proprep™

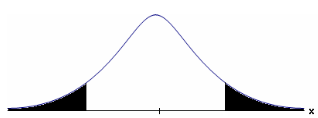
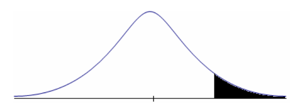
Workbook

Table of Contents

Hypothesis Testing about a Mean

When the Population Variance is Known

Background

The Null Hypothesis: **The Alternative Hypothesis:**	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$
Conditions:	**1)** σ known **2)** $X \sim N$ or a large enough sample		
Decision Criterion: Rejection region of H_0:	$Z_{\bar{X}} > Z_{1-\frac{\alpha}{2}}$ or $Z_{\bar{X}} < -Z_{1-\frac{\alpha}{2}}$  $-Z_{1-\frac{\alpha}{2}}$ $Z_{1-\frac{\alpha}{2}}$ ■ - Reject H_0.	$Z_{\bar{X}} < -Z_{1-\alpha}$  $-Z_{1-\alpha}$ ■ - Reject H_0.	$Z_{\bar{X}} > Z_{1-\alpha}$  $Z_{1-\alpha}$ ■ - Reject H_0.

The Test Statistic:

$$Z_{\bar{X}} = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

An alternative to the decision criterion:

H_0 **is rejected if:**	$\bar{X} > \mu_0 + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ or $\bar{X} < \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$

Example (solution in the recording)

The average weight of Army recruits 20 years ago was 65 kg.
A study seeks to test whether the current average weight of recruits is bigger.
Assume that the weight of recruits has a normal probability distribution with a standard-deviation of 12 kg. A random sample of 16 recruits had an average weight of 71 kg.

a. What is the p-value of the results?
b. What conclusion is drawn if the level of significance is 5%, and what conclusion is drawn for 1%?

Questions:

1) The average score on a state exam is 72 with a standard deviation of 15 points. A teacher claims that she has developed a new teaching method that will raise the average score. The Department of Education decided to assign 36 random students to the teacher, and after studying according to her method, their average mark was 75.5. Assuming the standard deviation under the new method is also 15, what is your conclusion at a 5% level of significance?

2) The average weight of athletes in a certain event is 200 lbs, with a standard deviation of 18 lb. In order to test a certain diet's effectiveness, a random sample of 50 athletes is taken. After one year, the average weight of the athletes in the sample is 190 lbs. Test whether the diet causes weight loss at a 10% level of significance.

3) The average monthly salary in 2012 was $\$8,800$ with a standard deviation of $\$2,000$. A sample of 100 employees had an average salary of $\$9,500$. The purpose of the study is to test whether salaries have risen since 2012. At what level of significance will the researcher decide that the average salary has risen?

4) A person suspects that a candy company is misrepresenting the weight of their candy bar, and that its weight is less than the advertised 100 grams. The standard deviation of the bar's weight is known to be 12 grams. The person plans to weigh 100 randomly selected bars and obtains an average weight of 98.5 grams. What conclusion do we reach at a 5% level of significance?

Answer Key:

1) We can't reject the null hypothesis (H_0).

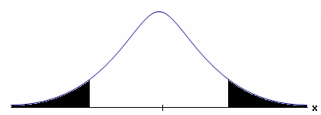
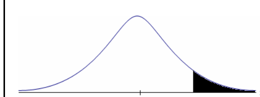
2) The diet helps people to lose weight (i.e. we reject H_0 and accept H_1).

3) $\alpha = 0.001$

4) We can't reject the null hypothesis (H_0).

Unknown Population Variance

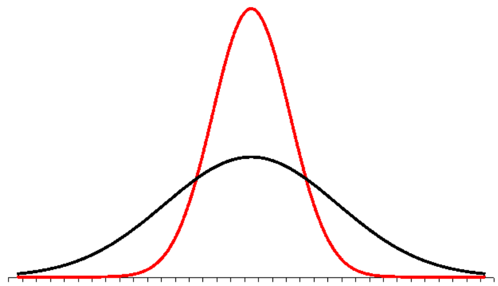
Background

The Null Hypothesis: **The Alternative Hypothesis:**	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$
Conditions:	**1)** σ unknown **2)** $X \sim N$ or a large enough sample		
Decision Criterion: Rejection region of H_0:	$t_{\overline{x}} > t^{(n-1)}_{1-\frac{\alpha}{2}}$ or $t_{\overline{x}} < -t^{(n-1)}_{1-\frac{\alpha}{2}}$  $-t_{1-\frac{\alpha}{2},n-1}$ $t_{1-\frac{\alpha}{2},n-1}$ █ - Reject H_0	$t_{\overline{x}} < -t^{(n-1)}_{1-\alpha}$  $-t_{1-\alpha,n-1}$ █ - Reject H_0	$t_{\overline{x}} > t^{(n-1)}_{1-\alpha}$  $t_{1-\alpha,n-1}$ █ - Reject H_0
Alternative to the Decision Criterion: **Reject** H_0 **if:**	$\overline{X} > \mu_0 + t^{n-1}_{1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$ *or* $\overline{X} < \mu_0 - t^{n-1}_{1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$	$\overline{X} > \mu_0 - t^{n-1}_{1-\alpha} \cdot \frac{S}{\sqrt{n}}$	$\overline{X} > \mu_0 + t^{n-1}_{1-\alpha} \cdot \frac{S}{\sqrt{n}}$

The test statistics is: $t_{\overline{x}} = \dfrac{\overline{x} - \mu_0}{S / \sqrt{n}}$

$$S^2 = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{n-1} = \frac{\sum_{i=1}^{n} X^2{}_i - n\overline{X}^2}{n-1}$$

T-Probability Distribution



The T probability distribution is a symmetric bell-shaped probability distribution with an expectation of 0. The probability distribution is similar to that of Z, but is wider, and its values are therefore higher.
The T probability distribution depends on a concept call 'degrees of freedom'.

The degrees of freedom are $df = n-1$.
Higher degrees of freedom mean that the probability distribution becomes higher and narrower.

When the degrees of freedom tend to infinity, the T probability distribution converges to the Z probability distribution.

Example (solution in the recording)

A factory receives an order to produce pallets with a thickness of $0.1\ \mathrm{cm}$.
In order to test whether the factory is fulfilling the requirement, 10 pallets were sampled, and their average thickness was $0.104\ \mathrm{cm}$.
The estimate for the standard deviation was $0.002\ \mathrm{cm}$.

a. What are the study hypotheses?
b. What assumption is necessary for finding a solution?
c. Test at a 5% level of significance.

Questions

1) An existing medicine has an average recovery time of 120 hours with an unknown standard deviation. We want to test whether a new drug reduces this recovery time. A sample of 5 patients who took the new drug had the following recovery times: 90, 95, 100, 80, and 125 hours. At a 5% level of significance, what conclusion can be made, if any?

2) IQ scores follow a normal probability distribution. Suppose the average score among American college students is 100. In a sample of 23 randomly chosen Japanese students, the average score was 104.5 with a sample standard deviation of 16. What conclusions at a 5% significance level can be made about differences in average IQ scores between American and Japanese college students?

3) On a production line, the weight of sugar packages is normally distributed and is supposed to be 1 kg. Five random packages are sampled every day to test whether the production line is functioning properly. Suppose one day the weights are: 1008, 1024, 996, 1005, 997 grams. At a 5% level of significance, what is the conclusion?

4) A researcher tests the assertion that employees working on the night shift work more slowly than day employees. Assume that production time follows a normal distribution, and the average production time during the day is 6 hours. It was found for a random sample of 25 employees on the night shift that the average time needed to produce the same product was 7 hours, with a standard deviation of 2 hours. At a 5% level of significance, what is the conclusion?

Answer key:

1) We can reject the null hypothesis (H_0) and accept H_1.
2) There is no difference.
3) We can't reject H_0, i.e. the machine is working well that day.
4) We can't reject H_0, i.e. there is no difference between day and night shifts.