Statistical Bulletin
Reliability & Variation Research

DETROIT RESEARCH INSTITUTE
15224 KERCHEVAL AVE. • GROSSE POINTE PARK, MICHIGAN 48230 • (313) 499-1150

Leonard G. Johnson, EDITOR

Vol. 2 May, 1972
Bulletin 2 Page 1

SAMPLE COMPARISON VS. HYPOTHESIS TESTING

In comparing two samples we calculate the confidence C that one of the samples represents a population which is superior to the population represented by the other sample.

In testing the hypothesis that a given sample belongs to a population superior to some standard population we specify so-called α and β errors, and determine whether under the (α, β) pair chosen we should accept or reject the hypothesis that the sample belongs to the standard population.

The question which now can be asked is "How are these two statistical techniques related (i.e., sample comparison and hypothesis testing)?" More specifically, we can ask "What is the relationship between α, β, and C?"

A comparison of two samples is generally done by comparing a particular type of population parameter or quantile level in the two samples. The parameter chosen could be the MEAN in the case of NORMAL DISTRIBUTION, or CHARACTERISTIC VALUES in the case of WEIBULL DISTRIBUTIONS.

The corresponding hypothesis which is tested is then the one which states that the MEAN of an observed sample belongs in POPULATION_1 with MEAN_1, or that the characteristic value θ of an observed sample belongs in POPULATION_1 with $\text{CHARACTERISTIC VALUE}_1 = \theta_1$ (θ_1 and θ_2 being the characteristic values of the two possible populations from which the sample could come).

In order to make a comparison of $PARAMETER_2$ vs. $PARAMETER_1$ we need distribution functions for both $PARAMETER_1$ (say, the estimated MEAN of $POPULATION_1$) and $PARAMETER_2$ (the estimated MEAN of $POPULATION_2$)

Once these distribution functions are known we construct an INTERFERENCE DIAGRAM, as in FIGURE 1.

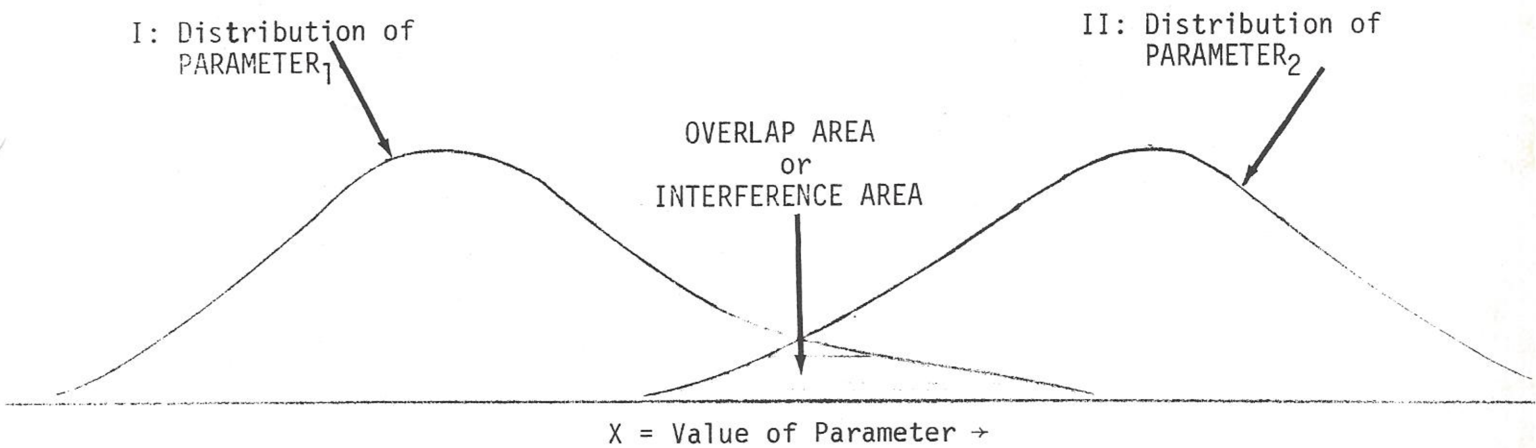

FIGURE 1

From FIGURE 1 it is possible to calculate the PROBABILITY THAT A RANDOM ELEMENT FROM II EXCEEDS A RANDOM ELEMENT FROM I. This is then defined to be the CONFIDENCE that $PARAMETER_2 > PARAMETER_1$.

Let the PDF (Probability Density Function) of I be $f_1(X)$.

Let $f_2(X)$ = PDF of II

Let $F_1(X)$ = CDF (Cumulative Distribution Function) of I

Let $F_2(X)$ = CDF of II

Then

$$C = \int_{-\infty}^{+\infty} F_1(X)\, f_2(X)\, dX \qquad [1]$$

According to formula [1] the confidence C increases as the OVERLAP AREA between $f_1(X)$ and $f_2(X)$ is made smaller, until when there is no overlap, the confidence becomes 1 (i.e., 100% or certainty) that whatever is selected in II will surely exceed whatever is selected in I. On the other hand, if $f_1(X)$ and $f_2(X)$ are identical (100% overlap), it follows that there is a 50-50 chance (i.e., 50% confidence or C = .5) that the selection from II will exceed the selection from I.

HYPOTHESIS TESTING (GENERAL CASE)

Once $f_1(X)$ and $f_2(X)$ are known it follows that all possible (α, β) pairs for the hypothesis test (that a selection comes from I or II) are completely specified. In fact, we can write for any selected CRITICAL VALUE X_0 that

$$\begin{cases} \alpha = 1 - F_1(X_o) \\ \beta = F_2(X_o) \end{cases}$$

With this (α, β) pair as a criterion we then conclude

(1) If any selection $X < X_o$ accept the hypothesis H_1 that X belongs to I.

(2) If any selection $X > X_o$ accept the hypothesis H_2 that X belongs to II.

MAKING α AND β EQUAL (SPECIAL CASE I)

For the special case when the (α, β) criterion has $\beta = \alpha$ we have the situation shown in FIGURE 2.

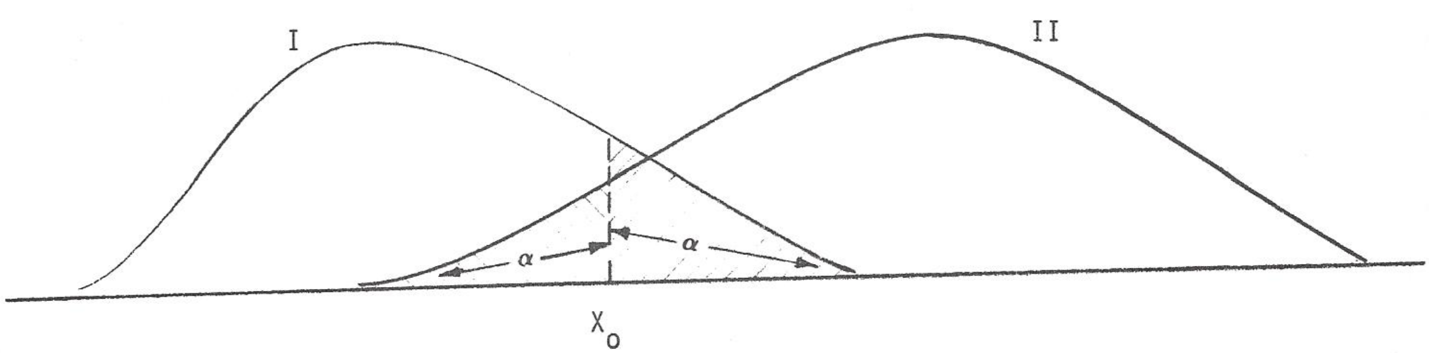

FIGURE 2

In FIGURE 2 the critical value X_0 is at a location such that

AREA OF I TO RIGHT OF X_0 = AREA OF II TO LEFT OF X_0 = α.

This is equivalent to having two non-overlapping confidence bands, each of width $W = 1 - 2\alpha$. For any pair of unimodal distributions it has been shown that when their neighboring confidence bands, each of width W just touch, the confidence that a random slection from II exceeds a random selection from I can be accurately calculated by using the formula

$$C = \frac{\text{LOG}\left(\frac{1-W}{2}\right)}{\text{LOG}\left(\frac{1-W}{2}\right) + \text{LOG}\left(\frac{1+W}{2}\right)}$$

[THIS FORMULA IS DERIVED FROM [1] BY TAKING NEIGHBORING WEIBULLS WITH THE SAME SLOPE AND MINIMUM VALUE BUT DIFFERENT CHARACTERISTIC VALUES.]

Now, in FIGURE 2, we have

$$W = 1 - 2\alpha$$

or $\frac{1-W}{2} = \alpha$ and $\frac{1+W}{2} = 1 - \alpha$

Hence, for the situation in FIGURE 2,

$$C = \frac{\text{LOG } \alpha}{\text{LOG } \alpha + \text{LOG } (1-\alpha)} \qquad [2]$$

We can take this formula [2] as the analytical relation between C and the pair (α, β) whenever $\beta = \alpha$.

Another empirical relation not involving logarithms is

$$C = 1 - \frac{5}{12}\alpha - \frac{7}{6}\alpha^2 \qquad [3]$$

(Only for $\beta = \alpha$)

<u>MAKING β = .5 (SPECIAL CASE II)</u>

For the special case in which $\beta = .5$ we have the situation illustrated in FIGURE 3. (NOTE: We are dealing only with unimodal distributions.)

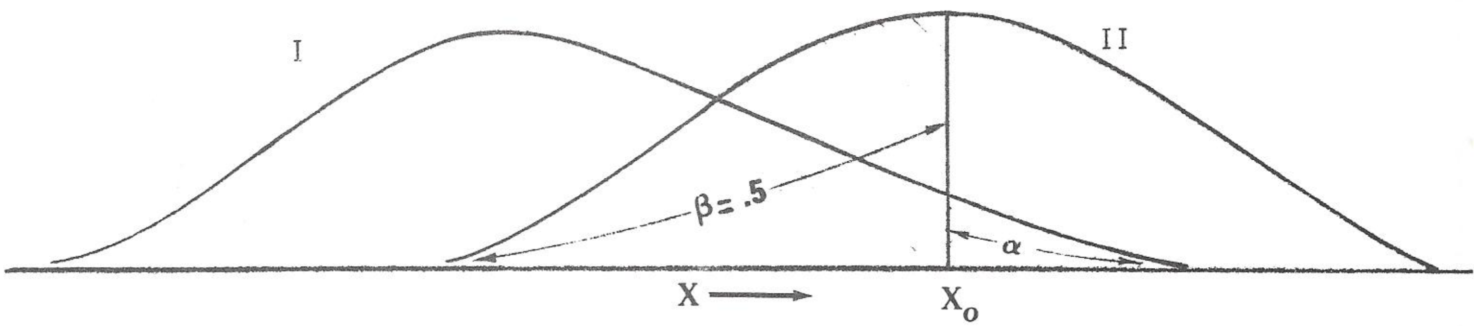

FIGURE 3

What is the confidence C that a random element from II will exceed a random element from I for the situation shown in FIGURE 3? The answer to this question depends on the ratio of standard deviations

$$\rho = \frac{\sigma_2}{\sigma_1} \text{, where}$$

σ_1 = standard deviation of I

σ_2 = standard deviation of II

In general, for unimodal distributions, C can be closely approximated by

$$C = N\left(\frac{t_{(1-\alpha)}}{\sqrt{1+\rho^2}}\right)$$

[THIS FORMULA IS MORE ACCURATE IF BOTH I AND II HAVE BEEN NORMALIZED BY A COMMON MONOTONIC TRANSFORMATION ON X.]

$t_{(1-\alpha)}$ = $(1-\alpha)$ t-score in a NORMAL DISTRIBUTION.

$N_{(v)}$ = Cumulative area under a NORMAL curve from $t = -\infty$ to $t = v$.

Note that when $\rho = 0$, i.e., when $\sigma_2 = 0$ the formula

$$C = N\left(\frac{t_{(1-\alpha)}}{\sqrt{1+\rho^2}}\right) \quad \text{becomes}$$

$$C = N\,(t_{(1-\alpha)}) = 1 - \alpha \quad \text{(By definition)}$$

At the other extreme, when $\sigma_2 = \infty$, we obtain

$$C = N\left(\frac{t_{(1-\alpha)}}{\infty}\right) = N\,(0) = .5$$

QUESTION: If $\beta = .5$ has $\alpha = \alpha_{.5}$, what value of α would make $\beta = \alpha$?

ANSWER

Let α_α = α-level when $\beta = \alpha$.

Let $\alpha_{.5}$ = α-level when $\beta = .5$

Then

$$\alpha_\alpha = \frac{-5 + \sqrt{697 - 672\, N\left(\dfrac{t_{(1-\alpha_{.5})}}{\sqrt{1+\rho^2}}\right)}}{28}$$

[THIS WAS DERIVED BY SOLVING [3] FOR α IN TERMS OF C, AND REPLACING C BY $N\left(\dfrac{t_{(1-\alpha_{.5})}}{\sqrt{1+\rho^2}}\right)$]

THE MOST GENERAL RELATION BETWEEN α, β, AND C

For any critical value X we define

$$\alpha(X) = 1 - F_1(X) \qquad \text{(a)}$$

$$\beta(X) = F_2(X) \qquad \text{(b)}$$

Furthermore, the general definition of C is

$$C = \int_{-\infty}^{+\infty} F_1(X)\ f_2(X)\ dX \qquad \text{(c)}$$

From (a) : $F_1(X) = 1 - \alpha(X)$

From (b) : $f_2(X) = F_2'(X) = \beta'(X)\ ;\ \left(\beta'(X) = \frac{d\beta}{dX}\right)$

Substituting these functions into (c) yields the most general relation between α, β, and C. This is

$$C = \int_{-\infty}^{+\infty} [1 - \alpha(X)]\ \beta'(X)\ dX$$

This general formula indicates that it is necessary to know fully the distribution functions of both α and β over all possible critical values in order to relate α and β to the single confidence number C for the confidence that II is better than I.

FINDING THE OPTIMUM (α, β) PAIR FOR A DICHOTOMOUS HYPOTHESIS TEST

A dichotomous hypothesis test is one involving two possible choices of decision in assigning a population to an observed sample parameter X. In other words, we have the situation illustrated in FIGURE 4, with two choices.

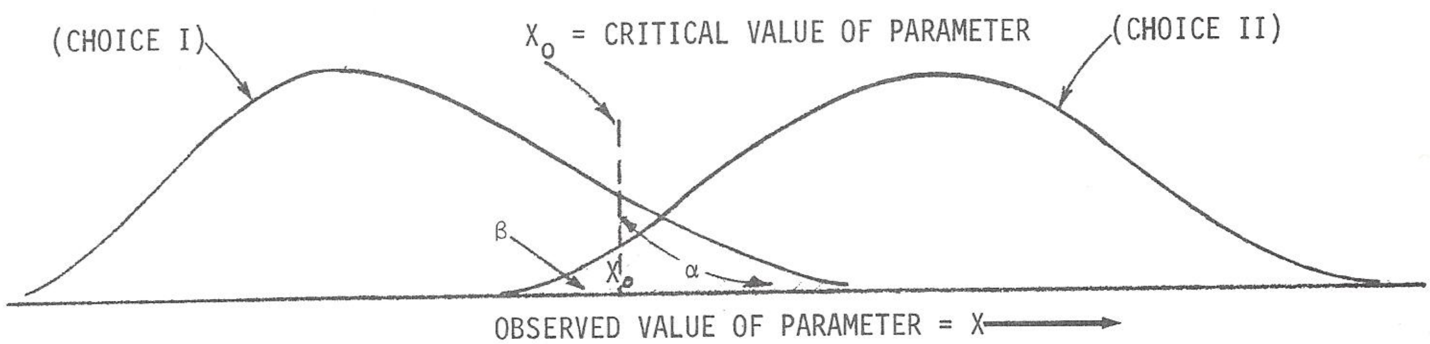

FIGURE 4

QUESTION: What is the optimum (α, β) pair, or, equivalently, the optimum critical value X_0 for this hypothesis testing situation?

SOLUTION

The two possible types of errors are

(1) The α-error, i.e., THE PROBABILITY OF MAKING CHOICE II WHEN WE SHOULD MAKE CHOICE I.

(2) The β-error, i.e., THE PROBABILITY OF MAKING CHOICE I WHEN WE SHOULD MAKE CHOICE II.

These two errors are mutually exclusive. Hence, we can write

PROB. [of wrong choice] = PROB. [α-error or β-error] = $\alpha + \beta$

From the standpoint of possible errors in our choice the optimum (α, β) pair is the one which will minimize the TOTAL PROBABILITY OF MAKING A WRONG CHOICE. In other words, we should choose α and β such that $S(X) = \alpha(X) + \beta(X)$ is a minimum.

Now, $\alpha(X) = 1 - F_1(X)$

and $\beta(X) = F_2(X)$

$$\therefore \quad S(X) = 1 - F_1(X) + F_2(X)$$

To minimize $S(X)$, find $\frac{dS}{dX}$ and equate it to zero:

$$\frac{dS}{dX} = -f_1(X) + f_2(X) = 0$$

$$\therefore \quad f_1(X) = f_2(X)$$

This last result tells us that the optimum critical value x is the abscissa at which the two frequency curves intersect (i.e., the X at which the ordinates are equal.)

Hence, $X_{opt.}$ = ABSCISSA OF THE INTERSECTION POINT

$\therefore\ \alpha_{opt.}$ = AREA IN I TO THE RIGHT OF INTERSECTION POINT

$\beta_{opt.}$ = AREA IN II TO LEFT OF INTERSECTION POINT

In summary, THE PROBABILITY OF MAKING A WRONG CHOICE is minimized when we make the critical value to be the abscissa of the intersection point.