.

6.2 Hypothesis Testing with Unknown Variance

The theory is analogous to the one for the case of known variance,

the difference being that instead of

- the standard normal distribution $\mathcal{N}(0,1)$ and
- the variance σ^2

we have to consider

- t_{n-1}, the t-distribution with $n-1$ degrees of freedom
- the estimator S^2

<u>Example 82</u>: A worried neighbor claims that students drink on average 3 litres of beer every night.

To investigate this claim, 25 randomly selected students are observed. The observations yield:

- a sample mean of 2.91 l
- a sample standard deviation of 0.47 l.

How can we verify the claim?

Example 82: A worried neighbor claims that students drink on average 3 litres of beer every night.

To investigate this claim, 25 randomly selected students are observed. The observations yield:

- a sample mean of 2.91 l
- a sample standard deviation of 0.47 l.

How can we verify the claim?

Null hypothesis: $H_0 : \mu = 3$ (i.e., $\mu_0 = 3$)

Test statistic: $v = \sqrt{n} \frac{(\bar{X} - \mu_0)}{S} = 5 \frac{0.09}{0.47} = \frac{0.45}{0.47} = 0.9574$

p-value: $p_v = P[\,|T_{24}| > v\,] = 2\,P[\,T_{24} > v\,] = 0.3479$

The experiment is consistent with the hypothesis

Test Statistics if variance is unknown

- Two-sided test: $$v = \sqrt{n}\,\frac{|\bar{x} - \mu_0|}{S}$$

  Acceptance if $v \leq t_{\alpha/2,\, n-1}$

- One-sided tests: $$v = \sqrt{n}\,\frac{\bar{x} - \mu_0}{S}$$

  Acceptance that $\mu \leq \mu_0$ if

  $$v \geq t_{\alpha,\, n-1}$$

  Acceptance that $\mu \leq \mu_0$ if

  $$v \geq t_{\alpha,\, n-1}$$