M358K: October 22nd, 2021.

Tests of Significance [cont'd].

Note: There is a relationship between

- a C-confidence interval

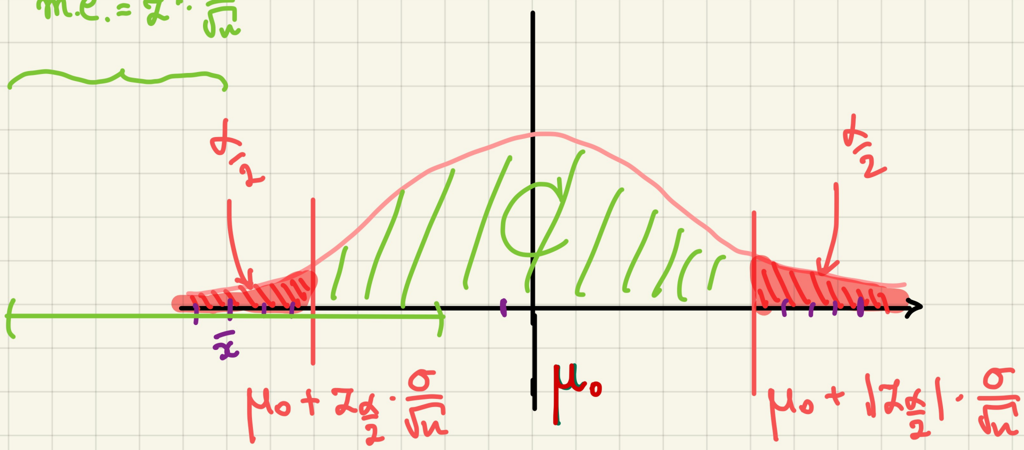
&

- the two-sided test of significance w/ $\alpha = 1 - C$

$C = 1 - \alpha$

$z^* = \Phi^{-1}\left(\frac{1+C}{2}\right)$

$m.e. = z^* \cdot \frac{\sigma}{\sqrt{n}}$



Reject the null $\Leftrightarrow$ the confidence interval does not contain μ_0

UNIVERSITY OF TEXAS AT AUSTIN

Problem Set # 11

Test of significance.

Problem 11.1. A test of significance can be used to test differences in categorical data. *True or false? Why?*

Problem 11.2. Confidence intervals and two-sided significance tests are linked in the sense that a two-sided test at a significance level α can be carried out in the form of a confidence interval with confidence level $1-\alpha$. *True or false?*

Problem 11.3. In a test of statistical hypotheses, what does the $p-$value tell us?

a. If the null hypothesis is true.
b. If the alternative hypothesis is true.
c. The largest level of significance at which the null hypothesis can be rejected.
d. The smallest level of significance at which the null hypothesis can be rejected

---

Complete the following statements:

Problem 11.4. When computing $p-$values, if the $p-$value is smaller than the chosen significance level α, we say that the results are statistically significant.

Provide your **complete solution** for the following problems.

Problem 11.5. You perform 2000 significance tests using a significance level 0.10. Under the assumption that all of the null hypotheses for the 2000 significance tests are true, how many of the 2000 significance tests would you expect to be statistically significant?

a. 200
b. 1800
c. 2000
d. 0
e. None of the above.

UNIVERSITY OF TEXAS AT AUSTIN

Problem Set # 12

Hypothesis testing: The normal case.

Problem 12.1. *Source: Ramachandran, Tsokos.*
The management of the local health club claims that its members lose on average 15 pounds or more within the first three months of their membership. A consumer agency took a simple random sample of 45 members and found the sample average of 13.8 in pounds lost. Assume that we model the weight loss as normal with an **unknown** mean μ and the **known** standard deviation of 4.2 pounds. What is the p−value corresponding to the gathered data? What would your decision be at the 0.05 significance level?

→ : The population model.

X ... weight loss of a randomly chosen gym member

$X \sim \text{Normal}(\text{mean} = \mu, \boxed{sd = \sigma = 4.2})$

Our hypotheses:

$H_0: \boxed{\mu = \mu_0 = 15}$ vs. $H_a: \mu < \mu_0 = 15$

The observed value of the z-statistic, i.e., your z-score

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{13.8 - 15}{\frac{4.2}{\sqrt{45}}} = -1.9166$$

The p-value is:

$$\mathbb{P}[Z < z] = \Phi(-1.9166) = \text{pnorm}(-1.9166) = 0.02764438$$

Less than $\alpha = 0.05$, so we reject the null hypothesis.

Problem 12.2. *Source: Ramachandran, Tsokos.*
It is claimed that sports-car owners drive on the average 20,000 miles per year. A consumer firm believes that the mean annual mileage is actually lower. To check, the consumer firm decided to test this hypothesis.

The modeling assumptions are that the annual mileage is normally distributed with an unknown mean μ and with the standard deviation of 1200.

The consumer firm obtained information from 36 randomly selected sports-car owners that resulted in a sample average of 19,530 miles. What is the decision of this hypothesis test at the signficance level of 0.01?

$\rightarrow$:

$$H_0: \boxed{\mu = 20000} \quad \text{vs.} \quad H_a: \mu < 20000$$

The observed value of the z-statistic:

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{19530 - 20000}{\frac{1200}{\sqrt{36}}} = -2.35$$

$$z_{0.01} = \Phi^{-1}(0.01) = \text{qnorm}(0.01) = -2.326348$$

the upper bound of the RR in std units

$\Rightarrow$ Reject the null hypothesis since

$$-2.35 = z < z_{0.01} = -2.33$$

■