

STATISTICS FOR DATA SCIENCE HYPOTHESIS and INFERENCE

Dr. Deepa Nair

Department of Science and Humanities



UNIT-4 HYPOTHESIS and INFERENCE

Session-2

Large-Sample Tests for a Population

Mean

Dr. Deepa Nair

Department of Science and Humanities

Large-Sample Tests for a Population Mean

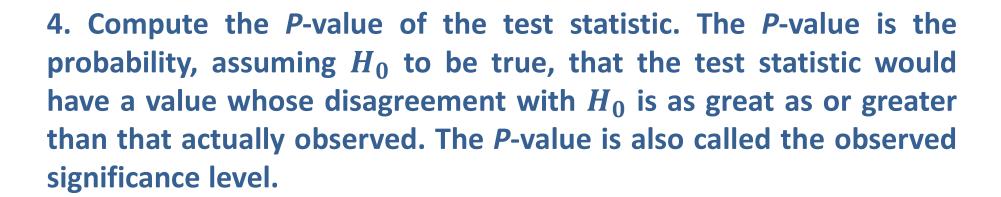




2. Assume H_0 to be true.

3. Compute a test statistic. A test statistic is a statistic that is used to assess the strength of the evidence against H_0 .

Large-Sample Tests for a Population Mean



5. State a conclusion about the strength of the evidence against H_0 .



Large-Sample Tests for a Population Mean



In any hypothesis test, we are calculating conditional probabilities based on the assumption that the null hypothesis is true.

Introduction

PES UNIVERSITY ONLINE

Type of Hypothesis Tests:

SI No	Type of Hypothesis Test	Alternate Hypothesis	Null Hypothesis
1	Left Tailed Test	<	≥
2	Right Tailed Test	>	≤
3	Two Tailed Test	≠	=

Large-Sample Tests for a Population Mean



Different Hypothesis Tests

Hypothesis Test	Test Statistic
z-test	z- statistic
t-test	t- statistic
Chi square test	Chi square statistic

Large-Sample Tests for a Population Mean



Note:

Hypothesis Test	Population S.D. known	Population S.D. unknown
n < 30 Small sample drawn from Normal population	Use z- test	Use t-test
$n \ge 30$ Large sample	Use z- test	Use z- test Or t- test

Large-Sample Tests for a Population Mean



Note:

A test that uses z-score as a test statistic is called z-test.

Large-Sample Tests for a Population Mean



Example:

- The article "Wear in Boundary Lubrication" (S. Hsu, R. Munro, and M. Shen, *Journal of Engineering Tribology*, 2002: 427–441) discusses several experiments involving various lubricants.
- In one experiment, 45 steel balls lubricated with purified paraffin were subjected to a 40~kg load at 600~rpm for 60~minutes.

Large-Sample Tests for a Population Mean



Example:

- The average wear, measured by the reduction in diameter, was $673.2~\mu m$, and the standard deviation was $14.9~\mu m$.
- Assume that the specification for a lubricant is that the mean wear be less than $675~\mu m$.
- Find the P -value for testing $H_0: \mu \geq 675$ versus $H_1: \mu < 675$.

Large-Sample Tests for a Population Mean



Solution:

- The null hypothesis is that the lubricant does not meet the specification, and that the difference between the sample mean of 673.2 and 675 is due to chance.
- The alternate hypothesis is that the lubricant does indeed meet the specification.

$$z = \frac{673.2 - 675}{2.22} = -0.81$$

Large-Sample Tests for a Population Mean



Soution:

- The P v alue is 0.209.
- Therefore if H_0 is true, there is a 20.9% chance to observe a sample whose disagreement with H_0 is as least as great as that which was actually observed.
- Since 0.209 is not a very small probability, we do not reject H_0 .
- Instead, we conclude that H_0 is plausible.

Large-Sample Tests for a Population Mean



Example:

- A scale is to be calibrated by weighing a 1000 g test weight 60 times.
- The 60 scale readings have mean 1000.6 g and standard deviation 2 g.
- Find the *P*-value for testing $H_0: \mu = 1000 \ versus \ H_1: \mu \neq 1000.$

Large-Sample Tests for a Population Mean



Solution:

$$H_0: \mu = 1000 \ versus \ H_1: \mu \neq 1000.$$

We assume H_0 is true

$$z = \frac{1000.6 - 1000}{0.258}$$
$$= 2.32$$

Large-Sample Tests for a Population Mean



Solution:

- The P-value is the sum of the areas in both of these tails, which is 0.0204.
- Therefore, if H_0 is true, the probability of a result as extreme as or more extreme than that observed is only 0.0204.
- The evidence against H_0 is pretty strong. It would be prudent to reject H_0 and to recalibrate the scale.



Dr. Deepa Nair

Department of Science and Humanities

deepanair@pes.edu