

REVIEW - HYPOTHESIS TESTING

A **hypothesis** is a statement, which can either be true or not. We want to decide on it's validity based on available data. Denote

- H_0 the null hypothesis,
- H_A the alternative hypothesis.

A **statistical test** is a decision rule based on observed data, with two possible outcomes:

- **reject** H_0 in favor of H_A (*if our data stand against H_0 and for H_A*),
- **do not reject** H_0 (*based on our data H_0 cannot be rejected, the data do not contradict H_0*).

In the decision process we can make an **error**:

Reality	Decision	
	reject H_0	don't reject H_0
H_0 holds	type I error	OK
H_0 doesn't hold	OK	type II error

Type I error is more severe (we are falsely incriminating) → its probability must be controlled:

- We choose a **level of significance** of the test α as the maximum possible probability of making type I error (most often $\alpha = 0.05$ or $\alpha = 0.01$).
- Denote $\beta = P(\text{type II error})$. Then $1 - \beta$ is called the *power* of the test.

p-value (p -value) = highest achieved level

- The result of the test often given by statistical software.
- Probability of getting a result which stands against H_0 even more than the data at hand, if H_0 is actually valid.
- Rule of thumb: $p\text{-value} \leq \alpha \Rightarrow$ reject H_0 in favor of H_A

Decision rule

- Reject H_0 if the tested value is not in the corresponding $(1 - \alpha)\%$ confidence interval.
- Equivalently: Reject H_0 if the p -value of the test is lower than α .
- Equivalently: Reject H_0 if the *test statistic* lies in the corresponding α -critical region.

EXERCISES 11 - HYPOTHESIS TESTING

1. A safety sensor regularly scans a room. If there is no movement in the room, the sensor gives a signal $X = W$, where W is a random variable with expected value 0. If there is a movement, the sensor gives a signal $X = W + \theta$, where $\theta > 0$ is an unknown constant. We have 35 observations and we computed symmetric confidence intervals for $\mu = EX$ as follows:

Symmetric 90% confidence interval: (0.405, 5.394)
 Symmetric 95% confidence interval: (-0.072, 5.872)

Test the hypothesis $H_0 : \mu = 0$ versus the alternative $H_A : \mu > 0$, so that the probability of type I error is at most 5%. Which interval do you use and why?

2. Before a storm, the variance of the measured wind speeds usually grows. If a storm is coming, the variance σ^2 exceeds 4.5. Suppose that the wind speed measurements are normally distributed. We have computed the confidence intervals for σ^2 based on 200 measurements as:

Symmetric 99% confidence interval: (2.64, 4.43)
 Symmetric 98% confidence interval: (2.7, 4.31)

Is there a storm coming? Test the hypothesis $H_0 : \sigma^2 \leq 4.5$ versus $H_A : \sigma^2 > 4.5$ so that the probability of type I error is at most 1%. Which interval have you used and why?

3. We have records from pizza sales from previous 9 days:

41, 38, 56, 38, 63, 59, 52, 49, 46.

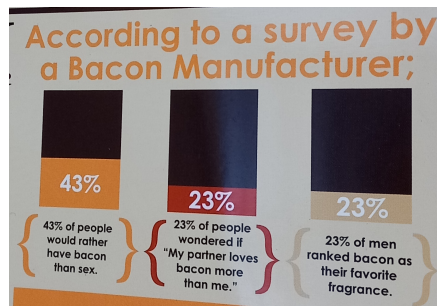
- a) Based on the data estimate how many pizzas we are likely to sell on the next day?
- b) Suppose that data come from the normal distribution. Find symmetric 90% and 95% confidence intervals for the expected value.
- c) Why is the 95% confidence interval wider than the 90% one?
- d) Can we reject the hypothesis that the expected value is 45?

4. Suppose we observe a random sample of $n = 16$ values from the normal distribution. The sample mean and sum of second powers are

$$\bar{X}_{16} = 13 \quad \text{and} \quad \sum_{i=1}^{16} X_i^2 = 2708.$$

- a) Test the hypothesis $H_0 : \mu = 15$ versus $H_A : \mu \neq 15$ on level of significance 5%.
- b) Test the hypothesis $H_0 : \sigma^2 = 0.5$ versus $H_A : \sigma^2 \neq 0.5$ on level of significance 5%.

5. A Bacon Manufacturer conducted a survey asking people whether they considered bacon a source of jealousy in their relationship. Out of 97 people 14 answered positively. They then published the study stating that 23% of people considered bacon a potential romantic rival.



- a) Can we reject their statement on level of significance 5%? And of level of significance 1%?
- b) What if the statement was "more than 23%" (again for $\alpha = 5\%$ and $\alpha = 1\%$)?

ADDITIONAL EXERCISES - HYPOTHESIS TESTING

One sample test of the expectation and variance

6. A car manufacturer states that the average fuel consumption of a new model is 6 liters per 100 kilometers. We have tested 20 cars and observed that the average consumption was $\bar{X}_n = 6.8$ liters per 100 km with a sample variance of $s_n^2 = 2.56$.

- a) On level of significance $\alpha = 5\%$, perform a test to determine whether the manufacturer's statement is true or whether the actual consumption is significantly higher.
- b) Repeat for $\alpha = 1\%$.

7. A production line makes steel parts. The variance of their lengths shouldn't be larger than $300 \mu\text{m}^2$. Based on a sample of 15 parts we have computed the sample variance as $s_n^2 = 580 \mu\text{m}^2$.

- a) On the level of significance $\alpha = 5\%$, perform a test to determine whether the production line is operating with sufficient precision, or whether it should be recalibrated because the variance is too large.
- b) Repeat for $\alpha = 1\%$.

8. We are doing a research to determine what is the actual volume of beer in one glass served in one specific bar. We bought ten beer and the volume was (in liters):

0.510, 0.462, 0.491, 0.466, 0.451, 0.503, 0.475, 0.487, 0.512, 0.505.

Suppose that the volume of beer in one glass follows the normal distribution $N(\mu, \sigma^2)$ and the measurements are independent.

- a) Estimate the expected volume μ of beer in one glass.
- b) Estimate the probability of receiving an undersized beer (less than 0.5 l).
- c) As customers, we would like to know whether the bartender isn't cheating on us. Perform a test of the hypothesis whether the real expected volume of beer truly is 0.5 liters, against the one-sided alternative that the real expected volume is less than 0.5 l. Use $\alpha = 5\%$.
- d) Repeat for $\alpha = 1\%$.