University of Jordan
Chemical Engineering Department
Process Modeling by Statistical Methods– 905331

Lecture 07: Testing of Hypothesis – Single Population

Dr. Ali Khalaf Al-Matar

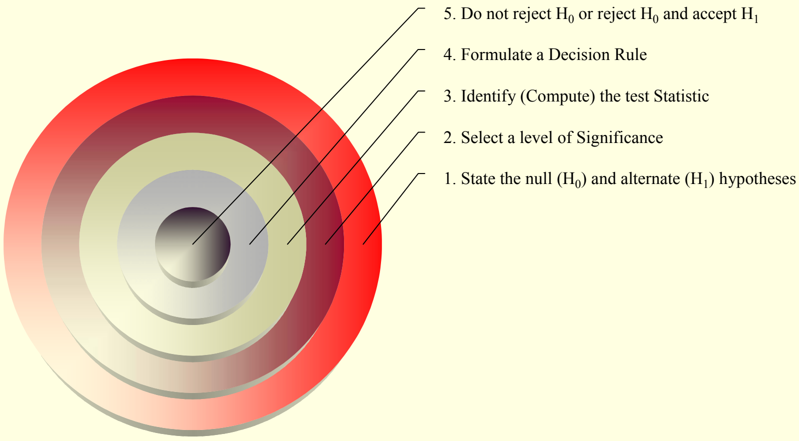
What is a Hypothesis?

- 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



Step 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.

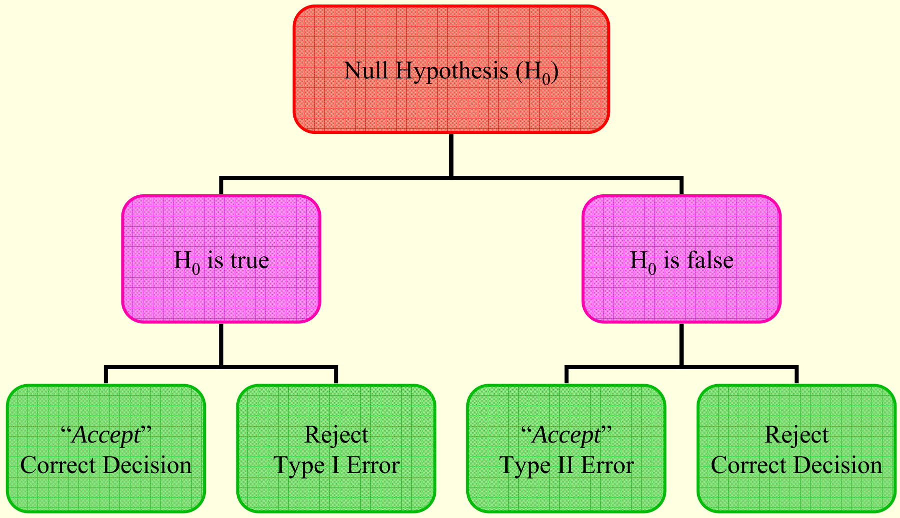
- Alternate hypothesis is a statement that is accepted if the sample data provide enough evidence that the null hypothesis is false.
- Remember always that the null hypothesis will always contain the equal sign. We turn to the alternate hypothesis only if we prove the null hypothesis to be untrue.

Step 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.
- 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.



Step 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 the mean, which is normally distributed.

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

Step 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

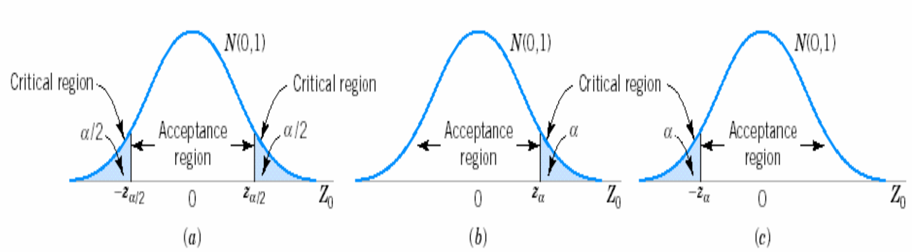
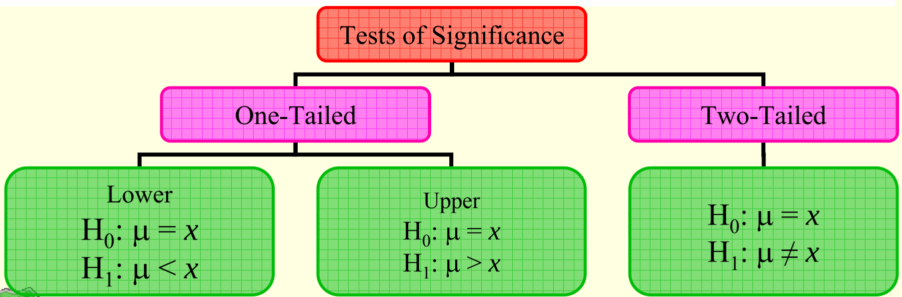


Figure 9-6 The distribution of Z_0 when H_0: $\mu = \mu_0$ is true, with critical region for (a) the two-sided alternative $H_1 : \mu \neq \mu_0$, (b) the one-sided alternative $H_1 : \mu > \mu_0$, and (c) the one-sided alternative $H_1 : \mu < \mu_0$.

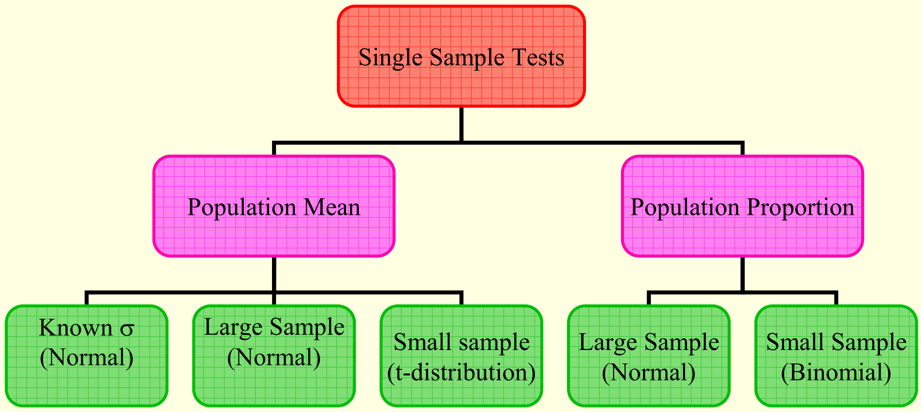


P-value

- 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.

Step 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.



Test on the Mean with Known σ

Null hypothesis $H_0 : \mu = \mu_0$

Test Statistic value: $z_0 = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \mu > \mu_0$	$z_0 \geq z_\alpha$
$H_1 : \mu < \mu_0$	$z_0 \leq -z_\alpha$
$H_1 : \mu \neq \mu_0$	either $z_0 \geq z_{\alpha/2}$ or $z_0 \leq -z_{\alpha/2}$

- Aircrew escape systems are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specifications require that the mean burning rate must be 50 centimeters per second. We know that the standard deviation of burning rate is 2 cm/s. The experimenter decides to specify a type I error probability or significance level of 0.05 and selects a random sample of n=25 and obtains a sample average burning rate of 51.3 cm/s. What conclusions should be drawn?

1. The parameter of interest is μ, the mean burning rate.
2. H_0: $\mu = 50$ centimeters per second
3. H_1: $\mu \neq 50$ centimeters per second
4. $\alpha = 0.05$
5. The test statistic is

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

6. Reject H_0 if $z_0 > 1.96$ or if $z_0 < -1.96$. Note that this results from step 4, where we specified $\alpha = 0.05$, and so the boundaries of the critical region are at $z_{0.025} = 1.96$ and $-z_{0.025} = -1.96$.
7. Computations: Since $\bar{x} = 51.3$ and $\sigma = 2$,

$$z_0 = \frac{51.3 - 50}{2/\sqrt{25}} = 3.25$$

8. Conclusion: Since $z_0 = 3.25 > 1.96$, we reject H_0: $\mu = 50$ at the 0.05 level of significance. Stated more completely, we conclude that the mean burning rate differs from 50 centimeters per second, based on a sample of 25 measurements. In fact, there is strong evidence that the mean burning rate exceeds 50 centimeters per second.

For the foregoing normal distribution tests it is relatively easy to compute the P-value. If z_0 is the computed value of the test statistic, the P-value is

$$P = \begin{cases} 2[1 - \Phi(|z_0|)] & \text{for a two-tailed test: } H_0\text{: } \mu = \mu_0 \quad H_1\text{: } \mu \neq \mu_0 \\ 1 - \Phi(z_0) & \text{for a upper-tailed test: } H_0\text{: } \mu = \mu_0 \quad H_1\text{: } \mu > \mu_0 \\ \Phi(z_0) & \text{for a lower-tailed test: } H_0\text{: } \mu = \mu_0 \quad H_1\text{: } \mu < \mu_0 \end{cases} \quad (9\text{-}15)$$

Here, $\Phi(z)$ is the standard normal cumulative distribution function defined in Chapter 4. Recall that $\Phi(z) = P(Z \leq z)$, where Z is $N(0, 1)$. To illustrate this, consider the propellant problem in Example 9-2. The computed value of the test statistic is $z_0 = 3.25$ and since the alternative hypothesis is two-tailed, the P-value is

$$P\text{-value} = 2[1 - \Phi(3.25)] = 0.0012$$

Thus, H_0: $\mu = 50$ would be rejected at any level of significance $\alpha \geq P\text{-value} = 0.0012$. For example, H_0 would be rejected if $\alpha = 0.01$, but it would not be rejected if $\alpha = 0.001$.

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: $z_0 = \dfrac{\bar{X} - \mu_0}{s / \sqrt{n}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \mu > \mu_0$	$z_0 \geq z_\alpha$
$H_1 : \mu < \mu_0$	$z_0 \leq -z_\alpha$
$H_1 : \mu \neq \mu_0$	either $z_0 \geq z_{\alpha/2}$ or $z_0 \leq -z_{\alpha/2}$

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: $t_0 = \dfrac{\bar{X} - \mu_0}{s / \sqrt{n}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \mu > \mu_0$	$t_0 \geq t_{\alpha, n-1}$
$H_1 : \mu < \mu_0$	$t_0 \leq -t_{\alpha, n-1}$
$H_1 : \mu \neq \mu_0$	either $t_0 \geq t_{\alpha/2, n-1}$ or $t_0 \leq -t_{\alpha/2, n-1}$

Example

- The increased availability of light materials with high strength has revolutionized the design and manufacture of golf clubs, particularly drivers. Clubs with hollow heads and very thin faces can result in much longer tee shots, especially for players of modest skills. This is due partly to the "spring-like effect" that the thin face imparts to the ball. Firing a golf ball at the head of the club and measuring the ratio of the outgoing velocity of the ball to the incoming velocity can quantify this spring-like effect. The ratio of velocities is called the coefficient of restitution of the club. An experiment was performed in which 15 drivers produced by a particular club maker were selected at random and their coefficients of restitution measured. In the experiment the golf balls were fired from an air cannon so that the incoming velocity and spin rate of the ball could be precisely controlled. It is of interest to determine if there is evidence (with $\alpha = 0.05$) to support a claim that the mean coefficient of restitution exceeds 0.82. The observations follow:
- **0.8411 0.8191 0.8182 0.8125 0.8750 0.8580 0.8532 0.8483 0.8276 0.7983 0.8042 0.8730 0.8282 0.8359 0.8660**
- The sample mean is 0.83725 and s = 0.02456. The normal probability plot of the data in Fig. 9-9 supports the assumption that the coefficient of restitution is normally distributed. Since the objective of the experimenter is to demonstrate that the mean coefficient of restitution exceeds 0.82, a one-sided alternative hypothesis is appropriate.

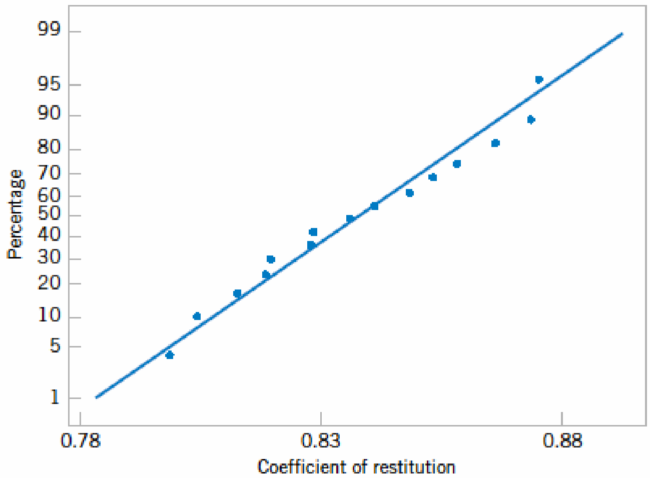


Figure 9-9. Normal probability plot of the coefficient of restitution data from Example 9-6.

1. The parameter of interest is the mean coefficient of restitution, μ.
2. H_0: $\mu = 0.82$
3. H_1: $\mu > 0.82$. We want to reject H_0 if the mean coefficient of restitution exceeds 0.82.
4. $\alpha = 0.05$
5. The test statistic is

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

6. Reject H_0 if $t_0 > t_{0.05,14} = 1.761$
7. Computations: Since $\bar{x} = 0.83725$, $s = 0.02456$, $\mu_0 = 0.82$, and $n = 15$, we have

$$t_0 = \frac{0.83725 - 0.82}{0.02456/\sqrt{15}} = 2.72$$

8. Conclusions: Since $t_0 = 2.72 > 1.761$, we reject H_0 and conclude at the 0.05 level of significance that the mean coefficient of restitution exceeds 0.82.

To illustrate, consider the t-test based on 14 degrees of freedom in Example 9-6. The relevant critical values from Appendix Table IV are as follows:

Critical Value:	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
Tail Area:	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005

Notice that $t_0 = 2.72$ in Example 9-6, and that this is between two tabulated values, 2.624 and 2.977. Therefore, the P-value must be between 0.01 and 0.005. These are effectively the upper and lower bounds on the P-value.

Example 9-6 is an upper-tailed test. If the test is lower-tailed, just change the sign of t_0 and proceed as above. Remember that for a two-tailed test the level of significance associated with a particular critical value is twice the corresponding tail area in the column heading. This consideration must be taken into account when we compute the bound on the P-value. For example, suppose that $t_0 = 2.72$ for a two-tailed alternate based on 14 degrees of freedom. The value $t_0 > 2.624$ (corresponding to $\alpha = 0.02$) and $t_0 < 2.977$ (corresponding to $\alpha = 0.01$), so the lower and upper bounds on the P-value would be $0.01 < P < 0.02$ for this case.

HYPOTHESIS TESTS ON THE VARIANCE AND STANDARD DEVIATION OF A NORMAL POPULATION

Null hypothesis $H_0 : \sigma^2 = \sigma_0^2$

Test Statistic value: $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \sigma^2 > \sigma_0^2$	$\chi_0^2 \geq \chi_{\alpha,\text{n-1}}^2$
$H_1 : \sigma^2 < \sigma_0^2$	$\chi_0^2 \leq -\chi_{\alpha,\text{n-1}}^2$
$H_1 : \sigma^2 \neq \sigma_0^2$	either $\chi_0^2 \geq \chi_{\alpha/2,\text{n-1}}^2$ or $\chi_0^2 \leq -\chi_{\alpha/2,\text{n-1}}^2$

- An automatic filling machine is used to fill bottles with liquid detergent. A random sample of 20 bottles results in a sample variance of fill volume of $s^2 = 0.0153$ (fluid ounces)^2. If the variance of fill volume exceeds 0.01 (fluid ounces)^2, an unacceptable proportion of bottles will be underfilled or overfilled. Is there evidence in the sample data to suggest that the manufacturer has a problem with underfilled or overfilled bottles? Use $\alpha = 0.05$, and assume that fill volume has a normal distribution.

1. The parameter of interest is the population variance σ^2.
2. H_0: $\sigma^2 = 0.01$
3. H_1: $\sigma^2 > 0.01$
4. $\alpha = 0.05$
5. The test statistic is

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

6. Reject H_0 if $\chi_0^2 > \chi_{0.05,19}^2 = 30.14$.
7. Computations:

$$\chi_0^2 = \frac{19(0.0153)}{0.01} = 29.07$$

8. Conclusions: Since $\chi_0^2 = 29.07 < \chi_{0.05,19}^2 = 30.14$, we conclude that there is no strong evidence that the variance of fill volume exceeds 0.01 (fluid ounces)^2.

Using Appendix Table III, it is easy to place bounds on the P-value of a chi-square test. From inspection of the table, we find that $\chi^2_{0.10,19} = 27.20$ and $\chi^2_{0.05,19} = 30.14$. Since $27.20 < 29.07 < 30.14$, we conclude that the P-value for the test in Example 9-8 is in the interval $0.05 < P < 0.10$. The actual P-value is $P = 0.0649$. (This value was obtained from a calculator.)

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: $z_0 = \dfrac{X - np_0}{\sqrt{np_0(1 - p_0)}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : p > p_0$	$z_0 \geq z_\alpha$
$H_1 : p < p_0$	$z_0 \leq -z_\alpha$
$H_1 : p \neq p_0$	either $z_0 \geq z_{\alpha/2}$ or $z_0 \leq -z_{\alpha/2}$

Test on Proportions: Small Sample

- Based directly on the binomial distribution.

P-value Rules

- If the *P*-value is less than
  1. 0.10, we have **some** evidence that H_0 is not true.
  2. 0.05, we have **strong** evidence that H_0 is not true.
  3. 0.01, we have **very strong** evidence that H_0 is not true.
  4. 0.001, we have **extremely strong** evidence that H_0 is not true.