

Reference: Chapter 7 of Devore

CONFIDENCE INTERVAL ESTIMATORS

Maghsoodloo

An interval estimator of a parameter θ is of the form $\theta_L \leq \theta \leq \theta_u$ at a confidence Pr (or a confidence coefficient) of $1 - \alpha$ before a random sample of size n is drawn from the target population. When $\theta_L = -\infty$, $-\infty < \theta \leq \theta_u$ is called a one-sided upper confidence interval, and when $\theta_u = +\infty$, $\theta_L \leq \theta < \infty$ is called a lower one-sided confidence interval. The question arises as to when we should develop a 2-sided CI and when a one-sided CI is appropriate. To answer this question, we need to determine whether the parameter θ is of the Smaller-the-Better (STB), Nominal-the-Best (NTB), or Larger-the-Better (LTB) type.

Generally, there are 3 types of continuous rvs: STB, NTB and LTB. Examples of STB quality characteristics (QCHs) are "rate of wear", failure rate, noise level of a compressor, tire imbalance, braking distance, and queuing time. All STB QCHs have two common properties. First their ideal target value is zero, and secondly they all have a single upper consumer specification limit $USL = x_u$. Note that if the QCH, X , is being manufactured (unlike queuing time), then the ideal target of $m = 0$ can never be reached!

Examples of LTB QCHs are tensile or welding strength, reliability (or TTF), fuel efficiency, tape adhesiveness, and percent yield. The ideal target for any LTB rv is $+\infty$ and there is only a single lower consumer specification limit $LSL = x_L$.

Finally, all nominal dimensions have both a lower consumer specification limit $LSL = m - \Delta$ and an upper specification limit $USL = m + \Delta$, where m is the ideal target value and is generally different from zero. Examples are "clearance", output voltage of a TV set, magnetic tape edge weave, and chemical content (such as pH) level.

In general but not always, this author recommends a 2-sided CI for all NTB type parameters, an upper one-sided CI: $0 \leq \theta < \theta_u$ for nearly all STB parameters, and a lower one-sided confidence interval $\theta_L \leq \theta < \infty$ if θ is of an LTB type parameter. Further, the shorter the length $\theta_u - \theta_L$ of a 2-sided CI is, the more precise estimator of the parameter θ the CI is.

In general, to develop a parametric CI for a parameter θ , the sampling distribution (SMD) of its point estimator, $\hat{\theta}$, must be known and then used appropriately. If the SMD of $\hat{\theta}$ is unknown and $n < 30$, a nonparametric procedure must be used to obtain a CI for θ . If n is large (say $n > 50$), then the CLT (central limit theorem) can be applied in order to assume that the SMD of $\hat{\theta}$ should be approximately Gaussian. Further, to obtain a one-sided CI, the value of α should generally (but not always) be placed at the same tail of the SMD that corresponds to rejecting of the null hypothesis. This concept will be clarified later.

CONFIDENCE INTERVAL FOR THE MEAN OF A NORMAL PROCESS WITH KNOWN VARIANCE σ^2

Example 36. A normal process manufactures ropes for climbing purposes with breaking strength $LSL = x_L = 1200$ psi and variance $\sigma^2 = 625$ psi². A random

sample of size $n = 25$ ropes gave $\sum_{i=1}^{25} X_i = 31,500$ psi. Our objective is to obtain the

point and interval estimators ($1 - \alpha = 0.95$) for the process mean μ , where $1 - \alpha = 95\%$ is called the confidence coefficient (or the confidence level).

The point estimator of the process mean breaking strength is

$$\hat{\mu} = \bar{x} = \frac{1}{25} \sum_{i=1}^{25} x_i = 1260 \text{ psi} .$$

Since μ is an LTB type parameter, then we develop a 95% lower CI for μ because there are no concerns on the high side. Recall that we have to use the sampling distribution (SMD) of the point estimator, \bar{x} , in order to develop the required CI. Further, recall that \bar{x} is $N(\mu, \sigma^2/n)$ as depicted in Figure 9. Figure 9 clearly shows that the $\Pr(-\infty < \bar{x} \leq \mu + Z_{0.05} \times 5) = 0.95 \rightarrow$

$\Pr(-\infty < -\mu \leq -\bar{x} + 1.645 \times 5) = 0.95 \rightarrow \Pr(\infty > \mu \geq \bar{x} - 1.645 \times 5) = 0.95 \rightarrow$

$$\Pr(\bar{x} - 8.225 \leq \mu < \infty) = 0.95 \tag{33}$$

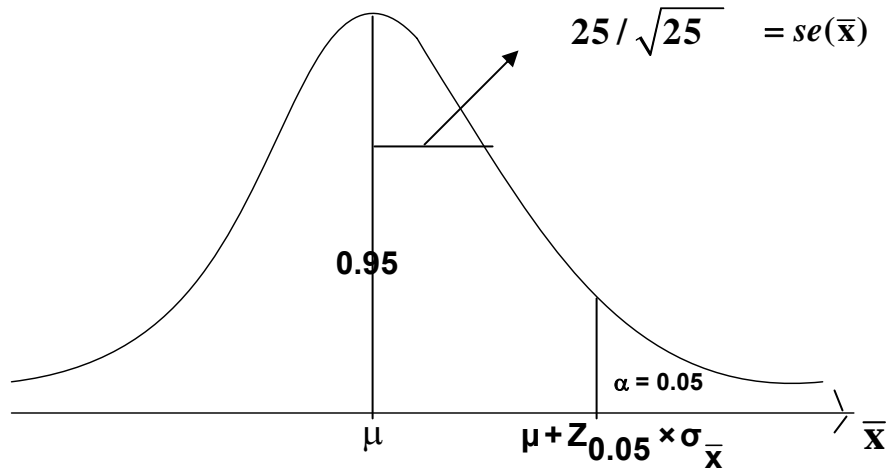


Figure 9. The Sampling Distribution (SMD) of \bar{x}

It is clear from Eq. (33) that $\mu_L = \bar{x} - 8.225$ and hence the 95% lower CI for μ , using $\bar{x} = 1260$, is $1251.7750 \leq \mu < \infty$. The 95% realized CI $[1251.775, \infty)$, which is not a statement about a rv, does not have a Pr of 0.95 to contain the true value of μ because that Pr is either 0 or 1. However, we do have 95% confidence that μ lies within the number interval $[1251.7750, \infty)$. This implies that if we repeat the same above experiment 100,000 times and obtain 100,000 distinct CIs such as $[1251.7750, \infty)$, then roughly 95,000 of such number CIs of differing μ_L will contain the true value of μ .

Note that the Pr for the confidence (number) interval, $[1251.7750, \infty)$, is not 0.95 to contain the true value of μ but rather 0 or 1 because the interval is not a statement about a rv. However, the interval $[\bar{x} - 8.225, \infty)$ does have a 95% Pr of containing the true value of μ before the sample is drawn because prior to the drawing of the n sample values, \bar{x} is a rv with SMD depicted in Figure 9 above.

Exercise 55. (a) Obtain the 99% CI for the population mean of the Example 36 above. (b) Work the appropriate parts of Exercises 7.1, 7.4, 7.5, and 7.6

on pages 290- 291 (286-287 of the 5th edition) of your text.

CONFIDENCE INTERVAL FOR THE MEAN OF A NORMAL POPULATION WITH UNKNOWN σ^2

When the process variance σ^2 is unknown, we make use of its unbiased

estimator $S^2 = (USS - CF)/(n - 1) = [\sum x_i^2 - (\sum x_i)^2/n]/(n - 1) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} =$

$CSS/(n - 1)$ in order to develop a CI for μ . We resort to the fact that the sampling distribution of the statistic $(\bar{x} - \mu)\sqrt{n} / S$ is that of the Student's t with $(n - 1)$ df as depicted in Figure 10.

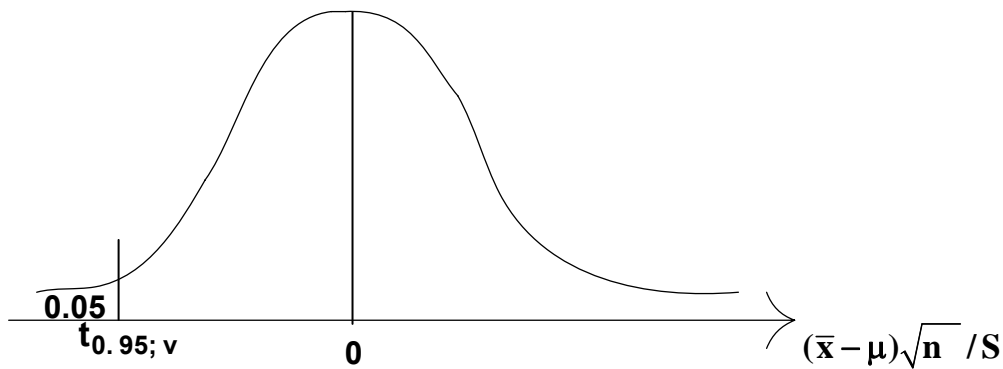


Figure 10. The Sampling Distribution of $(\bar{x} - \mu)\sqrt{n} / S \sim T_{n-1}$

For the sake of illustration, we develop an upper 1-sided 95% CI for the mean loudness of compressors from a certain manufacturer. Suppose a random sample of size $n = 13$ compressors were tested for their loudness with the sampling results $\bar{X} = 59.7$ decibels and $S^2 = 17.64$ dB². Since the rv $(\bar{x} - \mu)\sqrt{n} / S$ has to be

multiplied by (-1) in order to solve for μ , we have to place the entire $\alpha = 0.05$ on the lower tail of the t-distribution as shown in Figure 10 in order to obtain an upper one-sided CI for μ . Table A.5 on page 743 and Figure 10 clearly show that the

$P[(\bar{x} - \mu)\sqrt{n} / S \geq -1.782] = 0.95$ because the $P(T_{12} > -1.782) = 0.95$. Thus, the

$$P[\bar{x} - \mu \geq -1.782S/\sqrt{n}] = 0.95 \rightarrow P(\mu \leq \bar{x} + 1.782 S/\sqrt{n}) = 0.95 \rightarrow$$

$$P(0 < \mu \leq \bar{x} + 1.782S/\sqrt{n}) = 0.95.$$

This last inequality clearly indicates that the 95% upper one-sided CI for the population mean loudness is

$$0 < \mu \leq \bar{x} + 1.782S/\sqrt{n}.$$

Inserting the sample results $\bar{x} = 59.7$ and $S = 4.20$ into this last expression, we then have 95% confidence that the process mean lies in the interval $0 < \mu \leq 61.776$ dB, where μ is an STB type parameter. Note that with this interval the null hypothesis $H_0 : \mu = 61$ cannot be rejected while $H_0 : \mu = 62$ versus $H_1 : \mu < 62$ can be rejected at the 5% level.

Exercise 56. Work Exercises 32, 33, 34, 37 and 38 on pages 306-307 (302-303 of the 5th edition) of Devore.

CONFIDENCE INTERVAL FOR THE VARIANCE OF A NORMAL UNIVERSE

Since variance is a component of quality that must always be minimized, the appropriate CI for σ^2 is usually of the type $0 < \sigma^2 \leq \sigma_u^2$. However, 2-sided CIs for the process variance σ^2 are not uncommon.

To illustrate the procedure, consider the data of Exercise 45 on page 311 of your Devore. Assuming that the underlying distribution of the $X =$ Fracture

toughness is $N(\mu, \sigma^2)$, the objective is to obtain the upper 95% one-sided CI for σ^2 (i.e., $1 - \alpha = 0.95$).

We use the fact that the rv $(n - 1)S^2/\sigma^2$ has a χ^2 sampling distribution with $v = n - 1 = 21$ df. Since the expression $(n - 1)S^2/\sigma^2$ has to be inverted in order to solve for σ^2 , the entire $\alpha = 0.05$ would have to be placed in the lower tail of χ^2_{21} distribution as depicted in Figure 11. Table A.7, on page 745 of Devore, gives

$$\chi^2_{0.95; 21} = 11.5913.$$

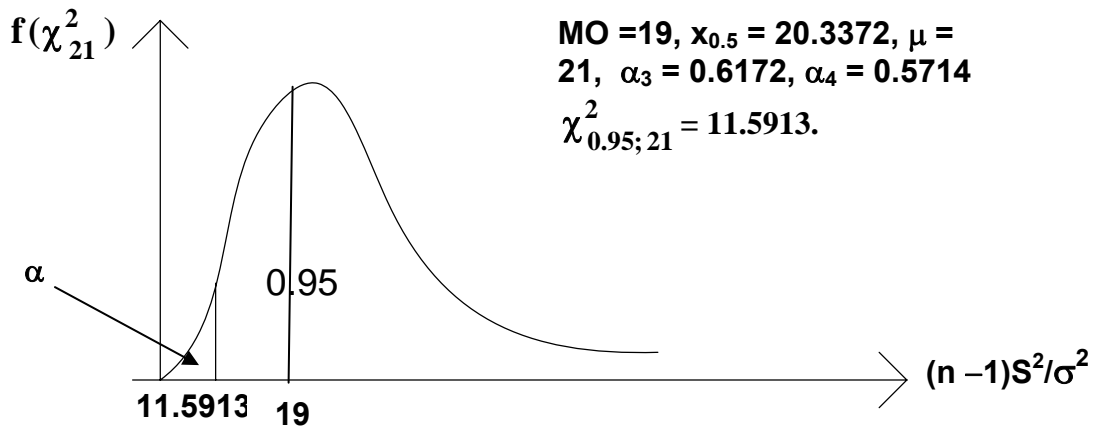


Figure 11. The SMD of $(n - 1)S^2 / \sigma^2 \sim \chi^2_{21}$

Figure 11 clearly shows that (note that $USS = 132097.35$ and $\sum_{i=1}^{22} x_i = 1701.30$)

$$\Pr[11.5913 \leq \frac{(n-1)S^2}{\sigma^2} < \infty] = 0.95 \rightarrow \Pr[\frac{1}{11.5913} \geq \frac{\sigma^2}{(n-1)S^2} > \frac{1}{\infty}] = 0.95$$

$$\Pr[\frac{(n-1)S^2}{11.5913} \geq \sigma^2 > 0] = 0.95 \rightarrow \Pr[0 < \sigma^2 \leq (n-1)S^2 / 11.5913] = 0.95 \rightarrow$$

$$\Pr[0 < \sigma^2 \leq \sigma_u^2] = 0.95.$$

This last confidence Pr statement indicates that $\sigma_u^2 = (n - 1) S^2 / 11.5913 =$

$CSS/\chi^2_{0.95; 21} = 21(25.3680)/11.5913$. Therefore, we are 95% confident that $0 < \sigma^2 \leq$

45.9593. Again, the 95% confidence interval $(0, 45.9593]$ does not have a 0.95 Pr of containing the true value of σ^2 because that Pr is either 0 or 1. However, in the long-run (i.e., in repeated sampling), roughly 95% of such number intervals will contain the true value of σ^2 . The upper 95% CI for σ is given by $0 < \sigma \leq \sqrt{45.959253} = 6.77933$, i.e., $0 < \sigma \leq 6.77933$ with 95% confidence.

Exercise 59. (a) Obtain the 2-sided 99% CI for the above Exercise as required by the problem statement. (b) Work Exercises 44 and 46 on page 311 of Devore.

Note that statistical inferences (i.e., CIs and Tests of Hypotheses) on two parameters are all covered in Chapter 9 of Devore while those on a single parameter are all covered in Chapters 7 and 8 of Devore.

LARGE-SAMPLE CONFIDENCE INTERVAL FOR A POPULATION PROPORTION

Let p denote the proportion of a population that possesses a certain characteristic, such as O-negative blood type. First, you must be cognizant of the fact that the more rare the characteristic is in the population, the larger the sample size n must be in order to obtain an accurate estimate of p . In fact, it can be shown that the required sample size for the estimation of p within certain specified error is inversely proportional to p . To illustrate the procedure for obtaining a CI for p , I will work a problem out of Chapter 7 of Devore and simultaneously describe the relevant procedure.

As an example, consider the Exercise 21 on page 298 of Devore. Here, the size of the sample is $n = 539$ (which is considered very large). First, one must consider the 539 households in the sample as 539 independent Bernoulli trials, and success occurs when a household has at least one Firearm in its possession. Recall that the Bernoulli rv, X_i , consisting of one randomly selected household, takes on the value of 1 if success occurs and the value of $X_i = 0$ ($i = 1, 2, \dots, 539$) when failure (= no firearm in the house) occurs. That is, $X_i = 1$ only if the i^{th} household in the sample has at least one Firearm. The result of our sample of $n = 539$ households can be

written as

$$X = X_1 + X_2 + \dots + X_n = \sum_{i=1}^{539} X_i$$

where each X_i has a Bernoulli pmf (with range space $R_i = 0, 1$) and unknown parameter p and as a result the rv, X , has a Bin($n = 539, p$) pmf [Bin stands for the binomial]. Further, recall that $E(X_i) = p$ and $V(X_i) = pq$, where p is the Pr that a randomly selected household has at least one Firearm. Therefore, the 1st two moments of the Bin rv, X , are $E(X) = np$ and $V(X) = npq$. Since X represents the sum of n independent and identical rvs, due to the CLT the sampling distribution (SMD) of X approaches a normal curve with mean np and standard deviation \sqrt{npq} . Because the sample proportion is given by $\hat{p} = X/n$, the SMD of \hat{p} also approaches normality with mean $E(\hat{p}) = E(X/n) = E(X)/n = np/n = p$ and $V(\hat{p}) = V(X/n) = V(X)/n^2 = npq/n^2 = pq/n$. Clearly, the sample proportion X/n is a point unbiased estimate of p , i.e., $\hat{p} = 133/539 = 0.24675$ is a point estimate of the proportion, p , of households in the entire city that are gun-owners. Our objective is to make use of the CLT to obtain a 2-sided 95% CI for the unknown parameter p . The approximate sampling distribution of \hat{p} is depicted Figure 12.

Figure 12 clearly shows that the $P[p - 1.96\sqrt{pq/n} \leq \hat{p} = X/n \leq p + 1.96\sqrt{pq/n}] = 0.95$. This 2-sided inequality implies that $P[p - 1.96\sqrt{pq/n} \leq \hat{p}]$ and simultaneously $[\hat{p} \leq p + 1.96\sqrt{pq/n}]$, or $P[-\hat{p} - 1.96\sqrt{pq/n} \leq -p]$ and at the same time $[-p \leq -\hat{p} + 1.96\sqrt{pq/n}]$. If we put these last two joint inequalities together, we obtain $P[-\hat{p} - 1.96\sqrt{pq/n} \leq -p \leq -\hat{p} + 1.96\sqrt{pq/n}]$. Therefore, $P[\hat{p} + 1.96\sqrt{pq/n} \geq p \geq \hat{p} - 1.96\sqrt{pq/n}]$. As a result, the lower large-sample

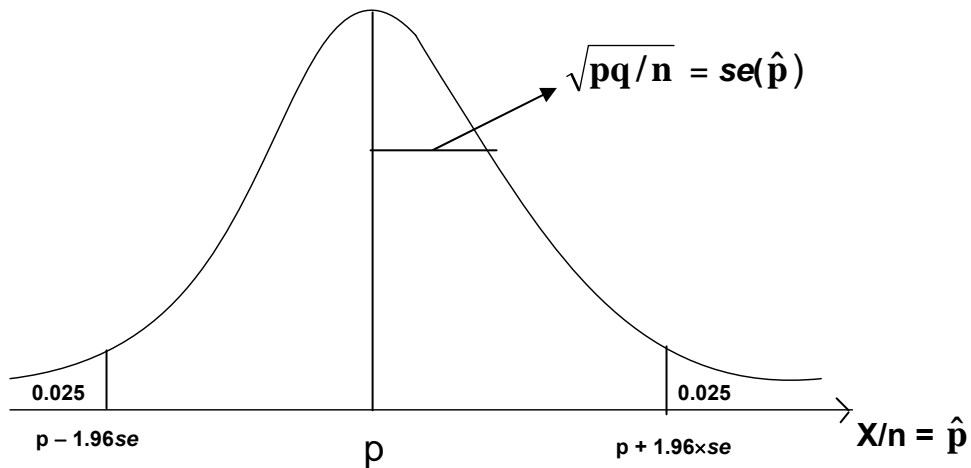


Figure 12. The Approximate SMD of $\hat{p} = X/n$

confidence limit for p is $p_L = \hat{p} - 1.96\sqrt{pq/n}$ and similarly $p_u = \hat{p} + 1.96\sqrt{pq/n}$.

Unfortunately, the above confidence limits depend on the unknown population parameter p itself and are therefore impossible to compute. The only recourse we have is to replace the value of p by its point estimate $\hat{p} = X/n = 0.24675$. Note that in

so doing the estimate of the $se(\hat{p}) = \sqrt{pq/n}$ becomes $\widehat{se}(\hat{p}) = \sqrt{\hat{p}\hat{q}/n}$

$= \sqrt{0.24675 \times 0.75325 / 539} = 0.01857$. Thus, the 95% half-CI length (HCIL) is given by $1.96 \times se(\hat{p}) = 0.0364$, and thus, $p_L = 0.24675 - 0.0364 = 0.21036$, and $p_u = 0.28315$.

As a result, we are 95% confident (again the Pr is not 95%) that the value of population proportion p lies in the interval $0.21036 \leq p \leq 0.28315$. Note that Devore gives a complicated more exact expressions for p_L and p_u in his equation (7.10) on page 295, but I would refrain from its use because it does not improve the values of p_L and p_u unless $n \ll 50$. I would recommend that you use the procedure that I have outlined above to obtain the requisite CI for p .

Exercise 60. Work exercises 47 and 49 on page 311 of Devore.

THE LAST TOPIC FOR STAT 3600 (TOLERANCE INTERVALS FOR a NORMAL UNIVERSE)

For starters, you should study pp. 305-306 of Devore on this topic. Below, I will also explain what we mean by a $(\gamma, 1 - \alpha)$ tolerance interval.

NATURAL TOLERANCES

Natural tolerance limits are the actual capabilities of a manufacturing process and can be considered as the dimensions within which all but a tolerable fraction, α , of the items manufactured will fall. For example, if $LNTL_{0.99}$ and $UNTL_{0.99}$ represent the lower and upper 99% (tolerance level = $\alpha = 0.01$) natural tolerance limits of a process, then 99% of the manufactured items have their dimensions within the 99% tolerance interval ($LNTL_{0.99}, UNTL_{0.99}$).

Exercise 61. (a) A manufacturing process produces steel pipes whose lengths are $N(12.00", 0.0009)$. Obtain the 99% natural tolerance limits ($\alpha = 0.01$) and interpret. (b) Obtain the 6-sigma natural tolerance limits and determine the value of α . ANS: (a) $LNTL_{0.99} = 11.9227$, (b) (11.91, 12.09).

DESIGN TOLERANCES OR (CUSTOMER SPECIFICATIONS)

Customer specifications (LSL, USL) are limits that are set (or specified) by the designer (in order to meet consumer needs) without regard to process capability (or natural tolerances). Design tolerances on a nominal dimension (nominal-the best = NTB) are usually given as $m \pm \Delta$, where m is the ideal target value, i.e., $LSL = m - \Delta$ and $USL = m + \Delta$. Some authors use T or τ for target, but m is more common because it implies "middle" of tolerances. Note that if a QCH (quality characteristic) is of the smaller-the-better (STB) type, then there is only a single USL and the ideal target is always $m = 0$. On the other hand, if a QCH is of the LTB type, then there is only a single LSL and the ideal target is $m = \infty$. Further, all nominal dimensions have both an LSL and an USL.

Exercise 62. The design (or consumer) specifications for the steel pipes of Exercise 61 are $12.00'' \pm 0.075$. Compute the fraction nonconforming (FNC), p , of the process and decide if the 99% natural tolerances are capable of conforming to specifications, i.e., determine if $p < \alpha$. ANS: No, because $p = 0.01242 > \alpha$. (b) Determine the value of σ such that the 99% natural tolerances barely meet design specifications. (c) For the value of σ found in (b), compute the 99% process capability ratio, whose desired value is greater than 1, and its definition is given below:

$$PCR = \frac{USL - LSL}{(2z_{\alpha/2})\sigma} \quad (34)$$

If $PCR < 1$, then the machining process is not generally capable of meeting customer specifications. Note that a simple way to evaluate the quality capability of a machining process is to estimate the FNC, p , and compare it against the company-wide tolerance level α . If $p > \alpha$, then the process fails to meet customer specifications. When $\alpha = 0.0027$ so that $Z_{\alpha/2} = 3$ and $PCR = 2$, then the process is said to operate at the Motorola's 6-Sigma Quality. The Motorola's 6-Sigma Quality, however, can be deceiving when the process is not centered. Because, when a process is centered (i.e., $\mu = m$) and $USL - LSL = 12\sigma$ (or $PCR = 2$), then the Motorola's Gaussian FNC is equal to $p_M = 0.001973175401$ parts per million (ppm), or roughly 2 parts per billion. However, if a normal process mean shifts as much as one-and-a-half sigma (i.e., $\mu = m \pm 1.5\sigma$), its FNC increases to $p_M = 3.3976731564911$ ppm instead of the perceived amount of 0.001973175401 parts per million.

NATURAL TOLERANCE LIMITS WHEN μ AND σ ARE UNKNOWN

In practice, the mean μ and standard deviation σ of a manufacturing process are generally unknown constants and have to be estimated from a random sample

of size n , respectively, by $\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, and by $S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = CSS/$

$(n - 1)$, where $CSS = USS - CF$, the $USS = \sum(x_i^2)$, $CF = (\sum_{i=1}^n x_i)^2/n$, and thus

$$s^2 = \frac{1}{n-1} \left[\sum x_i^2 - \frac{(\sum x_i)^2}{n} \right] = \frac{1}{n-1} [USS - CF] \quad (35)$$

However, the tolerance interval $\bar{x} \pm (1.96 \times S)$ no longer has $\gamma = 100\%$ probability of containing 95% of the manufactured items since \bar{x} and S are both sample statistics (i.e., random variables) and change from one sample of size n to the next sample of size n . Fortunately, it is possible to determine a constant k called the tolerance factor, for a normal universe, such that the interval $\bar{x} \pm k S$ contains at least $(1 - \alpha)$ proportion of the population with a confidence Pr equal to γ .

Tables A.6 on page 744 of Devore gives the values of k for $n = 2$ to 300, $\gamma = 0.950$, 0.99 and $1 - \alpha = 0.90$, 0.95, and 0.99. Note that the tolerance factor $k \geq Z_{\alpha/2}$ and k very slowly approaches $Z_{\alpha/2}$ as $n \rightarrow \infty$. Therefore, for a random sample of size n from a $N(\mu, \sigma^2)$ universe, we are $\gamma \times 100\%$ confident that at least $1 - \alpha$ proportion of the population lies within the $(\gamma, 1 - \alpha)$ tolerance interval $\bar{x} \pm kS$. The original article by Wald and Wolfowitz (1946) and an MS thesis by Hsin-Cheng Chiu (directed by this author) contains the procedure for obtaining the values in tables A.6.

Example 37. A random sample of $n = 25$ tires had an average tread length $\bar{x} = 75.10$ " with $S = 0.070$. (a) Determine a tolerance interval that contains $1 - \alpha = 99\%$ of the total output with a confidence pr of $\gamma = 0.95$. (b) Given that the design specifications are 75 ± 0.25 " and the tolerable fraction nonconforming (or tolerance level) is $\alpha = 0.01$, determine if the capability of the process conforms to specifications.

Solution. (a) Since $\alpha = 0.01$, $\gamma = 0.95$ and $n = 25$, Table A.6 gives $k = 3.457$. Thus the $(0.95, 0.99)$ tolerance interval is $75.10 \pm 3.457 \times 0.07 = (74.858, 75.342)$ ". Thus, we are 95% sure that 99% of the product output have dimensions within the tolerance interval $(74.858, 75.342)$ ". (b) Since the sample indicates the process may not be centered (i.e., $\mu \neq 75.00$), it is not appropriate to measure process capability

at $\alpha = 0.01$ with PCR = 1.3863 defined in equation (34) and compare it against 1.

We observe that on the lower side, the $LTL_{.99} = 74.858$ conforms well to $LSL = 74.75$, but $UTL_{.99} = 75.342$ does not conform to the $USL = 75.25$. Therefore, we must use the capability index

$$C_{pk} = \frac{Z_{\min}}{Z_{\alpha/2}} = \frac{(75.25 - 75.10)/0.07}{Z_{0.005} (=2.57583)} = 0.83191$$

where $Z_{\min} = \text{Min}(Z_L, Z_U) = \text{Min}[(\bar{X} - LSL)/S, (USL - \bar{X})/S]$, and compare its value against 1. Since $C_{pk} < 1$, then altogether the TL's do not conform to customer specifications and thus $p > 0.01$. An estimate of p may be obtained from

$$\hat{p} = \Phi(-5) + \left[1 - \Phi\left(\frac{75.25 - 75.10}{0.07}\right) \right] = 0.0161, \text{ which exceeds } \alpha = 0.01.$$

Exercise 63. Work Exercises 37 (a and c) and 38(b) on page 307 of Devore.

ANS to 38 (b): $0 < 95\%$ of the X 's ≤ 0.080615 .

Final Exam for STAT3600