Matematisk Statistik och Disktret Matematik, MVE051/MSG810, VT19

Föreläsning 9

Nancy Abdallah

Chalmers - Göteborgs Universitet

Definitions

- The other form of inferential statistics is **hypothesis testing**.
- A **hypothesis** may be defined simply as a statement about one or more populations.
- A hypothesis concerns parameters of the population (mean, variance, etc.)
- With hypothesis testing, one determines whether the hypothesis is compatible with the given data.
- There are two hypotheses:
    - The **alternative** or **research hypothesis** which is the hypothesis proposed by the experimenter. It is denoted by H_1 or H_A.
    - The **null hypothesis** which is the negation of the alternative hypothesis. It is denoted by H_0.

- If θ is the parameter of our concern, the null hypothesis will be of the form $\theta = \theta_0$, $\theta \geq \theta_0$ or $\theta \leq \theta_0$ (the indication of equality must appear in the null hypothesis). In this case, the alternative hypothesis will be $\theta \neq \theta_0$, $\theta < \theta_0$ or $\theta > \theta_0$ respectively.
- The null hypothesis is the hypothesis to be tested using a given set of data.
- What we hope from hypothesis testing is to reject the null hypothesis and therefore accept the alternative hypothesis.
- The null hypothesis is sometimes considered as $\theta = \theta_0$ instead of $\leq, \geq$.

Example

1. "Are the majority of students at Chalmers from Gothenburg?" Let p be the proportion of the students from Gothenburg. Then,

$$H_0: \ p = 0.5 \quad \text{and} \quad H_1: \ p > 0.5$$

2. A political news company believes the national approval rating for the current president has fallen below 40%. Let r be the national approval rate. Then,

$$H_0: \ r = 0.4 \quad \text{and} \quad H_1: \ r < 0.4$$

Example

3. The length of a certain lumber from a national home building store is supposed to be 8.5 feet. A builder wants to check whether the shipment of lumber she receives has a mean length different from 8.5 feet.
   Let l be the length. Then,

   $$H_0: \ l = 8.5 \quad \text{and} \quad H_1: \ l \neq 8.5$$

Results for a hypothesis test

- The possible end results for a hypothesis test:
  1. H_0 is rejected when it was true: **type I error**.
  2. H_0 is rejected when it was not true: Correct action.
  3. H_0 is not rejected when H_1 was true: **type II error**.
  4. H_0 is not rejected when H_1 was not true: Correct action.
- The probability of type I error is called the level of significance and is denoted by α.
- The probability of type II error is denoted by β.

Condition of Null Hypothesis

		True	False
Possible Action	Fail to reject H_0	Correct action	Type II error
	Reject H_0	Type I error	Correct action

Remark: Usually type I error can be controlled, and α is chosen to be small. However, β is usually unknown and cannot be controlled.

- The **power** of the test is given by $1 - \beta$ and is equal to the probability that the null hypothesis will be rejected when the alternative hypothesis is true (H_0 is false).

Examples 8.3.3, 8.3.4, 8.3.5 give a better understanding of type I and II errors, and power of a test but these won't be our main concern.

Steps for hypothesis testing

1. **Data:** The data of a sample used in order to do the test.
2. **Assumptions:** The assumptions made about the population, i.e. the type of distribution, the given parameters (μ or σ, etc.).
3. **Hypotheses:** Decide which one is the null hypothesis and which is the alternative hypothesis.
4. **Test statistic:** A statistic that may be computed from the data of the sample. The test statistic will be used to accept or reject the null hypothesis. An example of a test statistic is

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Steps for hypothesis testing

5. **Distribution of the test statistic:** The probability distribution of the test statistic has to be specified *under the assumption that θ_0 is the true value of θ*

6. **Decision rule:** The graph of the distribution of the test statistic is divided into two groups: **the rejection** or **critical region** and **the nonrejection region**. The decision rule tells us to reject the null hypothesis if the value of the test statistic that we compute from our sample is in the rejection region and to not reject the null hypothesis if the computed value of the test statistic is in the nonrejection region with a certain *level of significance* (defined below).

Steps for hypothesis testing

7. **Calculation of the test statistic:** Use the given data to compute the test statistic.

8. **Statistical decision:** Reject or accept the null hypothesis by checking if the statistic computed belongs or not to the rejection region.

9. **Conclusion:** If H_0 is rejected, we conclude that H_1 is true. If H_0 is not rejected, we conclude that H_0 **may** be true.

Hypothesis testing on the mean
Sampling from Normally Distributed Population: Population Variance Known

Suppose now we want to test the population mean μ. Assume that the population is normally distributed and the population variance σ^2 is known.

- When sampling from a normally distributed population, the sample mean is normally distributed with mean μ and variance σ^2/n.

- The test statistic in this case is

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

- Z is standard normally distributed.

Example

We explain the process of hypothesis testing on the mean using an example.

Researchers are asking if the mean age of a certain population is different from 30 years old. They choose a random sample of 10 individuals and computed the sample mean $\overline{x} = 27$. It is given that the random variable giving the age of the individuals is approximately normally distributed with variance $\sigma^2 = 20$.

Hypotehses:

$$H_0 : \ \mu = 30 \quad \text{and} \quad H_1 : \ \mu \neq 30$$
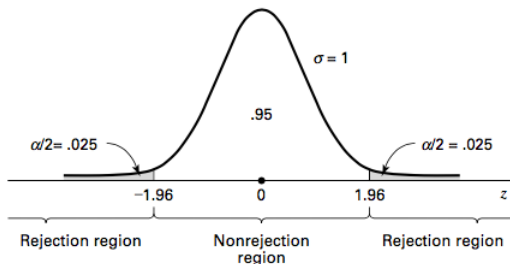
Test statistic:

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

Example

Distribution of the test statistic: Z has a standard normal distribution.

Decision rule: Suppose that the researchers chose the level of significance $\alpha = 0.05$. The null hypothesis is false if the test statistic is sufficiently large or sufficiently small. We consider then $\alpha/2$ to be associated with large values and $\alpha/2$ to be associated with small values. We compute then the **critical values**, i.e. the values z_1 and z_2 for which $P(Z < z_1) = 0.025$ and $P(Z > z_2) = 0.025$. By symmetry we get $z_2 = z_{1-\alpha/2}$ and $z_1 = -z_2$. We call this test a **two-tailed test**.

Example

Calculation of the test statistics: $z = \frac{27-30}{\sqrt{20/10}} = -2.12$.

Statistical decision: Since $-2.12 \leq 1.96$ (z is in the rejection region), we reject H_0 with a 0.05 level of significance.

Conclusion: We conclude that μ is not equal to 30.

Hypothesis Testing versus Confidence Intervals

If we consider the previous example and we compute a 95% confidence interval based on the given sample, we get

$$(27 - 1.96\sqrt{20/10}, 27 + 1.96\sqrt{20/10})$$

$$(24.2286; 29.7714)$$

We can see the the mean $\mu = 30$ does not belong to the interval, so we can reject the hypothesis.

One-tailed test
Left-tailed test

Consider that in the previous example, researchers are asking instead if the mean age of the population is less than 30.

**Hypotheses:** $H_0 : \mu \geq 30$ and $H_1 : \mu < 30$

The test statistic, its distribution and the value of the test statistic that corresponds to the given sample are the same as before. However the rejection region and therefore the statistical decision change.

Suppose we take again $\alpha = 0.05$. In this case, only small values of the test statistic will be rejected.

One-tailed test
Left-tailed test

Therefore, we compute only one critical z_1 such that all the values $z \leq z_1$ are rejected. The value of z_1 is such that $P(z \leq z_1) = 0.05$ and is denoted $z_{\alpha/2}$. In our case, $z_{\alpha/2} = z_{0.95} = -1.645$
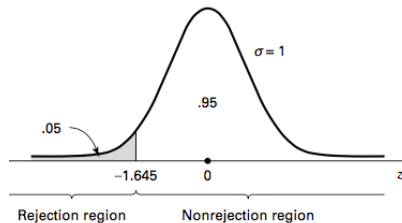
This test is called a **left-tailed test**.
Since the test statistic $z = -2.12$ is in the rejection region, we reject H_0 and we can say that $\mu < 30$ with a level of significance 0.05

One-tailed test
Right-tailed test

In the same way, we can define a **right-tailed test** which is a hypothesis test that correspond to the research hypothesis H_1 : $\theta > \theta_0$. For instance, if we consider the previous example with

$$H_0 : \ \mu \leq 30 \quad \text{and} \quad H_1 : \ \mu > 30$$

the rejection region will be the region on the right of $z_{1-\alpha} = 1.645$. That is, we reject H_0 if the test statistic gives a value greater than 1.645.

In our example, since $z = -2.12$, we do not reject H_0, and hence, the mean may be less than or equal to 30.

One or two tailed test?

$H_0 : \mu = \mu_0$	$H_0 : \mu \geq \mu_0$	$H_0 : \mu \leq \mu_0$
$H_1 : \mu \neq \mu_0$	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$
Two-tailed test	Left-tailed test	Right-tailed test

It is common to always write H_0 as an equality since we always assume the parameter μ ot be equal to μ_0. Therefore, another way of writing hypotheses is:

$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$
$H_1 : \mu \neq \mu_0$	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$
Two-tailed test	Left-tailed test	Right-tailed test

P-value

Another way to decide whether we reject or not the null hypothesis, is to use the **P-value**. It is called significance testing.

- The P-value for a test is the smallest value of α for which the null hypothesis can be rejected.
- If $\alpha \geq P$ we reject the null hypothesis.
- If $\alpha < P$, we do not reject the null hypothesis.
- The P-value depends on the value of the test statistic, and therefore it changes if the sample is changed.
- How to compute the P-value?

Computation of the P-value

Let z_0 be the value of the test statistic.

- **Two-tailed test:** In a two tailed test, the P-value is given by

$$P = \begin{cases} 2P(z \leq z_0) & \text{if } z_0 < 0 \\ 2P(z \geq z_0) & \text{if } z_0 > 0 \end{cases}$$

- **Left-tailed test:** In a left-tailed test, the P-value is given by

$$P = P(z \leq z_0)$$

- **Right-tailed test:** In a right-tailed test, the P-value is given by

$$P = P(z \geq z_0)$$

Example

Consider the first example in hypothesis testing (the two-tailed test). The P-value is given by
$P = 2P(z \leq -2.12) = 2(0.017) = 0.034$. Since $\alpha = 0.05 > P$, then we reject the hypothesis.
In the left-tailed test, $P = P(z \leq -2.12) = 0.017$ which is again less than the P-value. Hence, H_0 is rejected.

Hypothesis testing on the mean
Sampling from a Normally Distributed Population: Population Variance Unkown

- If the population is normally distributed with unknown variance, the test statistic used in hypothesis testing is

$$t = \frac{\overline{x} - \mu}{s/\sqrt{n}}$$

- The test statistic follows a T-distribution.
- The critical values for a two-tailed test are $t_{\alpha/2}$ and $t_{(1-\alpha/2)}$
- The critical values for a left-tailed test and right-tailed test are t_{α} and $t_{(1-\alpha)}$ respectively.
- All the other steps of hypothesis testing remain the same as in the first case.

Example

The maximum acceptable level for exposure to microwave radiation in the US is an average of 10 microwats per cm^2. It is a feared that a large tv transmitter may be polluting the air nearby by pushing the level of microwave radiation above the safe limit. A random sample of 25 readings is obtained. It is found that $\overline{x} = 10.3$ and $s = 2$. Use a hypothesis test with level of significance 10% to determine if a large tv may pollute the air.

$$H_0 : \mu \leq 10 \quad H_1 : \mu > 10$$

- We have a right-tailed test.
- Since the population variance is not known, we use the T-distribution.
- The test statistic is $t = \frac{\overline{X} - 10}{2/\sqrt{25}}$.

Example

- The critical value is $t_{1-\alpha} = t_{0.9} = 1.318$.
- The value of the test statistic is $t_0 = \frac{10.3-10}{2/5} = 0.75$
- Since $t_0 < t_{0.1}$ then t is in the nonrejection region and we cannot reject H_0. This means that the mean may be less than or equal to 10. In other words, the sample does not support the research hypothesis.
- The P-value is given by $P = P(t > 0.75)$. The table of the T-distribution does not give all the possible values of t. In the row $\gamma = 24$ (24 degrees of freedom), the value 0.75 i between 0.685 and 1.318. We can deduce that $0.75 < P(t < 0.75) < 0.9$. Therefore $0.1 < P = 1 - P(t < 0.75) < 0.25$. Hence, $P > \alpha = 0.1$ so we do not reject H_0.

Hypothesis testing on the mean
Sampling from a nonnormally distributed population

- If the distribution of the population is unknown, and n is large, by the central limit theorem, $\overline{X}$ is normally distributed with mean μ and variance σ^2/n.
- In this case, the hypothesis testing is done with the same steps as mention in the first case (population normally distributed with known variance).