Dimensionless Technologies Private Limited
Visit us at: www.dimensionless.in
info@dimensionless.in
9923170071, 8108094992

Notes For Students - Lesson 9

Hypothesis Testing

2 friends race against each other. One of the friends A won the game for two days in a row. The other friend speculates if he is cheating by taking some energy enhancing drug.

Q. How likely is it that A won the games by chance?

Probability that A wins for two times in a row = 0.25

So, there is around 25% chance that A won by random chance. Would you think he took drugs?

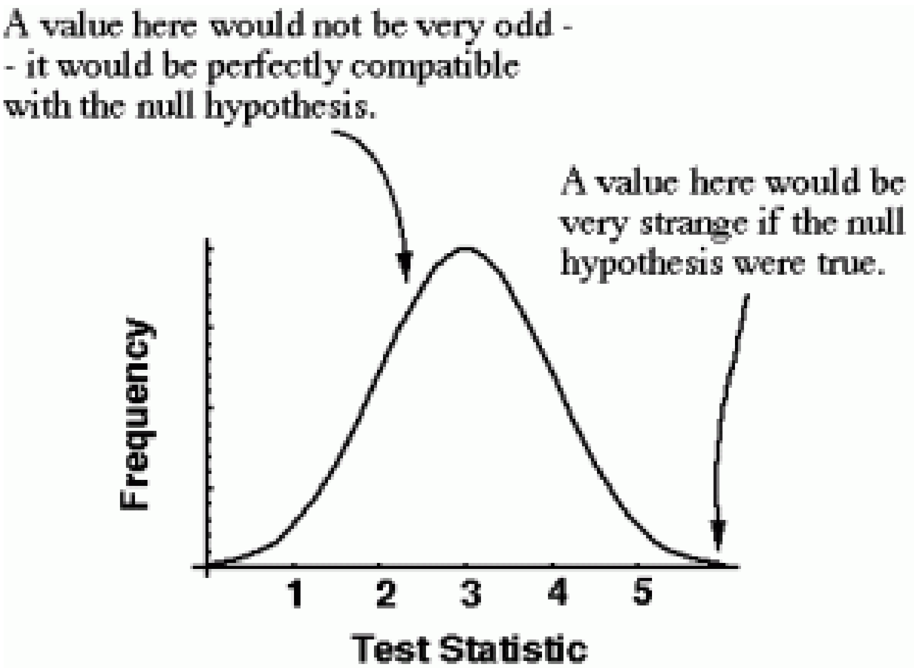
Suppose A didn't lose for 7 days straight.

Q. How likely is it that A didn't lose by random chance?

Probability that A didn't lose for seven times = 0.0078

So, there is 0.78% probability that A won by chance.

Q. Would you conclude that A had taken some drug or any other measure?



Q. Experts have designed a new additive to increase the time taken by a packed juice bottle to expire. Let μ denote the true average time when the bottle expires.

n = 25 **μ = 70.8 min** **σ = 9 min**

A random sample from this sampling distribution is tested for the additive and the sample mean was found to be 75 min.

Q. What will be the null and alternate hypothesis?

$H_o : \mu = 70.8$ Null Hypothesis is the claim that is initially assumed to be true

$H_a : \mu > 70.8$ The alternate hypothesis is the assertion that is contradictory to the null hypothesis

Null Hypothesis : **$H_o : \mu = \mu_o$**

Alternate Hypothesis

$H_a : \mu \neq \mu_0$	**two tailed**
$H_a : \mu > \mu_0$	**right tailed**
$H_a : \mu < \mu_0$	**left tailed**

Observation → **Hypothesis** → **Prediction**

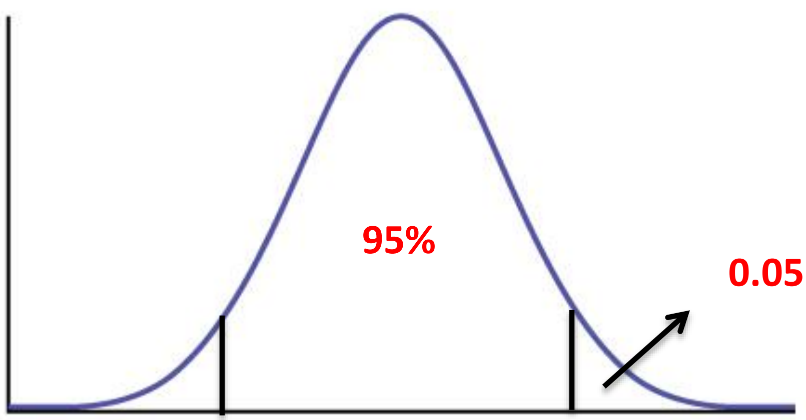
Q. What is the z score value?

Test Statistic value :

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

= 2.33

Q. What is the z critical value for a significance level of 5%?



1.96

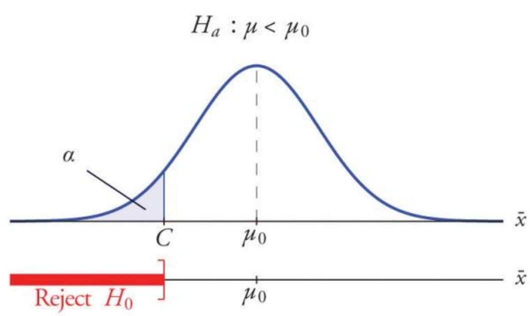
Q. Should we retain or reject Ho for a significance level of 5%?

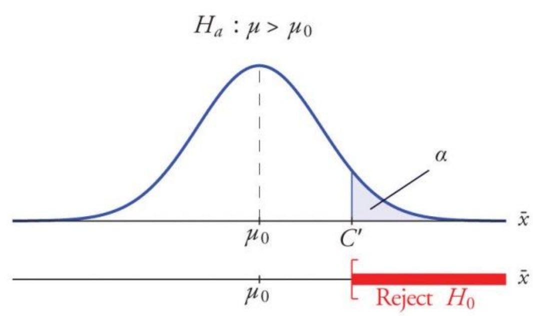
Reject

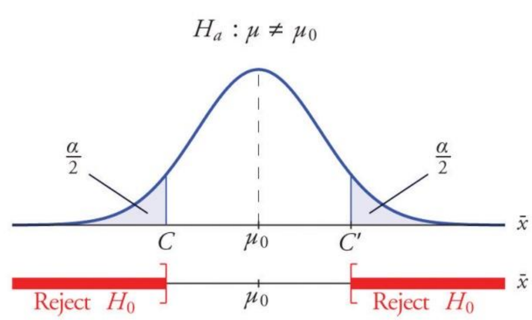
Do Not Reject

Q. What conclusion can we draw from this observation?

Alternate Hypothesis	**Rejection Region**	
Ha : $\mu > \mu_0$	**$z \geq z_\alpha$**	**(upper-tail)**
Ha : $\mu < \mu_o$	**$z \leq z_\alpha$**	**(lower-tail)**
Ha : $\mu \neq \mu_o$	**$z \geq z_\alpha$ or $z \leq z_\alpha$**	**(two-tail)**







Q. A cigarette manufacturer claims that the average nicotine content µ of brand B cigarettes is at most 1.5 mg. Govt. wants to decide if it is true or not by analyzing a random sample of 32 cigarettes.

σ = 0.2

$\overline{X}$ = **1.6**

Q. What will be the null and alternate hypothesis?

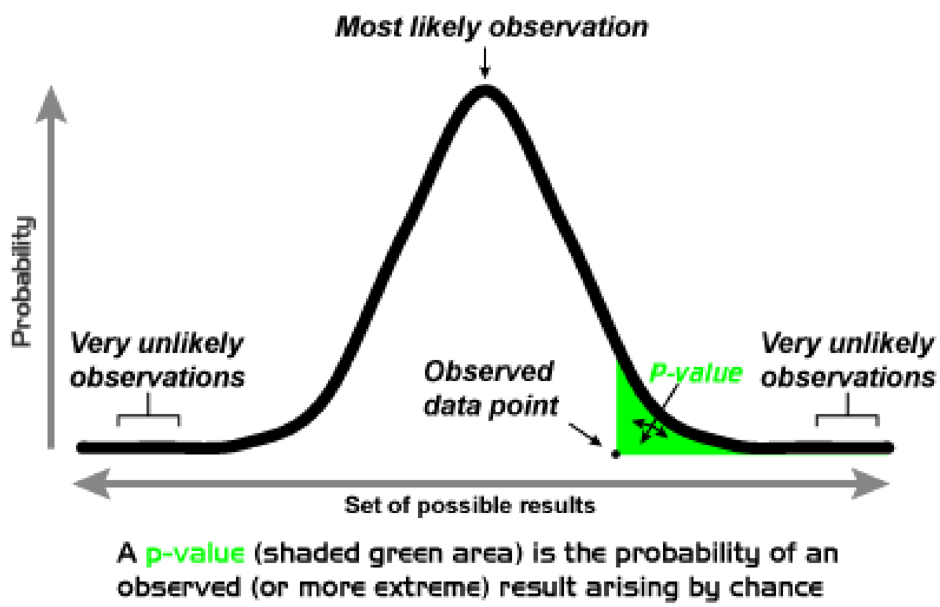
H_o : μ = 1.5

H_a : μ > 1.6

P value

Earlier we tested our hypothesis by computing value of test statistic, H_o being rejected if the value falls in the rejection region.

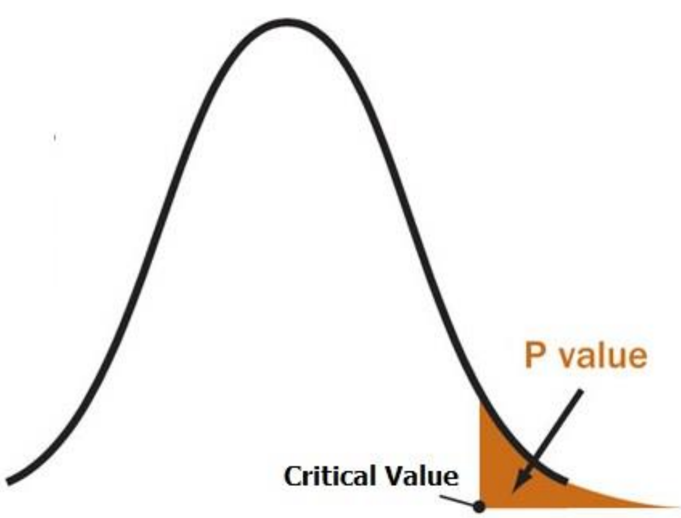
We now use an alternative approach of reaching a conclusion in a hypothesis testing analysis by calculation of certain probability called P-value.



Q. What is the value of z statistic?

z = 2.8248

Q. What is the P value?



= 0.002

Q. Would you reject or retain the null hypothesis?

Reject

Retain

P value

Reject H_o if P-value ≤ α

Do not Reject H_o if P-value > α

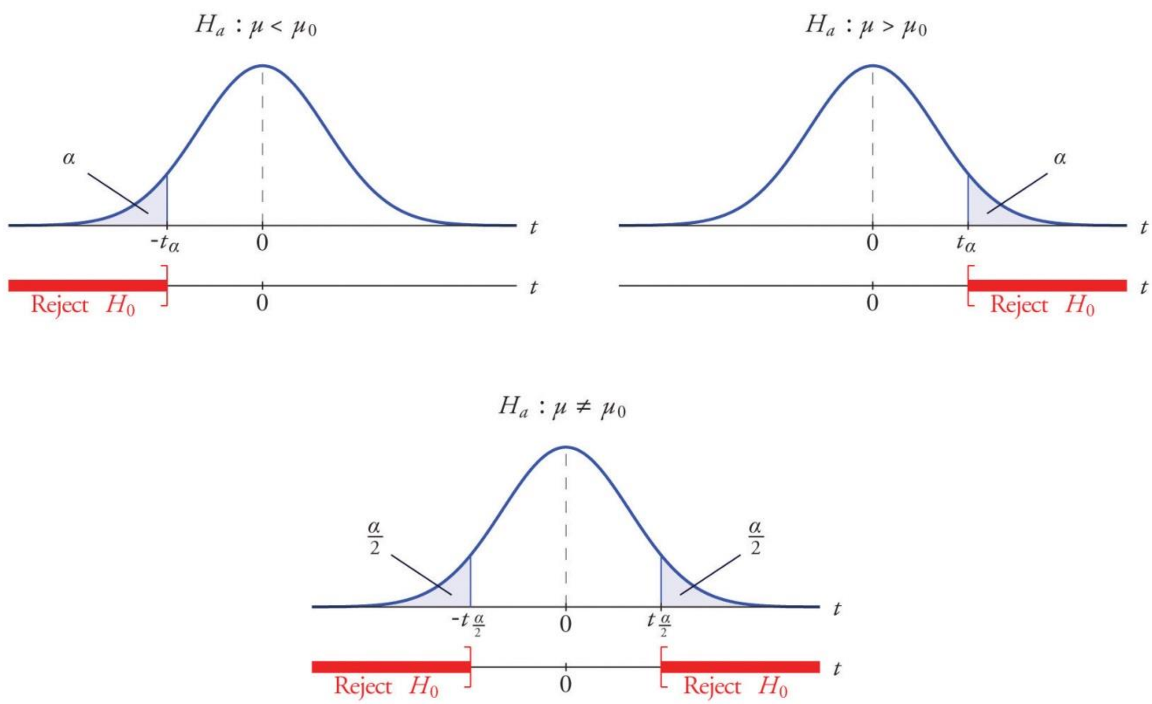
Calculating P value on the basis of z test

$1 - \phi(z)$ (upper tail)

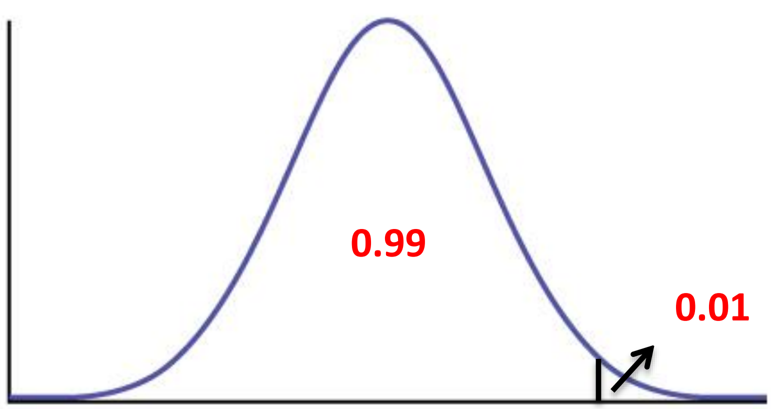
$\phi(z)$ (lower tail)

$2[1 - \phi(z)]$ (two-tail)

Dimensionless Technologies Private Limited
Visit us at: www.dimensionless.in
info@dimensionless.in
9923170071, 8108094992



Q. Would you reject it at a confidence level of 99% as well?



Use of Excel Functions for Hypothesis Testing

Use NORM.DIST / NORMDIST to calculate the P-value given the values of a normal distribution.

Use NORM.S.DIST / NORMSDIST to calculate the P-value given the z value.

Use NORM.INV / NORMINV to calculate x variable given the P-value.

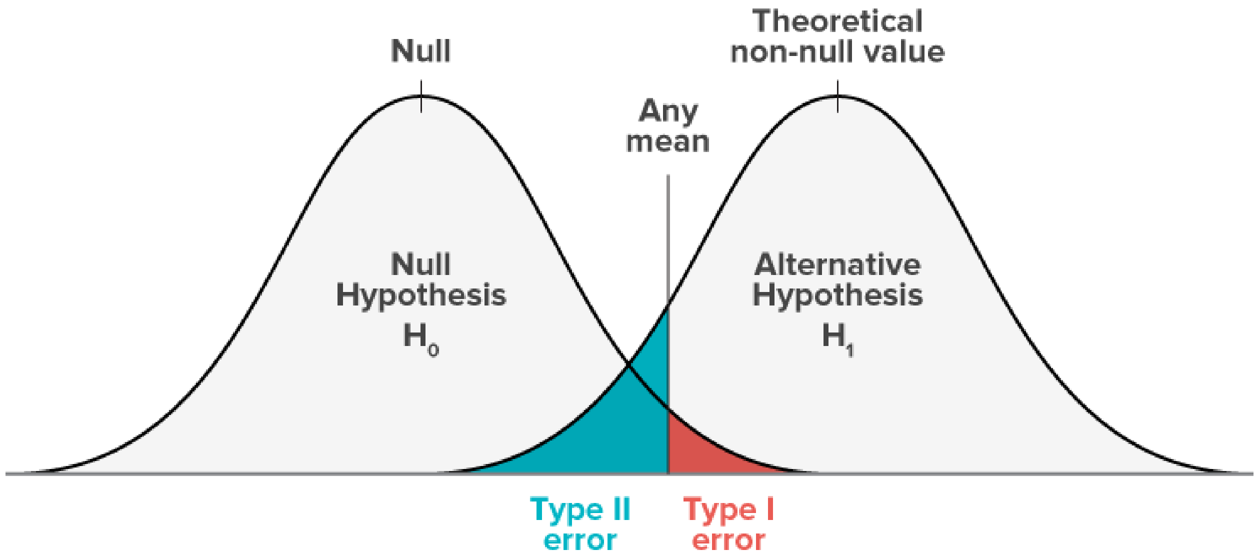
Use NORM.S.INV / NORMSINV to calculate the z value given the probability that the variable is within a certain distance of the mean.

Errors in Hypothesis Testing

Consider a rejection region $z \geq 1.645$. Even when $H_o : \mu = 10$ is true, it might happen that an unusual sample results in $z = 2$, so H_o is erroneously rejected. On the other hand, even when $H_a : \mu < 10$ is true, an unusual sample might yield $z = 1.4$, in which case H_o would not be rejected. Thus, it is possible that H_o may be rejected when it is true or may not be rejected when it is false.

Type I Error consists of rejecting H_o when it is true (**False Positive**)

Type II Error consists of not rejecting H_o when it is false (**False negative**)



Test procedures for neither type of error could be achieved only when we examine the entire population which is often not possible.

So, instead of demanding error free procedures, we must seek procedures for which the probability of making either type of error is small.

Q. What do you think which error would be more serious between the two?

Decision

	Reject H_o	**Retain H_o**
H_o is True	**Type I Error**	**No Error**
H_o is False	**No Error**	**Type II Error**

Dimensionless Technologies Private Limited
Visit us at: www.dimensionless.in
info@dimensionless.in
9923170071, 8108094992

Q. A sample of 51 batteries were checked to see how much zinc mass they release.

$\overline{X}$ **= 2.06 g**

Sigma = 0.141 g

Does this data provide compelling evidence for concluding that the population mean zinc mass exceeds 2.0 g?

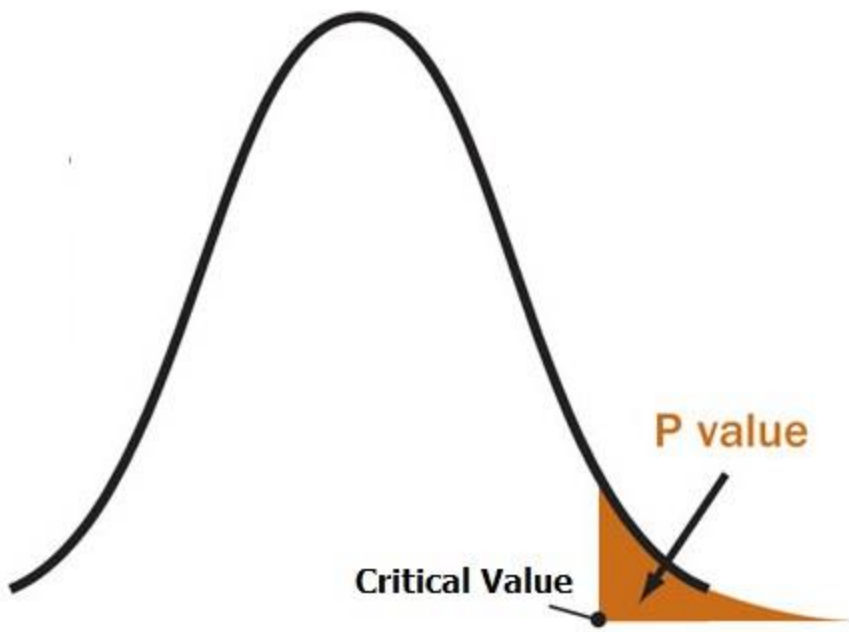
$\mathbf{H_o}$: $\mu = 2$

$\mathbf{H_a}$: $\mu > 2$

Q. What is the z value?

3.04

Q. What is the P value?



P value = P (z ≥ 3.04 when μ = 2)

= 1 – ϕ (3.04)

= 0.0012

Q. Would you reject or retain the null hypothesis?

Reject

Do not reject

Q. Determine if you reject or retain the null hypothesis for a significance level of 99% on the basis of following information.

$H_0 : \mu = 245$

$H_a : \mu \neq 245$

n = 50 , $\overline{X}$ = 246.18 , sigma = 3.6

Q. Would your decision be the same if significance level was 95%?

Yes

No

Q Give the significance level for each of the following:

a) $H_a : \mu > \mu_o$, rejection region $z \geq 1.88$

b) $H_a : \mu > \mu_o$, rejection region $z \leq -2.75$

c) $H_a : \mu \neq \mu_o$, rejection region $z \geq 2.88$ or $z \leq -2.88$

Q. For which of the given P values would the H_o be rejected when performing a level 0.05 test?

a) **0.001**

b) **0.021**

c) **0.078**

Q. Find the P value associated with the z test statistic.

a) **1.42**

b) **0.90**

c) **1.96**

Q. To investigate if the power plants are complying with the regulation of the water discharged in the river to be at most 150^0 F, 50 water samples are taken (at different times) and temperature is recorded to make an analysis.

Q. What will be the null and alternate hypothesis?

Q. Which type of error do you consider serious for the above problem?

Q. What is the range of values that P value can take?

0 to 1

Q. Consider a case as follows-

$H_o : \mu = 100$

$H_a : \mu > 100$

sigma = 10

Suppose a true value of $\mu = 101$ would not represent a serious departure from $H_{o.}$ So, with a sample mean average of 101, we would not want the null hypothesis to be rejected.

Q. For n = 25, what is the P value?

Q. Will you reject it at 5% significance level?

Q. Would you reject it for n = 400 as well?

Q. And for n = 2500?

When sample size is large, any small departure from H_o will almost surely be detected by a test, yet such a departure may have little practical significance.