PDHonline Course P210 (2 PDH)

Statistical Methods for Process Improvement - Part 4: Using Data to Make Decisions

Instructor: Davis M. Woodruff, PE, CMC

2020

PDH Online | PDH Center

5272 Meadow Estates Drive
Fairfax, VA 22030-6658
Phone: 703-988-0088
www.PDHonline.com

An Approved Continuing Education Provider

Part 4: Using Data to Make Decisions

An introduction to statistical inference or hypothesis testing

He uses statistics as a drunken man uses lampposts—for support rather than for illumination. (Source: More funny Andrew Lang quotes)

The product of an arithmetical computation is the answer to an equation; it is not the solution to a problem. Ashley-Perry

Sample Statistics

Previously, the subject of descriptive statistics was discussed (PDH course 207). The purpose is fairly self-explanatory -- description of a group of data. Now we turn our attention to an area with a different focus -- statistical inference. With statistical inference, the objective is to infer conclusions about populations by using information obtained from samples. Statistical inference involves mathematical procedures which allow us to draw these conclusions. This course presents an overview of the subject and gives practical information and examples.

Let's clarify a few terms. A parameter is a mathematical characteristic of a population, e.g. mean or variance. Because the exact value of a parameter is seldom known, we need an accurate estimation. A statistic is a value calculated from sample data and used to estimate a population parameter. Thus, $\bar{X}$ or the sample mean is a statistic and an estimator of μ, the population mean.

Hypothesis Testing

The mathematical methods used in statistical inference are referred to as hypothesis tests. A hypothesis is a statistical statement made for the purpose of rejecting or not rejecting (accepting). Hypotheses can be stated about means, variances or shapes of distributions. Some examples of hypotheses are:

- The rejection rate for resin #125 in Demopolis is the same as St. Paul.
- The mean dimension is 2.105 inches.
- The variability on this product is no different in Greenville or Hayneville.

The test is simply a means of deciding whether the hypothesis is false or not, depending upon the sample observations. There are two hypotheses for any statistical test. The most important is called the null hypothesis, H_o. The alternative hypothesis, H_a, is automatically accepted if the test shows that the null hypothesis should be rejected. An example of a null and alternative hypothesis is shown below:

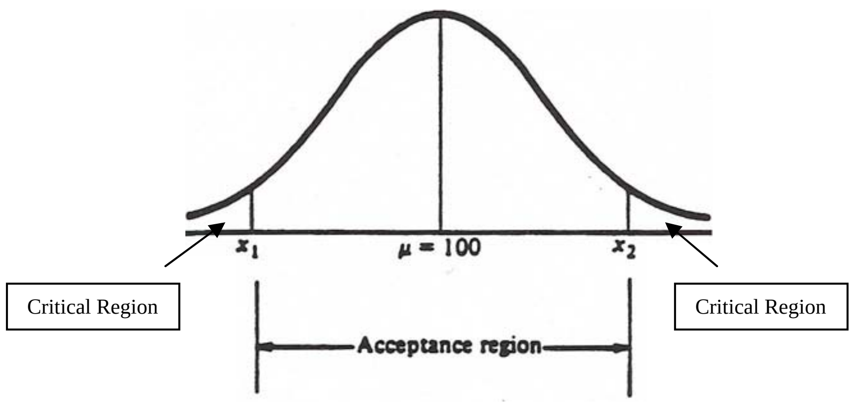
$$H_o: \sigma = 1.85$$
$$H_a: \sigma \neq 1.85$$

The objective is to not reject H_o. The terminology of 'not reject' is correct since it is not possible to establish it to be true. Even so, it is possible to calculate the probability of its being false. If that probability is high enough, we can reject it. *What this means is that we are putting the burden of proof on the alternative hypothesis*. However, for the ease of communication, we will use the term 'accept H_o' instead of 'not rejecting' it.

In order to accept the null hypothesis, the result of the statistical test must fall into an acceptance region. Values outside of this region fall into the critical region and require the rejection of H_o. Suppose a normal population has a true mean of 100. We form the hypotheses:

$$H_o: \sigma = 100$$
$$H_a: \sigma \neq 100$$

Two values, shown as x_1 and x_2, usually known as critical values must be determined which separate the acceptance and critical regions. The figure below shows this.



Acceptance Region

Errors in Testing

There are four possible outcomes from testing. These are based on the fact that we could draw the correct conclusion and we might not. These are:

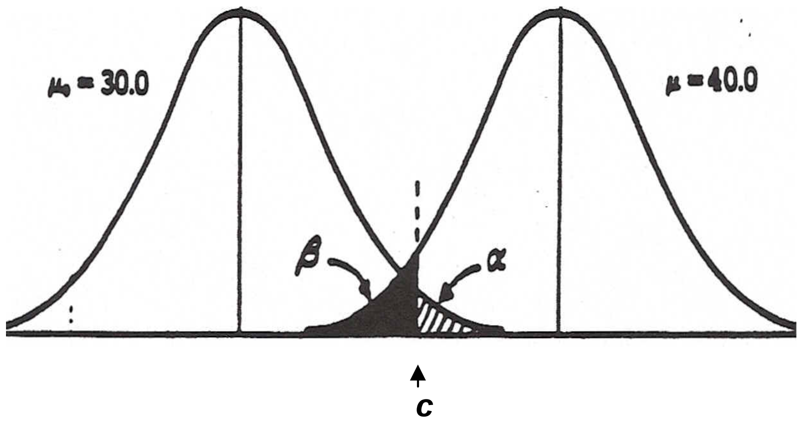
		True Situation	
	Decision	**Ho True**	**Ho False**
Results of Testing	Accept H_o	Correct	Type II Error
	Reject H_o	Type I Error	Correct

The probability of these errors is important. It determines the sensitivity and power of the test in distinguishing true from false hypotheses. These probabilities are symbolized by the Greek letters α and β.

α = the probability of rejecting a true H_o

β = the probability of accepting a false H_o

These are sometimes referred to as producer's risk and consumer's risk, respectively. The diagram below shows these errors graphically.



Graphical Presentation of α and β Errors

If **c** represents the critical value that determines the boundary of the acceptance region, you can see that it is not possible to eliminate the chances of both errors for this case. Any movement of **c** will decrease the chance of one error but will also increase the chance of the other.

Usually the situation is that we form a hypothesis about the population, but we do not know the location of the alternate hypothesis. So we cannot determine the error. It is for this reason that we normally work with the error only.

Type I error, or , is also known as the significance level of the test. Significance level is the amount of uncertainty about the hypothesis. This level is usually set before the test is run. Common levels are 0.10, 0.05, and 0.01, with 0.05 being the most common. If $\alpha = 0.05$, the significance level is 5 percent. This means that on the average, a type I error (rejecting a true H_o) will happen 5 times out of 100 times that the hypothesis is tested.

Confidence level is also an important term in hypothesis testing. It is the degree of assurance that a hypothesis is correct. Confidence level = $1 - \alpha$. For $\alpha = 0.05$, the confidence level is 95%. If we reject H_o, we are 95% confident that we are making the right decision.

One-Sided and Two-Sided Test

Whether or not we are concerned with a one-sided or two-sided test depends on how the hypothesis is stated. Let's begin with the two-sided test. Suppose the null hypothesis was that $\mu = 15$ and the alternative hypothesis was that $\mu \neq 15$. The test focuses on the mean being different from 15 regardless of whether it is smaller or larger. This is called a two-sided or two-tailed test because the difference, if detected, can be on either side of the parameter under study. With a two-sided test, the area is divided equally on both ends of the distribution curve.

A one-sided test has the total area in only one tail of the curve. The null hypothesis is usually stated in a pessimistic manner. For example, changes have been made in a process intended to increase the mean value of a key variable. The null hypothesis for this case would state that the process had not been improved. This will then force that data to show that an improvement has occurred. For example, changing vendors of a raw material is hoped to improve the average yield. The hypotheses would be

$$\mathbf{H_0:}\ \mu_{new} \leq \mu_{old}$$
$$\mathbf{H_a:}\ \mu_{new} > \mu_{old}$$

If we were trying to reduce that variability, the hypothesis would read

$$\mathbf{H_0:}\ \sigma^2_{new} \leq \sigma^2_{old}$$
$$\mathbf{H_a:}\ \sigma^2_{new} > \sigma^2_{old}$$

Selection of one-sided or two-sided test also impacts the β risk involved. The appropriate hypothesis should be determined before performing the test.

General Hypothesis Testing Procedure

Making statistical inference generally follows a standard procedure. The steps are as follows:

1. State the hypotheses, both the null and alternative. Choose between a one-sided or two-sided test.
2. Select α, the significance level.
3. Select the appropriate test statistic. The formula used to test the null hypothesis is called the test statistic.
4. Sample to collect the necessary data. Determine degrees of freedom, if appropriate.
5. Determine the rejection criteria. Draw a picture.
6. Calculate the value of the test statistic.
7. Accept or reject H_a. If the calculated test statistic is within the critical region, accept H_a; if it is in the acceptance region, reject H_a and accept the null hypothesis.

Now let's look at some specific test statistics and how they can be used.

Testing Variances

The F statistic is used to test for the equality of variances from two normal populations.

$$F = s_a^2/s_b^2$$

The parameters of the F distribution are the degrees of freedom associated with s_a and s_b.

Degrees of freedom for numerator $= n_a - 1$
Degrees of freedom for denominator $= n_b - 1$

Test of a Single Variance

At times we may wish to compare a population with a known or standard variance.

Procedure for Two-Sided Test

Step 1. State the Hypotheses:

$$\mathbf{H_0}: \sigma^2_{new} = \sigma^2_{old}$$
$$\mathbf{H_a}: \sigma^2_{new} \neq \sigma^2_{old}$$

Step 2. Select , the significance level.

Step 3. The test statistic is F:

$$F = s^2/\sigma_o^2 \quad \text{or} \quad F = \sigma_o^2/ s^2, \text{ whichever is larger}$$

Step 4. Take a random sample and calculate s. Determine degrees of freedom df = n-1.

Step 5. Determine the critical region from

$$Fcrit = F_{(dfn,dfd,\alpha/2)}$$

Step 6. Calculate the F statistic.

Step 7. Make a decision: (Using the F distribution tables in any of the reference texts provided on the reading list)

If $F_{calc} > F_{crit}$, accept H_a
If $F_{calc} \leq F_{crit}$, reject H_a and accept H_o

Example: The standard deviation for TOC analysis has been long established as 0.21 with the equipment at your lab. In evaluating new equipment, a short study showed the standard deviation to be 0.27 for 10 analyses. Is the variability different?

Following the steps outlined above:

Step 1: **H_0:** $\sigma^2_{new} = \sigma^2_{old} = (0.21)^2$
H_a: $\sigma^2_{new} \neq \sigma^2_{old}$

Step 2: Let $\alpha = 0.05$

Step 3: $F = s^2/\sigma_o^2$

Step 4: Data given $\sigma_o = 0.21$, $n = \infty$ (historical data); $df = \infty$
σ_{new} or $s_{new} = 0.27$, $n = 10$; $df = 9$

Step 5: $Fcrit = F_{(9,\infty,0.025)} = 2.11$ (from F distribution tables in most any statistics textbook or online at the following link: **[PDF]** **F-distribution table** - The F Distribution

Step 6: $Fcalc = (0.27)^2/(0.21)^2 = 1.65$

Step 7: Decision: $F_{calc} < F_{crit}$ (1.65 < 2.11) thus we Reject H_a and accept H_o.

We conclude that the variability of the new type is not different from the standard type at the 95% confidence level.

Tests of Variances of Two Populations

Comparing the variances of two populations are done very similarly. The difference is in the use of s for the other population rather than a known σ.

Should a one-sided test be conducted, the hypotheses would be different and α rather than α /2 would be used in determining the critical values of F.

Comparing Several Variances

On occasion it may be of interest to know if several variances may be considered equal. This is sometimes referred to as the homogeneity of variances. Special tables have been compiled to simply this consideration. Of course, it may be more effective to just use excel® and do an analysis of variance, or ANOVA test which is beyond the scope of this course. Thus, we will limit this course to comparing two variances as shown above.

However, we will state the basic procedure for this comparison without ANOVA which is as follows:

1. Take a random sample of size n from each of the k lots or populations.
2. Calculate

$$g = \frac{\text{largest } s^2}{s_1^2 + s_2^2 + \ldots s_k^2}$$

3. Chose α , the significance level.
4. Determine $g_{critical}$ (from specialized tables found in some statistics books) for α , n, and k values.
5. If $g_{calc} < g_{crit}$, accept the hypothesis that the variances are the same. Otherwise, reject this hypothesis.

Testing Means (n ≥ 30)

One of the more common uses of hypothesis testing is to determine if the means of samples are equal to, greater than, or less than each other or some specified value. These tests are conducted using either the Z statistic or the t statistic. We will examine six of these tests.

In order to test the means, we must have some information about the variability of the two populations. Different test statistics are required depending on whether the variances are equal or not equal. The Z statistic is usually used for large sample sizes ($n \geq 30$) and the t statistic for small sample sizes.

In comparing two samples to determine if their population means are different the standard deviation of the populations are not usually known. In these cases, s is used as an estimator. To determine if the variances can assumed to be equal, an F test must be conducted. The results will then determine which t test statistic should be used.

The parameter of the t distribution is the degrees of freedom associated with it. These must be determined in order to locate the critical value of t from the tables. Again, depending on the particular test statistic used, the calculation of the degrees of freedom will be different.

The following examples show calculations required for comparing means.

Example: Comparing a Mean to a Standard with σ^2 Known
(Z Test)

Info: A particular part has a target length of 3.5. Historically, the standard deviation has been 0.25. For the last few days, the length has averaged 3.70. The USL is 3.9 and LSL is 3.1. The quality engineer insists this is different from previous production. Based on a sample of 25, does she have a case?

Step 1: Hypothesis: **H_0:** $\mu_{new} = \mu_{old} = 3.5$
H_a: $\mu_{new} \neq \mu_{old}$

Step 2: Select the significance level, $\alpha = 0.05$

Step 3: Calculate the test statistic

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

Step 4: Using our data, $\bar{X} = 3.7$; $\sigma = 0.25$; $n = 25$

Step 5: Determine acceptance criteria

$Z_{crit} = Z_{0.25} = \pm 1.96$ (from the table at

Step 6: Calculate Z using the data above: $Z_{calc} = 4.0$

Step 7: Since $Z_{calc} > Z_{crit}$ ($4.0 > 1.96$) accept H_a, Therefore we conclude at the 95% confidence level that the length is different from the target and the quality engineer is correct in their assessment.

Example: Comparing a Mean to a Standard with σ^2 Unknown
(t Test)

Info: The lab recently analyzed five different lots from a new vendor. According to their sales literature, their product color index is 8.0. The five lots ran 7.9,8.0, 8.0,8.1,8.1. It appears that these lots came from a population that averaged more than 8.0. What conclusion can be drawn at the 95% confidence level?

Step 1: Hypothesis: $\mathbf{H_0}$: $\mu_{new} \leq \mu_{old} = 8.0$ One sided test
$\mathbf{H_a}$: $\mu_{new} > \mu_{old}$

Step 2: Select the significance level, $\alpha = 0.05$

Step 3: Calculate the test statistic

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} \quad df = n-1$$

Step 4: Using our data, $\bar{X} = 8.02$; $\sigma = 0.084$; $n = 5$

Step 5: Determine acceptance criteria

$$t_{crit} = t_{\alpha,df,} = t_{0.05,4} = 2.132$$

Step 6: Calculate t using the data above: $t_{calc} = 0.53$

Step 7: Since $t_{calc} < t_{crit}$ (0.53 < 2.132) reject H_a and accept H_o. There is insufficient evidence to say the color is > 8.0.

Note: In excel®, go to f(x) and use the statistical functions for TTEST and follow the instructions. It returns the probability, which can be translated into a statement at the 95% (or an exact) confidence level.

Example: Comparing Means of Two Populations with σ^2 Known
(Z Test)

Info: Analysis over the past quarter has shown that operating results seem to vary according to plant. For instance, in Greensboro the variability (std. dev.) in rejects (ppm) is 93.4 but at Decatur has “dropped” to 82.4. Is the apparent difference in the means of 274.6 ppm versus 299.5 ppm real? The values were calculated from 750 and 725 data points over the past six months of operations.

Step 1: State the hypothesis, let A = Greensboro and B = Decatur

H_0: $\mu_A = \mu_B$ Two sided test
Ha: $\mu_A \neq \mu_B$

Step 2: Select the significance level, in this case use α = 0.05

Step 3: Determine the test statistic

$$Z = \frac{(\bar{X}a - \bar{X}b)}{\sqrt{\sigma_a^2/n_a + \sigma_b^2/n_b}}$$

Step 4: Using the process data given

Data	Day Shift (A)	Night Shift (B)
$\bar{X}$	274.6	299.5
σ	93.4	82.4
n	750	725

Step 5: Determine the acceptance criteria

$Z_{crit} = Z_{0.25} = \pm 1.96$ (from the table at

Step 6: Calculate Z using the formula and data given: $Z_{calc} = \pm 5.43$

Step 7: Since $Z_{calc} > Z_{crit}$ (5.43 > 1.96 and -5.43 < -1.96) accept H_a, Therefore we conclude at the 95% confidence level that the averages differ. Decatur reject ppm is lower.

Note: Again, it is easy to use the statistical functions in excel®.

Example: Comparing Means of Two Populations With σ^2 Unknown But Equal
(t Test)

Info: Δ G values on the injection molding unit had been stable. Due to a minor electrical problem, there was a process upset. The period prior to the upset was characterized by an average of 2.7 and s = 0.4 and the period following by 3.l and s = 0.7 for 10 samples each. Have the Δ G values shifted?

Step 1: State the hypothesis; let A = After the upset and B = Before the upset

Now, we are moving into an area that requires us to first establish if the variances are the same or different in order to know which means test to use. Remember, when dealing with small samples it is essential to first establish if the variability of the population has changed before attempting to test for mean shifts.

$$H_0: \sigma^2_a = \sigma^2_b$$
$$H_a: \sigma^2_a \neq \sigma^2_b$$

Two Sided test, the question is have the values shifted?

Step 2: Select the significance level, $\alpha = 0.05$

Step 3: Test statistic:

$$F = \sigma_a^2 / \sigma_b^2$$

Step 4: Data given (note the symbol σ is used even though it technically is s, the sample standard deviation). Remember, σ^2 unknown simply means we are dealing with samples and not the population which is generally the case in manufacturing, thus the df in the denominator for the std. deviation calculation is n-1.

Data	After (A)	Before (B)
σ	0.7	0.4
n	10	10
df	9	9

Step 5: From an F distribution table, determine $F_{crit} = F_{(9,9,.025)} = 4.03$

Step 6: Calculate the F statistic, $F_{calc} = 3.06$

Step 7: Make a decision. Reject H_a and accept H_o. We conclude that the variability before and after the upset is not different, therefore, we can use the t Test (Test 4) to compare the mean before and after the upset.

Example: Comparing Means of Paired Observations
(Paired t Test)

Info: Some discrepancies seem to exist with shipment weights for barges. The last II barges as determined by receiving were compared to the bill of lading had an average discrepancy of I5,000 lbs. and the standard deviation was 300 lbs. Are the barge weights as determined by the customer really different?

Step 1: Hypothesis

H_o: d = 0
H_a: d ≠ 0

Step 2: α = 0.05

Step 3: t = $\bar{d}/sd\sqrt{n}$; where sd = standard deviation of the differences

Step 4: $\bar{d}$ = 15,000 lbs.
sd = 300 lbs.
n = 11 shipments

Step 5: $t_{crit} = t_{\alpha/2,df}$ = $\mathbf{t}_{.0025,10}$ = **2.228**

Step 6: t_{calc} = (15000)$\diagup$300$\sqrt{11}$ = 15000/995 = 15.07

Step 7: Since $t_{calc} > t_{crit}$ (15.07 > 2.228), Accept Ha, the barge weights are different at the 95% confidence level.

So, what's the next step? Perhaps it is to check meters or scales at both locations. If this involved liquids, perhaps loading temperatures could have an effect. You get the idea. The statistical conclusion is simply that and no more. The real work begins as you troubleshoot based on the facts.

Note: In excel® there is a specialized function for the paired t test, and the information is simple to input and the results easy to interpret. This is found in the statistical analysis package.

Chi Square Goodness of Fit

So far the statistical tests we have discussed have involved population parameters such as means or variances. Now we are going to make inferences about a population distribution. Again we will use a null and alternative hypothesis. Generally,

H_o: the sample is from a specified distribution
H_a: the sample is not from a specified distribution

The test utilizes a Chi square distribution. Although it can be used for continuous distributions, we will focus on discrete distributions. The test statistic is

$$X^2 = \Sigma \frac{(O-E)^2}{E}$$

where O = the observed frequency value
E = the expected frequency value

The degrees of freedom are given by

$$df = k - r - 1$$

where k is the number of classes the data is divided into and r is the number of parameters of the hypothesized distribution estimated from the sample data.

The following example illustrates the test:

The HR manager contends that absenteeism is twice as bad on Mondays as any other day of the week for the Greensboro plant. He wants you to review the last 3 month's data and verify this if possible.

Day of the week	M	T	W	Th	F
Days absent	304	176	139	141	130

If twice as many absences occur on Mondays, then this is a ratio of 2:1:1:1:1. This means that 2/6 of the absences occur on Monday and 1/6 of the absences occur on each of the other days of the week. The null hypothesis is the absences are distributed as noted.

Step 1. H_o : Absences are distributed as described above
H_a : Absences are not distributed this way

Note: Chi Square (χ^2) test is always a one sided test.

Step 2. $\alpha = 0.05$

Step 3.The correct test is $X^2 = \Sigma \frac{(O-E)^2}{E}$

Step 4. & Step 6.

The following table shows the calculations of the Chi square value:

Day	Absences Observed	Absences Expected	$(O - E)^2$	$\frac{(0-E)^2}{E}$
M	304	296.7	53.29	0.18
T	176	148.3	767.29	5.17
W	139	148.3	86.49	0.58
Th	141	148.3	53.29	0.36
F	130	148.3	334.89	2.25
	890	889.9		$8.54 = \chi^2$

Step 5. With = 0.05, and the degrees of freedom at 4 (df = 5 _ 1 = 4 because there were no parameters estimated from the sample) the critical value of χ^2 is 9.49. This test is always a one-sided test.

Step 7. Since χ^2_{calc} is less than χ^2_{crit} (8.54 < 9.49) we can reject H_a and accept H_o. It appears that the absenteeism rate is twice as high on Mondays.

One *note* of caution: all the expected frequency values, E, used in this test must have a minimum frequency of 5 for validity. If necessary, groups may be combined to guarantee all E's > 5.

χ^2 Test for Independence of Factors

Sometimes samples may be classified according to different factors, each factor having two or more levels. A Chi square test can help determine if the observed values are independent of the factors. The data below shows the observed values of weekly defects by three shifts and two production lines. These are usually referred to as contingency tables. This test will allow us to statistically determine if defects occur independently of shift and line.

	Shift		
Line	1	2	3
1	10	12	13
2	14	9	12

The null hypothesis is that the defects are independent of the shift and line. The alternate hypothesis is that they are not independent; <u>but this test does not determine how or to what degree any dependence exists</u>. It is simply a quick way to check for independence.

The table can be rewritten to show the expected number of defects as well. These are found by totaling the rows and columns, then multiplying the probability of a defect from a row by the probability of a defect from a column by the total number of defects. For 1st shift and line 1 the expected number of defects would be

$$\frac{\text{Row total}}{\text{Grand Total}} \quad X \quad \frac{\text{Column total}}{\text{Grand Total}} \quad X \ \text{Grand Total}$$

Or,

$$(24/70) \times (35/70) \times (70) = 12$$

This can be done for all entries and is shown in the table below with expected values in parentheses.

Line	Shift 1	Shift 2	Shift 3	Row Total
1	10	12	13	35
	(12)	(10.5)	(12.5)	
2	14	9	12	35
	(12)	(10.5)	(12.5)	
Column Total	24	21	25	70

Summing $(O - E)^2/E$ for the six values gives $\chi^2 = 1.14$.

The *degrees of freedom* for contingency tables are found by

$$df = (\#\text{ of rows} - 1)(\#\text{ of columns} - 1)$$

So for this example, $df = (2 - 1)(3 - 1) = 2$.

With $\alpha = 0.05$, $\chi^2_{crit} = 5.99$.

Since $\chi^2_{calc} < \chi^2_{crit}$ then we accept the null hypothesis and conclude that the defects are independent of shift and production line.

Summary of Hypothesis Test

This has simply been an overview of a branch of statistics known as statistical inference where one can draw conclusions about process improvements or other factors by using the data. However, it is just what it says: an overview.

In depth training for this subject alone would be a couple of days, so just be aware of the tools available and determine the best tool to answer the question.

Of course, the most important factor is to ask the right question! Only when you understand the right question can the right tool be selected for the task. Oh yes, don’t forget it is all based upon and underlying normal distribution of the population.

Glossary of Statistical Terms

Accuracy - The closeness of agreement between and observed value and an accepted reference value.

Alternative Hypothesis - The hypothesis that is accepted if the null hypothesis is disproved.

Attribute Data - Qualitative data that typically shows only the number of articles conforming and the number of articles failing to conform to a specified criteria. Sometimes referred to as Countable Data.

Average - The sum of the numerical values in a sample divided by the number of observations.

Bar Chart - A chart that uses bars to represent data. This type of chart is usually used to show comparisons of data from different sources.

Batch - A definite quantity of some product or material produced under conditions that are considered uniform.

Bias - A systematic error which contributes to the difference between a population mean of measurements or test results and an accepted reference value.

Bimodal Distribution - A distribution with two modes that may indicate mixed data.

Binomial Distribution - A distribution resulting from measured data from independent evaluation, where each measurement results in either success or failure and where the true probability of success remains constant from sample to sample.

Cells - The bars on a histogram each representing a subgroup of data.

Common Cause - A factor or event that produces normal variation that is expected in a given process.

Confidence - The probability that an interval about a sample statistic actually includes the population parameter.

Control Chart - A chart that shows plotted values, a central line and one or two control limits and is used to monitor a process over time.

Control Limits - A line or lines on a control chart used as a basis for judging the significance of variation from subgroup to subgroup. Variation beyond a control limit shows that special causes may be affecting the process. Control limits are usually based on the 3 standard deviation limits around an average or central line.

Countable Data - The type of data obtained by counting -attribute data.

Data - Facts, usually expressed in numbers, used in making decisions.

Data Collection - The process of gathering information upon which decisions to improve the process can be based.

Detection - A form of product control, not process control, that is based on inspection that attempts to sort good and bad output. This is an ineffective and costly method.

Distribution - A group of data that is describable by a certain mathematical formula.

Frequency Distribution - A visual means of showing the variation that occurs in a given group of data. When enough data have been collected, a pattern can usually be observed.

Histogram - A bar chart that represents data in cells of equal width. The height of each cell is determined by the number of observations that occur in each cell.

k - The symbol that represents the number of subgroups of data. For example, the number of cells in a given histogram.

Lower Control Limit (LCL) - The line below the central line on a control chart.

Mean - The average value of a set of measurements, see Average.

Measurable Data - The type of data obtained by measurement. This is also referred to as variables data. An example would be diameter measured in millimeters.

Median - The middle value (or average of the two middle values) of a set of observations when the values have been ranked according to size.

Mode - The most frequent value in a distribution. The mode is the peak of a distribution.

n - The symbol that represents the number of items in a group or sample.

np - The symbol that represents the central line on an np chart.

Nonconformities - Something that doesn't conform to a drawing or specification; an error or reason for rejection.

Normal Distribution - A symmetrical, bell-shaped frequency distribution for data. This is a distribution that is often seen in industry.

Null Hypothesis - The hypothesis tested in test of significance that there is no difference (null) between the population of the sample and the specified population (or between the populations associated with each sample).

Out of Control - The condition describing a process from which all special causes of variation have not been eliminated. This condition is evident on a control chart by the presence of points outside the control limits or by nonrandom patterns within the control limits.

p - The symbol on a p chart that represents the proportion of nonconforming units in a sample.

Parameter - A constant or coefficient that describes some characteristics of a population (e.g. standard deviation, average, regression coefficient).

Pareto Charts - A bar chart that arranges data in order of importance. The bar representing the item that occurs or costs the most is placed on the left-hand side of the horizontal axis. The remaining items are placed on the axis in descending (most to least) order.

Population - All members, or elements, of a group of items. For example, the population of parts produced by a machine includes all of the parts the machine has made. Typically in SPC we use samples that are representative of the population.

Prevention - A process control strategy that improves quality by directing analysis and action towards process management, consistent with the philosophy of continuous quality improvement.

Process - Any set of conditions or set of causes working together to produce an outcome. For example, how a product is made.

Product - What is produced; the outcome of a process.

Quality - Conformance to requirements.

Random Sampling - A data collection method used to ensure that each member of a population has an equal chance of being part of the sample. This method leads to a sample that is representative of the entire population.

Range - The difference between the highest and lowest values in a subgroup.

Repeatability - The variation in measurements obtained when one operator uses the same test for measuring the identical characteristics of the same samples.

Reproducibility - The variation in the average of measurements made by different operators using the same test when measuring identical characteristics of the same samples. (In some situations it is the combination of operators, instruments and locations.)

Run Chart - A line chart that plots data from a process to indicate how it is operating.

Sample - A small portion of a population.

Sampling - A data collection method in which only a portion of everything made is checked on the basis of the sample being representative of the entire population.

Significance Level () - The risk we are willing to take of rejecting a null hypothesis that is actually true.

Skewed Distribution - A distribution that tapers off in one direction. It indicates that something other than normal, random factors are effecting the process. For example, TIR is usually a skewed distribution.

Special Cause - Intermittent source of variation that is unpredictable, or unstable; sometimes called an assignable cause. It is signaled by a point beyond the control limits or a run or other nonrandom pattern of points within the control limits. The goal of SPC is to control the special cause variation in a process.

Specification - The extent by which values in a distribution differ from one another; the amount of variation in the data.

Standard Deviation (σ) - The measure of dispersion that indicates how data spreads out from the mean. It gives information about the variation in a process.

Statistic - A quantity calculated from a sample of observations, most often to form an estimate of some population parameter.

Statistical Control - The condition describing a process from which all special causes of variation have been eliminated and only common causes remain, evidenced by the absence of points beyond the control limits and by the absence of nonrandom patterns or trends within the control limits.

Statistical Methods - The means of collecting, analyzing, interpreting and presenting data to improve the work process.

Statistical Process Control (SPC) - The use of statistical techniques to analyze data, to determine information, and to achieve predictability from a process.

Statistics - A branch of mathematics that involves collecting, analyzing, interpreting and presenting masses of numerical data.

Subgroup - A group of consecutively produced units or parts from a given process.

Tally or Frequency Tally - A display of the number of items of a certain measured value. A frequency tally is the beginning of data display and is similar to a histogram.

Tolerance - The allowable deviation from standard, i.e., the permitted range of variation about a nominal value. Tolerance is derived from the specification and is NOT to be confused with a control limit.

Trend - A pattern that changes consistently over time.

Type I Error (α) - The incorrect decision that a process is unacceptable when, in fact, perfect information would reveal that it is located within the "zone of acceptable processes".

Type II Error(β) - The incorrect decision that a process is acceptable when, in fact, perfect information would reveal that it is located within the "zone of rejectable processes".

u - The symbol used to represent the number of nonconformities per unit in a sample which may contain more than one unit.

Upper Control Limit (UCL) - The line above the central line on a control chart.

Variables - A part of a process that can be counted or measured, for example, speed, length, diameter, time, temperature and pressure.

Variable Data - Data that can be obtained by measuring. See Measurable Data.

Variation - The difference in product or process measurements. A change in the value of a measured characteristic. The two types of variation are within subgroup and between subgroup. The sources of variation can be grouped into two major classes: Common causes and Special Causes.

X - The symbol that represents an individual value upon which other subgroup statistics are based.

$\bar{X}$ (x bar) - The average of the values in a subgroup.

$\bar{\bar{X}}$(x double bar) - The average of the averages of subgroups.

Suggested Readings and/or References:

Duncan, A.J., *Quality Control and Industrial Statistics*, 5th ed. Irwin.

Grant, E.L. and Leavenworth, R.S., *Statistical Quality Control*, 5th ed., McGraw Hill.

Blank, L.T., *Statistical Procedures for Engineering, Management and Science*, 1st ed, McGraw Hill.

Wadsworth, H.M, Stephens, K.S. and Godfrey, A.B., *Modern Methods for Quality Control and Improvement*, 1st ed., Wiley.

Juran, J.M., et al, *Quality Control Handbook*, McGraw Hill

Ott, E.R., *Process Quality Control: Troubleshooting and Interpretation of Data*, McGraw hill

Shewhart, W.A., *Economic Control of Quality of Manufactured Products*, Van Nostrand.

Levin, R.I., *Statistics for Management*, 3rd ed., McGraw Hill.

Feigenbaum, A.V., *Total Quality Control*, 3rd ed., McGraw Hill

Statistical Quality Control Handbook, Western Electric.

Charbonneau, Harvey C. and Webster, Gordon L., *Industrial Quality Control.*