A First Course in Quality Engineering – 2nd edition

K.S. Krishnamoorthi

With V. Ram Krishnamoorthi

Chapter 2 : Statistics for Quality - Part 3

Inference of Population Quality from Sample

Definitions

A statistic is a function of observations from a sample.

Sample average $\overline{X}$, range *R*, standard deviation *S*, and the largest value X_{max}, etc., are examples of a statistic.

A statistic is a random variable because it is a function of random variables.

When we say we want to make inference about the quality of a population, it means we want to know its distribution.

Inference of Population Quality from Sample (contd)

If we assume that the population we are dealing with is normally distributed, we just need to know the two parameters μ and σ of the population.

The exact values of the parameters are never known so we estimate them using the statistics from samples.

A statistic that is chosen to estimate a parameter is called an <u>estimator</u> for the parameter.

Certain criteria are used to select a "good" estimator from available candidates of estimators for a parameter.

Selecting an Estimator

Unbiasedness:

An estimator that, on the average, equals the values of the parameter is called an unbiased estimator for that parameter.

$E(\overline{X}) = \mu$, $E(M) = \mu$, and $E(S^2) = \sigma^2$

So, the sample average $\overline{X}$ and the sample median M are unbiased estimators for μ and the sample variance S^2 is unbiased for σ^2 .

Minimum Variance:

If there are several unbiased estimators available to estimate a parameter, then choose the one that has the least amount of variance among all the unbiased estimators.

Selecting an Estimator (contd.)

The MVUB
The UB estimator that has the least amount of variability among All UB estimators for a parameter is called the Minimum Variance Unbiased estimator (MVUB) for that parameter.

It is known that, for normal populations, the MVUBs for the two Parameters are as follows:

Parameter	MVUB
μ	$\bar{X}$
σ^2	S^2

It is for this reason that we use $\bar{x}$ and S^2 to estimate μ and σ^2 respectively in the methods we discuss below.

The Point Estimate

The value of an estimator observed in a single sample is called a point estimate.

The point estimate, being one observation of a random variable, is not of much value, especially from small size samples.

So we resort to interval estimation where we create an interval (using the point estimate) such that the probability the parameter of interest lies in that interval is $(1 - \alpha)$.

We call such an interval $(1 - \alpha)100\%$ confidence interval (C.I.) for the parameter.

Estimating a Parameter using Confidence Intervals

Suppose we want to estimate the mean μ of a normal population using sample average $\overline{X}$ as the estimator.

Using the observed value of $\overline{X}$ an interval is created such that:

$$\mathrm{P}\left(\bar{x} - k \le \mu \le \bar{x} + k\right) = 1 - \alpha$$

The interval $\left(\bar{x} - k, \bar{x} + k\right)$ is called a 100% confidence interval for μ.

$\left(1 - \alpha\right)$ is called the confidence coefficient.

The value of k is determined from the distribution of the estimator $\overline{X}$.

C.I. for the μ of a Normal Population when σ is known

If the population has $\mathrm{N}(\mu,\sigma^2)$, the average has $\mathrm{N}(\mu,\sigma^2/\mathrm{n})$ where *n* is the sample size.

Then $$\frac{(\overline{\mathrm{X}}-\mu)}{\sigma/\sqrt{\mathrm{n}}} \sim N(0.1)$$

Using this result, the $(1-\alpha)$ 100% C.I. for μ is obtained as:

$$\left[\bar{x}-z_{a/2}\,\sigma/\sqrt{n},\ \ x+z_{a/2}\,\sigma/\sqrt{n}\right]$$

Where σ is the known standard deviation and $z_{\alpha/2}$ is such that $P(Z > z_{\alpha/2}) = \alpha/2$. (See figure in next slide)

C.I. for the μ of a Normal Population when σ is known

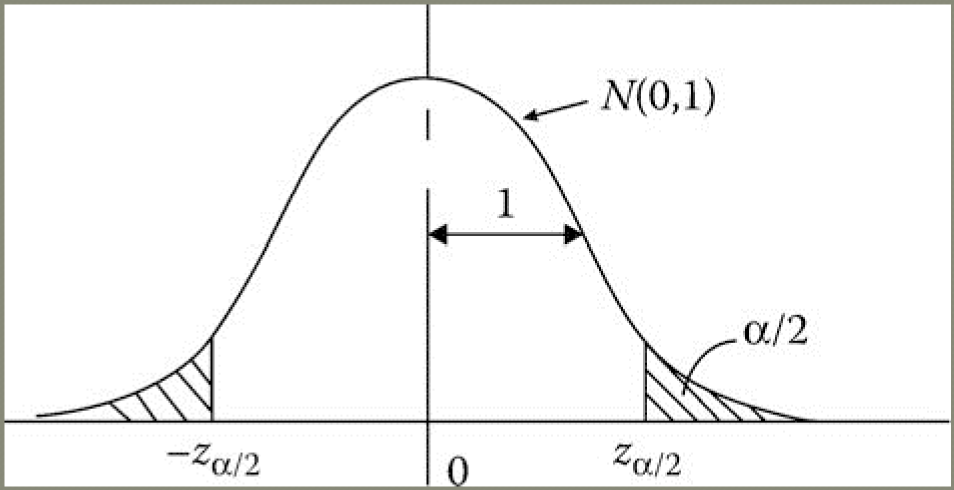


FIGURE 2.15 Definition of $z_{\alpha/2}$

C.I. For population mean – an example

Need a 99% C.I. for the mean turbidity in (all) the bottles of fabric softener filled in a line. σ is known to be = 0.3.

A sample of 4 bottles gave: 12.6, 13.4, 12.8, 13.2 ppm.

$$(1-\alpha)=0.99 \rightarrow \alpha/2=0.005 \qquad z_{\alpha/2}=2.575$$

$$\bar{x}=52/4=13.0, \sigma=0.3$$

99% C.I. for $\mu=\left[13.0-2.575\left(0.3/\sqrt{4}\right),\ 13.0+2.575\left(0.3/\sqrt{4}\right)\right]$

$$=[12.61,\ 13.39]$$

What is the interpretation for this confidence interval?

C.I. for μ when σ is not known

the C.I. is created from the fact that the statistic $\dfrac{\left(\overline{X}-\mu\right)}{S/\sqrt{n}}$

has the *t*-distribution with (*n*-1) degrees of freedom.

A $(1-\alpha)$100% C.I. for μ of a population that is normally distributed is given by:

$$\left[\bar{x}-t_{\alpha/2}(s/\sqrt{n}),\bar{x}+t_{\alpha/2}(s/\sqrt{n})\right]$$,where

$t_{\alpha/2}$ is such $P\left(t_{n-1}>t_{\alpha/2}\right)=\alpha/2$

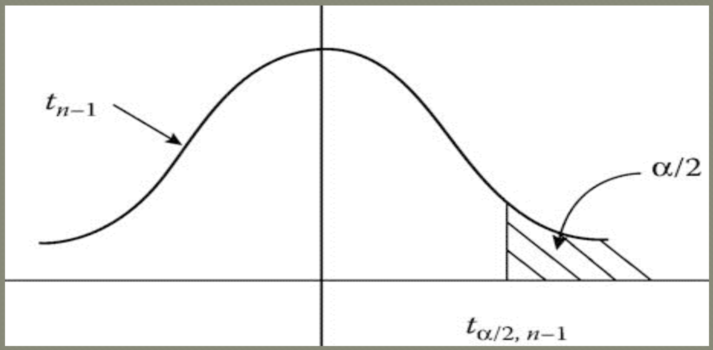


Figure 2.16 Example of a *t*-distribution

C.I. for μ when σ is not known-an example

Example 2.37

Four measurements of turbidity in bottles of fabric softener from a filling line are: 12.6, 13.4, 12.8, 13.2. Set up a 99% C.I. for average turbidity in the bottles of fabric softener. Assume normal distribution. The σ is not known.

$n - 1 = 3$ $\quad \alpha/2 = 0.005$ $\quad t_{0.005,\,3} = 5.841$ $\quad \bar{X} = 13.0$

$s = 0.366$ (calculated from sample observations)

99% C.I. = $\left[13.0 \pm 5.841\left(0.366/\sqrt{4}\right)\right]$ = [11.93, 14.06]

Compare this interval with the one obtained for the σ known case.

C.I. for σ^2 of a normal population

A (1- α)100% C.I. for the variance σ^2 of a normal population is given by:

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}} \right)$$

where S^2 is the sample variance, n the sample size, and $\chi^2_{\alpha/2,n-1}$ is such that

$$P\left(\chi^2_{n-1} > \chi^2_{\alpha/2,n-1}\right) = \alpha/2$$

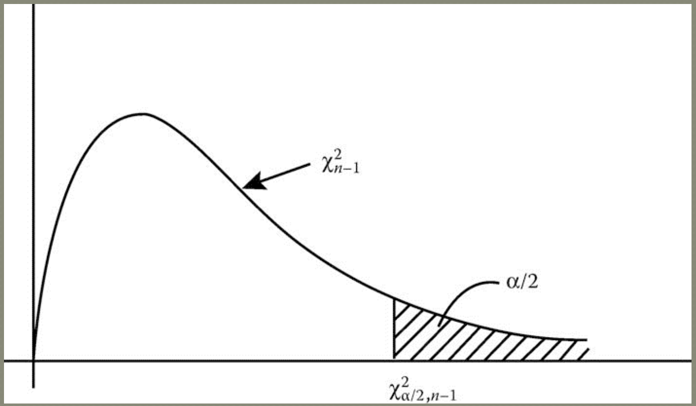


Figure 2.17 Example of a Chi-Square Distribution

C.I. for σ^2 of a normal population-example

Set up a 99% C.I. for the standard deviation σ of turbidity in bottles of cloth softener, if a sample of 4 bottles gave measurements: 12.6, 13.4, 12.8, and 13.2. Assume normality.

$\alpha/2 = 0.005$ $s^2 = 0.133$ (calculated from the sample)

From the tables, $\chi^2_{0.005,3} = 12.838$ $\chi^2_{0.995,3} = 0.0717$

99% C.I. for $\sigma^2 = \left[\frac{(3)(.133)}{12.838}, \frac{(3)(.133)}{0.0717}\right]$ = [0.031, 5.56]

99% C.I. for σ = $\left[\sqrt{.031}, \sqrt{5.56}\right] = [0.176, 2.36]$

Compare this with a 95% C.I. For σ

Hypothesis Testing

Two hypotheses are proposed:

Null hypothesis denoted by H_0.

Alternative hypothesis denoted by H_1

The hypotheses are complementary to each other.
If one is true, other is not true and vice versa.

The statement to be affirmed is placed in the Alt. Hyp. H_1

Possible Errors

Type-I error occurs if H_0 is declared false when in fact it is true.

Type-II error occurs if H_0 is declared true when in fact it is false.

Outcomes of a Statistical Test

		The test declares H_0 true	The test declares H_0 not true
In reality H_0 is	true	OK	Error-Type I
	not true	Error-Type II	OK

Probability of Type I error is denoted by α and is also called the <u>level of significance</u>.
Probability of Type II error is denoted by β, and $(1- \beta)$ is called the <u>power</u> of the test.

Designing a test

The test procedure is designed in such a way that the probability of the Type I error occurring is contained within a specified, small value α.

For every test, we choose a test statistic that provides the relationship between an estimator and the parameter about which hypotheses are proposed. The distribution of the test statistic will be known.

Designing the test consists of identifying the Critical Region (CR), the Critical Region being the set of observed values of the test statistic that will lead to rejection of the null hypothesis.

The Testing Procedure

The steps in hypothesis testing:

1. Set up H_0
2. Select an appropriate H_1
3. Choose a level of significance
4. Choose an appropriate test statistic
5. Design the test by specifying the critical region (C.R.)
6. Select a sample from the population(*s*) and compute the value of the test statistic
7. If the observed value of the test statistic is in CR
   reject H_0; otherwise do not reject
8. Interpret the results

An Example of Hypothesis Testing

Test Concerning the Mean μ of a normal population when σ is known

H_0: $\mu = \mu_0$ (Hypothesize that the mean equals a number μ_0)
H_1: $\mu > \mu_0$ (The alternate hypothesis is: if the mean is not equal to μ_0 it must be greater than μ_0)

Test-Statistic: $$\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$$

This test-statistic relates the sample average $\overline{X}$ to the population mean μ, and its distribution is known to be the $N(0, 1)$. distribution

Designing a Test - an example

Choosing the Critical Region:

The alternate hypothesis determines where the C.R. is located.

With the given H_1, we will reject H_0 in favor of H_1 if Z_{obs} is too large.

We need to draw a line to decide how large a value of Z_{obs} is too large.

We draw the line at $Z\alpha$ so that the probability of Type I error is limited to α.

Choosing the Critical Region

Z_α is called the critical value of the statistic.

The values in the distribution of the test statistic beyond Z_α constitute the Critical Region

$$H_0: \mu = \mu_0 \qquad H_1: \mu > \mu_0$$

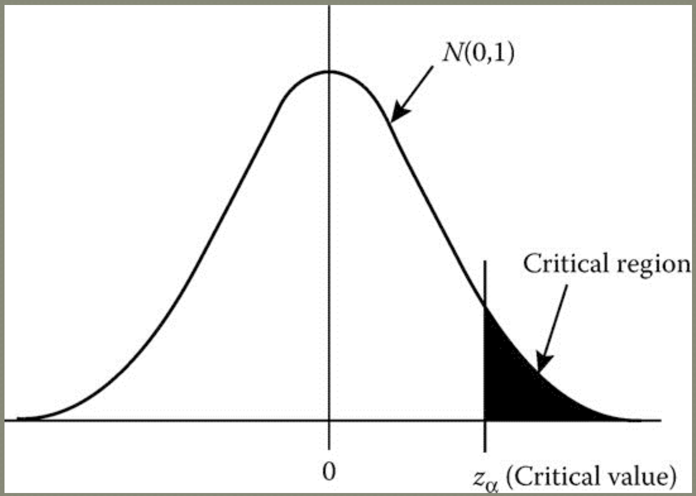


Figure 2.19 Critical region in the distribution of the test statistics

Three Possible Alternate Hypotheses

Case 1.	Case 2.	Case 3.
H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$
H_1: $\mu < \mu_0$	H_1: $\mu > \mu_0$	H_1: $\mu \neq \mu_0$

The test statistic will be the same for all three cases; however the location of C.R. will differ as shown in Figure 2.22.

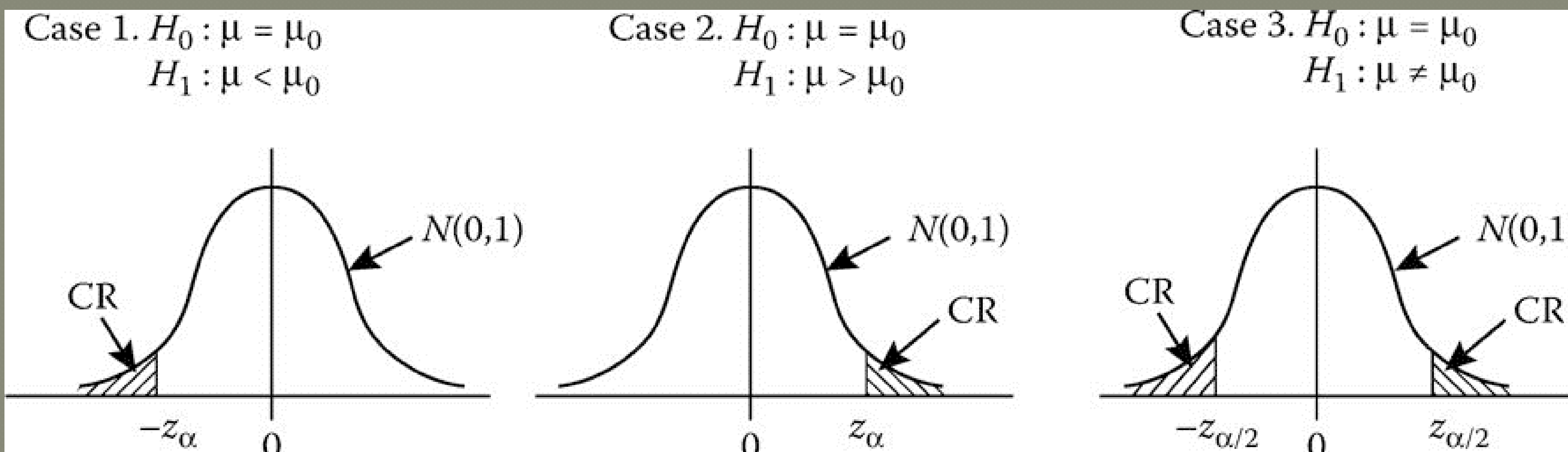


Figure 2.22 : Alternate Hypotheses and corresponding Critical Regions

Test for the Mean-An Example

Example:

A supplier of nylon rope claims their new product has average strength greater than 10 kg. A sample of 16 rope pieces gave an average of 10.2 kg. If the standard deviation of the strength is known to be 0.5 kg, test the hypothesis μ = 10 vs. μ > 10. Use α=0.01.

H_0: $\mu = 10$ H_1: $\mu > 10$ Test Statistic: $\dfrac{\bar{X} - 10}{\sigma/\sqrt{n}} \sim Z$

Critical region : $Z_{obs} > Z_{\alpha} = Z_{.01} = 2.326$.

The observed value of the test statistics: $Z_{obs} = \dfrac{10.2 - 10}{0.5/\sqrt{16}} = 1.6$

The Z_{obs} is not in the C.R. and the null hypothesis is not rejected. The mean strength of the population of nylon ropes is not greater than 10 kg. The supplier's claim is not valid.

Test Concerning the Mean μ When σ is Not Known

There are again three possible alternate hypotheses.

Case 1.	Case 2.	Case 3.
$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$
$H_1: \mu < \mu_0$	$H_1: \mu > \mu_0$	$H_1: \mu \neq \mu_0$

Test Statistic: $\dfrac{\overline{X} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$

The critical regions:

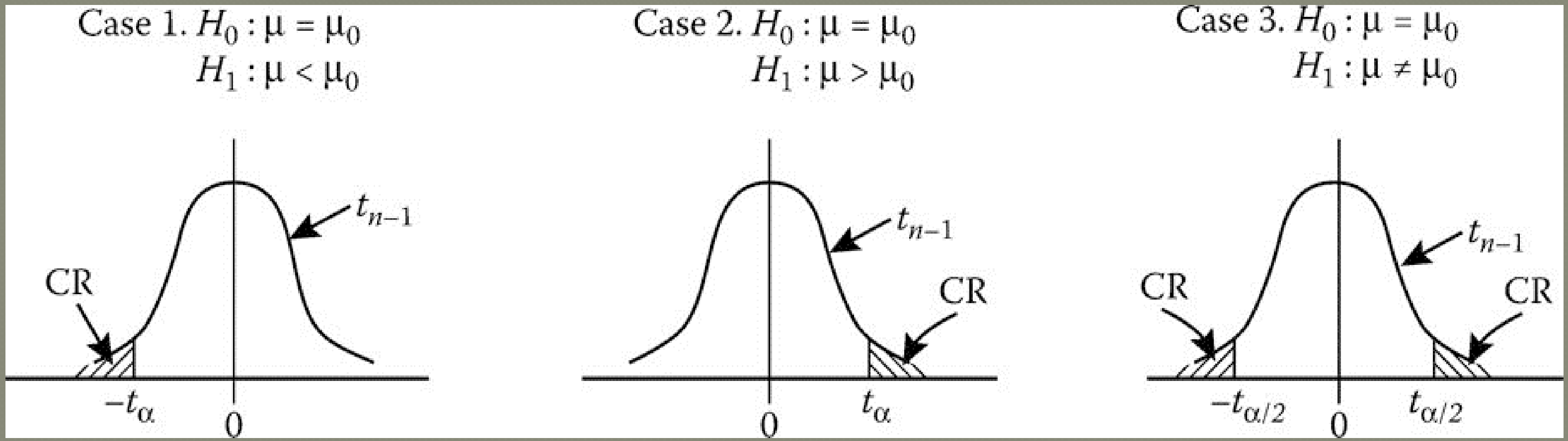


Figure 2.23: Critical Regions for tests for μ when σ is not known

Test for the Mean μ When σ is Not Known-

Example: The amount of ash in a box of sugar should be less than 2 grams according to a manufacturer's claim. Lab analysis of 5 boxes gave the following results: 1.80, 1.92, 1.84, 2.02, 1.76 grams. Is the ash content in the boxes is less than 2 gms as claimed? Use $\alpha = 0.05$.

H_0: $\mu = 2$

H_1: $\mu < 2$ (Notice that the claim is in the alternative hypothesis.)

$\bar{X} = 1.868$ $s = 0.104$ (computed from sample data)

Test Statistic: $\dfrac{\bar{X} - 2}{S/\sqrt{5}} \sim t_4$ C.R.: *Observed* $t_4 < -t_{0.05,4} = -2.132$

The observed value of t_4: $t_{obs} = \dfrac{1.868 - 2}{0.104/\sqrt{5}} = -2.838$

t_{obs} ***is*** in C.R. => Reject H_0

$\Rightarrow$Mean ash content *is* than 2 grams.

$\Rightarrow$ The manufacturer's claim is valid.

Test for Difference of Two Means- σ's are Known

H_0: $\mu_1 - \mu_2 = 0$ (i.e., no difference between the two pop. means)

Case 1:	Case 2:	Case 3:
H1: $\mu_1 - \mu_2 < 0$	$\mu_1 - \mu_2 > 0$	$\mu_1 - \mu_2 \neq 0$

Test Statistic: $$\frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim Z$$

Critical Regions corresponding to the three cases of alternate hypotheses are shown in Figure 2.22.

Case 1: Case 2: Case 3:

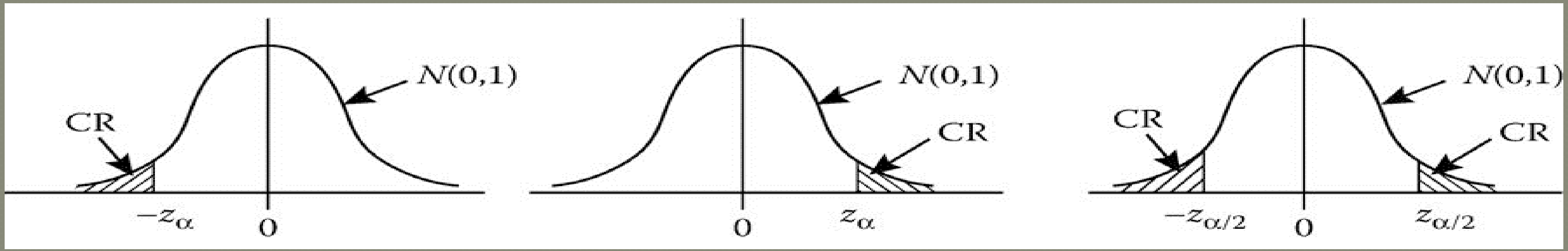


Figure 2.25 Critical regions for hypotheses about difference of two means

Test for Difference of Two Means- σ's Known - Example

Test the hypothesis that male and female workers in this factory earn equal pay against the alternate that they don't, at $\alpha = 0.01$.

Male workers	Female workers
$\bar{x}_1 = 35{,}000$	$\bar{x}_2 = 34{,}600$
$s_1 = 1200$	$s_2 = 1800$
$n_1 = 10$	$n_2 = 8$

$$H_0: \ \mu_1 - \mu_2 = 0$$

$$H_1: \ \mu_1 - \mu_2 \neq 0$$

The two-sided alternate hypothesis is chosen because there is no reason to believe the difference in average will be positive or negative.

Test for Difference of Two Means- σ's Known – Example (contd.)

Test Statistic: $$\frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}}$$

C.R.: Values of $Z_{obs} > z_{\alpha/2}$ or $Z_{obs} < -z_{\alpha/2}$

$Z_{\alpha/2} = Z_{0.005} = 2.575$ (from Normal tables)

$$Z_{obs} = \frac{400}{\left[\frac{1200^2}{10} + \frac{1800^2}{8}\right]^{1/2}} = \frac{400}{740} = 0.541$$

Z_{obs} is not in the critical region. Hence do not reject H_0. There is no significant difference between the salaries of male and female workers.

Many Other Models for Testing Hypothesis

There are models that can be used to test hypotheses for:

Difference of two means when population standard deviations are not known and sample sizes are small. We then use a test statistic that has *t* distribution.

The model to test the hypothesis about population variance uses a test statistic that has $\chi 2$ distribution.

The model to test the hypothesis about ratio of two variances uses a test statistic that has F distribution.

There are also testing procedures that do not require normality assumption. These are known as distribution-free tests or non-parametric tests.

Testing for the Distribution of a Population

Use of Normal Probability Plot

Normal probability plotting involves plotting the Cum. Distr. of the data on specially designed Normal Probability Paper (NPP).

The NPP has been designed in such a way that if the data had come from a normal population, the Cum. Distr. will plot as a st. line.

Conversely, if the Cum. Distr. of a set of data plots as a straight line on a NPP, then we conclude that the data comes from a normal population.

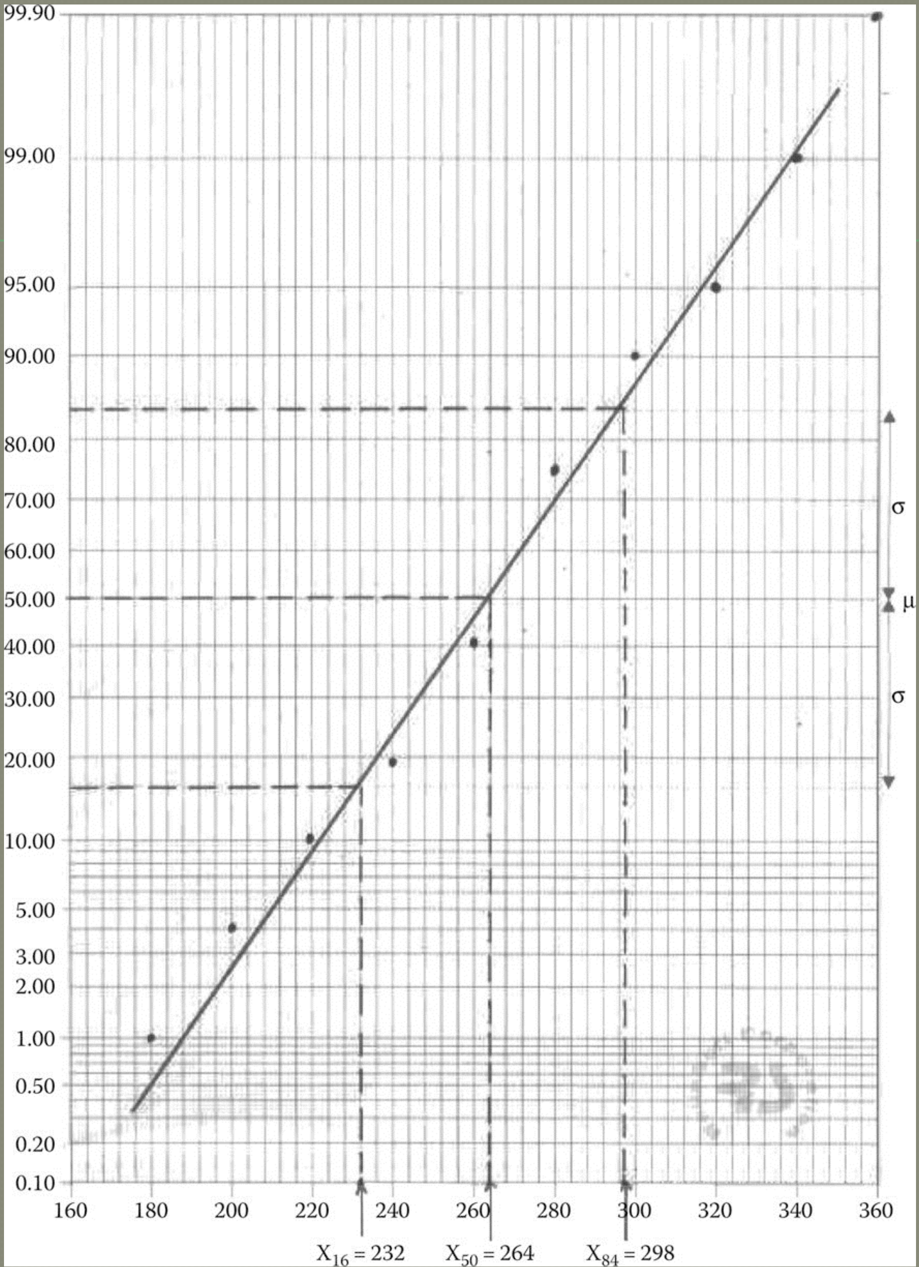
The procedure is described below using an example.

Table 2.3 Frequency distribution, Cum. Freq. distribution

1

	Cell Limits	Midpoint	Frequency (σ_i)	% Frequency	% Cumulative frequency
1	160–180	170.5	1	1	1
2	181–200	190.5	3	3	4
3	201–220	210.5	6	6	10
4	221–240	230.5	9	9	19
5	241–260	250.5	22	22	41
6	261–280	270.5	34	34	75
7	281–300	290.5	15	15	90
8	301–320	310.5	5	5	95
9	321–340	330.5	4	4	99
10	341–360	350.5	1	1	100
		Total	100	100	

Figure 2.27 Example of a Normal Probability Plot using commercially available NPP From www.Weibull.com



Normal Probability Plot on the Computer

A computer program calculates the cumulative probabilities for individual values using what are called the mean rank of the value in the data.

To calculate the mean rank, the data is first arranged in an ascending order. Then the mean rank of the observation that has the *i*th rank in the data = $\dfrac{i}{n+0.5}$.

Another formula used to estimate the cum. probs. is the median rank, calculated using the formula $\dfrac{i-0.3}{n+0.4}$

Value	Rank in data	Mean rank(x100)	Median rank (x100)
175	1	0.995	0.697
187	2	1.99	1.69
197	3	2.985	2.69
----	----	-----	-----

Normal Probability Plot on the Computer

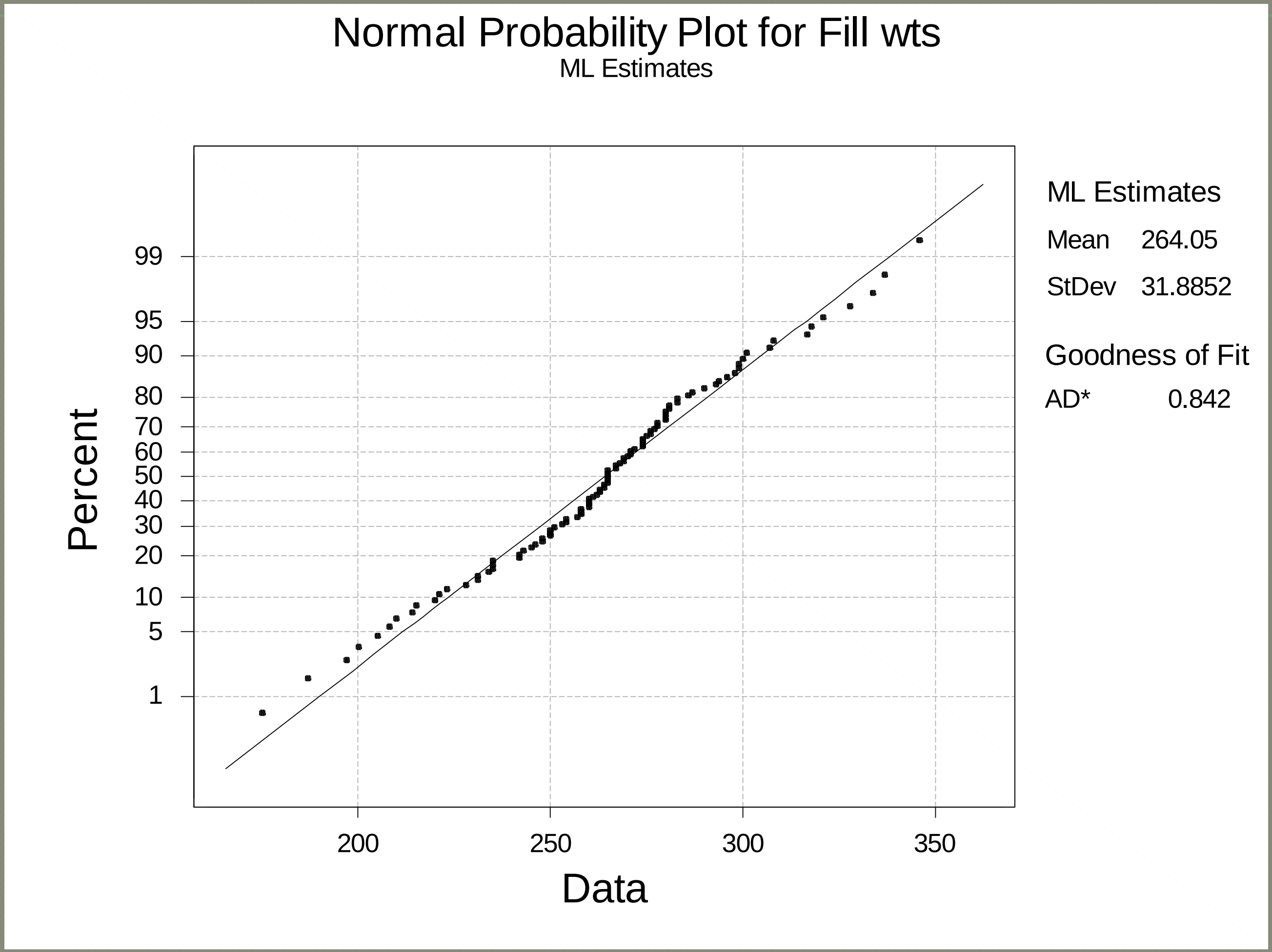


Figure 2.28 Normal probability plot produced by Minitab software for data in Table 2.1

Goodness-of-fit Tests

There are several goodness of fit tests such as the Chi-square g.f. test, Kolmogorov-Smirnov g.f. test and Anderson-Darling g.f. test, etc.

A goodness of fit in general obtains a quantity to measure the total deviation of the actual cumulative distribution of the data from the hypothesized cumulative distribution.

This is done by computing the total of the distances between the actual and hypothesized distributions at several points within the range of data.

If the total deviation is "too large," then the hypothesis about that particular distribution is rejected.

Chi-squared Goodness-of-fit Test

As an example, the Chi-squared goodness-of-fit test is done by grouping the data into cells, as for making a histogram, and calculating a quantity to measure the total deviation.

$$\sum_i \frac{(a_i - e_i)^2}{e_i}$$

In this, a_i represents actual number of observations in each cell and e_i represents the expected number of observations in each cell if the distribution were normal.

This quantity is known to follow the Chi-squared distribution with (*p* - 1) degrees of freedom, where *p* is the number of cells into which the data have been tallied.

Chi-squared Goodness-of-fit Test - Example

	Cell Limits	Midpoint	Frequency (a_i)		% Frequency	% Cumulative Frequency	% Expected Frequency (e_i)		$\frac{(e_i - a_i)^2}{e_i}$
1	161–180	170.5	1		1	1	0.372		
2	181–200	190.5	2	9	2	3	1.85	8,32	0.056
3	201–220	210.5	6		6	9	6.17		
4	221–240	230.5	10		10	19	14.21		1.247
5	241–260	250.5	18		18	37	22.35		0.846
6	261–280	270.5	35		35	72	24.12		4.91
7	281–300	290.5	17		17	89	17.82		0.038
8	301–320	310.5	6		6	95	9.02		
9	321–340	330.5	4	11	4	99	3.13	12.9	0.278
10	341–360	350.5	1		1	100	0.745		
		Total	100		100		100		7.375

P-Value

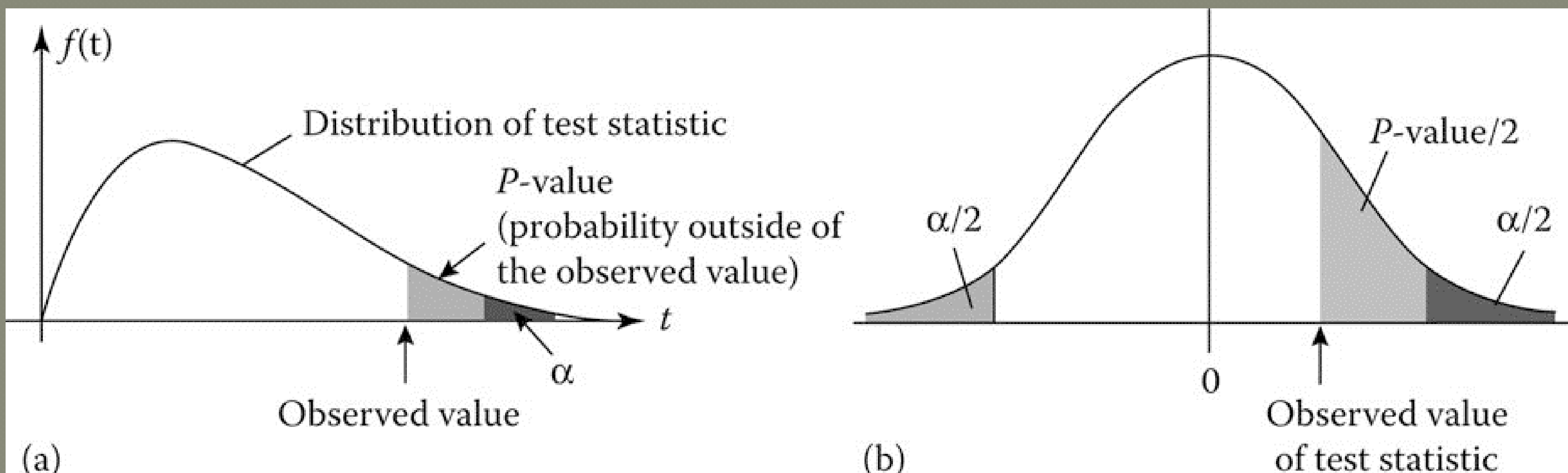


Figure 2.29 Meaning of the *P*-value