A SunCam online continuing education course

WHAT EVERY ENGINEER SHOULD KNOW ABOUT ENGINEERING STATISTICS II

by

O. Geoffrey Okogbaa, Ph.D., PE

Table of Contents

Introduction

The focus of this course is on an important area in engineering analyses and design, namely how we analyze data and use the information to make decisions about the engineering problem. The whole process of explicating the complexities of the data to yield information that would eventually be used to make design or mission decisions is known as inference or more appropriately statistical inference. If we examine the relationship between the population and the sample (as we did in the first course) we note that there is sort of a symbiotic (parent-population, offspring-sample) relationship between the two. Probability deals with the population with its parameters (parent values) while statistical inference deals with the sample and its statistic (values computed from the sample and used to estimate the population or universe parameters). Thus, while probability and statistics both deal with questions involving parameters and statistic, they do so in an "inverse manner" as shown in figure 1.

From the point of view of the population, information about the sample can be obtained by probability analyses. On the other hand, given some sample statistic, one can parley those values into making inference about the nature of the population parameters. Statistics really is about statistical inference, namely, trying to infer whether the sample statistic (such as sample mean $\overline{X}$ versus the population mean μ or the sample standard deviation S^2 versus the population variance σ^2) can reasonably be assumed to be good estimates of the parameters of the parent population. Statistical inference is the idea of assessing the properties of an underlying distribution through the method of <u>deductive analyses</u>. Inferential statistical analysis infers properties about a population through the following major schemes, namely; a). <u>Estimates</u> (point estimates, interval estimates), and b). <u>Tests of Hypotheses</u>. The population or the universe is assumed to be infinite and thus is larger than the sample from which a data set is drawn. Thus, probability projects information from the population to the sample via <u>inductive reasoning</u>, while we use the sample statistic as a means of understanding the nature of the population parameter using deductive reasoning.

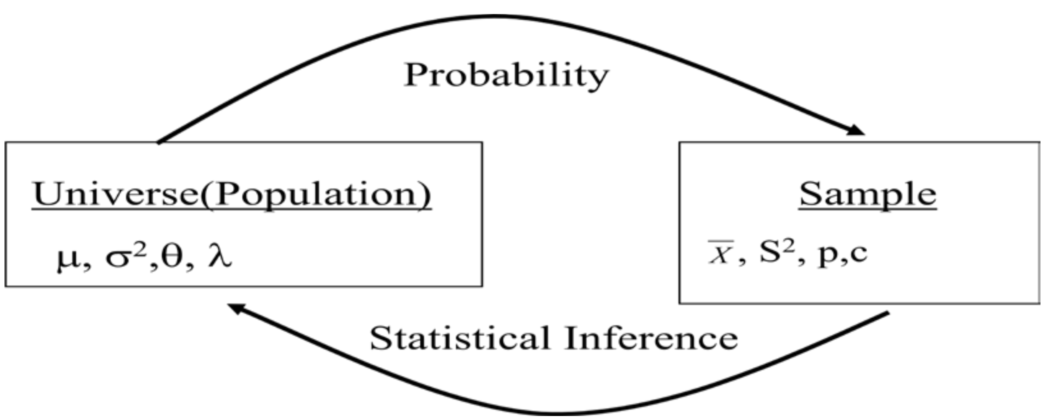


Figure 1: Relationship between Population and Sample

1.1 Mean and the Dispersion

In any data analyses scenario, the focus has always been on finding some way to understand what the data means, what it is saying about the underlying process and if there is anything that can be gleaned about the trend or characteristics. Unfortunately, parameters are not easy to come by and for that matter neither is the population itself. **For engineers, it is important to understand that we do not conduct experiments for the sake of the estimates but for eventually assessing the parameters**. The goal of any experiment is beyond just computing the statistic from the sample but to go over and above that to understand how those statistics explain away the population parameters. The measures or the sample estimate or statistic are never the goal but a pathway to the goal which in this case is the population parameter. That is why we always seek the best estimators for the parameter that we want to estimate so we can **get as close as** possible to the real thing. Given the **nature of a random experiment** it is expected that each realization of the experiment may very well produce different statistic(s). Hence each realization of the experiment may produce different estimates of the same statistic(s). For example, the Diametral-Pitch (DP) for a sample taken from a lot or population of spur-gears, the first sample n_1 could yield a mean value $\overline{X}_1 = 4\,inches.$ A second sample n_2 from the same lot could yield a DP of $\overline{X}_1 = 6\,inches.$ Thus the statistic from the two samples would be numerically different, even though as we would show later, they could be statistically the same.

The overarching goal in any experimental situation is not to estimate the different statistic for their own sake because they are useful only to the extent that we need them to discover the true value of the population. The ultimate goal is therefore to get a sense of the population parameter(s) value (μ or σ in this case) or what is generally called the true mean or the true variance. These measures or statistic are used as a pathway to access the population parameters. No one sets out in an experiment hoping that the statistic they obtained from a sample is the end all. Even for those without a background in mathematics and/or statistics it is generally understood that samples, everything else being equal, are mere representatives (and in some cases true representatives) of the parent population and hence any decisions made will hopefully reflect the parent population. Thus statistic(s) from samples are mere surrogates because we really do not have access to the population (very rarely do we have access to the population) in most cases and so the closest thing is the sample. In essence, all we are doing is to estimate the parameter value from the sample statistic because they (sample statistics are the only things we have access to.

Statistical Inference

In many engineering settings (manufacturing, chemical, electronic, computing, service, etc.), we encounter many random quantities. Often, we do not know the probability structure of these variables or their underlying characteristics. Still we do want to determine these quantities to have better control of the system operation. This is usually accomplished by taking random samples or observations on the random variables. Based on the classical definition of probability, the determination of the probability or the expected value associated with the random variables would require an 'infinite number of observations. However, since in some cases (the case of the spur-gear described earlier) we have a very large but finite population, we can usually estimate the values in question in the form of sample statistics computed from the samples.

Ultimately, at the end of a statistical inference analyses, the decision is always to act or not to act. In some instance, the decision could be to accept the observed or computed value of the estimator as the unknown parameter without requiring that it be exactly the true value. On the other hand, we may decide to reject or not reject the assumptions about certain distribution without conceding that such a statement is true beyond doubt. Thus, the use of statistical inference enables us to control the possible errors that could arise as a result of our decisions and to ensure that these errors, while inevitable, are as small and as economically as possible.

As indicated earlier, inferential statistics can be divided into three main branches, **namely point estimation, interval estimation, and test of hypotheses.**

Estimation (Point Estimation)

For a good estimation, a fairly large sample is needed. In some cases, only very limited samples may be all that is available. Such limitation could result in a situation where the distribution is assumed beforehand since the sample size is limited and thus the ensuing analysis is only meant to verify that the distribution has not changed.

There are two types of estimators, namely point estimators and interval estimators. Two methods are generally used in generating estimators of parameters, namely, the methods of moments and maximum likelihood. For some problems both the method of moments and maximum likelihood lead to the same estimators and for others they do not. When the two methods do not agree, the maximum likelihood estimator is usually preferred.

Let X be a random variable with probability density function (pdf). The pdf has a known form and is based on an unknown parameter θ that belongs to the parameter space Ω. This means that we have a family of distributions whose values all lie in the parameter space. Thus, to each value of $\theta \in \Omega$, we have one member of the possible family of distributions. In most cases the experimenter wants to choose only one member of the family to represent the pdf of the random variable of interest. Since we have a family of distributions whose parameter values belong in the parameter space Ω, the problem becomes one of defining a statistic that will be a good point estimator of the parameter of interest.

3.1 Point Estimates for the Mean, Median and Variance

A point estimates is a single value or number, a point on the real line, which we feel is a good guess for the unknown population parameter value that is being sought. They are statistics obtained from the sample that we then use to estimate the population parameter. The motivation for conducting an experiment stems from the understanding that in most cases it is impractical to obtain the value of the parameter that we seek because that would require the almost impossible task of observing the outcome of an infinite population. This being the case, the problem then reduces to one of attempting to extract as much information as possible about the parameter from the sample(s) based on the sample statistic.

In other words, point estimates are summary statistics that capture the essence of the parameter being sought. However, there are several ways in which a parameter can be represented. As an example, in estimating the central tendency, which is a population parameter, it is generally agreed that the mean and the median are both reasonable quantities with which to measure such a parameter. Also in estimating the variance of a random variable, the sample variance and the range are both used as estimators. Obviously, only one of estimates can be used or employed at any one time. Thus, there needs to be a set of criteria, standards, or properties by which to judge or characterize the estimators. The properties of unbiasedness, and efficiency are two of the commonly sought-after properties that are desired in a good estimator.

A statistic $\overline{X}$ is called 'best unbiased estimator (BUE) for the parameter θ if the statistic is unbiased and efficient, i.e., if $E(\overline{X}) = \theta$ and if the variance of $\overline{X}$ is less than or equal to the variance of every other unbiased statistic. The issue of the efficiency of an estimator has to do with its variance. In terms of the BUE, the smaller the variance of an estimator, the more efficient the estimator.

In the case of the sample mean and median as estimators of the population central tendency, both are considered unbiased estimators, i.e., $E(\overline{X}) = \theta$ and $E(\tilde{X}) = \theta$. The variance of the sample mean and that of the median are as shown.

$$\textit{For the Mean, } V\left(\overline{X}\right) = \sigma_{\overline{X}}^2 = \frac{\sigma^2}{n}, \quad \textit{For the Median, } V\left(\tilde{X}\right) = \sigma_{\tilde{X}}^2 = \frac{1.57\sigma^2}{n}$$

The variance of the median is 1.57 times the variance of the mean. Therefore, using the criteria for BUE, the sample mean is considered the BUE because it has the minimum variance with respect to all the estimators of θ. As noted previously, both the Mid-range and the Mode are also unbiased estimators of the population mean but they are not BUE.

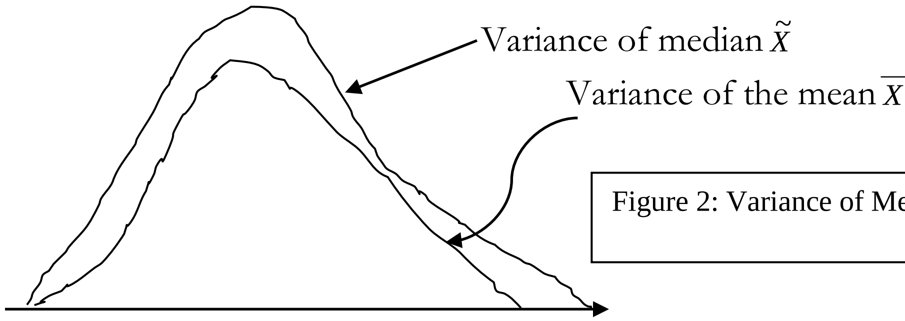


Figure 2: Variance of Mean & Median

3.1.1 Point Estimates for the Mean and Variance of the Population

The following are the point estimates for the mean and variance. For the mean, we have

$$\mu_{\bar{x}} = \frac{\sum \overline{X}}{n}, \quad and \quad \mu_{\bar{x}} = \frac{\sum_{i=1}^{k} \mu_x}{k}$$

Where k =number of subgroups and n is the sample size

$$\sigma_{\bar{x}}^{\ 2} = \frac{\sigma_x^{\ 2}}{\sqrt{n}}, \ where: \sigma_{\bar{x}}^{\ 2} = \frac{\sum(\overline{X} - \overline{\overline{X}})}{k-1}$$

3.1.2 Central Limit Theorem

The central limit theorem (CLT) is a statistical theory that states that given a sufficiently large sample size from a population with a finite variance, the mean of all samples from that population would be approximately equal to the mean of the population. Let $\overline{X}_1, \overline{X}_2, \overline{X}_3, \ldots\ldots \overline{X}_n$ denote the measurements or output of a random sample of size n from any distribution having finite variance σ^2 and mean μ, then the random variable $\dfrac{\sqrt{n}\left(\overline{X} - \mu\right)}{\sigma_X}$ has a limiting normal distribution with zero mean and variance equal to unity. In other words, even though the individual measurements have a distribution that is not the normal distribution, the distribution of the sample means $\overline{X}_1, \overline{X}_2, \overline{X}_3, \ldots\ldots \overline{X}_n$ as $n \rightarrow \infty$, tends to be approximately normally distributed. In other words, the sampling distribution of the sample means is the normal distribution. When this condition is true it would be possible to use this property to compute approximate probabilities concerning the distribution and to find an approximate confidence interval for μ as well as test certain hypotheses without knowing the exact distribution of μ in every case or situation.

The central limit theorem (CLT) establishes that, for the most commonly studied scenarios, when independent random variables are added, their sum tends toward a normal distribution even if the original variables themselves are not normally distributed. This is very important especially because it is often difficult to determine the underlying parent distribution which is needed to determine the probabilities of event occurrence to enable engineering decisions to be made in an informed manner

3.1.3 Sampling Distribution for the mean

The sampling distribution of the sample mean is the normal distribution based on the CLT. In other words, the distribution of the sample mean $\overline{X}$ is the normal distribution with the mean and variance as follows: $\mu_{\overline{X}} = \overline{\overline{X}}$, and $\sigma_{\overline{X}}^2 = \dfrac{\sigma^2}{n}$.

For example, the Diametral-Pitch (DP) for a sample taken from a lot or population of spur-gears, the first sample n₁ could yield a mean value $\overline{X}_1 = 4\,inches.$ A second sample n₂ from the same lot could yield a DP of $\overline{X}_1 = 6\,inches.$ The larger the DP, the higher the stress on the gear tooth.

Assume that on the average, the DP is 4 inches with S= 0.55 inches. A sample of 25 spur gears is taken from the lot with the following measurements as in table 1. Find the probability that some of the gear-spurs will not meet requirement, that is $P(\overline{X} > \mu_0)$.

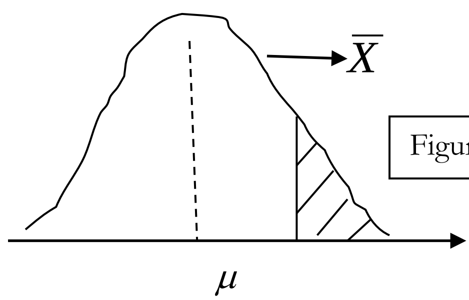


Figure 3: Sketch of Probability Distribution for X-bar

$$\mu_{\overline{X}} = 4,\ \sigma_{\overline{X}} = \frac{0.55}{\sqrt{25}} = 0.11, \overline{X}_0 = 4.149909(\,from\,table1)$$

$$P(\overline{X} > \mu_0) = P\left(\overline{X} > \frac{\overline{X}_0 - \mu_0}{\sigma_{\overline{X}}}\right) = 1 - \Phi\left(\frac{4.149909 - 4.0}{0.11}\right) = 1 - \Phi(1.363)$$

$$\Phi(1.363) = 0.9136$$

$$Shaded\ \ area = (1 - 0.9136) = 0.0864 \cong 9\%$$

There is only a 9% chance that the spur-gears from that population will not meet the design requirements. Note that we did not use the standard deviation we computed for the data for the problem. Why? You will recall that we are focused on the sampling distribution of the sample mean so the mean is the random variable in this particular case.

S/N	Diametral Pitch (DP) inches	S/N	Diametral Pitch (DP) inches
1	2.442966	14	4.616154
2	5.870707	15	3.914669
3	4.127012	16	5.484784
4	2.060597	17	3.387145
5	5.96805	18	2.976296
6	3.022355	19	4.135116
7	3.301695	20	4.422885
8	5.172247	21	5.738772
9	5.341773	22	4.407002
10	4.402271	23	4.690581
11	2.806244	24	3.997262
12	4.831229	25	3.40529
13	3.224622		
n	25		
Mean	4.149909		
Std Dev	1.09433		

Table 1: Diametral Pitch (DP) Measurements (in)

Later we will consider the sampling distribution for the variance based on the values from the data. The variance is considered a random variable because each sample realization (each sample we take) results in a variance estimate or statistic just like we have a mean estimate for each sample we take. Due to the unbiased nature of the sample mean as an estimator of the population mean, the sampling distribution of two or more means is normally distributed. sum of the means is also normally distributed.

3.1.4 Sampling Distribution for the Mean-The Student-t Distribution

The student-t arises when estimating the mean of a normally distributed population in those cases where the sample size is small, and/or the population variance or standard deviation is unknown.

The Student-t distribution is like the standard normal distribution when the sample size is small, typically n<30. Some of the characteristics of the student-t are the following:

1). The probability distribution appears to be symmetric about $t = 0$ just like the standard normal distribution
2). The probability distribution appears to be bell-shaped.

3). The density curve looks like a standard normal curve, but the tails of the *t*-distribution are "heavier" than the tails of the normal distribution. That is, we are more likely to get extreme *t*-values than extreme z values.

The nice thing about the student- t or the t distribution is that we can use it in the case where the sample size does not justify the use of the standard normal, that is when n<30.

Recall that in the case of the Standard Normal Variable, the random deviate $Z = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}}$

In the case of the student-t, the random variable t is give as

$t = \dfrac{\overline{X} - \mu}{S / \sqrt{n}}$ with υ=(n-1) degrees of freedom. Just like the Z-score, this is also called the t-score

3.1.5 Sampling Distribution for the Sample Variance

From the central limit theorem (CLT), we know that the distribution of the sample mean is approximately normal. Unfortunately, unlike the sample mean, there is no CLT analog for variance. However, when the individual observation X_is are from a normal distribution, there is a special condition under which we can consider the sampling distribution of the sample variance as follows. Suppose as indicated earlier, $X_1, X_2, \ldots, X_n$ are from a normal distribution $N(\mu,\sigma^2)$, and we will recall that the CLT applies to any arbitrary distributions. **The distribution of the sample variance is the Chi-Square distribution**. Note the following. For the $X_1, X_2, \ldots, X_n$,

$\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ is the mean, and $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(X - \overline{X}\right)^2$ is the sample variance then $\dfrac{(n-1)S^2}{\sigma^2}$ is the Chi-square distribution with (n-1) degrees of freedom. The Chi-square is available in most basic statistics texts.

3.1.6 Sampling Distribution for Two Variances

When we are concerned about the variances from two populations, the resulting sampling distribution of the combined variance of the two populations, follows the Snedecor's F-distribution or simply the F-Distribution. The sampling distribution for two variances is used to test whether the variances of two populations are equal. The F distribution is given as:

$F = \dfrac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2}$ with (υ_1, and υ_2) where $\upsilon_1 = n_1 - 1$ and $\upsilon_2 = n_2 - 1$; where the notation of 1 or 2 is perfunctory and depends on which variance is larger. Please note that for ease of computation, it is recommended that when taking ratios of sample variances, we should put the larger variance in the numerator and the smaller variance in the denominator. We will see how this is done with a numerical example later. To use this test, the following must hold:

- Both populations are normally distributed

- Both samples are drawn independently from each other.
- Within each sample, the observations are sampled randomly and independently of each other.

Interval (Confidence interval) Estimators

4.1 Error of Estimation

In practical situations, there are usually two types of estimation problems. In one case, we may have a constant ϕ which represents a theoretical quantity that has to be determined by means of measurements. For example, the time it takes to complete a machining operation, the amount of yield from a given reaction, the number of material handling moves required for a certain material handling type, and so on. The result Y of the measurement activity is a random variable whose distribution function depends on the constant ϕ (and perhaps other quantities).

The parameter or the unknown constant has to be estimated from the measurements taken namely; $X_1, X_2, X_3, ..., X_n$. In the other case, the quantity itself is a random variable, for example, the weights in a filling operation, the length of pistons from a given machine, and so on. In these types of cases we are interested in the average value and/or the dispersion of X, where X is the random variable of interest. Hence, we have to compute E(X) for the mean, and or $D^2(X)$, where $D^2(X) = \sigma^2$.

If we want a single number to use in place of the unknown constant or parameter, then point estimation is the appropriate method. If we are using a good estimator, based some of the criteria we discussed earlier, then it is understood that the resulting estimate should probably be close to the unknown true value. <u>We know that an estimator is subject to error of measurement (in the case of the constant) and variability (in the case of the random variable)</u>. In other words, the single number (or statistic) does not include any indication as to probability that the estimator has taken on a value close to the unknown parameter value. Consequently, it is instructive to have some information or some knowledge about the amount of deviation of the computed statistic from the true value (in our case the true mean or the true deviation). This is where confidence intervals come in because due to the variability or the error in measurement, we want to establish an interval within which we would reasonably expect the parameters value we seek to lie. In other words, in repeated sampling and using the same method to select the different samples, we would expect the true parameter value to fall within the specified interval a given percent of the time. **Thus, confidence intervals are established for parameter values NOT for sample statistics.**

For example, a 95% confidence interval means that in repeated sampling and using the same sampling method, we would expect the true parameter value to fall or lie within our confidence interval 95% of the time.

Let us do some housekeeping before we delve deep onto the area of Confidence intervals. First let us look at the error associated with the estimate $\overline{X}$.

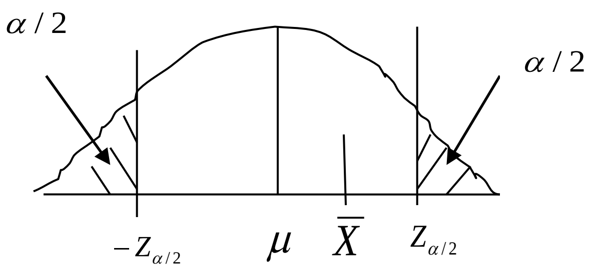


Figure 4: Confidence Interval for the Mean

We know that $Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ is true Z for either positive or negative depending on where Z is relative to the mean μ.

$$-Z_{\alpha/2} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < Z_{\alpha/2}$$

We want to solve for

$$-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \overline{X} - \mu < +Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \Rightarrow\Rightarrow -\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < -\mu < -\overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$multiply\ \ by\ \ (-1)$$

$$\overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} > \mu > \overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Thus, the confidence interval for the mean which is a probability statement is given by:

$$P\left(\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

4.2 Determination of Sample Size

If we examine at the error associated with the mean X-bar say E where E is given by

$E = \overline{X} - \mu$, we can re-express Z as follows

$$-Z_{\alpha/2} < \frac{E}{\sigma/\sqrt{n}} < Z_{\alpha/2}, \Rightarrow E = \pm\left(Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) \Rightarrow |E| = \left(Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

$$n = \left[\frac{Z_{\alpha/2}\sigma}{E}\right]^2 \ \ if\ \ \sigma\ \ is\ \ known$$

$$n = \left[\frac{t_{\alpha/2}S}{E}\right]^2, \ if\ \ \sigma\ \ is\ \ unknown\ \ \ but\ \ estimated\ \ from\ \ S$$

The question you might have is what does this all mean or why do we need n. Well the problem is that typically no one will give you the value of n to use as your sample size. What usually happens is a company may have a policy on the size of the error for a process which they have determined historically. Given that value and the level of confidence specified based on the data, then the sample size needed to cover that error is computed. A company may say that it is comfortable with an error of ±10% being the error between the true mean and the estimated mean.

Example: A company is willing to accept an error of **± 15% with a 90% confidence**.

a). Assuming that the variance is known and σ =1.5 units. What sample size is needed to guarantee this level of protection? b) Assuming that variance is unknown and that somehow the company has a value of the process sample standard deviation that was estimated from experience with a value of S=2 with 90% confidence, what sample size will be required?

a). $n = \left[\frac{Z_{\alpha/2}\sigma}{E}\right]^2$ *if σ is known*, α =(1-0.90)=0.1, α/2=0.05, E=0.15

from the standard normal Table $Z_{0.05}=Z_{0.95}=1.645$, $n = \left[\frac{Z_{\alpha/2}\sigma}{E}\right]^2 = \left[\frac{(1.645)1.5}{0.15}\right]^2 = 270.6 \approx 271$

b) $n = \left[\frac{t_{\alpha/2}S}{E}\right]^2$ *if σ is unknown*

Strictly speaking, there is no way we can evaluate this without knowing the sample size. Remember that to evaluate the t-statistic we need the degrees of freedom equal to n-1. So even though the variance is unknown, we do have the estimate of S (where S is an estimate of σ) determined historically from the sample, we can use the Z distribution in place of the t-distribution to evaluate the sample size. Note that the t-statistic and the Z-statistic is identical when n= infinity. So, in this case we will use the value of t-statistic with υ =∞. **From the formula for n, the smaller the error E, the larger the sample (n) required to detect the error and conversely, the larger the error, the smaller the sample size required to detect it.**

$$n = \left[\frac{t_{\alpha/2}S}{E}\right]^2 = \left[\frac{(1.645)2}{0.15}\right]^2 = 480$$

4.3 Confidence Intervals for the Mean

The method of confidence intervals is meant to provide an indication of both the actual numerical value of the parameter and also the level of confidence, based on the sample information, that we have a correct indication of the possible value of the unknown parameter or constant. We have three different scenarios, namely Case I, Case II, and Case III

4.3.1 Case I

Confidence Interval for the population mean μ with σ^2 (the population variance) known or assumed. The sampling distribution is the normal distribution and the test statistic is the standard normal deviate. The (1-α) Confidence Limits for μ is:

$$P\left(-Z_{\frac{\alpha}{2}} < \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < Z_{\frac{\alpha}{2}}\right) = 1-\alpha \Rightarrow P\left(\overline{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

4.3.2 Case II

Confidence Interval for the population mean μ with σ^2 (the population variance) unknown and n >30. The sampling distribution is again the normal and the test statistic is the standard normal deviate. The (1-α) Confidence Limits for μ is computed by replacing or estimating **σ** using sample standard deviation. Note: The limits of the confidence interval are referred to as the Upper Confidence Limit (UCL) and the lower as the Lower Confidence Limit or the (LCL)

$$P\left(-Z_{\frac{\alpha}{2}} < \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} < Z_{\frac{\alpha}{2}} \right) = 1-\alpha \;\Rightarrow\; P\left(\overline{X} - Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \overline{X} + Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \right) = 1-\alpha$$

4.3.3 CASE III

Confidence Interval for the population mean μ with σ^2 (the population variance) unknown and n<30. Replace σ with sample standard deviation s. The sampling distribution is the student t distribution and the test statistic is the student T statistic. Hence replace Z with the t with degrees of freedom df = υ, where υ = n-1.

$$P\left(-t_{\frac{\alpha}{2},\nu} < \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} < t_{\frac{\alpha}{2},\nu} \right) = 1-\alpha \;\Rightarrow\; P\left(\overline{X} - t_{\frac{\alpha}{2},\nu} \frac{s}{\sqrt{n}} < \mu < \overline{X} + t_{\frac{\alpha}{2},\nu} \frac{s}{\sqrt{n}} \right) = 1-\alpha$$

If the individual values are normally distributed, then the confidence interval for μ is the following:

$$P\left(X - Z_{\frac{\alpha}{2}}\sigma < \mu < X + Z_{\frac{\alpha}{2}}\sigma \right) = 1-\alpha$$

EXAMPLE: CASE I

For a grinding operation A, assume that n = 25, $\overline{X}$ = 75 minutes, σ =10 minutes. Find a two-sided 95% CI (Confidence Interval) for μ.

Since σ is given, we will assume that the sampling distribution is the normal distribution with the sample statistic equal to the Z.

$$Z_{\frac{\alpha}{2}} = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

$$P\left(\overline{X} + Z_{\frac{\alpha}{2}}\left(\sigma/\sqrt{n}\right) < \mu < \overline{X} - Z_{\frac{\alpha}{2}}\left(\sigma/\sqrt{n}\right) \right) = 1-\alpha$$

$$Z_{0.975} = 1.96 = Z_{0.025}$$
$$(1.96(10))/5 = 3.92$$
$$CI = 75 \pm 3.92$$
$$\begin{Bmatrix} UCL = 75 + 3.92 = 78.92 \\ LCL = 75 - 3.92 = 71.08 \end{Bmatrix} \quad \boxed{P(\ 71.08 < \mu < 78.92\ ) = 0.95}$$

This says that in repeated sampling and under the same sampling scheme, we will expect the mean of the population μ to lie in the interval: [71.08, 78.92] 95% of the time.

Example Case II

For the grinding time (in sec) for another product B, Let n = 36, $\overline{X}$ = 70 seconds, S =8 seconds. Find a two-sided 90% CI (Confidence Interval) for μ

Since n>30, we will assume that the sampling distribution is the normal distribution with the sample statistic equal to the Z.

$$Z_{\frac{\alpha}{2}} = \frac{\overline{X} - \mu}{s/\sqrt{n}}$$
$$P\left(\overline{X} + Z_{\frac{\alpha}{2}}\left(s/\sqrt{n}\right) < \mu < \overline{X} - Z_{\frac{\alpha}{2}}\left(s/\sqrt{n}\right)\right) = 1 - \alpha$$
$$Z_{0.95} = 1.645 = Z_{0.05}, \Rightarrow (1.645(8))/6 = 2.193, \quad CL = 70 \pm 2.193$$
$$UCL = 70 + 2.193 = 72.193,\ LCL = 70 - 2.193 = 67.807$$

Example Case III

For yet another product C, let n = 25, $\overline{X}$ = 15 minutes, S =1.5minutes. Find a 95% CI (Confidence Interval) for μ. Since n<30, we will assume that the sampling distribution is the Student-t distribution with the sample statistic equal to the t.

$$t_{\frac{\alpha}{2},\upsilon} = \frac{\overline{X} - \mu}{S/\sqrt{n}} \textit{ with } \upsilon(= n-1) \text{ deg}\textit{rees of freedom}$$
$$P\left(\overline{X} + t_{\frac{\alpha}{2},\upsilon}\frac{S}{\sqrt{n}} < \mu < \overline{X} - t_{\frac{\alpha}{2},\upsilon}\frac{S}{\sqrt{n}}\right) = 1 - \alpha,\ \upsilon = n-1$$

Note that student-t is not symmetric so

$$t_{0.025,14} = 2.14$$
$$(2.14)(1.5)/5 = 0.642$$
$$CI = 15 \pm 0.642, \begin{Bmatrix} UCL = 15.0 + 0.642 = 15.642 \\ LCL = 15.0 - 0.642 = 14.358 \end{Bmatrix} \quad \boxed{P(\ 14.358 < \mu < 15.642\ ) = 0.975}$$

4.4 Confidence Intervals for One Variance

The sampling distribution of the variance (one variance) is the Chi-square distribution with (n-1) degrees of freedom. Please note that unlike the symmetric normal distribution, the Chi-square is not symmetric so the α values corresponding to the tails of the distribution are different. As a matter of fact, the Chi-square is a skewed distribution. Consider the following 6 data points (in inches) from a welding process: {24, 28, 21, 23, 32, 22}

$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ =25 inches, the median $\breve{X} = 23.5\ inches$, the Range =(32-21)=11 inches

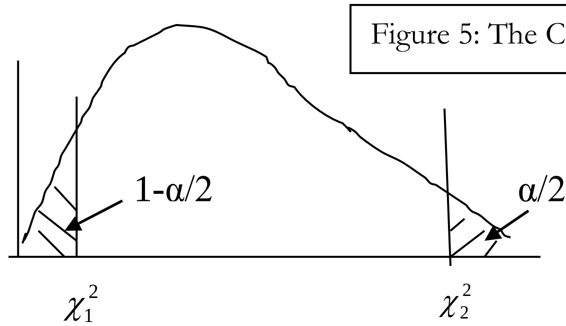
$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X - \overline{X}\right)^2$ is the sample variance= $\sqrt{\frac{1^2+3^2+4^2+2^2+7^2+3^2}{5}} = 4.195\ inches$

We indicated earlier that that statistic $\frac{(n-1)S^2}{\sigma^2}$ is approximately Chi-square distribution, that is

$\frac{(n-1)S^2}{\sigma^2} \approx \chi_{n-1}$. Thus, we can establish the confidence interval as

$$P\left(\chi_1^2 < \frac{(n-1)S^2}{\sigma^2} < \chi_2^2\right) = 1-\alpha$$

Figure 5: The Chi-square Distribution



Want: $P\left(LCL < \sigma^2 < UCL\right) = 1-\alpha$

From $P\left(\chi_1^2 < \dfrac{(n-1)S^2}{\sigma^2} < \chi_2^2\right) = 1-\alpha$

We have: $\dfrac{1}{\chi_1^2} > \dfrac{\sigma^2}{(n-1)S^2} \Rightarrow \dfrac{(n-1)S^2}{\chi_1^2} > \sigma^2$; and $\dfrac{(n-1)S^2}{\chi_2^2} < \sigma^2$

Hence: $P\left(\dfrac{(n-1)S^2}{\chi_2^2} < \sigma^2 < \dfrac{(n-1)S^2}{\chi_1^2}\right) = 1-\alpha$, hence $P\left(\dfrac{(n-1)S^2}{\chi^2_{\alpha/2,\,(n-1)}} < \sigma^2 < \dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2,\,(n-1)}}\right) = 1-\alpha$

Please note that for the same degrees of freedom (n-1), $\chi^2_{\alpha/2,\,(n-1)} > \chi^2_{1-\alpha/2,\,(n-1)}$

Example: Assume for the grinding example product D, the sample size n=25, s=4.9, establish a 95% confidence interval for the variance σ^2.

$P\left(\dfrac{24(4.9)^2}{\chi^2_{0.025,24}} < \sigma^2 < \dfrac{24(4.9)^2}{\chi^2_{0.975,24}}\right) = 0.95$, from the table $\chi^2_{0.025,24}$ =39.364, $\chi^2_{0.975,24}$ =12.401

$$P\left(\frac{24(4.9)^2}{39.364} < \sigma^2 < \frac{24(4.9)^2}{12.401}\right) = 0.95 \Rightarrow P\left(14.639 < \sigma^2 < 46.467\right) = 0.95$$

$$P\left(3.83 < \sigma < 6.82\right) = 0.95$$

4.4 Confidence Intervals for Two Variances (σ_1^2, σ_2^2)

The Test Statistic is:

$F = \dfrac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2}$, Note: Since $\boxed{F_{\alpha/2,\,\upsilon_2,\upsilon_1} > F_{1-\alpha/2,\,\upsilon_2,\upsilon_1}}$

Then: $P\left[ULC \le \dfrac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2} \le LCL\right] = 1-\alpha, \Rightarrow P\left[F_{1-\alpha/2,\,\upsilon_2,\upsilon_1} \le \dfrac{S_2^2\sigma_1^2}{S_1^2\sigma_2^2} \le F_{\alpha/2,\,\upsilon_2,\upsilon_1}\right] = 1-\alpha$

$$P\left[\frac{S_1^2}{S_2^2}F_{1-\alpha/2,n_2-1,n_1-1} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{S_1^2}{S_2^2}F_{\alpha/2,n_2-1,n_1-1}\right] = 1-\alpha$$

Example:

Suppose that in the Diametral Pitch example, the Spur-gears were supplied by two suppliers/clients, with the following data, Supplier 1: n_1=21, S_1=0.56 inches, Supplier 2, n_2=16 S_2=1.8 inches. Find a 95% confidence interval on the ratio of the two variances. Assume that the processes are independent, and the Spur-gear operations are normally distributed

$$P\left[\frac{S_1^2}{S_2^2}F_{0.975,15,20} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{S_1^2}{S_2^2}F_{0.025,15,20}\right] = 1-\alpha$$

$$\left[\frac{(0.56)^2}{(1.8)^2}\frac{1}{F_{0.025,\ 20,\ 15}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{(0.56)^2}{(1.8)^2}F_{0.025,\ 15,\ 20}\right] = \frac{(0.56)^2}{(1.8)^2}\frac{1}{2.76} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{(0.56)^2}{(1.8)^2}2.57$$

$$P\left[0.035 \le \frac{\sigma_1^2}{\sigma_2^2} \le 0.249\right] = 1-\alpha$$

4.6 One-Sided Confidence Interval

Under certain conditions only one-sided intervals may be of interest. For example, take the case of still bars where we want the measured strength to be as high as possible. Our major concern then is that the strength values do not go beyond a certain lower limit. In that case, we will be establishing a lower confidence (one-sided) interval rather than a two-sided interval. On the other hand, we may have a variable (say the number of defects) in which case we want the value to be as close to zero as possible. In that case, we only worry about how high the value can go. So, we want to establish a one-sided confidence interval. A one-sided confidence interval is looked at as a one-tailed interval (UCL or LCL but not both) unlike the two tails of the two-sided confidence Interval. That being the case we use α rather than α/2.

$$UCL_t:\ P\left(\mu < \overline{X} + t_{\alpha,\upsilon}\frac{S}{\sqrt{n}}\right) = 1-\alpha, \qquad LCL_t:\ P\left(\mu > \overline{X} - t_{\alpha,\upsilon}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

$$UCL_Z:P\left(\mu < \overline{X} + Z_{\alpha,}\frac{(S\ or\ \sigma)}{\sqrt{n}}\right) = 1-\alpha, \qquad LCL_Z:\ P\left(\mu > \overline{X} - Z_{\alpha}\frac{(S\ or\ \sigma)}{\sqrt{n}}\right) = 1-\alpha$$

$$UCL_\sigma:P\left(\sigma^2 < \frac{24(4.9)^2}{\chi^2_{0.975,24}}\right) = 1-\alpha, \qquad LCL_\sigma:P\left(\sigma^2 > \frac{24(4.9)^2}{\chi^2_{0.025,24}}\right) = 1-\alpha$$

Let us use the example for CASE III example to illustrate. Assume now that we want a 95% lower confidence interval (LCL) for the grinding duration of product C.

$$P\left(\mu < \overline{X} - t_{\alpha,\upsilon}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

$$t_{0.05,14} = 1.761\ \textit{from the student}-t\ \textit{table}$$

$$\boxed{\begin{array}{l} 1.761(1.5)/5 = 0.5283 \\ LCL = 15.0 - 0.5283 = 14.358 \\ P(\mu < 14.358) = 95\% \end{array}}$$

Test of Hypothesis

A test of hypothesis is a test on an assumption or statement that may or may not be true concerning the parameter of the population of interest. The truth or falsity of such a test can only be known if the entire population is examined. Since this is impractical in most situations, a random sample is taken from the population and the information used to deduce whether the hypothesis is likely true or not. Evidence from the sample that is inconsistent with the stated hypothesis leads to a rejection whereas evidence supporting the hypothesis leads to its acceptance. **The acceptance of a statistical hypothesis does not necessarily imply that it is true.** It does not necessarily mean that the hypothesis is true because if we have another set of data, the decision might be different. The acceptance of a statistical hypothesis is simply an indication that, the data on hand and only because of the data on hand, we have been led to accept the hypothesis. It does not necessarily mean that the hypothesis is true because if we have another set of data, the decision might be different. This is where the issue of variability comes in

The hypotheses that are formulated with the hope of rejecting are called null hypotheses and denoted by H_o. The rejection of H_o leads to the acceptance of an alternate hypothesis denoted by H_1. The decision to reject or not reject a hypothesis is based on the value of the test statistic. The test statistic is compared to a critical value. The critical value is based on the level of significance of the test and represents values in the critical region as defined by the significance level. Depending on the nature of the test, that is:

Less than ($\mu < \mu_o$)
Greater than ($\mu > \mu_o$)
Not Equal ($\mu \neq \mu_o$).

Based on the value of the test statistics as compared to the critical value (or the table value of the significance level of the test), a decision is made to reject or not reject the null hypothesis. In such test of hypothesis, if the test statistic falls in the acceptance region, then H_o is not rejected, else it is rejected. The hypothesis is then specified as:
The null is give as:

$$H_o: \mu = \mu_o$$

The alternative is given in the form of one of the following:

$$H_1: \mu < \mu_o$$
$$H_1: \mu > \mu_o$$
$$H_1: \mu \neq \mu_o$$

5.1 Errors Associated with Decisions on Test of Hypothesis

Decision to reject or not reject a test naturally leads to two possible types of errors. The reason for the error is that the decision is made based on information from a sample rather than the actual process population itself. The fact is that we are trying to ascertain the true state of nature using information from the sample. We of course do not know the true state of nature and would like to INFER it from the sample. This notion is perhaps one of the most important foundations of statistics, namely the fact that while we do in fact seek the population value we can only approach that value by way of the sample value which in and of itself is of limited value unless it points us to or gives us the population value. All samples are taken not for their own sake but to provide information or inference about the population value. The errors are the errors of **Type I (α)**, and **Type II (β)**.

5.1.1 Type I Error (α)

This type of error is committed when the null Hypothesis (H_0) is rejected.

5.1.2 Type II Error (β)

This is the type of error committed when the null Hypothesis (H_0) is not rejected. This is loosely referred to as accepting the null Hypothesis. It is a more consequential and less forgiving error than the type I or alpha error.

These errors are aptly demonstrated by the schematic in Table 2.

	TRUE STATE OF NATURE	
	H_0 True	**H_0 False**
DECISION Accept	NO ERROR	TYPE II ERROR
DECISION Do Not Accept	TYPE I ERROR	NO ERROR

Table 2: Schematic for Type I and Type II Errors

5.1.3 The Relationship Between the Type I (α) and Type II (β) errors

Note that α and β are always at the opposite side of the target or what we will later call the dividing line of criteria. However, it is important to note that we cannot talk about committing a type II error (β) if we do not know what the true mean value is. In order words, you can only have made a mistake when you know what the target or what the aimed at value is. If we look at the real implication of the type II error, it says that we are accepting the null hypothesis when it is false.

This is a very serious error that is not taken lightly. And to say that we committed such an error, we must know what the true state of nature is to say that we did commit the error of Type II. This says that given that we have μ_0 and the sample X or $\overline{X}$ we can look at the probability of type I error as the probability of rejecting the null hypothesis when indeed it is true. However, if we say we accept the null hypothesis then we must know the true mean value to say that indeed accepted something we should not have. That true mean value is denoted as μ_1. So μ_0 is related to type I error (α) and μ_1 is related to type II error (β). Note that sometimes rather than specifically talk about β, we talk about (1- β) which is also referred to as the power of the test.

5.1.4 Computation of the Required Sample Size (n) given (α) and (β)

In order to exert some control over a process, the engineer might specify the size of both Type I and Type II errors that the system can tolerate. The question then what value of n (sample) would help guarantee the level of protection based on these error levels.

When the underlying process is normally distributed or when our focus is on the mean of the process (as you may recall even if the process is not normally distribution according to the central limit theory, the means from the process follow the normal distribution).

Assume we have specified α and μ_0. If we also specify β then we must necessarily specify μ_1. A ketch of the relationship between these parameters will help explain the procedure

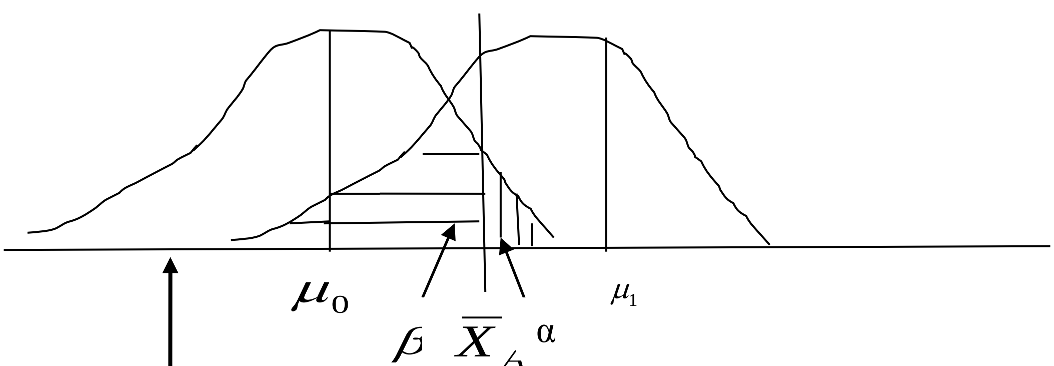
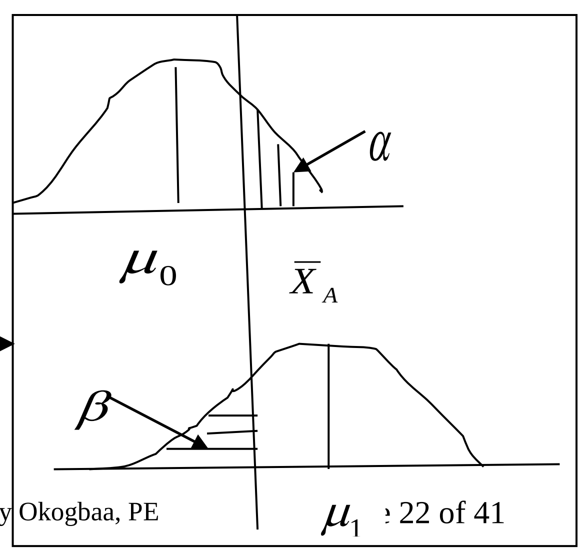


Figure 6a: Location of α and β, $\mu_1 > \mu_0$

Figure 6b: Exploded view of Fig 6a



$$\frac{\overline{X_A}-\mu_1}{\sigma/\sqrt{n}}=-Z_\beta \Rightarrow\Rightarrow \overline{X_A}=\mu_1-Z_\beta\left(\sigma/\sqrt{n}\right)\text{...........(1)}$$

$$\frac{\overline{X_A}-\mu_0}{\sigma/\sqrt{n}}=\ Z_\alpha \Rightarrow\Rightarrow \overline{X_A}=\mu_0+Z_\alpha\left(\sigma/\sqrt{n}\right)\text{..........(2)}$$

$$Solving:\ (1)\ -\ (2):\left(\mu_1-\mu_0\right)-Z_\beta\frac{\sigma}{\sqrt{n}}-Z_\alpha\frac{\sigma}{\sqrt{n}}=0\Rightarrow\frac{\sigma}{\sqrt{n}}\left(Z_\alpha+Z_\beta\right)=\left(\mu_1-\mu_0\right)$$

$$\frac{\left(Z_\alpha+Z_\beta\right)\sigma}{\left(\mu_1-\mu_0\right)}=\sqrt{n}\Rightarrow\frac{\left(Z_\alpha+Z_\beta\right)}{\left(\frac{\mu_1-\mu_0}{\sigma}\right)}\Rightarrow n=\frac{\left(Z_\alpha+Z_\beta\right)^2}{\left(\frac{\mu_1-\mu_0}{\sigma}\right)^2}\text{..........(3)}$$

$$\frac{\sigma}{\sqrt{n}}\left(Z_\alpha+Z_\beta\right)=\left(\mu_1-\mu_0\right)$$

$$Z_\beta=\frac{\left(\mu_1-\mu_0\right)}{\sigma}\sqrt{n}\ -Z_\alpha\Rightarrow Z_\beta=\frac{\left(\mu_1-\mu_0\right)\sqrt{n}}{\sigma}-Z_\alpha\quad\text{..........(4)}$$

Example

Let $\mu_0=100$, $\sigma=10$, $\alpha=0.05$. Let $\beta=0.1$ for $\mu_1=110$

Compute n that will provide the level of protection given by the type I and type II errors

$$n=\frac{\left(Z_{0.05}+Z_{0.10}\right)^2}{\left(\frac{110-100}{10}\right)^2}=\left(\frac{1.645-1.282}{1}\right)^2=\left(2.927\right)^2=8.57\approx 9$$

Note: if $\mu_1<\mu_0\Rightarrow\sqrt{n}=\dfrac{\left(Z_\alpha-Z_\beta\right)}{\left(\frac{\mu_1-\mu_0}{\sigma}\right)}\Rightarrow n=\dfrac{\left(Z_\alpha-Z_\beta\right)^2}{\left(\frac{\mu_1-\mu_0}{\sigma}\right)^2}$

5.1.5 Computation of β when $\mu_1>\mu_0$

Suppose $\mu_0=100$, $\sigma=10$, $\alpha=0.05$. Let $\beta=?$ for $\mu_1=110$

Previously we observed: $\left(\mu_1-\mu_0\right)-Z_\beta\frac{\sigma}{\sqrt{n}}-Z_\alpha\frac{\sigma}{\sqrt{n}}=0\Rightarrow\Rightarrow Z_\beta\frac{\sigma}{\sqrt{n}}=\left(\mu_1-\mu_0\right)-Z_\alpha\frac{\sigma}{\sqrt{n}}$

$$Z_\beta=\frac{\left(\mu_1-\mu_0\right)\sqrt{n}}{\sigma}-Z_\alpha$$

$Z_\beta=\Delta\sqrt{n}-Z_\alpha$, where $\Delta=\left(\dfrac{\mu_1-\mu_0}{\sigma}\right)$, Note: If $Z_\beta<0$, then $\beta>0.5$ or 50%

For our example: $\mu_0=100$, $\sigma=10$, $\alpha=0.05$. $\beta=?$ for $\mu_1=110$

$Z_{\beta} = \Delta\sqrt{n} - Z_{\alpha} \Rightarrow 1\sqrt{12} - 1.645 = 3.464\text{-}1.645 = 1.82$, $\Phi(1.82) = 0.9656$

From Figure 6, $\beta = 1 - \Phi(1.82) = 0.0344$ or 3%

5.1.6 Computation of β when $\mu_1 < \mu_0$

To complete this important example, let look at the case when $\mu_1 < \mu_0$

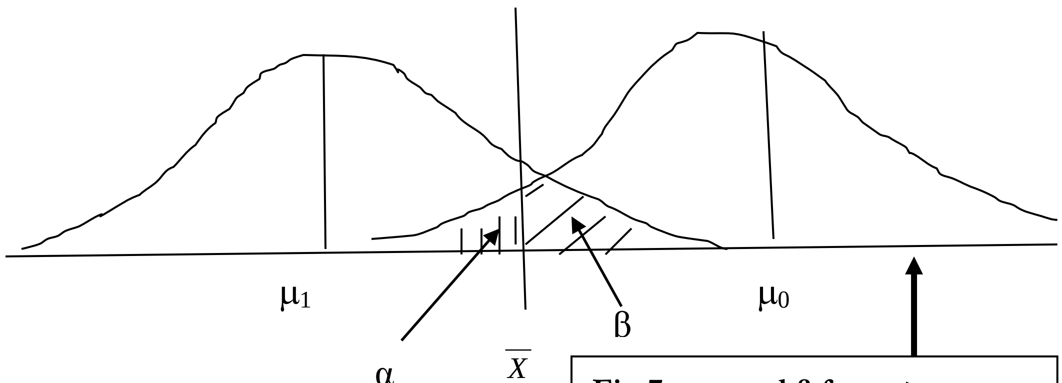


Fig 7a: α and β for $\mu_0 > \mu_1$

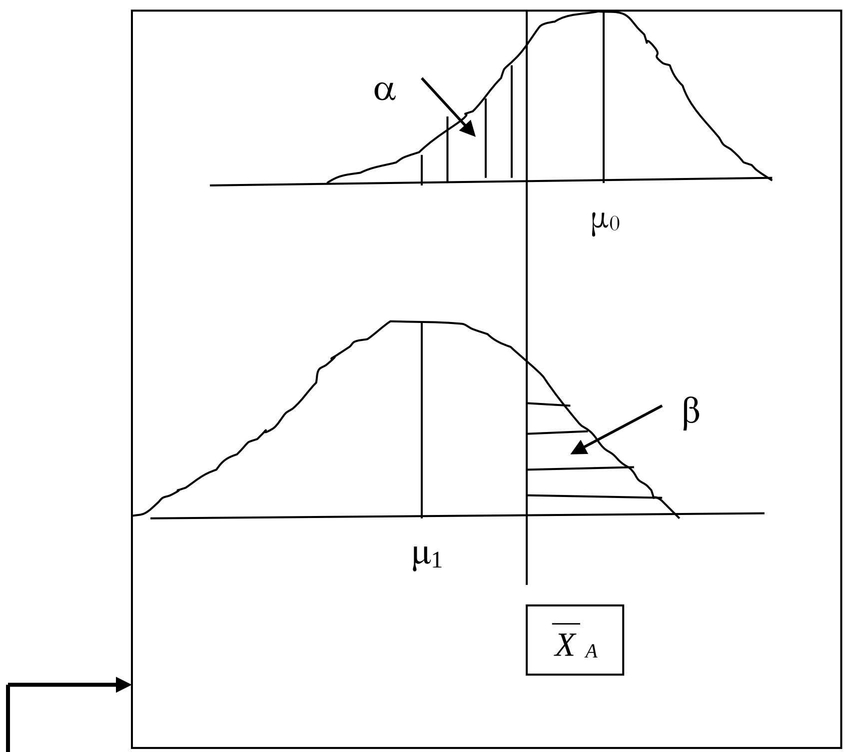


Fig 7b: Exploded view of α and β for $\mu_0 > \mu_1$

$$\frac{\overline{X}_A - \mu_0}{\sigma/\sqrt{n}} = Z_\alpha \quad \Rightarrow \overline{X}_A = \mu_0 + Z_\alpha \frac{\sigma}{\sqrt{n}} \text{ ....(1)}$$

$$\frac{\overline{X}_A - \mu_1}{\sigma/\sqrt{n}} = Z_\beta \quad \Rightarrow \overline{X}_A = \mu_1 + Z_\beta \frac{\sigma}{\sqrt{n}} \text{.................(2)}$$

$$(1)-(2): \mu_0 - \mu_1 = Z_\alpha \frac{\sigma}{\sqrt{n}} - Z_\beta \frac{\sigma}{\sqrt{n}} \Rightarrow \sqrt{n}\left(\frac{\mu_0 - \mu_1}{\sigma}\right) = Z_\alpha - Z_\beta$$

$$Let \quad \Delta = \left(\frac{\mu_0 - \mu_1}{\sigma}\right) \Rightarrow \sqrt{n}\Delta = Z_\alpha - Z_\beta \Rightarrow -Z_\beta = \sqrt{n}\Delta - Z_\alpha$$

$$Z_\beta = 1 - \Phi\left(-Z_\alpha + \sqrt{n}\Delta\right)$$

$$\sqrt{n} = \frac{(Z_\alpha - Z_\beta)}{\left(\frac{\mu_0 - \mu_1}{\sigma}\right)} = \frac{(Z_\alpha - Z_\beta)}{-\left(\frac{\mu_1 - \mu_0}{\sigma}\right)} \Rightarrow n = \frac{(Z_\alpha + Z_\beta)^2}{\left(\frac{\mu_1 - \mu_0}{\sigma}\right)^2}$$

For $\mu_1 > \mu_0$, Z_β, = $Z_\beta = \Delta\sqrt{n} - Z_\alpha$

For $\mu_1 < \mu_0$, Z_β, = $1 - \Phi\left(-Z_\alpha + \sqrt{n}\Delta\right)$

For $\mu_1 \neq \mu_0$, $Z_\beta = \Delta\sqrt{n} - Z_{\alpha/2}$

n= $\dfrac{(Z_\alpha + Z_\beta)^2}{\left(\frac{\mu_1 - \mu_0}{\sigma}\right)^2}$ for both $\mu_1 > \mu_0$ and $\mu_1 < \mu_0$

n= $\dfrac{(Z_{\alpha/2} + Z_\beta)^2}{\left(\frac{\mu_1 - \mu_0}{\sigma}\right)^2}$ for $\mu_1 \neq \mu_0$

5.2 Operating Characteristic Curve

Operating characteristic curves are useful tools for exploring the power of a control process. Typically used in conjunction with standard quality control plots, OC curves provides a mechanism to gauge how likely it is that a sample statistic is not outside of the control limits when, in fact, it has shifted by a certain amount? This probability is usually referred to as β or Type II error probability, that is, the probability of erroneously accepting the 'true state of nature' (e.g. mean, variance, etc.) as being "in control" when in fact it is not. The OC curve also provides another measure of the test in the context of its overall power, namely know the extent to which the test can detect the effect or shift in quality level of a given metric, often referred to as the 'power of the test' and is denoted by 1-β. Note that operating characteristic curves pertain to the false-acceptance probability using the sample-outside-of- control-limits criterion. The sample size for establishing an OC curve is determined by the cost of implementing the plan (e.g., cost per item sampled) and on the costs resulting from *not* detecting quality problems and thus passing unfit products. The OC curve provides the ability to assess the risk associated with each quality level when there is a shift in the process quality.

5.2.1 Computation of the Parameters of the OC Curve

β (or the Type II error) is the probability of accepting the original hypothesis H_0 when it is not true or when some alternative hypothesis, H_1 is true. Thus β is a function of the value of the test statistic that is less (or greater) than the hypothesized value. Suppose the critical Value (Y-bar) for the mean μ based on a 95% CI is 18.8. Also, let n=25, and σ=2. We can now examine how β varies for different values of μ.

μ	$Z = \left(\dfrac{18.8 - \mu}{\sigma / \sqrt{n}} \right)$	β	1-β (power of the test)
18	2	0.02	0.98
18.3	1.25	0.11	0.89
18.5	0.75	0.23	0.77
19	-0.5	0.69	0.31
19.2	-1.0	0.84	0.16
19.5	-1.75	0.96	0.04

Table 3: Computation of the parameters of the OC Curve

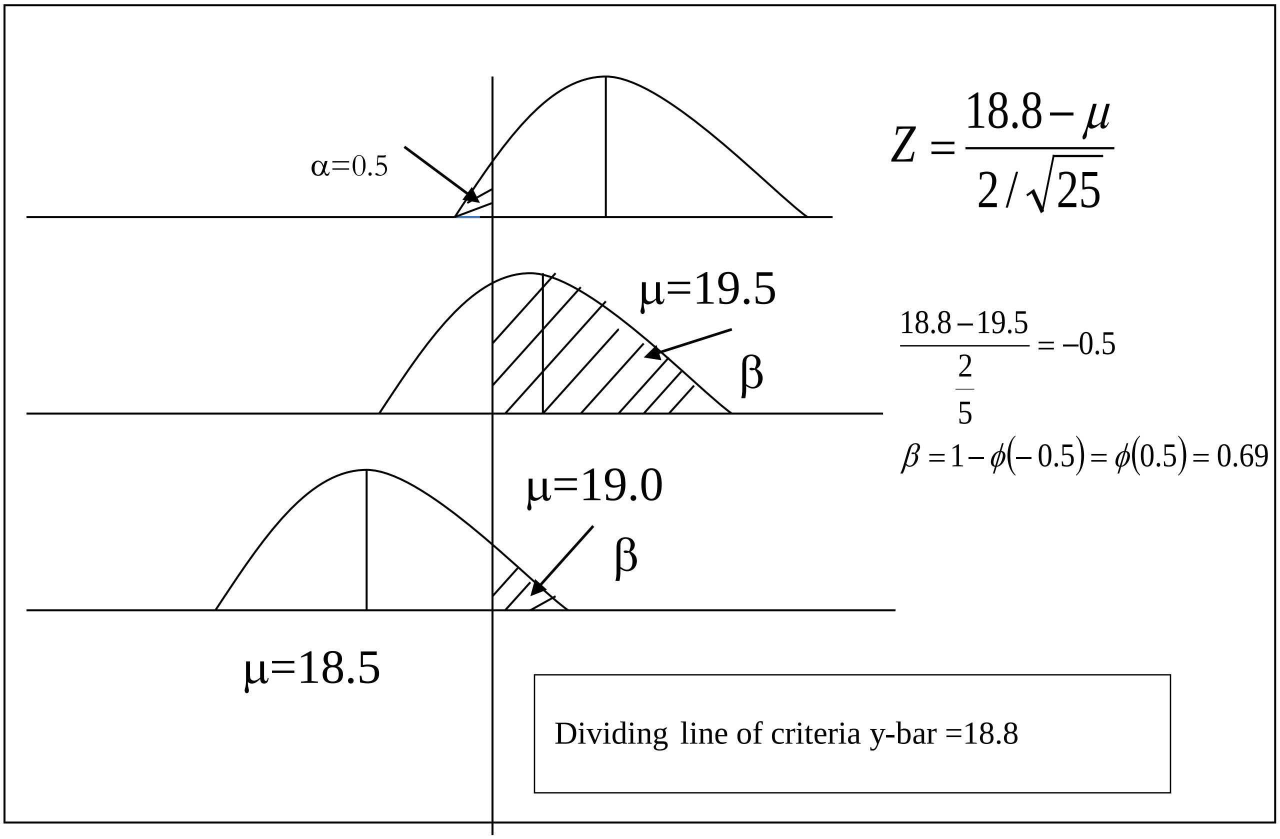


Figure 8: Computation of Operating Characteristic Curve Data Points

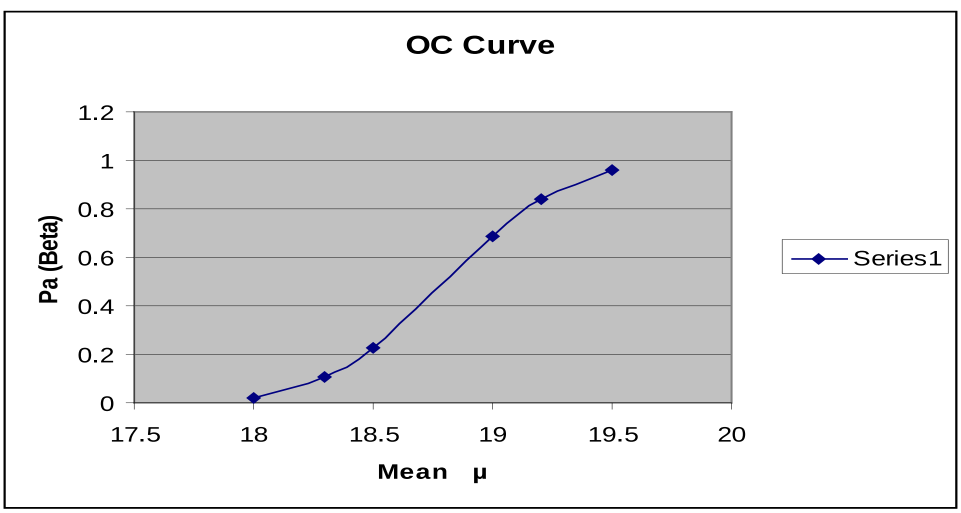


Figure 9: Plot of the Operating Characteristic Curve

5.3 Steps in Hypotheses Testing

5.3.1 Set up the Hypothesis and its alternative

Example H_o: $\mu = 19.5$ g
H_1: $\mu < 19.5$ g

5.3.2 Set the significance level of the test α and the sample size n.

Specify or compute σ

Example: $\alpha = 0.05$, $n = 25$, $\sigma = 2$

5.3.3 Determine a sampling distribution and the corresponding test statistic

Choose a sampling distribution and the corresponding test statistic to test H_0 with the appropriate assumptions.

Example: Assuming σ known, $\overline{X}$ is normally distributed with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$ or Z ε N (0,1). Also for the test statistic, we have $Z_C = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$

5.3.4 Set up a critical region for the test statistic

Set up a critical region for this test statistic where H_o will be rejected 100p percent of the samples when H_o: is true

Example: In our example where H_1: $\mu < 19.5$ g, the critical region would consist of all computed values of the test statistic (Z_C) less than the table or specified value ($-Z_\alpha$).

Thus, the decision would be to reject the null hypothesis H_0 if $Z_C < -Z_\alpha$.

Similarly, for H_1: $\mu > 19.5$ g, the critical region would consist of all computed values of the test statistic (Z_C) greater less than the table or specified value (Z_α). Thus, the decision would be to reject the null hypothesis H_0 if $Z_C > Z_\alpha$.

5.3.5 Perform the Experiment

Example: Choose a random sample of n observations, compute the test statistics and make a decision on H_o

5.3.6 Numerical Examples: Hypothesis Tests for the Mean, σ known or n >30

i) Hypothesis for: $\mu < \mu_0$

H_o: $\mu = 19.5$ g
H_1: $\mu < 19.5$ g

Reject H_0 if $Z_C < -Z_{table}$.,

Let: $\alpha = 0.05$, $n = 25$, $\sigma = 2$, $\overline{X} = 18.9$, $\mu = 19.5$. Also, $Z_\alpha = Z_{0.95} = 1.645$

With σ known the sampling distribution is the normal. The test statistic is the standardized Z, hence

$$Z_C = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \frac{18.9 - 19.5}{2 / \sqrt{25}} = -1.5$$

Critical Region: All values of the test statistic less than -1.5

Reject if $Z_C < -Z_{0.95}$

$Z_C = -1.5$, $-Z_{0.95} = -1.645$

But: $-1.5 > -1.645$

<u>Therefore, do not Reject H_0</u>. There is no evidence based on the data to suggest that the true mean of the population is not 19.0 grams.

ii) Hypothesis for: $\mu > \mu_0$

H_o: $\mu = 100$ g

H_1: $\mu > 100$ g

a). Let: $\alpha = 0.05$, $n = 9$, $\sigma = 10$, $\overline{X} = 106$. Also, $Z_\alpha = Z_{0.95} = 1.645$, $\beta = 0.1$ for $\mu_1 = 110$

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{106 - 100}{10/3} = 1.8$$

Critical Region: All values of $Z > Z_\alpha$, that is all $Z > 1.645$

But $1.8 > 1.645 \Rightarrow$ <u>Hence Reject H_0</u>

iii) Hypothesis for: $\mu \neq \mu_0$

H_o: $\mu = 100$ g, H_1 $\mu \neq 100$ g

Let: $\alpha = 0.01$, n =? , $\sigma = 16$, $\overline{X} = 106$. Also, $Z_{\alpha/2} = Z_{0.995} = 2.976$, $\beta = 0.20$ for $\mu_1 = 92$

Reject if $|Z| < Z_{\alpha/2}$, $n = \frac{\left(Z_{\sigma/2} + Z_\beta\right)^2}{\Delta^2}$, *where* $\Delta = \frac{|\mu_1 - \mu_0|}{\sigma}$

$$n = \frac{(2.576 + 0.84)^2}{(8/16)^2} = \frac{(2.576 + 0.84)^2 (16)^2}{64} = 46.67 = 47$$

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{106 - 100}{16/\sqrt{47}} = \frac{6\sqrt{47}}{16} = \frac{6(6.856)}{16} = 2.57$$

Since $|Z| < Z_{\alpha/2} \Rightarrow 2.57 < 2.976 \Rightarrow$ Reject H_0

5.3.6 Examples for Test of Hypothesis for the Mean, σ unknown and n <30

i) Hypothesis for: $\mu < \mu_0$

H_o: $\mu = 175$

H_1 $\mu < 175$

Let: $\alpha = 0.05$, $n = 6$, $s = 7.9$, $\overline{X} = 172.8$, $\mu_0 = 175$

Sampling distribution is the student-t (σ unknown and n <30)
t (0.05, 5) =2.015. **Reject if t<-2.015**

$$t = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{(172.8 - 175)\sqrt{6}}{7.6} = -0.68$$

Since -0.68>-2.015⇒ Do NOT REJECT H_0
Computation of n for this example:
α = 0.05, n = 6, assume σ =7.9, $\overline{X}$ =172.8, μ_0=175, μ_1=170 for β=0.2

$$n = \frac{(Z_\sigma + Z_\beta)^2}{\Delta^2},\ where\ \Delta = \frac{|\mu_1 - \mu_0|}{\sigma} = \frac{(1.645 + 0.84)^2}{\left(\frac{170 - 175}{7.9}\right)^2} = \frac{6.1745}{0.4006} = 15.41 \approx 16$$

5.4 Summary Tests for One Mean

5.4.1 Variance known

H_o: $\mu = \mu_o$;
H_1: $\mu < \mu_o$ Reject if: $Z < -Z_\alpha$
H_1: $\mu > \mu_o$ Reject if: $Z > Z_\alpha$
H_1: $\mu \neq \mu_o$ Reject if: $|Z| > Z_{\alpha/2}$

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

5.4.2 Variance unknown, but n > 30 (σ estimated from s)

H_o: $\mu = \mu_0$
H_1: $\mu < \mu_o$ Reject if: $Z < -Z_\alpha$
H_1: $\mu > \mu_o$ Reject if: $Z > Z_\alpha$
H_1: $\mu \neq \mu_o$ Reject if: $|Z| > Z_{\alpha/2}$

Test Statistic: $Z = \dfrac{\overline{X} - \mu}{s/\sqrt{n}}$

5.4.3 Variance unknown, n ≤ 30 (σ estimated from s)

H_o: $\mu = \mu_o$
H_1: $\mu < \mu_o$ Reject if: $t < -t_{\alpha,\nu}$
H_1: $\mu > \mu_o$ Reject if: $t > t_{\alpha,\nu}$
H_1: $\mu \neq \mu_o$ Reject if: $|t| > t_{\alpha/2,\nu}$
where df= ν = n-1

$$\boxed{Test\ Statistic: t = \frac{\overline{X} - \mu}{S/\sqrt{n}}}$$

Test on Means (More Than One Mean)

6.1 Variance known

a). For two Independent Samples, the difference between two means ($(\overline{X}_1 - \overline{X}_2)$The variance of the difference between two means for two independent samples from normal populations.

$$\sigma^2_{\overline{X}_1-\overline{X}_2} = \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}, and \;\; (\mu_1 - \mu_2) = 0 = \delta$$

H_o: $\mu_1 = \mu_2$; and the alternatives, namely:

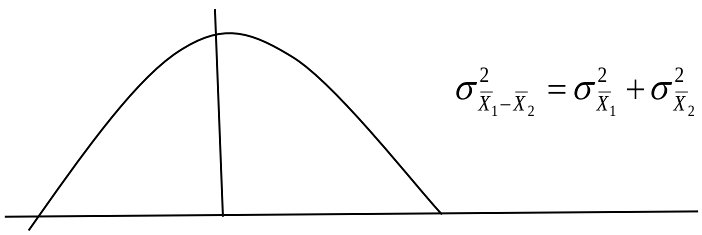
H_1: $\mu_1 < \mu_2$ Reject if: $Z < -Z_\alpha$,

H_1: $\mu_1 > \mu_2$ Reject if: $Z > Z_\alpha$

H_1: $\mu_1 \neq \mu_2$ Reject if: $|Z| > Z_{\alpha/2}$

The Test Statistic is given by Z as shown above

$$\boxed{Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}}$$



$$\sigma^2_{\overline{X}_1-\overline{X}_2} = \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

μ_1-μ_0

Figure 10: Differences between two variances

Example:

The manufacturing engineer has been tasked to determine the setup configuration for two contract broaching processes. The manufacturer of the broaching machine has historical data on the expected time to complete a broaching operation for each configuration. Assume that the population variances are also known and are as follows:

$$\overline{X}_1 = 45, n_1 = 25, \overline{X}_2 = 50, n_2 = 25, \sigma_1 = 10, \sigma_2 = 8, \alpha = 0.05$$

i) H_o: $\mu_1 = \mu_2$; , H_1: $\mu_1 \neq \mu_2$ Reject if: $|Z| > Z_{\alpha/2}$ Note: $Z_{\alpha/2} = Z_{0.975} = 1.96$

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{|50-45|}{\sqrt{\frac{100}{25} + \frac{64}{25}}} = \frac{5}{\sqrt{\frac{164}{25}}} = \frac{5(5)}{\frac{\sqrt{164}}{5}} = 1.952$$

Since Z (1.952) < $Z_{\alpha/2}$(1.96), Do not Reject. The two broaching configurations are essentially the same. However, because of the closeness of the critical (table) value to the computed value, additional analyses need to be carried out.

b) We use the Test Static above when we are sampling from normal populations. However, we can use a modified version when the population is not normally distributed, but the sample sizes are large enough (>30) in which case we can apply the central limit theorem(CLT) and approximate σ_1 and σ_2 with S_1 and S_2 respectively. That is

$$Z = \frac{(\overline{X}_1 - \overline{X}_{21}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \Rightarrow Z = \frac{(\overline{X}_1 - \overline{X}_{21}) - \delta}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Example

As way to strengthen its material properties, a company is considering annealing of a piece part and then measure the ductility. The project engineer claimed that annealing will increase ductility by 0.01in/in percent. After tensile testing the percent elongation as a measure of ductility for the annealed parts was $\overline{X}_1$= 0.211 in/in with standard deviation =0.0035 in/in and n=40. Values (percent elongation) obtained for the standard material without annealing was $\overline{X}_2$=0.187 in/in with standard deviation of 0.007 in/in and n=40. Set up the hypothesis and at α= 0.01, determine if the claim by the project engineer can be supported by the data.

1. H_0: μ_1-μ_2, H_1: μ_1-μ_2 >∂ (=0.01),
2. α= 0.01, n_1=n_2=40
3. Reject if Z> Z_α ($Z_{0.99}$=2.33)

$$Z = \frac{(\overline{X}_1 - \overline{X}_{21}) - \delta}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{0.211 - 0.187 - 0.01}{\sqrt{\frac{(0.0035)^2}{40} + \frac{(0.007)^2}{40}}} = 3.232$$

4. Decision: Since Z (3.232) is greater than the critical value of 2.33, then we must reject the null hypothesis. This means that the data supports the claim of the project engineer for the annealing process.

Example:

Let us consider a common example of mating or tolerance parts, where the focus is on the optimum fit that is, the optimum clearance. Our example is a shaft/bearing scenario where the clearance is zero for optimum fit. Please note that if the clearance is less than or equal to zero, then it would be difficult for the shaft to fit into the bearing. Given the following information about the mating parts (shaft and bearing), what is the probability that the shaft will not fit in the bearing?

Define $C = S - B, \mu_C = \mu_S - \mu_B, \sigma_C^2 = \sigma_S^2 + \sigma_B^2,$

> Note: if X=aY then Vax(X)=a^2Var(Y)
> This leads to:
> if D=X-Y, then Var(D)=Var(X)+Var(Y)

Let μ_B=0.732 inches, μ_S=0.698 inches, σ_B=0.004 inches, σ_S=0.016 inches

$$\sigma_C^2 = (0.004)^2 + (0.016)^2 = 2.72x10^{-4}, \mu_C = 0.034, \sigma_C = 0.0165$$

$$P\{(S-B)\le 0\} = P(C\le 0) = \frac{0-\mu_C}{\sigma_C}$$

$$Z = \frac{0-\mu_C}{\sigma_C} = \frac{0-0.034}{0.0165} = -2.06,\ \Phi(-2.06) = 1-\Phi(2.06) = 1-0.9803 = 0.0197 \approx 2\%$$

The probability that in the mating arrangement the shaft will not fit into the bearing is about 2%

6.2 Variance unknown but assumed equal and n<30

$$t = \frac{(\overline{X}_1 - \overline{X}_{21}) - 0}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

H_o: $\mu_1 = \mu_2$; or H_0: $\mu_1 - \mu_2 = 0$

H_1: $\mu_1 < \mu_2$ or H_1: $\mu_1 - \mu_2 < 0$ Reject if: $t < -t_{\alpha,\nu}$

H_1: $\mu_1 > \mu_2$ or H_1: $\mu_1 - \mu_2 > 0$ Reject if: $t > t_{\alpha,\nu}$

H_1: $\mu_1 \ne \mu_2$ or H_1: $\mu_1 - \mu_2 \ne 0$ Reject if: $|t| > t_{\alpha/2,\nu}$

Where $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$, $df = \upsilon = (n_1+n_2-2)$

For the previous problem assume that the variance is unknown and estimated from the data and that $\overline{X}_1 = 45, n_1 = 25, \overline{X}_2 = 50, n_2 = 25, S_1 = 12, S_2 = 9, \alpha = 0.05$, $\nu=48$, $t_{0.05,48}=1.68$

H_o: $\mu_1 = \mu_2$

H_1: $\mu_1 < \mu_2$ Reject if: $t < -t_{\alpha,\nu}$

$$S_p = \sqrt{112.5} = 10.6,\quad t = \frac{(45-50)-0}{10.6\sqrt{\frac{1}{25}+\frac{1}{25}}} = -\frac{5\sqrt{25}}{10.6\sqrt{2}} = -\frac{25}{15} = -1.67$$

Reject if t<-t. But -1.67>-1.68, hence Do NOT Reject. However, the values are close enough to warrant further investigation and analyses.

6.3 Variance unknown and unequal ($\sigma_1 \ne \sigma_2$)

H_o: $\mu_1 = \mu_2$; or H_0: $\mu_1 - \mu_2 = 0$

H_1: $\mu_1 < \mu_2$; or H_1: $\mu_1 - \mu_2 < 0$ Reject if: $t < -t_{\alpha,\nu}$

H_1: $\mu_1 > \mu_2$, or H_1: $\mu_1 - \mu_2 > 0$ Reject if: $t > t_{\alpha,\nu}$

H_1: $\mu_1 \ne \mu_2$, or H_1: $\mu_1 - \mu_2 \ne 0$ Reject if: $|t| > t_{\alpha/2,\nu}$

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - 0}{\sqrt{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)}}$$

$$where: \nu = \left\{ \frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}} \right\}$$

	Plant I(minutes)	Plant II(minutes)
	102	81
	86	165
	98	97
	109	134
	92	92
		87
		114
Mean	97.4	110
Variance	78.8	913.33
n	5	7
Table 4: Machining times for machines I and II		

The above data is from two different plants. The difference in the machining time of two identical operations at the two different plants of a multinational company is of concern to the Director of Engineering Services. It is believed that a difference of more than 10 minutes would cause a problem about cycle time which would require a major change in the system design. Determine what should be done based on a test of hypotheses at α=0.1.

i). H_0: μ_{II} -μ_I=10, H_1: μ_{II}-μ_I>10, Reject if $t > t_\alpha$

$$t = \frac{(\bar{X}_{II} - \bar{X}_I) - 10}{\sqrt{\frac{s_1^2}{5} + \frac{s_2^2}{7}}} = \frac{(110 - 97.4) - 10}{\sqrt{\frac{78.8}{5} + \frac{913.33}{7}}} = \frac{2.6}{12.1} = 0.21$$

$$\upsilon = \frac{((78.8/5) + (913.33/7))^2}{\frac{(15.76)^2}{4} + \frac{(130.48)^2}{6}} = 7.4 = 7$$

$t_{0.1,7}$=1.415Since t (0.21) < t_α, (1.47), therefore cannot reject H_0. Hence it is reasonable to suggest that the difference between the two machines is statistically not more than 10 minutes.

6.4 Paired Tests

In some situation, the samples for μ_1, μ_2 are not independent which is an assumption we have made or implied in most of the foregoing tests. In some applications, paired data are encountered. For example, while matching a cylindrical disk, it may be necessary to take measurements at two different reference points. In such circumstances, the difference between the measurements rather than the actual measurements becomes important. The difference test is sometimes referred to as the dependency test. The random variable of interest is the difference, d_j where: $d_j = X_{1j} - X_{2j}$, j = 1,2,...n

H_0: $\mu_d = 0$, H_1: $\mu_d \neq 0$

$$\bar{d} = \frac{\sum d_j}{n}, \quad s_d^2 = \frac{\sum (d_i - \bar{d})^2}{n-1} = \frac{\sum d_j^2 - \frac{(\sum d_j)^2}{n}}{n-1}$$

The Test Statistic is given by: $t = \dfrac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$, $with\ \nu = (n-1) = df$

Example: A group of 10 engineering students were pre-tested before instruction and post-tested after 6 weeks of instruction with the following results as shown in table 5.

$$H_0 : \mu_d = 0,\ H_1 : \mu_d > 0$$

$$\alpha = 0.05, \nu = 9(= n-1), t_{9,\,0.95} = 1.833, \quad \bar{d} = 2.5, s_d = \sqrt{\frac{\sum (d - \bar{d})^2}{n-1}} = 2.2236$$

$$t = \frac{\bar{d}}{s_d / \sqrt{n}} = \frac{2.5}{2.2236/\sqrt{10}} = 3.56, Since\ \ t > t_{9,\,0.95}\ \ , then\ \ will\ \ reject\ \ H_0$$

Conclusion: Based on the test results, there is not enough statistical evidence to suggest an improvement due to the intervention.

Student	Before	After Instruction	Difference 'd'
1	14	17	3
1	12	16	4
3	20	21	1
4	8	10	2
5	11	10	-1
6	15	14	-1
7	17	20	3
8	18	22	4
9	9	14	5
10	7	12	5

Table 5: Result of Post- Test Pre-Test

Test of Variance

7.1 Variance from One Population

$$\text{Test Statistic} = \chi^2 = \frac{(n-1)S^2}{\sigma_o^2}$$

Let: $H_o: \sigma^2 = \sigma_o^2$

H_1: $\sigma^2 < \sigma_o^2$, reject if: $\chi^2 < \chi^2_{1-\alpha,\, n-1}$, $\sigma^2 > \sigma_o^2$, reject if: $\chi^2 > \chi^2_{\alpha, n-1}$

$\sigma^2 \neq \sigma_o^2$, reject if: $\chi^2 > \chi^2_{\alpha/2, n-1}$ or $\chi^2 < \chi^2_{(1-\alpha/2),\, n-1}$

The sampling distribution for the variance of a population is the chi-square, where s^2 is computed from a random sample of n observations and σ_o^2 is the given or specified value. Note that the Chi-square unlike the normal is not symmetrical. Also for a specific α, and same degrees of freedom given by ν = (n-1):

$$\chi^2_{\alpha,\,\upsilon} > \chi^2_{1-\alpha,\,\upsilon}$$

Example: The population variance from a machining operation was given by the lathe manufacturer as σ^2 =30. A sample from the current machining operation was taken with the following values N=25, S =4.92. α=0.05

Test the hypothesis: H_0: σ^2=30 against H_1:σ^2 < 30

Reject if $\chi^2 < \chi^2_{1-\alpha,\,n-1}$

$$\chi^2 = \frac{(n-1)S^2}{\sigma_o^2} = \frac{24S^2}{30} = 0.8(4.9)^2 = 19.208$$

$$\chi^2_{0.95,\,24} = 13.848, \Rightarrow since\ \chi^2(19.208) > \chi^2_{0.95,241}(13.848)$$

Do not reject H_0. There is not enough evidence to believe that σ^2 is not statistically equal to 30.

7.2 Variance from Two Populations

$$\boxed{F_{1-\alpha,n_1-1,n_2-1} = \frac{1}{F_{\alpha,n_2-1,n_1-1}}}$$

H_o: $\sigma_1^2 = \sigma_2^2$, H_1: $\sigma_1^2 < \sigma_2^2$ Reject if: $F < F_{1-\alpha,n_1-1,n_2-1}$

$\sigma_1^2 > \sigma_2^2$ Reject if: $F > F_{\alpha,n_1-1,n_2-1}$

$\sigma_1^2 \neq \sigma_2^2$ Reject if: $F > F_{\alpha/2,n_1-1,\,n_2-1}$ *or* $F < F_{1-\alpha/2,\,n_1-1,n_2-1}$

Due to the difficulty of accessing some of the data from the F-table, we recast the Test statistic and critical region as follows:

$\sigma_1^2 < \sigma_2^2$, $F = \frac{S_2^2}{S_1^2}$ Reject if: $F < \frac{1}{F_{\alpha,n_2-1,n_1-1}}$

$\sigma_1^2 > \sigma_2^2$, $F = \frac{S_2^2}{S_1^2}$ Reject if: $F > F_{\alpha,n_1-1,n_2-1}$

$$\boxed{where: S_M^2 > S_m^2}$$

$\sigma_1^2 \neq \sigma_2^2$, $F = \frac{S_M^2}{S_m^2}$ Reject if: $F > F_{\alpha/2,n_M-1,\,n_m-1}$

The test statistic is the F-Distribution = s_1^2/s_2^2 Where $s_1^2 > s_2^2$

Example: We will use the Spur-gear example to demonstrate the F-test for two variances. Suppose that in the Diametral Pitch example, the Spur-gears were supplied by two suppliers/clients, with the following data, Supplier 1: $n_1=21$, $S_1=0.56$ inches, Supplier 2, $n_2=16$, $S_2=1.8$ inches. Perform a test hypothesis that says that the variance of supplier 2 is greater than that of supplier 1 at $\alpha=0.01$.

H_0: $\sigma_1^2=\sigma_2^2$, H_1: $\sigma_2^2>\sigma_1^2$, Reject if $F > F_{\alpha, v2, v1}$

$$F=\frac{S_2^2}{S_1^2},\ with\ \ \upsilon_2=15\ \upsilon_1=20$$

$$\boxed{F=\frac{(1.8)^2}{(0.56)^2}=10.331}$$

$F_{0.01,15,20}=2.20$, Since F (10.331)> $F_{0.01,15,20}$ (2.20), **Reject the null hypotheses**

Example: Suppose we are interested in testing for H_0: $\sigma_1^2=\sigma_2^2$, H_1: $\sigma_2^2<\sigma_1^2$

Let $n_1=21$, $n_2=13$, $S_1=4.2$, $S_2=2.7$, $\alpha=0.01$, Reject if $F<\dfrac{1}{F_{\alpha,n_2-1,n_1-1}}$

$$F_{table}=\frac{1}{F_{\alpha,n_2-1,n_1-1}}=\frac{1}{F_{0.01,\,12,20}}=\frac{1}{2.28}=0.4386$$

$F=\dfrac{(4.2)^2}{(2.7)^2}=2.42$, Hence, we cannot Reject H_0. No reason to believe that the variances are NOT equal based on the data we have.

Example: Testing for equality of the variances

Two types of production processes are under consideration. One is based machine enabled used a built-in mechanism, the other is robot enables. The production time in minutes is as follows: $n_1=10$, $s_1=20.5$, $n_2=8$, $S_2=15.3$, $\alpha=0.1$. Determine if the variances are equal.

H_0: $\sigma_1^2=\ \ \sigma_2^2$, H_1: $\sigma_2^2\neq\sigma_1^2$

Reject if: $F>F_{\alpha/2,n_M-1,\ n_m-1}$, $F_{0.05,9,7}=3.68$, ($F_{0.05,\ 10-1,\ 8-1}$)

$F=\dfrac{(20.5)^2}{(15.3)^2}=1.795$, Since $F<F_{table}$, Cannot Reject H_0

7.3 Why the F test for Two Variances

Let us examine why these hypothesis tests (especially the one about the equality of variances) are important. Recall that when we were carrying out the test of two means, we were concerned about the equality of variances, so we had to assume that the variances are either equal or not equal especially when we are not operating under the normal distribution. As a matter of fact, we used a

certain test statistic if the variances were assumed equal and another if we could not assume they are equal. What does all these mean now that we have a way to test for the equality of the variances. The implication of this going forward is that now that we have a way to determine whether the variances are equal or identical, we can no longer simply assume away the possibility. As engineers, we must operate from the standpoint point of knowledge and information.
So, this let us see how this would work in a more practical way. Recall that in our test for two means, we had two special cases, namely, variance unknown & equal, and variance unknown and unequal.

For the case of Variance unknown but assumed equal we have the following test statistic:

$$t = \frac{(\overline{X}_1 - \overline{X}_{21}) - 0}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \; or \; Z = \frac{(\overline{X}_1 - \overline{X}_{21}) - 0}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$, depending on the same size.

For the case where the variance is unknown and unequal, we have

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - 0}{\sqrt{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)}}$$ with degrees of freedom equal to:

$$where: v = \left\{ \frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}} \right\}$$

What we need to do now is rather than assume, we will first determine whether the variances are equal using the test of hypothesis and then based on that we make a decision on which of the test statistic to use. Let us assume that we have been given the times to perform a finishing operation by two different processes (X_1, and X_2).

$\overline{X}_1 = 68, \quad \overline{X}_2 = 25$

$S_1 = 28.5 \;\; S_2 = 12.3, n_1 = 12, \quad n_2 = 12$

Test: H_o: $\mu_1 = \mu_2$; H_1: $\mu_1 > \mu_2$, at α=0.01

Procedure:

1. Test for the equality of the variances, then
2. Test for the means based on the required hypothesis which in this case is: (H_1: $\mu_1 > \mu$)

Test for the variances

a.) H_0: $\sigma_1^2 = \sigma_2^2$; μ_2: H_1: $\sigma_1^2 \neq \sigma_2^2$

b). $n_1 = n_2 = 12$, α=0.01

c). Test Statistics: $F=(S_1^2/S_2^2)$, d). Reject if $F > F_{\alpha/2, n_M - 1, \; n_m - 1}$

Note: n_M= Sample size for the larger variance, while n_m is the sample size for the smaller variance

$$F = \frac{(28.5)^2}{(12.3)^2} = 5.369,\ F_{\alpha/2,12,\ 12} = F_{0.05,\ 12,12} = 2.69$$

Therefore, we Reject the null hypothesis of the equality of the variances

Now we test for the means given that we know (rather than assume the nature of the variances)

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

2. Test for the means: H_o: $\mu_1 = \mu_2$; H_1: $\mu_1 > \mu_2$, at α=0.01

The degrees of freedom df. is given by:

$$where: \nu = \left\{\frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}\right\}$$

ν=15, t (0.1,15) =2.602

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}} = \frac{68 - 25}{\sqrt{\frac{28.5}{12} + \frac{12.3}{12}}} = \frac{43}{8.961} = 4.799$$

Reject if: $t > t_{\alpha,\ 15}$

t= (4.799) > $t_{\alpha,\ 15}$(2.602), ∴ Reject H_0

Our decision then is to: **Reject the null hypotheses that the finishing operating times by the two different processes are statistically the same.**

Summary

The result of a statistical inference is always a decision to act or not to act. In some instance, the decision could be to accept, in place of the unknown parameter, the observed or computed value of the estimator without requiring that it be exactly the true value. On the other hand, we may decide to reject or not reject the assumptions about certain distribution without conceding that such a statement is true beyond doubt. The use of statistical inference enables us to control the possible errors that could arise because of our decisions and to ensure that these errors, while inevitable, are as small as economically possible.

Of interest is the determination of the type II error. We have provided several ways we can do that. Tables are available especially for the case of two means. However, those were not included because it requires the use of tables which unfortunately we could not include here but they are available in most basic probability and statistics books.

Finally, we have the case of testing two means and the assumptions of the nature of the variances. Because of the nature of the work we do as engineers we should never assume away anything especially if the data is available for us to determine the veracity of the information. Thus, in the case of the test of two means, if the we cannot assume that the process is normal then we can use the central limit theorem as a guide. If the variances are unknown, we must test to see whether they are equal or unequal in which case we can use the appropriate test statistic.

The materials present here are hardly exhaustive especially about hypothesis testing. For example, we did not look at test for proportions and the like. The idea is to provide a pathway that would lead to how we look at the problems of these type and then we can use the knowledge gained to extrapolate into related areas. For example, the test of hypothesis for proportions follows the same method, in terms of the null and alternative hypothesis and the reject criteria. The only deference is the Test Statistic which is determined by the underlying distribution of the random variable of interest.

REFERENCES

1 Johnson, R., Miller, I. (2015), Miller & Freund's Probability and Statistics, 9th Edition, Pearson Publishers, Boston, MA, USA

2. Vardeman, S. B. (1994), Statistics for engineering **problem solving, 1st ed.,** PWS Pub. Co., Boston MA

3. Montgomery, D.C., Runger, G., and Hubele, N. (1998), Engineering Statistics, 1st ed John Wiley and Sons, NY, USA.

4. Montgomery, D.C., and Runger, G. (2011), Applied Statistics and Probability for Engineers, 5th ed., John Wiley and Sons, NY.

5. Devore, J.L., (2012), Probability and Statistics for Engineering and the Sciences, 8th ed., Duxbury & Brooks/Cole, Boston, MA.

6. Ross, S. (1987), Introduction to Probability and Statistics for Engineers and Scientists, 1st ed., John Wiley & Sons, NY.

7. Walpole, R., Meyers, R.H. (2002), Probability and Statistics for Engineers and Scientists, 7th ed., Prentice- Hall Inc., Upper Saddle River, NJ.

8. Hogg, R. Craig, A. (2004), Introduction to Mathematical Statistics, 6th ed., Prentice-Hall, Englewood Cliffs, NJ.

9. Larsen R.J., Marx M.L. (2012), Introduction to Mathematical Statistics and Its Applications, 5th ed., Prentice-Hall Englewood Cliffs, NJ.

10. DeGroot M.H., Schervish M.J. (2011), Probability and Statistics, 4th ed., Addison Wesley, Boston, MA

11. "Data to Action," A Harvard Business Review (HBR) Insight Center White Paper, Sponsored by SAS, Inc., 2014, Harvard Business Publishing, Cambridge, MA

12. "Big Data Analytics, what it is and why it matters," International Institute for Analytics, Thomas H. Davenport and Jill Dyché, Copyright © Thomas H. Davenport and SAS, Institute Inc, 2013

13. International Federation of Robotics(IFR), Executive Summary, World Robotics 2013-2015.

14. Nuclear Power Plants World-wide, Source European Nuclear Society, November 2016.

15. Kelly Brown, 'A Terabyte of Storage Space: How Much is Too Much,' in The Information Umbrella, Musings on Applied Information Management, University of Oregon, 2014

.

