Statistical Inference About **Parameters of Two Normal Populations**

1. CONFIDENCE INTERVALS FOR DIFFERENCE IN MEANS OF TWO NORMAL PROCESSES WHEN BOTH VARIANCES ARE KNOWN

Reference: CHAPTER 9 OF Devore's 8th Edition **By S. Maghsoodloo**

Example 41 On a Nominal CI . A consumer group tested 2 major brands of radial tires, X and Y, to determine if there were significant differences in expected (or mean) tread-life measured in 1000 miles. The data (in 1000 miles) are given below:

X : 51.5, 53, 52, 47.0, 63.0, 51, 51, 51, 46, 56, 52.5, 42.5, 58, 52, 47, 49, 51.5
Y : 57.0 62, 48, 51.5, 54.5, 58, 57, 54, 63, 58, 62.0, 59.0, 56, 62, 58

From past experience it is known that the tread life X is N(μ_x, 22) and Y $\sim$ N(μ_y,18). Our objective is to develop a 95% (2-sided) CI for the mean difference $\mu_x - \mu_y$. Since a point estimator of $\mu_x - \mu_y$ is $(\hat{\mu}_x - \hat{\mu}_y) = \bar{x} - \bar{y} = 51.4118 - 57.3333 = -5.92157$, we have to make use of the sampling distribution of $\bar{x} - \bar{y}$ in order to obtain the requisite CI. Since $\bar{x}$ is N(μ_x, 22/17) and is independent of $\bar{y} \sim$ N(μ_y, 18/15), then $\bar{x} - \bar{y}$ is N($\mu_x - \mu_y$, 22/17 + 18/15 = 2.49412) as depicted below in Figure 18. Figure 18 clearly illustrates the SMD of $\bar{x} - \bar{y}$.
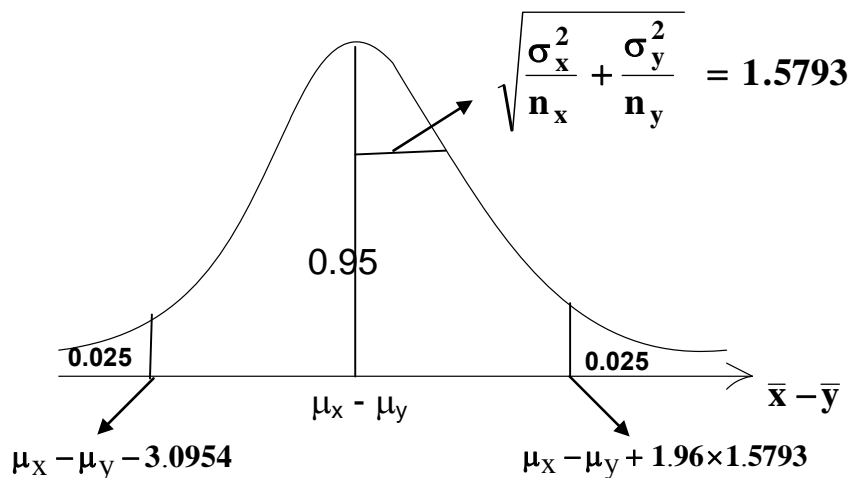

Figure 18. The Sampling Distribution of $\bar{x} - \bar{y}$

$$Pr(\mu_x - \mu_y - 3.09533 \leq \overline{x} - \overline{y} \leq \mu_x - \mu_y + 3.09533) = 0.95$$

or

$$Pr(-\overline{x} + \overline{y} - 3.09533 \leq -\mu_x + \mu_y \leq -\overline{x} + \overline{y} + 3.09533) = 0.95$$

$$Pr(\overline{x} - \overline{y} + 3.09533 \geq \mu_x - \mu_y \geq \overline{x} - \overline{y} - 3.09533) = 0.95$$

Rearranging this last Pr statement, we obtain the desired result:

$$Pr(\overline{x} - \overline{y} - 3.09533 \leq \mu_x - \mu_y \leq \overline{x} - \overline{y} + 3.09533) = 0.95.$$

Substituting for $\overline{x} - \overline{y} = -5.92157$ into the above confidence Pr statement, the 95%

CI for $\mu_x - \mu_y$ is given by

$$-9.01690 \leq \mu_x - \mu_y \leq -2.82624.$$

Since the above interval excludes 0, there is apparently a significant difference between the two

brand means, i.e., the null hypothesis $H_0 : \mu_x - \mu_y = 0$ must be rejected at the 5% LOS. Note that

the 95% CI for $\mu_y - \mu_x$ is given by $2.82624 \leq \mu_y - \mu_x \leq 9.01690$, which also excludes zero.

Exercise 71. A study was conducted to determine the impact of viscosity on the coating

thickness produced by a paint operation. Up to a certain paint viscosity, higher viscosities cause

thicker coatings. The data are provided below:

X = Low Viscosity : 1.09, 1.12, 0.83, 0.88, 1.62, 1.49, 1.59, 0.83, 1.04, 1.34, 1.83, 1.65, 1.71, 1.76

Y = High Viscosity : 1.46, 1.51, 1.59, 1.40, 0.94, 0.98, 0.89, 1.03, 2.05, 2.17, 2.06, 2.02, 1.51, 1.46,

1.42, 1.40, 1.53, 2.07

Assuming that X is N (μ_x, 0.130) and Y is also N(μ_y, 0.170) but independent of X, obtain the 95%

lower one sided CI for $\mu_y - \mu_x$. ANS: $-0.03934 \leq \mu_y - \mu_x < \infty$. Note that this interval does include

0, and therefore the null hypothesis $H_0 : \mu_y - \mu_x = 0$ cannot be rejected at the 5% level! (b) Work

Exercise 9 on page 356 of Devore, using the large-sample approximation to Student's t.

2. TEST OF HYPOTHESIS ON EQUALITY OF MEANS OF TWO De Moivre
   ## (or Normal) POPULATIONS WITH KNOWN VARIANCES

For the sake of illustration, consider the Exercise 6, page 355 of Devore (8e) where n_x = 40

with $\sigma_x = \sigma_1 = 1.60$ kgf/cm^2 for the Modified Mortar (the treatment group), and n_y = 32, $\sigma_2 = \sigma_y$ =

1.4 for the unmodified mortar (the control group). Our objective is to test if the modified mortar

(the treatment group) has a greater average bond strength than that of unmodified, i.e., we

wish to test H_0: $\mu_1 = \mu_2$ vs H_1: $\mu_1 > \mu_2$ at $\alpha = 0.01$. Since $\overline{x}_1 - \overline{x}_2$ is $N(\mu_1 - \mu_2, \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}) = N(\mu_1 -$

μ_2, 0.12525), then under the null hypothesis H_0: $\mu_1 - \mu_2 = 0$, the SMD of our test statistic $\overline{x}_1 - \overline{x}_2$

follows the normal distribution depicted in Figure 19. $\qquad A_U = (\overline{x}_1 - \overline{x}_2)_U$
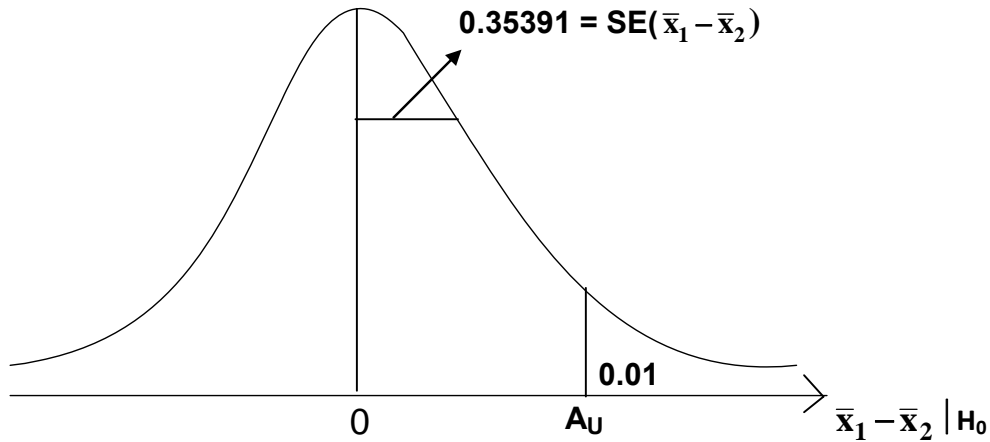

Figure 19. The Null Sampling Distribution of $\overline{x}_1 - \overline{x}_2$ Given that H_0 is True

The upper limit for the acceptance interval AI = $(-\infty, A_u)$ is $A_u = 0 + Z_{0.01} \times SE(\overline{x}_1 - \overline{x}_2) =$

$2.3263 \times 0.35391 = 0.82331$. Therefore, the critical (or rejection) region for the test statistic

$\overline{x}_1 - \overline{x}_2$ is $[0.8233, \infty)$, i.e., we must reject H_0 at the 1% level iff $\overline{x}_1 - \overline{x}_2 > 0.8233$. Because $\overline{x}_1 - \overline{x}_2$

= $18.12 - 16.87 = 1.25$ kgf/cm^2 is well inside the critical region $\overline{AI}$ = AI' = $0.8233 <, \overline{x}_1 - \overline{x}_2$ then we

do have sufficient evidence, at the 1% level, to reject H_0. Once H_0 is rejected at $\alpha = 0.01$ with the

given data, then the probability level (or *P-value*) of the test must be smaller than 0.01 as

computed next. Since the above test is right-sided, the *P-value* = $P(\overline{x}_1 - \overline{x}_2 \geq 1.25) = P(Z_0 \geq$

3.532) = 0.0002062, and hence $\hat{\alpha}$ = 0.0002062 < α = 0.01, as expected because the test statistic

$\overline{x}_1 - \overline{x}_2$ rejected H_0 at the 1% LOS. Further, the value of $Z_0 = 3.5320 > Z_{0.01} = 2.3263$, again

implying sufficient evidence to reject H_0 at the LOS of 1%.

　　　　To compute the type II error Pr, recall that 1st we have to assume H_0 false. Consequently,

suppose $\mu_1 - \mu_2 = 1 \neq 0$ (i.e., H_0 is assumed false) as illustrated in Figure 20. Then how do we compute the prior Pr of accepting H_0: $\mu_1 - \mu_2 = 0$ given that H_0 is now assumed false. Figure 20
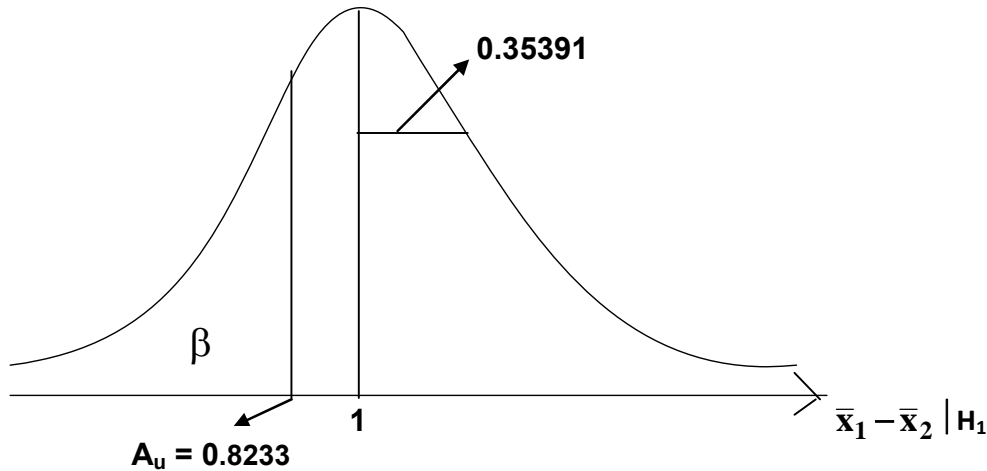

Figure 20. The (alternate) Sampling Distribution of $\overline{x}_1 - \overline{x}_2$ assuming H_0 is False

clearly shows that β(at $\mu_1 - \mu_2 = 1$) = $\Pr(\overline{X}_1 - \overline{X}_2 \leq A_u \mid \mu_1 - \mu_2 = 1)$ = $\Pr(-\infty < \overline{X}_1 - \overline{X}_2 \leq 0.8233 \mid \mu_1 - \mu_2 = 1)$ = $\Pr(Z \leq -0.4993) = \Phi(-0.4993) = 0.3088$.

 Exercise 72. (a) Work Exercise 6 on page 355 of Devore. For part (c) use the pertinent version of the formula page 124 of my notes. Note that Devore uses m for $n_1 = n_x$ and n for $n_2 = n_y$. (b) Draw the 5%-level OC curve for the Exercise 6 on page 355 of Devore by tabulating the values of β for $\mu_1 - \mu_2 = -0.50, 0, 0.50, 0.5821, 1,$ and 1.25. Note that you already know three points on the OC curve.

3. SAMPLE SIZE DETERMINATION GIVEN TWO SPECIFIC POINTS ON THE OC CURVE

 To illustrate the procedure, suppose we wish to test $H_0 : \mu_1 - \mu_2 = 0$ VS the alternative $H_1 : \mu_1 - \mu_2 \neq 0$, where $X \sim N(\mu_1, 0.0004)$ and $Y \sim N(\mu_2, 0.000625)$. Our objective now is to design a sampling procedure such that our OC curve goes thru the three points (0, 0.95) and ($\pm$ 0.02, 0.10), i.e., at $\delta = 0$, $\beta = 1-\alpha = 0.95$, but $\beta = 0.10$ when $\delta = \mu_1 - \mu_2 = \pm 0.02$. Note that Devore uses Δ for

the population mean difference $\mu_1 - \mu_2$, while I am using $\delta = \mu_1 - \mu_2$. We need the information only about 2 points; however, because of symmetry, the OC curve will also automatically go thru the 3rd point, and this in turn will determine the three unknowns n, A_L and A_u.

We will work with the two points ($\mu_1 - \mu_2 = 0$, $\beta = 0.95$) and (-0.02, 0.10). From the point (0, 0.95), we obtain $A_L = 0 - 1.96 \sqrt{\dfrac{0.02^2}{n_1} + \dfrac{0.025^2}{n_2}}$. Secondly, Figure 21 shows that $A_L = -0.02 +$ $1.28155 \times SE(\overline{x}_1 - \overline{x}_2)$. Equating these last 2 expressions, we obtain $A_L = 0 - 1.959964 \times$

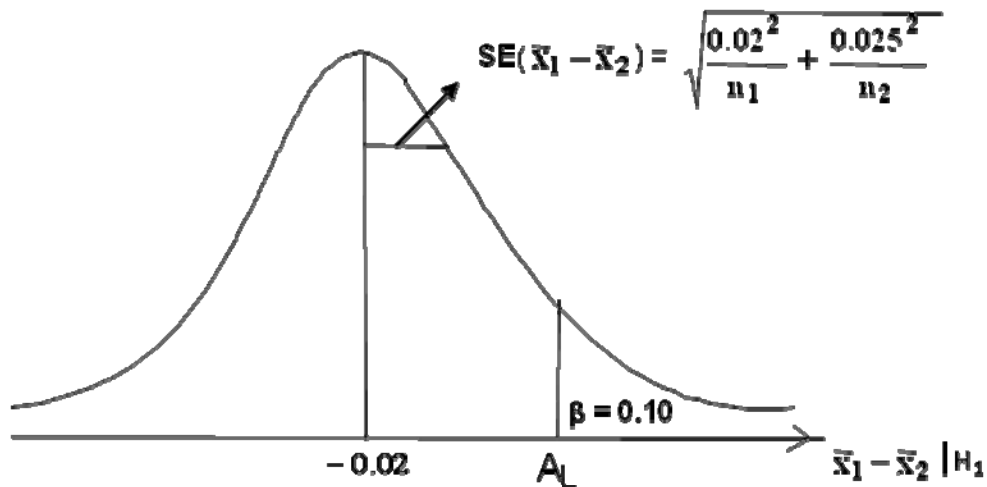$$\sqrt{\dfrac{0.02^2}{n_1} + \dfrac{0.025^2}{n_2}} = -0.02 + 1.281552 \times SE(\overline{x}_1 - \overline{x}_2), \text{ or}$$

Figure 21. The Sampling Distribution of $\overline{x}_1 - \overline{x}_2$ assuming that H$_0$ is false.

$$-1.96 \sqrt{\dfrac{0.02^2}{n_1} + \dfrac{0.025^2}{n_2}} = -0.02 + 1.281552 \times \sqrt{\dfrac{0.02^2}{n_1} + \dfrac{0.025^2}{n_2}}$$

$$0.02 = 3.2415 \times \sqrt{\frac{0.000400}{n_1} + \frac{0.000625}{n_2}} \quad . \tag{36}$$

Squaring both sides of equation (36) and simplifying gives rise to

$$0.0004 = \frac{0.00420297}{n_1} + \frac{0.00656714}{n_2}. \tag{37}$$

Equation (37) has two unknowns, and therefore, there are infinite number of solutions. For example, one possible solution is to take $n_1 = 15$ and $n_2 = 55$; another solution set is ($n_1 = n_x = m = 20$, $n_2 = n_y = n = 35$; note that Devore's notation m & n for sample sizes is not common statistically. In other words, if you specify the value of n_1 (or n_2), then I will in turn determine the other sample size such that $\dfrac{0.00420297}{n_1} + \dfrac{0.00656714}{n_2} \leq 0.0004$. The question arises how should we allocate our total resources $n_{Total} = n_1 + n_2$ between the two populations in order to minimize the $V(\overline{x}_1 - \overline{x}_2)$, which also automatically minimizes the $SE(\overline{x}_1 - \overline{x}_2)$. It can be proven that our total resources n_{Total} must be allocated proportional to the standard deviation of the two populations, i.e., the larger σ_i, $i = 1$ and 2 is, the larger the corresponding n_i should be. Partially differentiating $V(\overline{X}_1 - \overline{X}_2) = \sigma_1^2 / n_1 + \sigma_2^2 / n_2$ with respect to n_1 and n_2 and requiring that both partial derivatives be zero leads to the allocations $n_1 = (\dfrac{\sigma_1}{\sigma_1 + \sigma_2}) \times n_{Total}$ and $n_2 = n_{Total} - n_1$. Thus, our optimum solution for the above example is $n_1 = 0.444\overline{4} \times n_{Total}$ and $n_2 = 5555\overline{5} \times n_{Total}$. Inserting these into equation (36) yields $n_{Total} \geq 53.19383$ or $n_{Total} = 54$. Hence $n_1 = 0.4444\overline{4} \times 54 = 24$ and $n_2 = 30$.

It can be shown that in general the total sample size $= n_1 + n_2$ for a 2-sided test must be selected from the equation $n_{Total} = (Z_{\alpha/2} + Z_\beta)^2 (\sigma_1 + \sigma_2)^2 / (\delta_1 - \delta_0)^2$, where δ_0 is the hypothesized value of $\mu_1 - \mu_2$ under H_0, and δ_1 is a value of $\mu_1 - \mu_2$ under H_1 for which the type II error Pr is required to be β. Once n_{Total} is determined, then $n_1 = n_{Total}\sigma_1 / (\sigma_1 + \sigma_2)$. If the alternative is one-sided, then replace $Z_{\alpha/2}$ with Z_α.

4. CONFIDENCE INTERVAL ON THE MEAN DIFFERENCE OF TWO INDEPENDENT NORMAL POPULATIONS WHEN $\sigma_1^2 = \sigma_2^2 = \sigma^2$ IS UNKNOWN (THE COMPLETELY RANDOMIZED DESIGN)

For the sake of illustration consider the Experiment reported in the Journal of Waste and Hazardous Materials (Vol. 6, 1989), where X_1 = weight of calcium in standard cement (the control group), while X_2 = weight of calcium in cement doped with lead (the treatment group). Reduced levels of calcium would cause the hydration mechanism to get blocked and allow water to attack various locations of cement structure. Ten samples of standard cement gave $\bar{x}_1 = 90.0$ with $S_1 = 5.0$ while 15 samples of lead-doped cement resulted in $\bar{x}_2 = 87.0$ with $S_2 = 4.0$ Assuming that $X_1 \sim$ N(μ_1, unknown σ^2) and $X_2 \sim$ N (μ_2, σ^2), then $\bar{x}_1 - \bar{x}_2$ is N($\mu_1 - \mu_2, \dfrac{\sigma^2}{n_1} + \dfrac{\sigma^2}{n_2}$), where it is assumed that σ^2 is the common value of the unknown $\sigma_1^2 = \sigma_2^2$. It was shown in STAT 3600 that if $\hat{\theta}$ is an unbiased estimator of a parameter θ with standard error $se(\hat{\theta}) = S_{\hat{\theta}}$, then the sampling distribution of $(\hat{\theta} - \theta)/S_{\hat{\theta}}$ follows that of (Gosset's) Student's-t with df (degrees of freedom) equal to that of $S_{\hat{\theta}}$. Accordingly, let $\hat{\theta} = \bar{x}_1 - \bar{x}_2$; then $\bar{x}_1 - \bar{x}_2$ is an unbiased estimator of $\mu_1 - \mu_2$ with V($\bar{x}_1 - \bar{x}_2$) $= \sigma^2(\dfrac{1}{n_1} + \dfrac{1}{n_2})$. Since the common value of the process variances σ^2 is unknown, we must pool both unbiased estimators S_1^2 and S_2^2 of σ^2 to obtain one unbiased estimator of σ^2, which is given by their weighted average of the two sample variances based on their df (or DOF).

$$S_p^2 = \frac{v_1 S_1^2}{v_1 + v_2} + \frac{v_2 S_2^2}{v_1 + v_2} = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{CSS_1 + CSS_2}{n_1 + n_2 - 2}. \tag{38}$$

Note that $E(S_p^2) = \sigma^2$. Therefore, the $se(\bar{x}_1 - \bar{x}_2) = S_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$, and as a result the rv

$[(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)]/(S_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}})$ has a t-sampling distribution with $v = n_1 + n_2 - 2$ df. To

obtain the two-sided 95% CI for $\mu_1 - \mu_2$, we make use of the confidence Pr statement ($n_1 = 10$,

$n_2 = 15$, $t_{0.025,23} = 2.069$)

$$Pr(-2.0687 \leq T_{23} \leq 2.0687) = 0.95.$$

Substituting $[(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)]/\left(S_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}\right)$ for T_{23} in this last Pr statement and solving

for $(\mu_1 - \mu_2)$ leads to the desired 95% CI given below.

$$\bar{x}_1 - \bar{x}_2 - 2.06876 \times S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + 2.0687 \times S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \; .$$

To obtain the lower and upper confidence limits, we need to compute S_p^2 using $n_1 = 10$, $S_1{}^2$

= 25, $n_2 = 15$ and $S_2{}^2 = 16$. From equation (38), $S_p^2 = \dfrac{9(25)}{23} + \dfrac{14(16)}{23} = 19.52174$, and hence

$se(\bar{x}_1 - \bar{x}_2) = 4.41834 \sqrt{\dfrac{1}{10} + \dfrac{1}{15}} = 1.8038$, with $\bar{x}_1 = 90.0$ and $\bar{x}_2 = 87.0$. Thus, $L(\mu_1 - \mu_2) = 3 -$

$2.069 \times 1.8038 = -0.732$ and $U(\mu_1 - \mu_2) = 3 + 3.732 = 6.732$. Since this 95% CI, $-0.732 \leq \mu_1 - \mu_2 \leq$

6.732, includes zero, then there does not exist a significant difference between the two population

parameters μ_1 and μ_2 at the 5% level. This implies that doping cement with lead does not

significantly alter water hydration mechanism from a statistical standpoint at $\alpha = 0.05$. However,

the null hypothesis $H_0 : \mu_1 - \mu_2 = -1$ must be rejected at the 5% level because -1 lies outside the

95% CI.

Similar developments as above lead to the upper one-sided CI for the parameter difference

$\mu_1 - \mu_2$ given below:

$$-\infty < \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + t_{\alpha, n_1 + n_2 - 2} \times S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \; .$$

The lower one-sided CI for $\mu_1 - \mu_2$ with confidence coefficient $1 - \alpha$ is given by

$$\bar{x}_1 - \bar{x}_2 - t_{\alpha, n_1 + n_2 - 2} \times S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leq \mu_1 - \mu_2 < +\infty.$$

When the variances of the two independent normal populations, σ_1^2 and σ_2^2, are

unequal and unknown, then the CI for $\mu_1 - \mu_2$ must be obtained using the equation (9.2), on page

357 of Devore's 8th edition, but the *df* ν, where Min(n_1-1, n_2-1) < $\nu \leq n_1 + n_2 - 2$, is given by the

equation near the bottom of Devore's page 357. As a general rule of thumb, I would

recommend against using the pooled t procedure outlined on pp. 124-126 of my notes if

$S_1^2 > 2 S_2^2$, or vice a versa. When the variances of the two samples are significantly different, then

the two-sample t procedure (outlined in section 9.2. pp. 357-360 of Devore's 8th edition due to B.

L. Welch) must be applied for both obtaining a CI for $\mu_1 - \mu_2$ and testing the null hypothesis H_0 :

$\mu_1 - \mu_2 = \delta$.

To illustrate the two-independent B.L. Welch' t-test, we work exercise 21 on page 362 of

Devore (8e). Since $S_1^2 = (0.53)^2 = 0.2809$ and $S_2^2 = (0.87)^2 = 0.7569$, showing that $S_2^2 \gg 2 S_1^2$, then

the assumption of $\sigma_1^2 = \sigma_2^2 = \sigma^2$ is very tenuous at best! Therefore, we use the two-sample t

procedure, outlined in section 9.2 of Devore's 8th edition on page 357, which does not require the

assumption of $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

The sample statistics for the Exercise 21, p. 362, are for the case of W/O probing n_1 =

Devore's m = 8, $\bar{x}_1$ = 1.71 mm, S_1 = 0.53, and for the CTS (Carpal Tunnel Syndrome) subjects n_2 = n

= 10, $\bar{x}_2$ = 2.53 mm, S_2 = 0.87. We wish to test H_0: $\mu_2 - \mu_1$ = 0 against the alternative H_1: $\mu_2 - \mu_1$ > 0

at α = 0.01. Our test statistic from equation 9.2 of Devore (p. 357) is

$$t_0 = \frac{2.53 - 1.71 - 0}{\sqrt{\dfrac{0.87^2}{10} + \dfrac{0.53^2}{8}}} = 2.4634$$

with *df* ν = 15.11 (computed from the equation below Devore's Eq. (9.2) on page 357). Since

$t_{0.01,15}$ = 2.602 and $t_{0.01,16}$ = 2.583, then our critical threshold value is $t_{0.01,15.11} \cong 2.6003$. Since our

test statistic t_0 = 2.4634 does not exceed this threshold value, then the two samples provide

insufficient evidence, at the 1% level, to conclude that the true average gap detection for the CTS

population exceeds that of normal subjects. The *P-value* of the test is computed from $\hat{\alpha} \cong$

Pr($T_{15.11} \geq 2.4634$) = 0.01312, which exceeds α = 0.01 as expected. However, our test statistic t_0 =

2.4634 does provide sufficient evidence, at the 5% level, to conclude that $\mu_2 > \mu_1$.

Unfortunately, the SMD of the two-independent-sample t statistic given in equation (9.2) of Devore when H_0 is false, as far as I know, is intractable for small and moderate sample sizes, say $2 \leq n < 50$, and therefore, the power of the t-test cannot be estimated directly. For the pooled t-test, the OC curves of Table A.17 on page A-28 of Devore's 8th edition may be used only when $n_1 = n_2 = n$ with the abscissa d = $(\delta - \delta_0)/(2S_p)$ and the sample size for the OC curve as $n_{OC} = 2n - 1$.

Exercise 73. Work Exercises 18, 23(c) and 25 on pages 362-3 of Devore (8e).

5. CONFIDENCE INTERVAL ON $\mu_1 - \mu_2$ FOR PAIRED

OBSERVATIONS (I.E., THE RANDOMIZED COMPLETE BLOCKS)

As an example, consider the Example 9.9, on pp. 367-8 of Devore's 8th edition, where 16 subjects (or blocks) were selected at random and two responses (X_{1j}, X_{2j}) were obtained from each subject (or block). When j = 5, the paired responses from block number 5 are $(x_{15} = 90, x_{25} = 84)$. It is assumed that the random vector $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ has a bivariate normal distribution with unknown V(X_1) = σ_1^2 , V(X_2) = σ_2^2 , and unknown correlation coefficient ρ (note that X_{1j} and X_{2j} cannot possibly be independent because they originate from the same j^{th} block or j^{th} subject). Letting $D_j = X_{1j} - X_{2j}$, the rv D is also normally distributed with

$$E(D) = E(X_1 - X_2) = \mu_1 - \mu_2 = \mu_D , \qquad (39a)$$

and V(D) = $\sigma_1^2 + \sigma_2^2 - 2COV(X_1, X_2)$, or

$$\sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\sigma_{12}. \qquad (39b)$$

Clearly, all the parameters in equations (39) are unknown and have to be estimated by sample statistics. To obtain a CI for $\mu_D = \mu_1 - \mu_2$, we use the fact that the sampling distribution of $(\bar{d} - \mu_D)/ se(\bar{d})$ is that of W. S. Gosset's T rv with $\nu = (n - 1)$ *df*. For the Example 9.9 on pp. 367-368 of Devore's 8th edition, the number of blocks n = 16 and hence the *df* of $(\bar{d} - \mu_D) \sqrt{n} /S_d$ is $\nu = (n - 1) = 15$. The 16 paired differences for our example are provided in the Table in the middle of

page 367 of Devore. The n = 16 differences give rise to $\bar{d} = \dfrac{1}{16}\sum\limits_{j=1}^{n} d_j = 6.75$ and

$S_d^2 = \dfrac{1}{15}\sum\limits_{j=1}^{16}(d_j - \bar{d})^2 = 67.8 \rightarrow S_d = 8.2341 \rightarrow se(\bar{d}) = 2.05852$. Table A.5, p. A-9 of Devore's

8th ed., shows that the $Pr(-2.131 \le T_{15} \le 2.131) = 0.95$, or

$$Pr(-2.131 \le \dfrac{(\bar{d} - \mu_D)\sqrt{n}}{S_d} \le 2.131) = 0.95$$

Rearranging the above 2-sided inequality inside the parentheses yields

$$Pr[\bar{d} - t_{0.025;15} \times S_d / \sqrt{n} \le \mu_D \le \bar{d} + 2.131 \times S_d / \sqrt{n}\,] = 0.95$$

Therefore, $(\mu_D)_L = LB(\mu_1 - \mu_2) = 6.75 - 2.131(2.05852) = 2.3633$ and $UB(\mu_1 - \mu_2) = 11.1367$, where

LB stands for the lower bound and UB for upper bound. Since this 95% CI interval $2.3633 \le \mu_1 -$

$\mu_2 \le 11.1367$ excludes zero, it provides conclusive evidence that the 2 population averages are

significantly different, i.e., the null hypothesis $H_0 : \mu_1 - \mu_2 = 0$ must be rejected at the 5% level.

The testing procedure for the above Example 9.9 is well outlined by Devore at the top of

page 368. The critical level of the test is given by $\hat{\alpha} \cong 2 \times Pr(T_{15} \ge 3.27906) = 2(0.00253602) =$

0.0050721 (From Microsoft Excel function: tdist(3.27906,15,2), which is consistent with the

rejection of H_0 since the 95% CI excluded the hypothesized value of the population mean

difference equaling zero.

The power of the paired t test can be estimated using the OC curves of Table A.17 on page

A-28 of Devore's 8th edition with abscissa $d \cong \left| \mu_D - \mu_0 \right| / S_d$ and the sample size n equal to the

number of pairs (or blocks).

Exercise 74. (a) For the data of Example 9.9, pages 367-368 of Devore, obtain the point

unbiased estimates of $\mu_1 - \mu_2$ and σ_D^2. For σ_D^2 use equations (39b) and the unbiased estimator

$$\hat{\sigma}_D^2 = \dfrac{1}{n-1}\sum\limits_{j=1}^{n}(d_j - \bar{d})^2 = \dfrac{1}{n-1}\sum\limits_{j=1}^{n}\left[(x_{1j} - x_{2j}) - (\bar{x}_1 - \bar{x}_2)\right]^2$$

$$\hat{\sigma}_D^2 = \frac{1}{n-1} \sum_{j=1}^{n} [(x_{1j} - \bar{x}_1) - (x_{2j} - \bar{x}_2)]^2$$

$$\hat{\sigma}_D^2 = \frac{1}{n-1} \sum_{j=1}^{n} [(x_{1j} - \bar{x}_1)^2 + (x_{2j} - \bar{x}_2)^2 - 2(x_{1j} - \bar{x}_1)(x_{2j} - \bar{x}_2)]$$

$$= S_1^2 + S_2^2 - 2 \frac{1}{n-1} \sum_{j=1}^{n} (x_{1j} - \bar{x}_1)(x_{2j} - \bar{x}_2)]$$

$$= S_1^2 + S_2^2 - 2\hat{\sigma}_{12} = S_1^2 + S_2^2 - 2S_{12}/(n-1)$$

where

$$\hat{\sigma}_{12} = \frac{1}{n-1} \sum_{j=1}^{n} (x_{1j} - \bar{x}_1)(x_{2j} - \bar{x}_2) = \frac{S_{12}}{n-1} \text{ and } S_{12} = \sum_{j=1}^{n} (x_{1j} - \bar{x}_1)(x_{2j} - \bar{x}_2) = \text{Sum of Cross-}$$

$$\text{Products} = \sum_{j=1}^{n} x_{1j} x_{2j} - (\sum_{j=1}^{n} x_{1j} \sum_{j=1}^{n} x_{2j})/n = \sum_{j=1}^{n} (x_{1j} x_{2j}) - n\bar{x}_1 \bar{x}_2 = \sum_{j=1}^{n} (x_{1j} x_{2j} - \bar{x}_1 \bar{x}_2).$$

(b) Work Exercises 36, 37, 40, and 46 on pages 371-374 of Devore (8e).

6. (SIR Ronald A.) FISHER'S F DISTRIBUTION

A continuous rv, X, has the (Sir Ronald A. Fisher's) F distribution iff its pdf is given by

$$f(x) = C x^{(v_1 - 2)/2} (v_2 + v_1 x)^{-(v_1 + v_2)/2} , \qquad 0 \le X < \infty.$$

By now, surely you know the constraint on the normalizing constant C, that makes the above f(x) a

density function, and which leads to the specific value of $C = \dfrac{v_1^{v_1/2} v_2^{v_2/2} \Gamma(\dfrac{v_1 + v_2}{2})}{\Gamma(v_1/2)\Gamma(v_2/2)}$, where v_1

is called the *df* (or DOF) of the numerator and v_2 that of the denominator, for reasons that will be

explained in the following paragraphs. It can be shown that the modal point of the F distribution

occurs at MO = $v_2(v_1 - 2)/[v_1 (v_2 + 2)]$, $v_1 \ge 2$. It can also be verified that $0 \le$ MO < 1, and for very

large v_1 and v_2 the value of MO is close to 1 but always less than 1. Further, the 1st two moments

of F are given by E(F) = $v_2 / (v_2 - 2)$ for $v_2 > 2$, and

$$V(F) = \frac{2v_2^2\,(v_2 + v_1 - 2)}{v_1\,(v_2 - 2)^2\,(v_2 - 4)} \quad , \text{ only if } v_2 > 4.$$

These last two expressions imply that the mean of F does not exist for $v_2 \leq 2$ and the variance of F does not exist for $v_2 \leq 4$. The skewness of the F rv is given by

$$\alpha_3 = \frac{2v_1 + v_2 - 2}{v_2 - 6}\sqrt{\frac{8(v_2 - 4)}{v_1(v_1 + v_2 - 2)}} \quad , \quad v_2 > 6$$

so that the F distribution is positively skewed (MO < $x_{0.50}$ < μ). Its kurtosis, β_4, is given by

$$\beta_4 = \alpha_4 - 3 = \frac{12[(v_2 - 2)^2(v_2 - 4) + v_1(v_1 + v_2 - 2)(5v_2 - 22)]}{v_1(v_2 - 6)(v_2 - 8)(v_1 + v_2 - 2)} \quad , \quad v_2 > 8$$

where $\alpha_4 = E[(X - \mu)^4 / \sigma^4] = E[(X - \mu)^4] / \sigma^4 = E(Z^4)$.

Exercise 75. Graph the density functions of the Fisher's F rv for ($v_1 = 1$, $v_2 = 2$), ($v_1 = 2$, $v_2 = 3$), and for ($v_1 = 3$, $v_2 = 4$). Note that the graph of the F distribution resembles Figure 22 only for $v_1 > 2$. When $v_1 = 2$, the mode occurs at the origin and when $v_1 = 1$, f($F_{1,v2}$) has an asymptote at the origin.

The graph of pdf(F_{v_1, v_2}) for $v_1 > 2$ is given in Figure 22. Figure 22 clearly shows that the Pr($F_{v_1, v_2} > F_{\alpha, v_1, v_2}$) = α, i.e., F_{α, v_1, v_2} represents the 100×α percentage point of the F-dist with v_1 *df* for the numerator and v_2 *df* for the denominator. The percentage points of the F distribution are tabulated in A.9 on pp. A-14 through A-19 of Devore for $\alpha = $ 0.10, 0.05, 0.01 and 0.001, and I have provided more percentage points on my website under Finverse. The easiest way to obtain any percentage points is to use the Microsoft excel function Finv(α, v_1, v_2). For example, Table A.9 on pages A-14 shows that $F_{0.05,4,10} = 3.48$, i.e., the Pr($F_{4,10} > 3.48$) = 0.05, and $F_{0.01,4,10} = 5.99$, i.e., the Pr($F_{4,10} \leq 5.99$) = 0.99, while MS Excel gives Finv(0.05,4,10) = 3.47805.
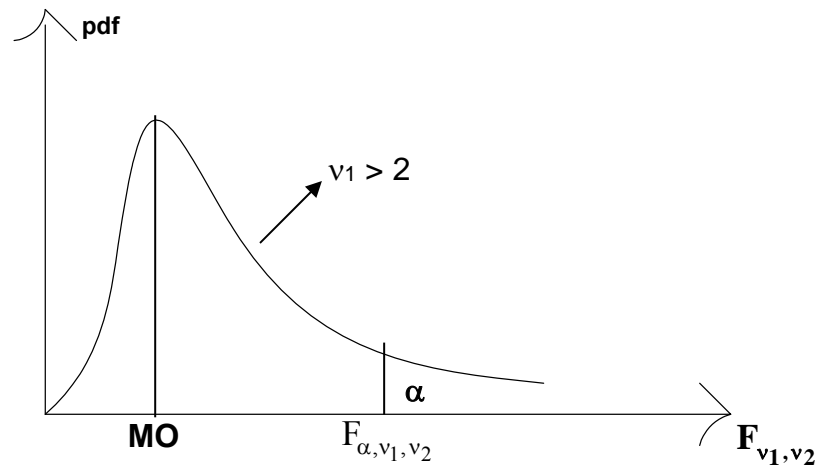
Figure 22. The Graph of the F Distribution for $v_1 > 2$

THE APPLICATION OF THE F DISTRIBUTION TO SAMPLING
OF Abraham De Moivre POPULATIONS

Statistical theory shows that an F rv can be generated through the ratio of two independent scaled (with respect to their *df*) χ^2 rvs as depicted in equation (40) below.

$$F_{v_1,v_2} = \frac{\chi^2_{v_1} / v_1}{\chi^2_{v_2} / v_2} \tag{40}$$

Equation (40) clearly shows that v_1 is the *df* of the χ^2 rv in the numerator and v_2 is the *df* of the χ^2 in the denominator, and as stated above both the numerator and denominator are scaled wrt (with respect to) their *df*.

As an application of equation (40), consider 2 independent normal (machining) processes $N(\mu_x, \sigma_x^2)$ and $N(\mu_y, \sigma_y^2)$. In order to compare the variances of two machines, we sample the two normal processes of sizes n_x and n_y, respectively. As you well know by now that the sampling distributions of both $(n_x - 1)S_x^2 / \sigma_x^2$ and $(n_y - 1)S_y^2 / \sigma_y^2$ follow χ^2 with $(n_x - 1)$ and $(n_y - 1)$ *df*, respectively. Substitution into equation (40) yields:

$$F_{n_x -1, n_y -1} = \frac{[(n_x -1)S_x^2 / \sigma_x^2]/(n_x -1)}{[(n_y -1)S_y^2 / \sigma_y^2]/(n_y -1)} = \frac{S_x^2 / \sigma_x^2}{S_y^2 / \sigma_y^2} \tag{41}$$

This above result shows that the ratio of two normal sample variances, S_x^2 / S_y^2, possesses a

sampling distribution which is the Fisher's F when the ratio S_x^2 / S_y^2 is multiplied by the scaling

factor σ_y^2 / σ_x^2. If it can be hypothesized that $\sigma_y^2 = \sigma_x^2$, then the sampling distribution of S_x^2 / S_y^2

$\sim F_{n_x -1, n_y -1}$. Put differently, under the null hypothesis $H_0 : \sigma_y^2 = \sigma_x^2$, the null SMD of the statistic

S_x^2 / S_y^2 follows the Fisher's F with $v_1 = n_y -1$ df for the numerator and $v_2 = n_x -1$ df for the

denominator.

Example 42. A random sample of size $n_x = 26$ is drawn from a N (μ_x, 25) process and one

of size $n_y = 21$ is selected from another N (μ_y, 49) process. Prior to the drawing of the two

samples, compute the $Pr(S_x^2 / S_y^2 \le 3.3712)$.

<u>Solution.</u>   Pr($S_y^2 / S_x^2 \le$ 3.3712) = $Pr[\frac{S_y^2 / \sigma_y^2}{S_x^2 / \sigma_x^2} \le \frac{3.3712 \times (25)}{49}]$ = Pr($F_{20,25} \le 1.72$) = 0.90,

which shows that the 90th percentile of the rv $F_{20,25}$ is equal to 1.72, or $F_{0.10,20,25} = 1.72$.

THE PERCENTAGE POINTS OF THE F DISTRIBUTION FOR $\alpha > 0.50$ (I.E., THE LEFT TAIL of F(v_1, v_2))

Table A.9 of Devore provides some percentage points of F(v_1, v_2) only for $\alpha < 0.50$, i.e.,

pages A-14 through A-19 of Devore tabulate only the right tail of the F distribution, and only for α

= 0.10, 0.05, 0.01, and 0.001. The question arises as to why the left tail such as $F_{0.99}(5, 10)$ is not

tabulated while the corresponding right tail $F_{0.01}(5, 10)$ is given to be 5.64 on page A-14?

Apparently there must exist a relationship between F_α and $F_{1-\alpha}$ as illustrated below for $v_1 = 5$ and

$v_2 = 10$.   Pr($F_{5, 10} \ge F_{0.99,5, 10}$) = 0.99 $\rightarrow$ $Pr[\frac{\chi_5^2 / 5}{\chi_{10}^2 / 10} \ge F_{0.99} (5,10)] = 0.99$

$$\longrightarrow \quad \Pr[\frac{\chi_{10}^2/10}{\chi_5^2/5} \le \frac{1}{F_{0.99,5,10}}] = 0.99 \qquad \longrightarrow \quad \Pr(F_{10,5} \le \frac{1}{F_{0.99,5,10}}) = 0.99$$

$$\longrightarrow \quad \Pr(F_{10,5} \ge \frac{1}{F_{0.99,5,10}}) = 0.01 \quad \longrightarrow \quad F_{0.01;10,5} = \frac{1}{F_{0.99,5,10}}. \quad \text{Thus, in general we have the}$$

result:

$$F_{\alpha,v1,v2} = \frac{1}{F_{1-\alpha,v2,v1}}, \quad \text{for all } 0 < \alpha < 1, \ v_1 \text{ and } v_2. \tag{42}$$

Exercise 76. A random sample of size $n_x = 13$ and one of size $n_y = 21$ are selected from 2 independent normal processes with $\sigma_x^2 = 6$ and $\sigma_y^2 = 10$. Compute the $\Pr(S_x^2/S_y^2 < 0.621)$. ANS: approximately 0.025. **(b)** Use Microsoft Excel to obtain $F_{0.995}(15, 20)$ and $F_{0.999,15,20}$. **(c)** Obtain the value of the constant C such that the $\Pr(C \le F_{5,10} \le 8.316C) = 0.80$.

7. CONFIDENCE INTERVALS FOR THE RATIO σ_x^2/σ_y^2 FROM TWO INDEPENDENT NORMAL (De Moivre) POPULATIONS

All three types of CIs for σ_x^2/σ_y^2 are of interest, i.e., a 2-sided CI: $L \le \sigma_x^2/\sigma_y^2 \le U$, a lower 1-sided: $L \le \sigma_x^2/\sigma_y^2 < \infty$, and an upper 1-sided: $0 < \sigma_x^2/\sigma_y^2 \le U$. Henceforth, σ_x^2 will represent the variance of process X (such as machine X) and σ_y^2 that of process Y.

Example 43. The diameter of steel rods manufactured on machines 1 and 2 are assumed $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively. Two random samples of sizes $n_1 = 15$ and $n_2 = 18$ are selected yielding the statistics $\bar{x}_1 = 8.73$ inches, $S_1^2 = 0.35$, $\bar{x}_2 = 8.68$ and $S_2^2 = 0.40$. The objective is to obtain a 95% 2-sided CI for σ_1^2/σ_2^2. We use the fact that the rv $(S_1^2/\sigma_1^2)/(S_2^2/\sigma_2^2)$ has the Fisher's $F_{14,17}$ sampling distribution depicted by Figure 23 given next page. Since the 2.5 percentage point of the rv $F_{14,17}$ is $F_{0.025,14,17} = 2.75$ and its 97.5 percentage point from Eq. (42) is

$F_{0.975,14,17} = 1/ F_{0.025,17,14} = 1/2.90 = 0.3448$, Figure 23 below clearly shows that

$$Pr(0.3448 \leq F_{14,17} \leq 2.7526) = 0.95.$$

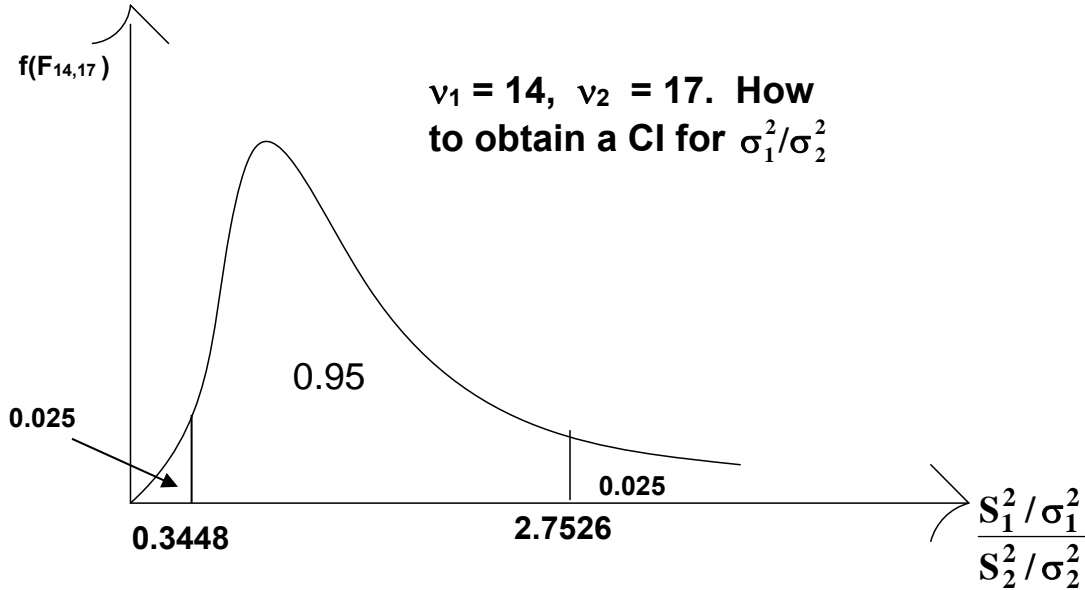$$\text{Or } Pr(0.3448 \leq \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \leq 2.7526) = 0.95$$

Figure 23. The Sampling Distribution of the rv $\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{14,\,17}$

$$Pr(0.3448 \leq \frac{S_1^2\,\sigma_2^2}{S_2^2\,\sigma_1^2} \leq 2.7526) = 0.95 \longrightarrow Pr(0.3448S_2^2/S_1^2 \leq \frac{\sigma_2^2}{\sigma_1^2} \leq 2.7526S_2^2/S_1^2) = 0.95$$

$$\longrightarrow Pr(\frac{S_1^2}{0.3448\,S_2^2} \geq \frac{\sigma_1^2}{\sigma_2^2} \geq \frac{S_1^2}{2.7526\,S_2^2}) = 0.95, \quad Pr(\frac{S_1^2}{2.7526\,S_2^2} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{0.3448\,S_2^2}) =$$

$$0.95 \quad \rightarrow \quad L(\sigma_1^2/\sigma_2^2) = \frac{S_1^2}{2.7526\,S_2^2}, \text{ and } U(\sigma_1^2/\sigma_2^2) = S_1^2/(0.3448\,S_2^2).$$

The use of the sample statistics $S_1^2 = 0.35$ and $S_2^2 = 0.40$ leads to $L(\sigma_1^2/\sigma_2^2) = 0.3179$ and

$U(\sigma_1^2/\sigma_2^2) = 2.5377 \rightarrow 0.3179 \leq \sigma_1^2/\sigma_2^2 \leq 2.5377$ at the 95 confidence level. Note that this CI

implies that the null hypothesis $H_0 : \dfrac{\sigma_1^2}{\sigma_2^2} = 1$ cannot be rejected at the 5% level because the

hypothesized value of $(\dfrac{\sigma_1^2}{\sigma_2^2})_0 \equiv 1$ is inside this CI: $0.3179 \leq \sigma_1^2/\sigma_2^2 \leq 2.5377$

Exercise 77. (a) Repeat the above Example, i.e., obtain the same CI : $0.3179 \leq$

$\sigma_1^2/\sigma_2^2 \leq 2.5377$, using the pdf of $\dfrac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2}$. (b) Obtain a 95% lower 1-sided CI for σ_y/σ_x for the

following data, where the rv X = Drying Time of White Paint and rv Y = Drying Time of Yellow

Paint. State the corresponding hypotheses H_0 and H_1. (c) Obtain a 95% lower 1-sided CI for $\mu_y -$

μ_x also for the following data and draw appropriate conclusions from your CIs.

X : 120, 112, 116, 122, 115, 110, 120, 107 minutes

Y : 126, 124, 116, 125, 109, 130, 125, 117, 129, 120 minutes.

ANS: (b) $0.6421 \leq \sigma_y / \sigma_x < \infty$, (c) $1.8536 \leq \mu_y - \mu_x < \infty$.

Exercise 78. Work Exercises 61, 62, 63, and 64 on pp. 385-6 of Devore (8e).

8. TEST OF HYPOTHESIS ABOUT THE VARIANCES OF
TWO INDEPENDENT NORMAL (De Moivre) POPULATIONS

Suppose we are interested in comparing the variability of machine X with that of machine

Y, i.e., we wish to test the null hypothesis $H_0 : \sigma_x^2 = \sigma_y^2$ versus one of the following three

alternatives: $H_1 : \sigma_x^2 \neq \sigma_y^2$, $H_1 : \sigma_x^2 < \sigma_y^2$, or $H_1 : \sigma_x^2 > \sigma_y^2$, at the nominal pre-

assigned LOS α = 5%. Note that, in general, the most prevalent alternative is $H_1 : \sigma_x^2 \neq \sigma_y^2$.

Assuming that the operations on the two machines are completely independent, then from

equation (41), the rv $F_{n_x-1, n_y-1} = \dfrac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2}$ has the Fisher's F distribution with *df* of the numerator

$\nu_1 = n_x - 1$ and that of the denominator $\nu_2 = n_y - 1$.   Equation (41) implies that under the null

hypothesis $H_0 : \sigma_x^2 = \sigma_y^2$, the sampling distribution of the statistic $S_x^2 / S_y^2 \sim F_{n_x-1, n_y-1}$.

As an example consider the data in Exercise 77 on page 136 of my notes, where X = Drying

Time of White Paint, and Y = Drying Time of Yellow Paint with $n_x = 8$ and $n_y = 10$. Our objective is

to test $H_0 : \sigma_x = \sigma_y$ VS the alternative $H_1 : \sigma_x \neq \sigma_y$ at the 5% LOS. Since the 2.5 percentage point

of $F_{7,9}$ is $F_{0.025,7,9} = 4.20$ and the 97.5 percentage point is $F_{0.975,7,9} = 1/F_{0.025,9,7} = 1/4.8232 = 0.2073$,

then our AI for the test statistic S_x^2/S_y^2 is AI = [0.2073, 4.1970] = $0.2073 \leq S_x^2/S_y^2 \leq 4.1970$ as

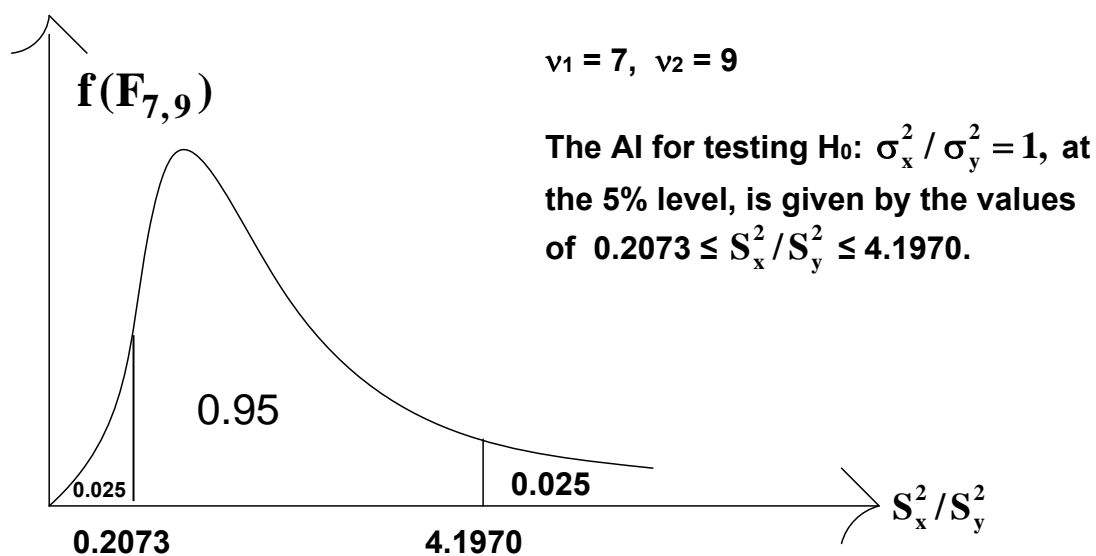depicted in Figure 24. Although Devore does not provide the 2&1/2 percentage points of the F

Figure 24. The Sampling Distribution of S_x^2 / S_y^2 Assuming H_0 is True

distribution, you may easily use Microsoft Excel to obtain FINV(0.025,7,9) = 4.197047 and

FINV(0.025,9,7) = 4. 823221. My website also lists the inverse function values of F. Because the

value of our test statistic $F_0 = S_x^2 / S_y^2$ = 28.2143/42.7666 = 0.65973 falls inside the AI = [0.2073,

4.1970], then the two data sets do not provide sufficient evidence to reject H_0, and therefore, we

cannot deduce that the variances of the two machines are significantly different at the 5% level.

By now you should be well cognizant of the fact that the *P-value* of the above test will be larger

than 5%. To compute the critical level $\hat{\alpha}$, we have to make use of our observed test statistic

$F_0 = 0.65973$. That is, even if the F distribution is skewed, the approximate *P-value* is $\hat{\alpha} = 2 \times \Pr(F_{7,9}$

$\leq 0.65973) = 2 \times 0.298484 \cong 0.597$. This *P*-value implies that if we decide to reject H_0, then the Pr

of committing a Type I error for such a decision is approximately 0.597.

Secondly, how do we approximate the Type II error Pr for the above example if H_0 is false,

say $\sigma_y = 1.5\sigma_x$? Unlike many Statistics texts, Devore does not provide OC curves for testing H_0:

$\sigma_x^2 = \sigma_y^2$, because the type II error Pr can be computed analytically. The abscissa of such OC

curves, only for the case $n_x = n_y = n$, is defined as $\lambda = \sigma_y/\sigma_x$ (or $\lambda = \sigma_x/\sigma_y$). For the example under

consideration, $\lambda = 1.50$ so that $\beta \cong 0.83$ (using $n = n_x = n_y \cong 8$ in Chart VI (o) on page 669 of

Montgomery and Runger (2003), third edition). Fortunately, the Pr of committing a type II error

for the F test can be computed analytically, as illustrated below, and most likely this is the reason

Devore does not provide the OC curves at the end of his text.

$\beta(\text{at } \lambda = 1.5) = \Pr(0.2073 < F_0 < 4.1970 \mid \sigma_y = 1.5 \ \sigma_x)$

$$= \Pr\left(0.2073 < \frac{S_x^2}{S_y^2} < 4.1970 \mid \sigma_y/\sigma_x = 1.5 \right)$$

$$= \Pr\left(0.2073 \times \frac{\sigma_y^2}{\sigma_x^2} < \frac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2} < 4.1970 \times \frac{\sigma_y^2}{\sigma_x^2} \mid \lambda = 1.5 \right)$$

$$= \Pr\left(0.2073 \times 2.25 < \frac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2} < 4.1970 \times 2.25\right)$$

$= \Pr(0.4665 < F_{7,9} < 9.4434) = \text{cdf}(9.4434) - \text{cdf}(0.4665);$ using MS. Excel

$= [1 - \text{FDIST}(9.4434, 7, 9)] - [1 - \text{FDIST}(0.4665, 7, 9)]$

$= \text{FDIST}(0.4665, 7, 9) - \text{FDIST}(9.4434, 7, 9) = 0.83634 - 0.00160$

$\beta(\text{at } \lambda = 1.5) = 0.8348$ (From MS Excel).

Exercise 79. (a) Repeat the above example using the sampling distribution of S_x^2/S_y^2.

(b) Redo the above example using the alternative $H_1: \sigma_y > \sigma_x$ at the LOS $\alpha = 5\%$. (c) Compute the

value of β for the above 2-sided test of hypothesis $H_0 : \sigma_x = \sigma_y$ if $\sigma_y = 2\sigma_x$. (d) For Exercises 72

on page 387 of Devore (8e), also test for variance equality of the two types of joints.

9. STATISTICAL INFERENCE ABOUT DIFFERENCES

BETWEEN TWO POPULATION PROPORTIONS

Devore covers this topic in section 9.4 (pp. 375-380). To illustrate the procedure, I will go thru the solution of Exercise 86 on page 389 of Devore in detail. Here our null hypothesis is $H_0 : p_1 = p_2$ VS the alternative $H_1 : p_1 \neq p_2$, or $H_0 : p_1 - p_2 = 0$ vs $H_1 : p_1 - p_2 \neq 0$, where p_1 is the proportion of eggs surviving at 11 ^oC and p_2 is the proportion of eggs surviving at 30^oC. Therefore, points unbiased estimates of p_1 and p_2 are, respectively, $\hat{p}_1 = 73/91 = 0.8022$ and $\hat{p}_2 = 102/110 = 0.9273$. We wish to ascertain if the sample difference of $\hat{p}_2 - \hat{p}_1 = 0.1251$ is significantly different from zero to warrant the rejection of $H_0 : p_2 - p_1 = 0$.

Recall that in Chapter 7 we used the fact that the SMD of $\hat{p} = X/n$ is approximately normal as long as $n > 50$ and $0.10 < p < 0.90$, with mean $E(\hat{p}) = p$ and $SE(\hat{p}) = \sqrt{pq/n}$. Since the two populations are independent, then the null SMD of $\hat{p}_2 - \hat{p}_1$ is also approximately Gaussian with $E(\hat{p}_2 - \hat{p}_1) = p_2 - p_1 = 0$ and

$$V(\hat{p}_2 - \hat{p}_1) = V(\hat{p}_2) + V(\hat{p}_1) = \frac{p_2 q_2}{n_2} + \frac{p_1 q_1}{n_1} \qquad (43)$$

as depicted in Figure 25. Since p_2 is hypothesized to be equal to p_1, then in Fig. 25 the $E(\hat{p}_2 - \hat{p}_1)$
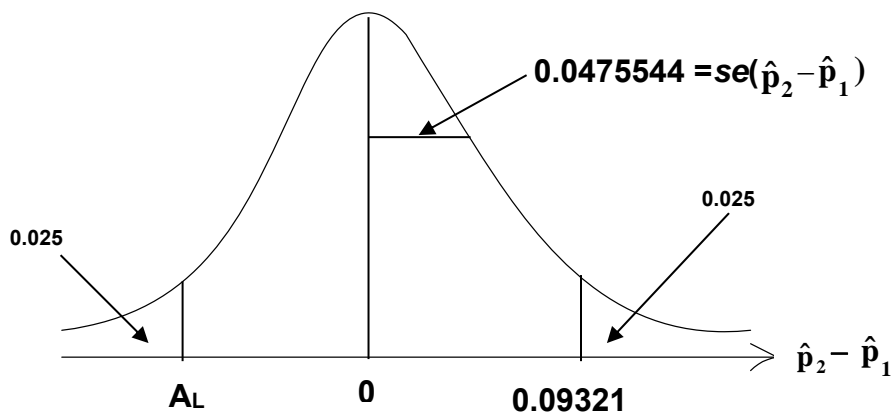

Figure 25. The Approximate Null SMD of $\hat{p}_2 - \hat{p}_1$

= 0 under H₀, and further, the estimate of the SE($\hat{p}_2 - \hat{p}_1$) will be obtained assuming that $p_2 = p_1 = p$. Thus, the pooled estimate of p from the combined samples is $\hat{p} = (73 + 102) / (91 + 110) =$

0.87065. Therefore, under H₀ $se(\hat{p}_2 - \hat{p}_1) = \sqrt{\dfrac{\hat{p}\hat{q}}{n_1} + \dfrac{\hat{p}\hat{q}}{n_2}} = \sqrt{\dfrac{0.112621}{91} + \dfrac{0.112621}{110}} = 0.0475544$.

The 5%-level AI = (− 0.09321, 0.09321) implies that we cannot reject H₀ iff − 0.09321 ≤ $\hat{p}_2 -$

$\hat{p}_1$ ≤ 0.09321. Since the value of our test statistic $\hat{p}_2 - \hat{p}_1 = 0.1251$ falls outside this last AI, the

data provide sufficient evidence at the 5% level to reject H₀ and to conclude that the two survival

rates are significantly different . The *P-value* of the test is given by $\hat{\alpha}$ = 2×Pr($\hat{p}_2 - \hat{p}_1 \geq 0.1251$) =

2×Pr(Z ≥ 2.63014) = 2 × Φ(− 2.63014) = 0.008535, which as expected is much less than α = 0.05.

Exercise 80. (a) Obtain the 95% CI for $p_2 - p_1$ of Exercise 86 on page 389 of Devore and

determine if your CI is consistent with my test above. Note that in deriving the 95% CI, you may

not assume $p_2 = p_1 = p$, and therefore the SE($\hat{p}_2 - \hat{p}_1$) will have to be estimated from (see equation

(43)) $se(\hat{p}_2 - \hat{p}_1) = \sqrt{\hat{p}_1\hat{q}_1 / n_1 + \hat{p}_2\hat{q}_2 / n_2}$. I have not obtained this CI, but Minitab gives (0.02993,

0.22023) as the answer. (b) Work Exercises 53 and 55 on pp. 381 of Devore's 8th edition.

Summary of Chapters 7, 8 and 9

We have finally come to the end of SI on parameters of one or two normal populations,

and therefore, we will provide a summary of statistics and their sampling distributions in

conducting statistical inference. In all cases except in the case of SI on proportion(s), the tacit

assumption was made that the underlying distribution (or the parental variable X) was normal (or

Laplace-Gaussian).

(i) The sampling distributions of the rvs $(\bar{x} - \mu)\sqrt{n} / \sigma$, $(\bar{d} - \mu_d)\sqrt{n} / \sigma_d$, and

$[(\bar{x} - \bar{y}) - (\mu_x - \mu_y)] / \sqrt{\sigma_x^2 / n_x + \sigma_y^2 / n_y}$, are N(0, 1), while those of $(\bar{x} - \mu_x)\sqrt{n} / S_x$,

$(\overline{d} - \mu_d)\sqrt{n} / S_d$ (for paired-samples), and $[(\overline{x} - \overline{y}) - (\mu_x - \mu_y)]/[S_p \sqrt{1/n_x + 1/n_y}]$ follow

the Student's-t with $(n-1)$, $(n-1)$, and $(n_x + n_y - 2)$ df, respectively. This last t-statistic also

requires the assumption that $\sigma_x = \sigma_y$. Further, you must be cognizant of the fact that the above six

sampling distributions are used only in conducting statistical inference(s) on one or two population

means. If the assumption $\sigma_x^2 = \sigma_y^2$ is not tenable (or rejected at the 20% level), then the two-

independent-sample t-statistic for testing H_0: $\mu_x - \mu_y = \delta_0$ is given by $t_0 = [(\overline{x} - \overline{y}) - \delta_0] / \sqrt{\dfrac{S_x^2}{n_x} + \dfrac{S_y^2}{n_y}}$,

which has an approximate Student's t-distribution with df , ν, given near the bottom of page 357

of Devore and repeated below:

$$\nu = \frac{(S_x^2 / n_x + S_y^2 / n_y)^2}{(S_x^2 / n_x)^2 / (n_x - 1) + (S_y^2 / n_y)^2 / (n_y - 1)} = \frac{\nu_y \nu_x (k^2 + 1)^2}{\nu_y k^4 + \nu_x},$$

where $k = (S_x \sqrt{n_y}) / (S_y \sqrt{n_x})$ is called the *se* ratio.

(ii) The sampling distribution of $(n-1)S^2/\sigma^2$ follows a χ_{n-1}^2 and was used to conduct

statistical inference on one process variance σ^2. However, the Chi-square distribution has

numerous other applications, a few of which will be discussed in Chapter 14.

(iii) The sampling distribution of $\dfrac{S_x^2 / \sigma_x^2}{S_y^2 / \sigma_y^2}$ follows the Fisher's F distribution with df of the

numerator $\nu_1 = n_x - 1$ and that of the denominator $\nu_2 = n_y - 1$. Thus far, the F distribution was

used only to conduct statistical inference on σ_x^2 / σ_y^2. Under H_0: $\sigma_x^2 = \sigma_y^2$, the SMD of S_y^2 / S_x^2 ~

$F_{n_y - 1, n_x - 1}$.

In Chapters 10, 12, and 13 you will, however, learn that the Fisher's F distribution has

extensive applications in almost all facets of sciences and engineering. For example, you will study

that the F-statistic is used to test the null hypothesis H_0 : $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_a$ VS the alternative

that at least two treatment means differ significantly. However, the F-test in this case will always

be right-tailed even if the alternative is not one-sided.

(iv) The sampling distribution of sample proportion, $\hat{p} = X/n$ where $X \sim B(n, p)$,

is approximately normal with population mean p and its SE given by SE($\hat{p}$) = $\sqrt{pq/n}$. The

normal approximation requires that n > 50, 0.10 < p < 0.90, and n×p > 10. To be on the

conservative side, it will be best to have np > 15. When np < 10, the Poisson approximation to the

binomial is appropriate; the region 10 < np < 15 is gray, and for which it is not clear-cut as to which

approximation (Poisson or normal) to the B(n, p) is more accurate.

Further, the SMD of $\hat{p}_1 - \hat{p}_2$ is also approximately normal with E($\hat{p}_1 - \hat{p}_2$) =

$p_1 - p_2$ and SE($\hat{p}_1 - \hat{p}_2$) = $\sqrt{\dfrac{p_1 q_1}{n_1} + \dfrac{p_2 q_2}{n_2}}$. Only under H_0: $p_1 = p_2 = p$ this last SE reduces to

SE($\hat{p}_1 - \hat{p}_2 | H_0$) = $\sqrt{\dfrac{pq}{n_1} + \dfrac{pq}{n_2}}$.

Finally, you should always bear in mind that all CIs are simply tests of hypotheses in

disguise! Therefore, to be consistent, the lower one-sided CI : $\theta_L \le \theta < \infty$ always corresponds to a

right-tailed test on the parameter θ, and vice versa for an upper one-sided CI. Thus, we must

conclude that if a population parameter, θ , is of the STB type, than the corresponding test must be

left-tailed from an Engineering standpoint, and vice a versa for an LTB type parameter. Further, the

student must not confuse a decision (or acceptance) interval (DI) for a test-statistic with the

corresponding CI for the parameter under the null hypothesis. The former are values of the test-

statistic that lead to a decision regarding H_0, while the latter consists of parameter values at a given

confidence level.