Statistical Research Report
Institute of Mathematics
University of Oslo

No 5
December 1968

A NOTE ON THE MINIMUM VARIANCE UNBIASED ESTIMATE OF THE FRACTION OF A NORMAL DISTRIBUTION BELOW A SPECIFICATION LIMIT

by

Visiting professor
from University of Wyoming
William C. Guenther

A NOTE ON THE MINIMUM VARIANCE UNBIASED ESTIMATE OF THE FRACTION OF A NORMAL DISTRIBUTION BELOW A SPECIFICATION LIMIT

by

WILLIAM C. GUENTHER
UNIVERSITY OF WYOMING

1. INTRODUCTION

Let $F(x)$ be the distribution function of a normal random variable X with unknown mean and unknown variance and let x_0 be a given specification limit. The minimum variance unbiased estimate of $F(x_0)$ has appeared in the literature in two different forms. In this note we show that the two are equivalent, comment on the relative merits of each, and discuss some methods of evaluating the estimate. (We note that if we assume that all derivations have been made without error, the equivalence is already established since the minimum variance unbiased estimate is unique.)

2. TWO EXPRESSIONS FOR THE ESTIMATE

In 1952 Bowker and Goode [3] gave without derivation the estimete of $F(x_0)$ as

$$\hat{p} = \begin{cases} 0 & kv \leqq -1 \\ \Pr(Y < \frac{1}{2} + \frac{1}{2}kv) & -1 < kv < 1 \\ 1 & kv \geqq 1 \end{cases} \qquad (2.1)$$

where $v = (x_0 - \bar{x})/s$, $k = \sqrt{n}/(n-1)$, Y has the beta density function

$$g(y) = \frac{1}{\beta(\frac{n-2}{2}, \frac{n-2}{2})} y^{\frac{n-2}{2}-1} (1-y)^{\frac{n-2}{2}-1} , \quad 0<y<1 \qquad (2.2)$$

and $\bar{x}$, s^2 are the usual unbiased estimates of the mean and variance. In 1955 Liberman and Resnikoff [7] gave a derivation yielding the same result and presented an extensive table giving $\hat{p}$ if $v < 0$, $1-\hat{p}$ for $v > 0$ for $|v| = .1(.1).3(.01)3.90$ and $n = 3,4,5,7,10(5)40,50,75,100,150,200$. This table was later incorporated in Military Standard 414 [8] and can also be found in a textbook by Bowker and Lieberman [4]. Of course, (2.1) can also be evaluated from the Pearson [11] table of the incomplete beta function which contains entries for $n - 2 = 1(1)22(2)100$. However, linear interpolation with the latter table is not as satisfactory as with the former table which contains more entries for each n .

In 1961 Barton [1] derived the estimate in another form which was later reproduced in 1964 by Basu [2] and in 1965 by Folks, Pierce and Stewart [5]. The latter authors showed that the Barton estimate may be evaluated from the cumulative t-distribution. Their result may be written as

$$\hat{p} = \begin{cases} 0 & kv \leqq -1 \\ \Pr\left(T < \dfrac{kv(n-2)^{\frac{1}{2}}}{(1-k^2v^2)^{\frac{1}{2}}}\right) & -1 < kv < 1 \\ 1 & kv \geqq 1 \end{cases} \tag{2.3}$$

where T has a t-distribution with $n-2$ degrees of freedom. Apparently none of the authors of these three papers was aware of the results mentioned in the previous paragraph.

3. RELATIONSHIP BETWEEN THE t AND F DISTRIBUTIONS WHICH YIELDS EQUIVALENCE OF THE ESTIMATES

If Y has the density (2.2) it is a routine exercise to show that $F = (1-Y)/Y$ has an F-distribution with $n-2$, $n-2$ degrees of freedom and that the middle line of (2.1) can be written as

$$\hat{p} = \Pr\left(F > \frac{1-kv}{1+kv}\right) = \Pr\left(F < \frac{1+kv}{1-kv}\right) \tag{3.1}$$

If T has a t-distribution with $n-2$ degrees of freedom and F is defined by

$$T = \frac{\sqrt{n-2}\,(1-F)}{2\sqrt{F}} \tag{3.2}$$

then F has an F-distribution with $n-2$, $n-2$ degrees of freedom. This follows by the change of variable technique and the relationship

$$\Gamma(r) = \frac{2^{r-1}\,\Gamma(\frac{r+1}{2})\Gamma(\frac{r}{2})}{\Gamma(\frac{1}{2})} \tag{3.3}$$

where in our problem $r = n-2$.

Using the transformation (3.2) on (3.1) yields the middle line of (2.3). Hence the equivalence of the two estimates is established.

Incidently, we have shown how transformation (3.2) relates the t-distribution with r degrees of freedom and the F-distribution with r, r degrees of freedom. (It is a standard textbook problem to show how $F = T^2$ relates the t-distribution and the F-distribution with 1, r degrees of freedom.) Hence, if we have a good cumulative t-table, we also have a good cumulative F-table for the case of equal degrees of freedom.

4. SOME OBSERVATIONS AND COMPARISONS

The obvious advantage of using (2.3) rather than (2.1) is that tables of the t-distribution are readily available and appear in most textbooks. Two good tables of this type appear in publications by Owen [9, pp. 28-30] and Hald [6, p. 39]. These tables, designed for testing hypotheses and obtaining confidence intervals, are not well adapted to evaluating (2.3) since it is necessary to interoplate over relatively long intervals of probability levels. Consequently, estimates so obtained are usually not accurate to more than two decimal places (if that).

As an alternative to the standard type t-table we could use the Pearson and Hartley [10, pp. 138-140] table of the cumulative t-distribution which contains entries for $n-2 = 1(1)24, 30, 40, 60, 120, \infty$. Then, however, we might just as well use the Lieberman-Resnikoff [7] table (except for the small n missing from the table) since the latter table contains considerably more entries.

The Lieberman and Resnikoff table is actually a good table of the cumulative t-distribution (in that it has a large number of entries) and the cumulative F-distribution with equal degrees of freedom. When entered with n and

$$v = \frac{(n-1)|t_0|}{[n(n-2+t_0^2)]^{\frac{1}{2}}} \tag{4.1}$$

the table gives for $n-2$ degrees of freedom

$$\begin{array}{ll} Pr(T < t_0) & \text{if } t_0 < 0 \\ Pr(T > t_0) & \text{if } t_0 > 0 \end{array} \tag{4.2}$$

When entered with n and

$$v = \frac{|1-F_0|(n-1)}{(1+F_0)\sqrt{n}} \tag{4.3}$$

it gives for $n-2, n-2$ degrees of freedom

$$\begin{array}{ll} Pr(F < F_0) & \text{if } F_0 < 1 \\ Pr(F > F_0) & \text{if } F_0 > 1 \end{array} \tag{4.4}$$

5. A NUMERICAL EXAMPLE

<u>Example 5.1</u>

As a numerical example illustrating the evaluation of $\hat{p}$ we will use the same one considered by Folks, Pierce, and Stewart [5,p45]. They had $\bar{x} = 14$, $x_0 = 18$, $n = 5$, $s^2 = 10$.

(Like those authors we will ignore the problem of significant figures.)

We enter the Lieberman and Resnikoff [7] table with $v = 4/\sqrt{10} = 1.2648$ and $n = 5$. We obtain with

$$v = 1.26 \qquad 1-\hat{p} = .0921$$

$$v = 1.2648 \qquad 1-\hat{p} = .0909 \quad \text{by linear interpolation}$$

$$v = 1.27 \qquad 1-\hat{p} = .0896$$

Hence $\hat{p} = .9091$.

To use a t-table we need (2.3) which becomes $Pr(T < 1.732)$. With the standard type table we find (as the above authors did)

$$Pr(T < 1.638) = .90$$

$$\hat{p} = Pr(T < 1.732) = .9066 \quad \text{by linear interpolation}$$

$$Pr(T < 2.353) = .95$$

From the Pearson and Hartley [10] table we find

$$Pr(T < 1.7) = .90615$$

$$\hat{p} = Pr(T < 1.732) = .9090 \quad \text{by linear interpolation}$$

$$Pr(T < 1.8) = .91516$$

To use the incomplete beta table of Pearson [11] we need (2.1) which becomes $Pr(Y < .8535)$. We get

$$Pr(Y < .85) = .9059398$$

$$\hat{p} = Pr(Y < .8535) = .9091 \quad \text{by linear interpolation}$$

$$Pr(Y < .86) = .9149054$$

REFERENCES

1. BARTON, D.E., 1961. Unbiased estimation of a set of probabilities. Biometrika, 48, 227-229.

2. BSAU, A.P., 1964. Estimates of reliability for some distributions useful in life testing. Technometrics, 6, 215-219.

3. BOWKER, ALBERT H. and GOODE, HENRY P., 1952. Sampling Inspection for Variables. McGraw-Hill, New York.

4. BOWKER, ALBERT H. and LIEBERMAN, GERALD J., 1959. Engineering Statistics. Prentice-Hall, Englewood Cliffs, New Jersey.

5. FOLKS, JOHN LEROY, PIERCE, DONALD A., and STEWART, CHARLES, 1965. Estimating the fraction of acceptable product. Technometrics, 7, 43-50.

6. HALD, A., 1952. Statistical Tables and Formulas. Wiley, New York.

7. LIEBERMAN, GERALD J. and RESNIKOFF, GEORGE J., 1955. Sampling plans for inspection by variables. Journal of the American Statistical Association, 50-1, 457-516.

8. Military Standard 414, 1957. Superintendent of Documents, U.S. Government Printing office, Washington, D.C.

9. OWEN, D.B., 1962. Handbook of Statistical Tables. Addison-Wesley, Reading, Mass.

10. PEARSON, E.S. and HARTLEY, H.O., 1966. Biometrika Tables for Statisticions (3rd ed.). Cambridge University Press, Cambridge.

11. PEARSON, KARL, 1968. Tables of the Incomplete Beta Function, (2nd ed.). Cambridge University Press, Cambridge.