SAMPLING DISTRIBUTIONS

Definition: The collection of all objects, animate or inanimate, under statistical study is called Population.

Examples: (i) The students of Bapatla Engineering College

(ii) Cars produced by Maruthi Suzuki in an year.

Definition: If the number of objects in a population is finite then it is called a finite population otherwise it is called an infinite population. In a finite population, the number of objects is denoted by N.

Definition: Any statistical quantity that is measured from the population is called a parameter.

Example: The population means, the population variance etc.

Definition: Any finite subset of the population is called a Sample.

Example: (i) The students of I B.Tech in BEC

(ii) Cars produced by Maruthi Suzuki in the month of January, 2020.

Definition: Any statistical quantity which is computed from the sample is called a Statistic.

Example: The sample mean, the sample variance etc.

Definition: The process of obtaining a sample from a population is called Sampling. The sample size is denoted by n. If $n \geq 30$, the sample is called large sample and if $n < 30$, the sample is called small sample.

Definition: A set of observations $X_1, X_2, \ldots, X_n$ is said to be a random sample of size n from a population of size N if each member of the population has equal probability of being included in the sample.

Definition: A set of observations $X_1, X_2, \ldots, X_n$ constitute a random sample of size n from the infinite population f(x) if

(i) Each X_i is a random variable whose distribution is given by f(x).

(ii) These ‘n’ random variables are independent.

The Sampling Distribution of the Mean (σ known)

Theorem: If a random sample of size n is taken from a population having the mean μ and the variance σ^2, then $\overline{X}$ is a random variable whose distribution has the mean μ.

For the samples from infinite population the variance is $\frac{\sigma^2}{n}$

For samples from a finite population of size N the variance is $\frac{\sigma^2}{n}\frac{N-n}{N-1}$.

Note: The quantity $\frac{N-n}{N-1}$ is called the population correction factor and is close to one.

Law of Large numbers: Let $X_1, X_2, \ldots, X_n$ be independent random variables each having the same mean μ and variance σ^2. Then

$$P\left(\left|\overline{X}-\mu\right| > \varepsilon\right) \to 0 \ \ as \ \ n \to \infty$$

As the sample size increases, the probability that the sample mean differs from the population mean by more than arbitrary amount converges to zero.

Central limit theorem: If $\overline{X}$ is the mean of a random sample of size n taken from a population having the mean μ and the finite variance σ^2, then

$$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

is a random variable whose distribution function approaches that of the standard normal distribution as $n \to \infty$

Note :

(i) Here the quantity $\sigma/\sqrt{n}$ is called standard error of the mean and is denoted by $\sigma_{\overline{X}}$.

(ii) If the random samples come from normal population, the sampling distribution of the mean is normal regardless of the size of the sample.

(iii) The normal distribution provides an excellent approximation to the sampling distribution of the mean $\overline{X}$ for n as small as 25 or 30.

Problem: Find the value of the finite population correction factor for n = 10 and N = 1000.

Solution: The population correction factor is given by $\frac{N-n}{N-1} = \frac{1000-10}{1000-1} = 0.991$

Problem: If a 1 – gallon can of paint covers on the average 513.3 square feet with a standard deviation of 31.5 square feet, what is the probability that the sample mean area covered by a sample of 40 of these 1 – gallon cans will be anywhere from 510.0 to 520.0 square feet.

Solution: Given μ = 513.3, n = 40, σ = 31.5

We have to find out P(510 < $\overline{X}$ < 520). We will use Central limit theorem.

Now, $P\left(510 < \overline{X} < 520\right) = P\left(\frac{510-513.3}{31.5/\sqrt{40}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{520-513.3}{31.5/\sqrt{40}}\right)$

= P(-0.66 < Z < 1.34)

= F(1.34) – F(-0.66)

= F(1.34) – (1-F(0.66))

= F(1.34) – 1+F(0.66)

= 0.9099 – 1 + 0.7454 (Using table 3)

= 0.6553.

Problem: When we sample from an infinite population, what happens to the standard error of the mean if the sample size is

(i) Increased from 50 to 200

(ii) Increased from 400 to 900

(iii) Decreased from 225 to 25

(iv) Decreased from 640 to 40.

Solution: Standard error of the mean is $\sigma/\sqrt{n}$

(i) When the sample size is 50, $\sigma/\sqrt{n} = \sigma/\sqrt{50}$

When the sample size is 200, $\sigma/\sqrt{n} = \sigma/\sqrt{200}$

$= \sigma/\sqrt{4*50}$

$= \sigma / 2\sqrt{50}$

$= (1/2)\ \sigma / \sqrt{50}$

When the sample size increased from 50 to 200 the standard error of the mean is divided by 2.

(ii) When the sample size is 400, $\sigma / \sqrt{n} = \sigma / \sqrt{400}$

$= \sigma / 20$

When the sample size is 900, $\sigma / \sqrt{n} = \sigma / \sqrt{900}$

$= \sigma / 30$

$= (20/30)\ \sigma / 20$

$= (2/3)\ \sigma / 20$

$= (\sigma / 20)/(3/2)$

When the sample size increased from 400 to 900, the standard error of the mean is divided by (3/2).

(iii) When the sample size is 225, $\sigma / \sqrt{n} = \sigma / \sqrt{225}$

$= \sigma / 15$

When the sample size is 25, $\sigma / \sqrt{n} = \sigma / \sqrt{25}$

$= \sigma / 5$

$= 3\ (\sigma / 15)$

When the sample size is decreased from 225 to 25 the standard error of the mean is multiplied by 3.

(iv) When the sample size is 640, $\sigma / \sqrt{n} = \sigma / \sqrt{640}$

$= \sigma / 8\sqrt{10}$

When the sample size is 40, $\sigma / \sqrt{n} = \sigma / \sqrt{40}$

$= \sigma / 2\sqrt{10}$

$= 4\ (\sigma / 8\sqrt{10})$

When the sample size is decreased from 640 to 40 the standard error of the mean is multiplied by 4.

Problem: What is the value of the finite population correction factor when

(i) n = 5 and N = 250

(ii) n = 10 and N = 500

(iii) n = 100 and N = 5000

Solution:

(i) When n = 5 and N = 250, $\frac{N-n}{N-1} = \frac{250-5}{250-1} = \frac{245}{249} = 0.9839$

(ii) When n = 10and N = 500, $\frac{N-n}{N-1} = \frac{500-10}{500-1} = 0.9819$

(iii) When n = 100 and N = 5000, $\frac{N-n}{N-1} = \frac{5000-100}{5000-1} = 0.9801$

Problem: Hard disks for computers must spin evenly, and one departure from level is called roll. The roll for any disk can be modeled as a random variable having mean 0.2250 mm and standard deviation 0.0042 mm. The sample mean roll $\overline{X}$ will be obtained from a random sample of 40 disks. What is the probability that $\overline{X}$ will lie between 0.2245 and 0.2260 mm.

Solution: Given μ = 0.2250, n = 40, σ = 0.0042

We have to find out P(0.2245 < $\overline{X}$ < 0.2260). We will use Central limit theorem.

Now, P(0.2245 < $\overline{X}$ < 0.2260) $= P(\frac{0.2245-0.2250}{0.0042/\sqrt{40}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{0.2260-0.2250}{0.00425/\sqrt{40}})$

= P(-0.753 < Z < 1.506)

= F(1.506) – F(-0.753)

= F(1.506) – (1-F(0.753))

= F(1.506) – 1+F(0.753) (average of .75 and .76

= 0.9332 – 1 + 0.7734 is .755)

= 0.7066

Problem: A wire – bonding process is said to be control if the mean pull strength is 10 pounds. It is known that the pull-strength measurements are normally distributed with a standard deviation of 1.5 pounds. Periodic random samples of size 4 are taken from this process and the process is said to be "out of control" if a sample mean is less than 7.75 pounds. Comment.

Solution: Given the wire-bonding process is under control if the mean pull strength is 10 pounds.

So, μ = 10 pounds, n = 4, σ = 1.5 pounds

Since the random sample is chosen from a normal population, even though the sample size is small, we can apply Central limit theorem.

We have to find out $P(\overline{X} < 7.75) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{7.75-10}{1.5/\sqrt{4}}\right)$

$= P(Z < -3)$

$= F(-3)$

$= 1 - F(3)$

$= 1 - 0.9987$ (Using table 3)

$= 0.0013$

Since the probability that the sample mean is less than 7.75 is very small, we may assume that the process is under control.

Problem: If the distribution of the weights of all men traveling by air between Dallas and EI Paso has a mean of 163 pounds and standard deviation of 18 pounds, what is the probability that the **combined gross weight** of 36 men traveling on a plane between these two cities is more than 6000 pounds.

Solution: Given μ = 163 pounds, σ = 18 pounds, n = 36

Let X_i denote the gross weight of i^{th} person, where i = 1,2,…,36.

We have to find out $P(X_1+X_2+\ldots+X_{36} > 6000) = P\left(\sum_{i=1}^{36} X_i > 6000\right)$

$$= P\left(\frac{\frac{(\sum_{i=1}^{36} X_i)}{n} - \mu}{\sigma/\sqrt{n}} > \frac{\frac{6000}{36} - 163}{18/\sqrt{36}}\right)$$

$$= P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} > \frac{166.666 - 163}{18/\sqrt{36}}\right)$$

$$= P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} > \frac{166.67 - 163}{3}\right)$$

$$= P(Z > 1.223)$$

= 1 – F(1.223)

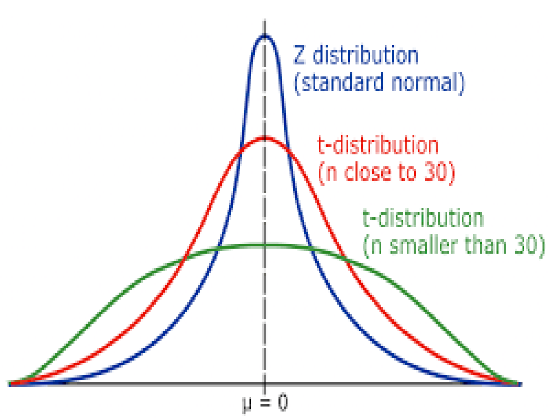
= 1- 0.8888 (avg of 1.22 & 1.23)

= 0.1112

The sampling Distribution of the Mean (σ unknown)

If the sample size, n is large then there won't be any problem even if we are not given σ. We can replace σ by s, the sample standard deviation. But for small samples (n < 30) the unknown σ can be substituted by s, provided the sample is drawn from a normal population.

Theorem: If $\overline{X}$ is the mean of a random sample of size n taken from a normal population having the mean μ and variance σ^2, and $S^2 = \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{n-1}$, then $t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$ is a random variable having the t distribution with the parameter υ = n-1.



The t- distribution curve is symmetric about the mean 0, bell shaped. It is similar to normal curve. The variance of the t-distribution depends on the parameter υ, called the number of degrees of freedom. The variance of the t-distribution exceeds 1 and it approaches to as n approaches to ∞. Here t_α is such that the area under the t distribution to its right is equal to α. By the symmetry of t distribution $t_{1-\alpha} = - t_\alpha$

Problem: A manufacturer of fuses claims that with a 20% overload, the fuses will blow in 12.40 minutes on the average. To test this claim, a sample of 20 of the fuses was subjected to a 20% overload, and the times it took them to blow had a mean of 10.63 minutes and a standard deviation of 2.48 minutes. If it can be assumed that the data constitute a random sample from a normal population, do they tend to support or refute the manufacturer's claim.

Solution: Given that Given μ = 12.40 minutes, $\bar{x}$ = 10.63 minutes, s = 2.48 minutes pounds and n = 20.

Since the data is taken from a normal population and the sample size is less than 30, we will use t distribution.

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

$$= \frac{10.63 - 12.40}{2.48/\sqrt{20}}$$

$$= -3.19$$

Which is a value of a random variable having t distribution with υ = 20-1=19 degrees of freedom.

We know that $P(t > t_\alpha) = \alpha$ and therefore $P(t < - t_\alpha) = \alpha$ (by the symmetric property)

From the table P(t > 2.861) = 0.005 for υ =19 degrees of freedom.

So, P(t <- 2.861) = 0.005. Since the calculated t value t = -3.19 is less than -2.861 and the probability that t will be less than -2.861 is 0.005 which is very small, we conclude that the data tend to refute the manufacturer's claim. If the calculated t value is either greater than t_α or less than - t_α we have to reject the claim.

Problem: The tensile strength (1000 psi) of a new composite can be modeled as a normal distribution. A random sample of size 25 specimens has mean 45.3 and standard deviation 7.9. Does this information tend to support or refute the claim that the mean of the population is 40.5.

Solution: Given that n = 25, $\overline{x}$ = 45.3, μ = 40.5 and s = 7.9. Since the sample size is small and is taken from normal population, we will use t distribution.

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

$$= \frac{45.3 - 40.5}{7.9/\sqrt{25}}$$

$$= \frac{45.3 - 40.5}{7.9/5}$$

$$= 3.0379$$

Which is a value of a random variable having t distribution with υ = 25-1=24 degrees of freedom.

We know that $P(t > t_\alpha) = \alpha$ and therefore $P(t < - t_\alpha) = \alpha$ (by the symmetric property)

From the table P(t > 2.797) = 0.005 for υ = 24 degrees of freedom. Since the calculated t value, t = 3.0379 exceeds 2.797 with probability 0.005, we conclude that the data tend to refute the claim that the mean of the population is 40.5. If the calculated t value is either greater than t_α or less than - t_α we have to reject the claim.

Problem: The following are the times between 6 calls for an ambulance and the patient's arrival at the hospital: 27, 15, 20, 32, 18 and 26 minutes. Use these figures to judge the reasonableness of the ambulance service's claim that it takes on the average of 20 minutes between the call for an ambulance and the patient's arrival at the hospital.

Solution: For the given data, n = 6, $\overline{x}$ = (27 + 15 + 20 + 32 + 18 + 26)/6 = 23 and s = 6.39, μ = 20. Since the sample size is small, we will use t distribution.

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

$$= \frac{23 - 20}{6.39/\sqrt{6}}$$

$$= 1.15$$

Which is a value of a random variable having t distribution with υ = 6-1= 5 degrees of freedom.

We know that P(t > t_α) = α and therefore P(t < - t_α) = α (by the symmetric property). From the table P(t > 1.476) = 0.10 for υ = 5 degrees of freedom. Since the calculated t value, t = 1.15 does not exceed 1.476 with probability 0.10, we conclude that the data fail to reject the claim that the ambulance takes on the average of 20 minutes between the call for an ambulance and the patient's arrival at the hospital.

Problem: A process for making certain bearings is under control if the diameters of the bearings have mean of 0.5000 cm. What can we say about this process if a sample of 10 of these bearings has a mean diameter of 0.5060 cm and a standard deviation of 0.0040 cm.

Solution: Given that n = 10, $\overline{x}$ = 0.5060, s = 0.0040, μ = 0.5000. Since the sample size is small, we will use t distribution.

$$t = \frac{\overline{X} - \mu}{S / \sqrt{n}}$$

$$= \frac{0.5060 - 0.5000}{0.0040 / \sqrt{10}}$$

$$= 4.74$$

Which is a value of a random variable having t distribution with υ = 10-1= 9 degrees of freedom.

We know that $P(t > t_\alpha) = \alpha$ and therefore $P(t < - t_\alpha) = \alpha$ (by the symmetric property)

From the table P(t > 3.250) = 0.005 for υ = 9 degrees of freedom. Since the calculated t value, t = 4.74 exceeds 3.250 with probability 0.005, we conclude that the data tend to refute the claim that the diameters of the bearings have mean of 0.5000 cm. If the calculated t value is either greater than t_α or less than - t_α we have to reject the claim.

The Sampling Distribution of the Variance

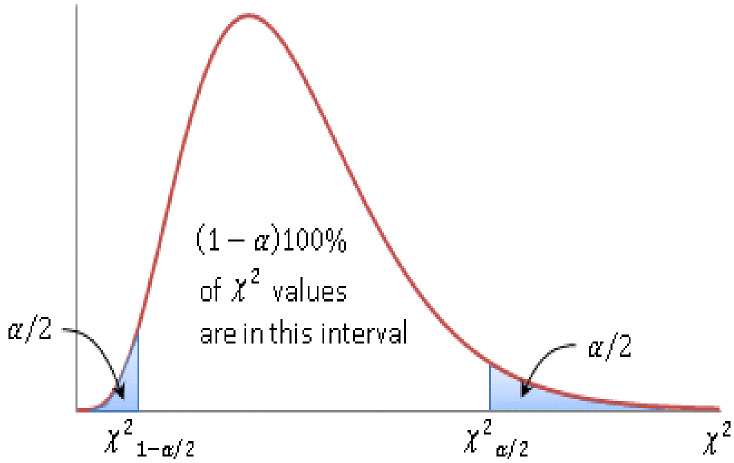
In the earlier discussion, we have used the sample mean to estimate the population mean. Similarly, the sample variance will be used to estimate the population variance.

Theorem: If S^2 is the variance of a random sample of size n taken from a normal population having the variance σ^2, then

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{\sigma^2}$$

is a random variable having the chi square distribution with parameter υ = n-1.

χ_α^2 is such that the area under the chi square distribution to its right is equal to α.



Note that chi square distribution is not symmetrical.

Problem: Hard disks for computers must spin evenly, and one departure from level is called pitch. Samples are regularly taken from production and each disk in the sample is placed in test equipment that yields a measurement of pitch. From many samples, it is concluded that the population is normal. The variance σ^2 = 0.065 when the process is in control. A sample of size 10 is collected each week. The process will be declared out of control if the sample variance exceeds 0.122. What is the probability that it will be declared out of control even though σ^2 = 0.065.

Solution: Given n = 10, s^2 = 0.122 and σ^2 = 0.065.

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

$$= \frac{(10-1)0.122}{0.065}$$

$$= 16.89$$

Then from table 5 that for 9 degrees of freedom, $\chi^2_{0.05} = 16.919$.

$P(S^2 > 0.122) = (\chi^2 > 16.89) = 0.05$.

Problem: A random sample of 10 observations is taken from a normal population having the variance σ^2 = 42.5. Find the approximate probability of obtaining a sample standard deviation between 3.14 and 8.94.

Solution: Given n = 10, σ^2 = 42.5.

We have to find out $P(3.14 < S < 8.94)$

But $P(3.14^2 < S^2 < 8.94^2) = P(9(3.14)^2/ 42.5< ((n-1)S^2/\sigma^2)< 9(8.94)^2/42.5)$

$= P(2.0879 < \chi^2 < 16.9249)$

$= P(2.088 < \chi^2 < 16.925)$

$= P(\chi^2 > 2.088) - P(\chi^2 > 16.925)$

$= 0.95 – 0.01$ (from Table 5 with $\upsilon = 9$)

$= 0.94$

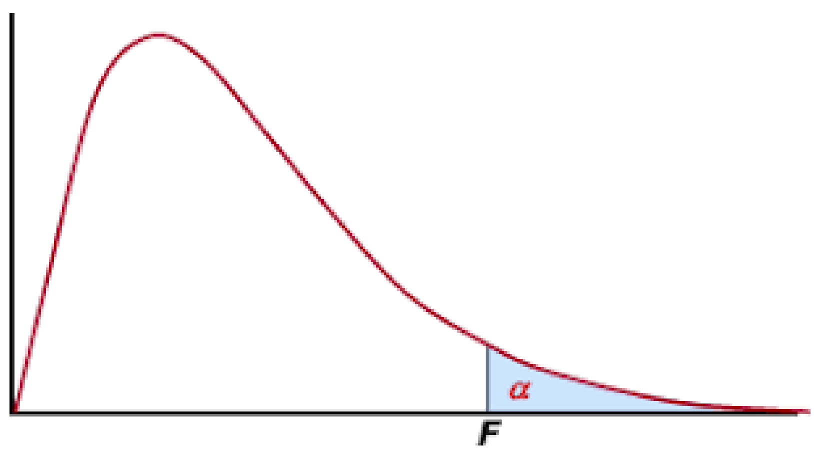
Here note that $P(\chi^2 > \chi^2_\alpha) = \alpha$ and $P(\chi^2 \leq \chi^2_\alpha) = 1- \alpha$.

A random variable having the F distribution:

Theorem: If S_1^2 and S_2^2 are the variances of independent random samples of size n_1 and n_2, respectively, taken from two normal populations having the same variance, then

$$F = \frac{S_1^2}{S_2^2}$$

Is a random variable having the F distribution with the parameters $\upsilon_1 = n_1-1$ and $\upsilon_2 = n_2-1$.



Note that $F_{1-\alpha}(\upsilon_1 , \upsilon_2) = (F_\alpha(\upsilon_2 ,\upsilon_1))^{-1}$

Problem: If two independent random samples of size $n_1 = 7$ and $n_2 = 13$ are taken from a normal population, what is the probability that the variance of the first sample will be at least three times as large as that of the second sample?

Solution: Given $n_1 = 7$ and $n_2 = 13$.

We have to find out $P(S_1^2 \geq 3S_2^2)$.

But, $P(S_1^2 > 3S_2^2) = P(F \geq 3)$

$= 0.05$ (From Table 6 with $\upsilon_1 = 7-1 = 6$ and $\upsilon_2 = 13-1 = 12$)

Problem: If independent random samples of size $n_1 = n_2 = 8$ come from normal populations having the same variance, what is the probability that either sample variance will be at least 7 times as large as the other.

Solution: Given $n_1 = n_2 = 8$. Both the samples come from normal populations having the same variance.

We have to find out $P(\frac{S_1^2}{S_2^2} > 7 \text{ or } \frac{S_2^2}{S_1^2} > 7)$

So, $P(\frac{S_1^2}{S_2^2} > 7 \text{ or } \frac{S_2^2}{S_1^2} > 7) = P(\frac{S_1^2}{S_2^2} > 7) + P(\frac{S_2^2}{S_1^2} > 7)$

$= 0.01 + 0.01 = 0.02$

Problem: Find the values of

(a) $F_{0.95}$ for 12 and 15 degrees of freedom;
(b) $F_{0.99}$ for 6 and 20 degrees of freedom.

Solution:

(a) $F_{0.95}(12, 15) = (F_{0.05}(15,12))^{-1}$
$= (2.62)^{-1}$
$= 0.38167.$

(b) $F_{0.99}(6, 20) = (F_{0.01}(20,6))^{-1}$
$= (7.40)^{-1}$
$= 0.13513.$

Problem: The chi square distribution with 4 degrees of freedom is given by

$$f(x) = \begin{cases} \frac{1}{4} x e^{-x/2} & \text{for } x > 0 \\ 0 & \text{for } x \leq 0 \end{cases}$$

Find the probability that the variance of a random sample of size 5 from a normal population with $\sigma = 12$ will exceed 180.

Solution: Let X be a random variable having Chi square distribution with 4 degrees of freedom. Given that X follow the probability density function,

$$f(x)=\begin{cases}\frac{1}{4}xe^{-x/2} & for\ x>0\\ 0 & for\,x\le 0\end{cases}$$

We have to find out $P(S^2 > 180)$

Now $P(S^2 > 180) = P(\chi^2 > 4(180)/12^2)$

$= P(\chi^2 > 5)$

$$=\int_5^\infty f(x)\,dx$$

$$=\int_5^\infty \frac{1}{4}xe^{-x/2}\,dx$$

= 0.2873 (by using integration byparts)

Point Estimation: The process of obtaining a single value for finding an approximate value of the population parameter θ from random samples of the population is called point estimation. A point estimator is a statistic used for estimating a population parameter θ and will be denoted by $\hat{\theta}$.

Example: The sample mean $\overline{X}$ is the point estimator for the population mean μ.

Unbiased estimator: A statistic $\hat{\theta}$ is said to be an unbiased estimator if and only if the mean of the sampling distribution of the estimator is equal to θ. That is $E(\hat{\theta}) = \theta$.

More efficient unbiased estimator: A statistic $\hat{\theta}_1$ is said to be a more efficient unbiased estimator of the parameter θ than the statistic $\hat{\theta}_2$ if

1. $\hat{\theta}_1$ and $\hat{\theta}_2$ are both unbiased estimators of θ.
2. $Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$.

Maximum error of estimate (σ known):

When we use a sample mean to estimate the mean of a population we will get an error as clearly both are not the same. The maximum error of estimate for large samples ($n \geq 30$) is given by the formula

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \text{ with probability } 1 - \alpha.$$

That is $P(|\overline{X} - \mu| \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1 - \alpha$

Problem 1: An industrial engineer intends to use the mean of a random sample of size n = 150 to estimate the average mechanical aptitude of assembly line workers in a large industry. If, on the basis of experience, the engineer can assume that σ = 6.2 for such data, what can he assert with probability 0.99 about the maximum size of his error.

Solution: Given n = 150, σ = 6.2 and 1 – α = 0.99. Then α = 0.01 and α/2 = 0.005.

The maximum error of estimate, $E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

$$= z_{0.005}\frac{6.2}{\sqrt{150}}$$

$$= 2.575\frac{6.2}{\sqrt{150}}$$

$$= 1.30.$$

Thus, the engineer can assert with probability 0.99 that his error will be at most 1.30.

One can obtain the sample size in such a way that the probability that the error will be at most $E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ is (1 – α)100%.

That is $\sqrt{n} = z_{\alpha/2}\frac{\sigma}{E} \Rightarrow n = (z_{\alpha/2}\frac{\sigma}{E})^2$

Problem 2: A research worker wants to determine the average time it takes a mechanic to rotate the tires of a car, and she wants to be able to assert with 95% confidence that the mean of her sample is off by at most 0.50 minute. If she can presume from past experience that σ = 1.6 minutes, how large a sample will she have to take.

Solution: Given E = 0.50, σ = 1.6, (1- α)100% = 95%.

1 – α = 0.95 and α = 0.05 so that α/2 = 0.025.

The size of the sample is given by n = $(z_{\alpha/2}\frac{\sigma}{E})^2$

$$= [z_{0.025}\frac{1.6}{0.50}]^2$$

$$= [1.96\frac{1.6}{0.50}]^2 = 39.3 \approx 40$$

Maximum error of estimate (σ unknown):

When σ is unknown, the maximum error of estimate is given by the formula $E = t_{\alpha/2}\frac{s}{\sqrt{n}}$.

That is we assert with (1 – α)100% confidence that the error made in using $\bar{x}$ to estimate μ is at most $E = t_{\alpha/2}\frac{s}{\sqrt{n}}$.

Problem 3: In six determinations of the melting point of tin, a chemist obtained a mean of 232.26 degrees Celsius with a standard deviation of 0.14 degree. If he uses this mean to estimate the actual melting point of tin, what can the chemist assert with 98% confidence about the maximum error.

Solution: Given n = 6, $\bar{x}$ = 232.26, σ = 0.14 and (1 – α)100% = 98%.

Than 1 – α = 0.98 so that α = 0.02 and α/2 = 0.01.

The maximum error of estimate, $E = t_{\alpha/2}\frac{s}{\sqrt{n}}$

$$= t_{0.01}\frac{0.14}{\sqrt{6}}$$

$$= 3.365\frac{0.14}{\sqrt{6}} \quad (t_{0.01}, 5 \text{ d.f} = 3.365)$$

$$= 0.19$$

Interval estimation: The process of obtaining a range of values with in which the the true value of the population parameter lies from random samples is called interval estimation.

Large sample confidence interval for μ (σ known)

If $\bar{x}$ is the sample mean of a random sample of size n (greater than or equal to 30) from a normal population with known variance σ^2 then the (1 – α)100% large sample confidence interval for the population parameter μ is given by

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

The end points of the above confidence interval are called the confidence limits.

Problem 4: A random sample of size n = 100 is taken from a population with σ = 5.1. Given that the sample mean is $\bar{x}$ = 21.6, construct a 95% confidence interval for the population mean μ.

Solution: Given n = 100, σ = 5.1, $\bar{x}$ = 21.6 and (1 – α)100% = 95%

Then α/2 = 0.025

The (1 – α)100% large sample confidence interval for the population parameter μ is given by

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow \quad 21.6 - z_{0.025}\frac{5.1}{\sqrt{100}} < \mu < 21.6 + z_{0.025}\frac{5.1}{\sqrt{100}}$$

$$\Rightarrow \quad 21.6 - 1.96\frac{5.1}{\sqrt{100}} < \mu < 21.6 + 1.96\frac{5.1}{\sqrt{100}}$$

$$\Rightarrow \quad 20.6 < \mu < 22.6$$

Small sample confidence interval for μ (σ unknown)

If $\bar{x}$ is the sample mean of a random sample of size n (less than 30) from a normal population with σ^2 unknown then the (1 – α)100% small sample confidence interval for the population parameter μ is given by

$$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}$$

Problem 5: The mean weight loss of n =16 grinding balls after a certain length of time in mill slurry is 3.42 grams with a standard deviation of 0.68 gram. Construct a 99% confidence interval for the true mean weight loss of such grinding balls under the stated conditions.

Solution: Given n = 16, $\bar{x}$ = 3.42, σ = 0.68 and (1 – α)100% = 99%

Then α/2 = 0.05

The (1 – α)100% small sample confidence interval for the population parameter μ is given by

$$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}$$

$$\Rightarrow \qquad 3.42 - t_{0.05}\frac{0.68}{\sqrt{16}} < \mu < 3.42 + t_{0.05}\frac{0.68}{\sqrt{16}}$$

$$\Rightarrow \qquad 3.42 - 2.947\frac{0.68}{\sqrt{16}} < \mu < 3.42 + 2.947\frac{0.68}{\sqrt{16}}$$

$$\Rightarrow \qquad 2.92 < \mu < 3.92$$

Problem 6: In a study of automobile collision insurance costs, a random sample of 80 body repair costs for a particular kind of damage had a mean of Rs.472.36 and a standard deviation of Rs.62.35. If $\bar{x}$ = Rs.472.36 is used as a point estimate of the true average repair cost of this kind of damage, with what confidence can one assert that the error does not exceed Rs.10.

Solution: Given n = 80, $\bar{x}$ = Rs.472.36, σ = 62.35 and the maximum error, E = 10

We know that $E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

$$\Rightarrow \quad z_{\alpha/2} = E\frac{\sqrt{n}}{\sigma}$$

$$\Rightarrow \quad z_{\alpha/2} = 10\frac{\sqrt{80}}{62.35}$$

$$\Rightarrow \quad z_{\alpha/2} = 1.43$$

We know that $P(Z > Z_\alpha) = \alpha$. Then $F(Z_\alpha) = 1 - \alpha$

Now $F(Z_{\alpha/2}) = 1 - \alpha/2$.

$\Rightarrow$ $F(Z_{\alpha/2}) = 1 - \alpha/2 = F(1.43)$

$\Rightarrow$ $1 - \alpha/2 = 0.9236$

$\Rightarrow$ $\alpha/2 = 1 - 0.9236$

$\Rightarrow$ $\alpha = 2(1 - 0.9236)$

$\Rightarrow$ $\alpha = 0.1528$

$\Rightarrow$ $1 - \alpha = 0.8472$

The probability that the error does not exceed Rs.10 is 0.8472.

Problem 7: If we want to determine the average mechanical aptitude of a large group of workers, how large a random sample will we need to be able to assert with probability 0.95 that the sample mean will not differ from the true mean hy more than 3.0 points. Assume that it is known from past experience that $\sigma = 20.0$.

Solution: Given $1 - \alpha = 0.95$, $E = 3$ and $\sigma = 20.0$. clearly $\alpha = 0.05$ and $\alpha/2 = 0.025$

We know that the maximum error of estimate

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow \quad 3 = z_{0.025}\frac{20}{\sqrt{n}}$$

$$\Rightarrow \quad 3 = 1.96\frac{20}{\sqrt{n}}$$

$$\Rightarrow \quad n = [1.96\frac{20}{3}]^2$$

$$\Rightarrow \quad n = 170.73 \approx 171$$

Problem 8: One novel process of making green gasoline takes biomass in the form of sucrose and converts it into gasoline using catalytic reactions. At one step in a pilot plant process, a chemical engineer measures the output of carbon chains of length three. Nine runs with same catalyst produced the yields (gal)

0.63, 2.64, 1.85, 1.68, 1.09, 1.67, 0.73, 1.04, 0.68

What can the chemical engineer assert with 95% confidence about the maximum error if she uses the sample mean to estimate true mean yield.

Solution: Given n = 9, (1 – α)100% = 95%, $\bar{x}$ = 1.334 and s = 0.674. Here

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad and \quad s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

α = 0.05 and α/2 = 0.025, $t_{0.025}$ 8 d.f = 2.306

For small samples the maximum error of estimate is given by

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$\Rightarrow \; E = t_{0.025} \frac{0.674}{\sqrt{9}}$$

$$\Rightarrow \; E = 2.306 \frac{0.674}{\sqrt{9}}$$

$$\Rightarrow \; E = 0.518$$

Problem 9: Ten bearings made by a certain process have a mean diameter of 0.5060 cm and a standard deviation of 0.0040 cm. Assuming that the data may be looked upon as a random sample from a normal population; construct a 95% confidence interval for the actual average diameter of bearings made by this process.

Solution: Given n = 10, (1 – α)100% = 95%, $\bar{x}$ = 0.5060 and s = 0.0040.

So, α = 0.05 and α/2 = 0.025.

We have to construct a confidence interval

$$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}$$

$$\Rightarrow \qquad 0.5060 - t_{0.025}\frac{0.0040}{\sqrt{10}} < \mu < 0.5060 + t_{0.025}\frac{0.0040}{\sqrt{10}}$$

$$\Rightarrow \qquad 0.5060 - 2.306\frac{0.0040}{\sqrt{10}} < \mu < 0.5060 + 2.306\frac{0.0040}{\sqrt{10}}$$

$$\Rightarrow \qquad 0.5031 < \mu < 0.5089$$

Problem 10: In an air-pollution study performed at an experiment station, the following amount of suspended benzene soluble organic matter was obtained for eight different samples of air:

2.2, 1.8, 3.1, 2.0, 2.4, 2.0, 2.1, 1.2

Assuming that the population sampled is normal, construct a 95% confidence interval for the corresponding true mean.

Solution: Given n = 8, (1 – α)100% = 95%, $\bar{x}$ = 2.1 and s = 0.5372. Here

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad and \quad s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

α = 0.05 and α/2 = 0.025, $t_{0.025}$ 7 d.f = 2.365

We have to construct a confidence interval

$$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}$$

$$\Rightarrow \qquad 2.1 - t_{0.025}\frac{0.0.5372}{\sqrt{8}} < \mu < 2.1 + t_{0.025}\frac{0.5372}{\sqrt{8}}$$

$$\Rightarrow \qquad 2.1 - 2.365\frac{0.0.5372}{\sqrt{8}} < \mu < 2.1 + 2.365\frac{0.5372}{\sqrt{8}}$$

$$\Rightarrow \qquad 1.6508 < \mu < 2.5492$$

Tests of Hypotheses

There are many physical problems in which instead of estimating the value of a parameter related to the population we need to decide whether to accept or reject a statement about the parameter. This statement is called a hypothesis and decision-making procedure about the hypothesis is called hypothesis testing.

Statistical hypothesis: A statistical hypothesis is a statement about the parameter of one or more populations.

Null hypothesis: A statistical hypothesis which is to be tested under the assumption that it is true is called a null hypothesis. It is denoted by H_o.

Alternate hypothesis: A statistical hypothesis that differs the null hypothesis is called alternate hypothesis. It is denoted by H_1.

Test of hypothesis: A procedure for deciding whether to accept or reject a particular hypothesis is called test of a hypothesis.

Type I error: Rejecting the hypothesis when it is true is called type – I error.

Type II error: Accepting the hypothesis when it is false is called type – II error.

Level of significance: The probability of committing type – I error is called level of significance and is denoted by α.

	Accept H	Reject H
H is true	Correct decision	Type-I error
H is false	Type-II error	Correct Decision

Null Hypothesis and Tests of Hypothesis

To approach problems of hypothesis testing systematically, it will help to proceed as outlined in the following five steps:

Step 1: We formulate a null hypothesis and an appropriate alternative hypothesis which we accept when the null hypothesis must be rejected.

Step 2: We specify the probability of a Type-I error. If possible, desired, or necessary, we may also specify the probabilities of Type-II errors for particular alternatives.

Step 3: Based on the sampling distribution of an appropriate statistic, we construct a criterion for testing the null hypothesis against the given alternative.

Step 4: We calculate from the data the value of the statistic on which the decision is to be based.

Step 5: We decide whether to reject the null hypothesis or whether to fail to reject it.

Problems on finding probabilities of type-I and type-II errors:

Problem 11: A paint manufacturer's claim that the average drying time, of is new "fast drying" paint is 20 minutes. It instructs a member of its research staff to paint each of 36 boards using a different 1-gallon can of the paint, with the intention of rejecting the claim if the mean of the drying time exceeds 20.75 minutes. Otherwise, it will accept the claim. It is known from the past experience that the standard deviation of such drying times can be expected to equal $\sigma = 2.4$ minutes. Find

(i) the probability of Type-I error when $\mu = 20$.

(ii) the probability of Type-II error when $\mu = 21$.

Solution:

(i) From the give data mean of the drying time $\mu = 20$. Population standard deviation, $\sigma = 2.4$.

Then the probability of type-I error is

$$\alpha = P(X > 20.75)$$

$$= P\left(\frac{X-\mu}{\sigma/\sqrt{n}} > \frac{20.75-\mu}{\sigma/\sqrt{n}}\right)$$

$$= P\left(Z > \frac{20.75-20}{2.4/\sqrt{36}}\right)$$

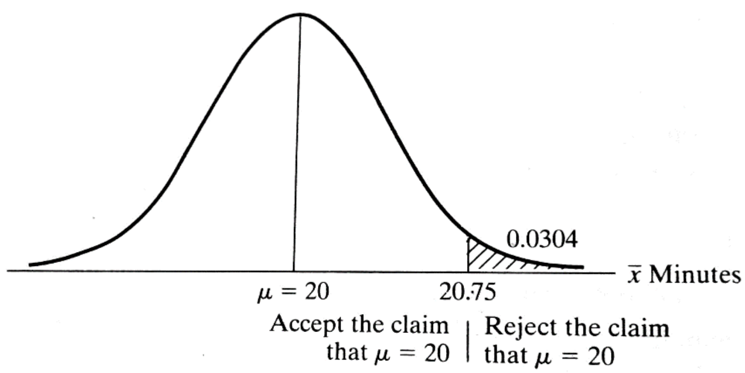
$$= P(Z > 1.875)$$

$$= 1 - F(1.875) = 1 - 0.9696 \qquad (\text{The average of } F(1.87) \text{ and } F(1.88))$$

$$= 0.0304$$

Hence the probability of rejecting the hypothesis $\mu = 20$ minutes is approximately 0.03.

The graph is



(ii) In finding the probability of type-II error, it is given that $\mu = 21$. The probability of Type-II error is

$$\beta = P(X < 20.75)$$

$$= P\left(\frac{X-\mu}{\sigma/\sqrt{n}} < \frac{20.75-\mu}{\sigma/\sqrt{n}}\right)$$

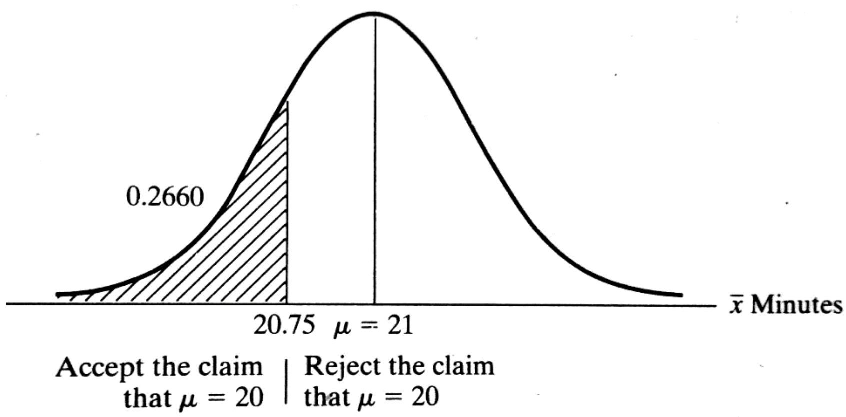
$$= P\left(Z < \frac{20.75-21}{2.4/\sqrt{36}}\right)$$

$$= P(Z < -0.625)$$

$$= F(-0.625) = 0.2660 \qquad (\text{The average of } F(-0.62) \text{ and } F(-0.63))$$

Hence the probability of accepting or failing to reject, the hypothesis $\mu = 20$ is approximately 0.27.

The graph is



Hypothesis concerning one mean (σ is known)

Let us suppose that we are dealing with a random samples of size n (≥ 30) taken from a normal population having means μ and known variances σ^2 and that we want to test the null hypothesis $\mu = \mu_0$ where μ_0 is a given constant, against one of the alternatives $\mu \neq \mu_0$, $\mu > \mu_0$, $\mu < \mu_0$. The test procedure is as follows:

Null Hypothesis, H_0 ; $\mu = \mu_0$

Alternate Hypothesis, H_1: $\mu < \mu_0$

or H_1: $\mu > \mu_0$

or H_1: $\mu \neq \mu_0$

Sample sizes: n (n is greater than or equal to 30)

Sample mean : $\bar{x}$

Sample variances: s^2

Level of significance, α (1% or 5%)

Test Statistic, $Z = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$

Critical regions for testing H_0: μ = μ_0

Alternate Hypothesis	Reject null hypothesis if
$\mu < \mu_0$	$Z < -z_\alpha$
$\mu > \mu_0$	$Z > z_\alpha$
$\mu \neq \mu_0$	$Z < -z_{\alpha/2}$ or $Z > z_{\alpha/2}$

Finally we have to write the decision whether to accept or reject the null hypothesis.

Problem 12: Suppose we want to establish that the thermal conductivity of a certain kind of cement brick differs from 0.340, the value claimed. To test the claim a random sample of 35 determinations yielded a mean of 0.343 and a standard deviation of 0.010. Use a 0.05 level of significance to test whether to accept or reject the claim. Assume that the probability distribution is normal.

Solution:

Null hypothesis , H_0: $\mu = 0.340$

Alternative hypothesis,H_1: $\mu \neq 0.340$

Level of Significance, $\alpha = 0.05$

Sample size, n = 35

$\overline{x}$ = 0.343

σ = 0.010

Test statistic, $Z = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$

$$= \frac{0.343 - 0.340}{0.01/\sqrt{35}}$$

$= 1.77$

$Z_{0.025} = 1.96$

Alternate Hypothesis	Reject null hypothesis if
H_1: $\mu \neq \mu_0$	$Z < -z_{\alpha/2}$ or $Z > z_{\alpha/2}$

Since Z = 1.77 is positive and not greater than $Z_{0.025} = 1.96$, we fail to reject the null hypothesis. So, the null hypothesis is accepted.

Problem 13: In 64 randomly selected hours of production, the mean and the standard deviation of the number of acceptable pieces produced by a automatic stamping machine are $\bar{x} = 1,038$ and $s = 146$. At the 0.05 level of significance, does this enable us to reject the null hypothesis $\mu = 1000$ against the alternative hypothesis $\mu > 1,000$.

Solution:

Null hypothesis, H_0: $\mu = 1000$

Alternative hypothesis, H_1: $\mu > 1000$

Level of Significance $\alpha = 0.05$

Sample size, n = 64

$\bar{x} = 1,038$

$s = 146$ (when $n \geq 30$ we will replace s by σ)

Test statistic, $Z = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$

$= \dfrac{1,038 - 1000}{146 / \sqrt{64}}$

$= 2.08$

$Z_{0.05} = 1.645$

Alternate Hypothesis	Reject null hypothesis if
$\mu > \mu_0$	$Z > z_\alpha$

Since $Z = 2.08$ is positive and is greater than 1.645, the null hypothesis must be rejected at level $\alpha = 0.05$. Hence accept the alternative hypothesis.

Problem 14: A trucking firm is suspicious of the claim that the average lifetime of certain tires is at least 28,000 miles. To check the claim, the firm puts 40 of these tires on its trucks and gets a mean lifetime of 27,463 miles with a standard deviation of 1,348 miles. What can it conclude if the probability of Type I error is to be at most 0.01?

Solution:

Null hypothesis, H_0: $\mu \geq 28,000$ miles

Alternative hypothesis, H_1: $\mu < 28,000$ miles

Level of Significance $\alpha \leq 0.01$

Sample size, n = 40

$\bar{x} = 27,463$

$\sigma = 1,348$

Test statistic, $Z = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$

$= \dfrac{27,463 - 28,000}{1,348 / \sqrt{40}}$

$= -2.52$

$Z_{0.01} = 2.33$

Alternate Hypothesis	Reject null hypothesis if
$\mu < \mu_0$	$Z < -z_\alpha$

Since $Z = -2.52$ is less than $-Z_{0.01} = -2.33$ the null hypothesis must be rejected at level $\alpha = 0.01$. So, accept the alternative hypothesis.

Hypothesis concerning one mean (σ is unknown)

Let us suppose that we are dealing with a random samples of size n (< 30) is taken from a normal population having means μ and known variances σ^2 (σ^2 is unknown) and that we want to test the null hypothesis $\mu = \mu_0$ where μ_0 is a given constant, against one of the alternatives $\mu \neq \mu_0$, $\mu > \mu_0$, $\mu < \mu_0$. The test procedure is as follows:

Null Hypothesis, H_0 ; $\mu = \mu_0$

Alternate Hypothesis, H_1: $\mu < \mu_0$

or H_1: $\mu > \mu_0$

or H_1: $\mu \neq \mu_0$

Sample sizes: n (n < 30)

Sample mean : $\bar{x}$

Sample variances: s^2

Level of significance, α (1% or 5%)

Test Statistic, $t = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}}$

Degrees of freedom, $\upsilon = n - 1$.

Calculate either t_α or $t_{\alpha/2}$ with n – 1 degrees of freedom.

Critical regions for testing H_0: $\mu = \mu_0$

Alternate Hypothesis	Reject null hypothesis if
$\mu < \mu_0$	$t < -t_\alpha$
$\mu > \mu_0$	$t > t_\alpha$
$\mu \neq \mu_0$	$t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$

Finally we have to write the decision whether to accept or reject the null hypothesis.

Problem 15: The specifications for a certain kind of ribbon call for a mean breaking strength of 180 pounds. If five pieces of the ribbon (randomly selected from different rolls) have a mean breaking strength 169.5 pounds with a standard deviation of 5.7 pounds, test the null hypothesis $\mu = 180$ pounds against the alternative hypothesis $\mu < 180$ pounds at the 0.01 level of significance. Assume that the population distribution is normal.

Solution:

Null hypothesis, H_0: $\mu = 180$ pounds

Alternative hypothesis, H_1: $\mu < 180$ pounds

Level of Significance, $\alpha = 0.01$

Sample size, n = 5

$\bar{x} = 169.5$

s = 5.7

Test statistic, $t = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}}$

$$= \frac{169.5 - 180}{5.7/\sqrt{5}}$$

$$= -4.12$$

$t_{0.01}$ (4 d.f) = 3.747

Alternate Hypothesis	Reject null hypothesis if
$\mu < \mu_0$	$t < -t_\alpha$

Since t = - 4.12 which is less than - $t_{0.01}$ (4 d.f) = - 3.747, we have to reject the null hypothesis. So, accept the alternative hypothesis.

Problem 16: A random sample of 6 steel beams has a mean compressive strength of 58,392 psi (pounds per square inch) with a standard deviation of 648 psi. Use this information and the level of significance $\alpha = 0.05$ to test whether the true average compressive strength of the steel from which this sample came is 58,000 psi. Assume normality.

Solution:

Null hypothesis, H_0: $\mu = 58,000$ psi

Alternative hypothesis, H_1: $\mu \neq 58,000$ psi

Level of Significance $\alpha = 0.05$

Sample size, n = 6

$\bar{x}$ = 58,392

s = 648

Test statistic, $t = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}}$

$$= \frac{58,392 - 58,000}{648/\sqrt{6}}$$

$$= 1.481$$

$t_{0.025}$ (5 d.f) = 2.571

Alternate Hypothesis	Reject null hypothesis if
$\mu \neq \mu_0$	$t > t_{\alpha/2}$ or $t < -t_{\alpha/2}$

Since t = 1.481which is positive and is not greater than $t_{0.025}$ (5 d.f) = 2.571, the null hypothesis cannot be rejected at level $\alpha = 0.01$.Hence compressive strength of the steel beams is 58,000 cannot be rejected.

Problem 17: A manufacturer claims that the average tar content of a certain kind of cigarette is $\mu = 14.0$. In an attempt to show that it differs from this value, five measurements are made of the tar content (mg per cigarette):

14.5, 14.2, 14.4, 14.3, 14.6

Show that the differences between the mean of this sample $\bar{x} = 14.4$, and the average tar claimed by the manufacturer $\mu = 14.0$ is significant at $\alpha = 0.05$. Assume normality.

Solution:

Null hypothesis, H_0: $\mu = 14.0$ psi

Alternative hypothesis, H_1: $\mu \neq 14.0$ psi

Level of Significance, $\alpha = 0.05$

Sample size, n = 5

Sample mean $\overline{X} = \dfrac{14.5 + 14.2 + 14.4 + 14.3 + 14.6}{5} = 14.4$

Sample variance

$$s^2 = \frac{(14.4 - 14.5)^2 + (14.4 - 14.2)^2 + (14.4 - 14.4)^2 + (14.4 - 14.3)^2 + (14.4 - 14.6)^2}{5 - 1}$$

$= 0.025$

s = 0.1581

Test statistic, $t = \dfrac{\overline{X} - \mu_0}{s / \sqrt{n}}$

$$= \frac{14.4 - 14.0}{0.1581 / \sqrt{5}}$$

$= 5.6574$

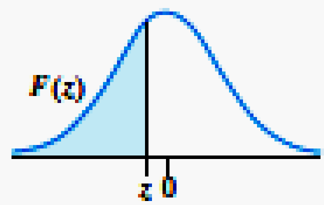
$t_{0.025}$ (4 d.f) = 2.776

Alternate Hypothesis	Reject null hypothesis if
$\mu \neq \mu_0$	$t > t_{\alpha/2}$ or $t < -t_{\alpha/2}$

Since $t = 5.6574$ is greater than 2.776, the null hypothesis must be rejected at level $\alpha = 0.05$. So, accept the alternative hypothesis.

Table 3 Standard Normal Distribution Function

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\, dt$$

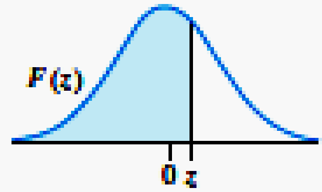


z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−5.0	0.0000003									
−4.0	0.00003									
−3.5	0.0002									
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

(continued on following page)

Table 3

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\,dt$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5973	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998									
4.0	0.99997									
5.0	0.9999997									