University of Jordan
Chemical Engineering Department
905509 Statistical Quality Control

Review of Test of Hypothesis

Dr. Ali Kh. Al-Matar
aalmatar@ju.edu.jo

Outline

- What is a hypothesis
- Procedure for test of hypothesis
- Types of errors
- Testing of hypothesis

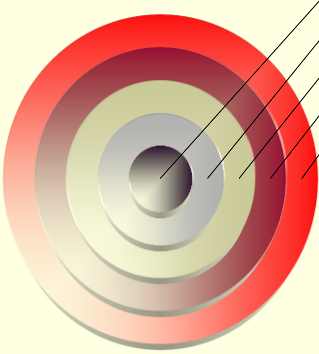
What is a Hypothesis?

- Hypothesis is a statement about a population developed for the purpose of testing.
  - In the legal system; a person is innocent till proven guilty. The judge and/or the jury subject this hypothesis to verification by reviewing the evidence and testimony before reaching a verdict.
  - A doctor observes certain symptoms on his/her patient and orders certain diagnostic tests and follow up with treatment.

What is Hypothesis Testing?

- Hypothesis Testing is a (systematic) procedure based on sample evidence and probability theory to determine whether the hypothesis is a reasonable statement.
  - Hypothesis testing does not provide proof that something is true in a mathematical sense.
  - It is in a sense close to the "Proof beyond reasonable doubt" used in the legal system.
  - From a philosophy of science point of view one can never prove anything, we can only disprove things.

Steps for Hypothesis Testing



5. Do not reject H_0 or reject H_0 and accept H_1
4. Formulate a Decision Rule
3. Identify (Compute) the test Statistic
2. Select a level of Significance
1. State the null (H_0) and alternate (H_1) hypotheses

1. State the Null (H_0) and Alternate (H_1) Hypotheses

- Null hypothesis is a statement about the value of a population parameter.
  - H stands for Hypothesis while the 0 stands for "no difference".
  - The null hypothesis is a statement that is not rejected if our sample data fail to provide convincing evidence that it is false.
  - Failing to reject the null hypothesis does not prove that H_0 is true, it means we have failed to disprove H_0.
  - To prove without any doubt that the null hypothesis is true, the population parameter would have to be known, which is not usually feasible.
- Alternate hypothesis is a statement that is accepted if the sample data provide enough evidence that the null hypothesis is false.
- Remember always that the null hypothesis will always contain the equal sign. We turn to the alternate hypothesis only if we prove the null hypothesis to be untrue.

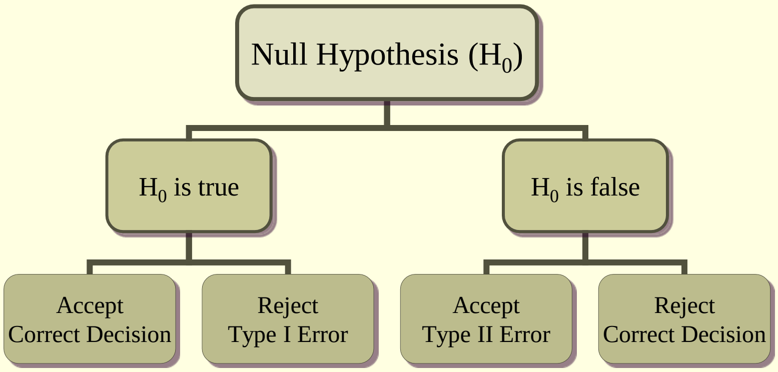
2. Select a Level of Significance

- Level of significance (level of risk) is the probability of rejecting the null hypothesis when it is true.
  - 0.05 level is used traditionally for consumer research projects,
  - 0.01 level is used traditionally for quality assurance
  - 0.10 level is used traditionally for political polling
- You as a researcher need to decide what is the level of significance before formulating the decision rule.

Type I and II Errors

- Type I Error
  - Occurs when rejecting the null hypothesis, H_0, when it is true.
  - The probability of committing a type I error is denoted α. Called “Producer’s Risk”.
    - Probability that a good lot will be rejected.
    - Probability of sending an innocent person to jail.
- Type II Error
  - Occurs when accepting the null hypothesis, H_0, when it is false.
  - The probability of committing a type II error is denoted β. Called “Consumer’s Risk”.
    - Probability that a poor lot will be accepted.
    - Probability of setting a person free al though he is guilty.

Types of Errors



Power of a Test

- The **Power** of a test of hypothesis is given by $(1 - \beta)$
- $(1 - \beta)$ is the probability of
  - Correctly rejecting the null hypothesis, or
  - the probability of rejecting the null hypothesis when the alternative is true.

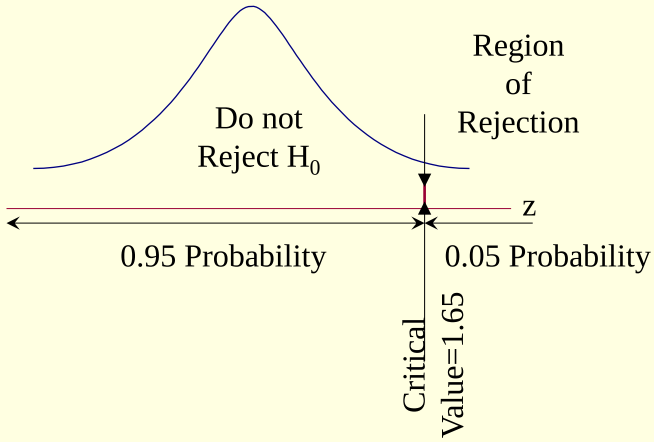
3. Compute the Test Statistic

- Tests statistic is a value, determined from sample information, used to determine whether to reject the null hypothesis.
  - Many statistics such as z, t, F and χ^2
  - z distribution as a test statistic is used to test for the mean (μ). The z value is based on the sampling distribution of X, which is normally distributed.

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

4. Formulate the Decision Rule

- Decision Rule is a statement of the conditions under which the null hypothesis is rejected and the conditions under which it is not rejected.
- Critical value is the dividing point between the region where the null hypothesis is rejected and the region where it is not rejected

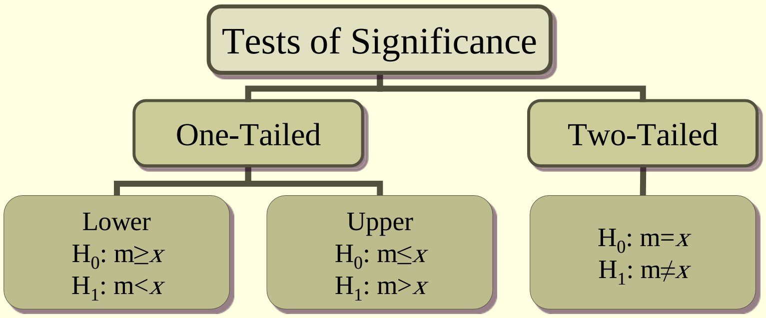


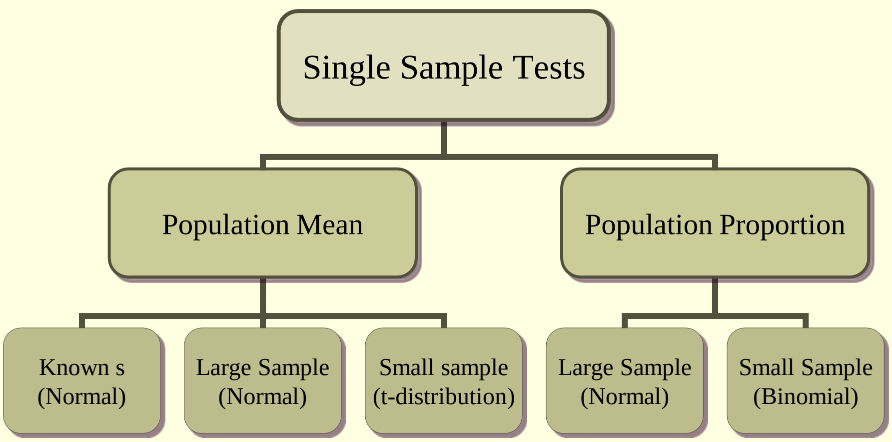
5. Make a Decision

- Depending on the z value of the sample either reject the null hypothesis or do not reject it.
- To phrase the “acceptance” of the null hypothesis, many researchers tend to use
  - Do not reject H_0.
  - We fail to reject H_0.
  - The sample results do not allow us to reject H_0.

One-Tailed and Two-Tailed Tests

- If H_1 specifies a direction, the test is one-tailed, otherwise it is two-tailed.





Test on the Mean with Known σ

Null hypothesis $H_0 : \mu = \mu_0$

Test Statistic value: $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \mu > \mu_0$	$z \geq z_{\alpha}$
$H_1 : \mu < \mu_0$	$z \leq -z_{\alpha}$
$H_1 : \mu \neq \mu_0$	either $z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$

Test on the Mean with Large Sample

- Replace population σ by the sample s (n>30).

Null hypothesis $H_0 : \mu = \mu_0$

Test Statistic value: $z = \dfrac{\bar{X} - \mu_0}{s / \sqrt{n}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \mu > \mu_0$	$z \geq z_{\alpha}$
$H_1 : \mu < \mu_0$	$z \leq -z_{\alpha}$
$H_1 : \mu \neq \mu_0$	either $z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$

Test on the Mean with Small Sample

- Normal distribution of the sample is no longer valid. The t-distribution is used instead

Null hypothesis $H_0 : \mu = \mu_0$

Test Statistic value: $t = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \mu > \mu_0$	$t \geq t_{\alpha,n-1}$
$H_1 : \mu < \mu_0$	$t \leq -t_{\alpha,n-1}$
$H_1 : \mu \neq \mu_0$	either $t \geq t_{\alpha/2,n-1}$ or $t \leq -t_{\alpha/2,n-1}$

Test on Proportions: Large Sample

- Valid for $np_0 \geq 5$ and $n(1-p_0) \geq 5$

Null hypothesis $H_0 : p = p_0$

Test Statistic value: $z = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : p > p_0$	$z \geq z_{\alpha}$
$H_1 : p < p_0$	$z \leq -z_{\alpha}$
$H_1 : p \neq p_0$	either $z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$

Test on Proportions: Small Sample

- Based directly on the binomial distribution.

P-value

- The P-value is the smallest level of significance at which H_0 would be rejected when a specified test procedure is used on a given data set.
- Once the P-value has been determined, the conclusions at any particular level α results from comparing the P-value to α:
  - P-value $\leq \alpha \Rightarrow$ reject H_0 at level α.
  - P-value $> \alpha \Rightarrow$ do not reject H_0 at level α.

P-value Rules

- If the P-value is less than
  - 0.10, we have **some** evidence that H_0 is not true.
  - 0.05, we have **strong** evidence that H_0 is not true.
  - 0.01, we have **very strong** evidence that H_0 is not true.
  - 0.001, we have **extremely strong** evidence that H_0 is not true.

- The output voltage of a power supply is assumed to be normally distributed. 16 observations are taken at random.

10.35	**9.3**	**10**	**9.96**
11.65	**12**	**11.25**	**9.58**
11.54	**9.95**	**10.28**	**8.37**
10.44	**9.25**	**9.38**	**10.85**

- Test the hypothesis that the mean voltage is 12V against a two-sided alternative that using α=0.05.

- Read about
  - Pooled t-test
  - Paired t-test
  - ANOVA

Calculation of probability of Type II error

- Assume the test of interest is

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

- P(Type II Error) is found to be

$$\beta = \Phi\left(Z_{\frac{\alpha}{2}} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-Z_{\frac{\alpha}{2}} - \frac{\delta\sqrt{n}}{\sigma}\right)$$

- The Power of the test is then 1 - β

Operating Characteristic Curves

- Operating Characteristic (OC) curve is a graph representing the relationship between β, α, δ and *n*.
- OC curves are useful in determining how large a sample is required to detect a specified difference with a particular probability.
- Calculate probability of Type II error by