NOTES

This section is devoted to brief research and expository articles, notes on methodology and other short items.

COMPUTATION OF FACTORS FOR TOLERANCE LIMITS ON A NORMAL DISTRIBUTION WHEN THE SAMPLE IS LARGE[1]

By Albert H. Bowker

Columbia University

In their paper [1], Wald and Wolfowitz discuss the problem of finding tolerance limits of the form $\bar{x} \pm \lambda s$ for a normal distribution. They propose the following large sample formula for λ which appears to be satisfactory for all practical purposes for $N \geqq 2$!

$$\lambda = \sqrt{\frac{n}{\chi_\beta^2}}\, r\left(\frac{1}{\sqrt{N}}, \gamma\right) \tag{1}$$

where N is the number of observations ($n = N - 1$), γ is the tolerance coefficient, β is the confidence coefficient, r is defined by

$$\frac{1}{\sqrt{2\pi}} \int_{(1/\sqrt{N})-r}^{(1/\sqrt{N})+r} e^{-t^2/2}\, dt = \gamma$$

and χ_β^2 has the property that $P(\chi^2 > \chi_\beta^2) = \beta$ for n degrees of freedom. To compute λ, tables [2] or known approximations [3] for χ_β^2 are customarily used, but the computation of r, even for large N, is tedious, involving an iterative procedure. The purpose of this note is to obtain an expansion of r in terms of $1/\sqrt{N}$ and to combine this expansion with a known one for χ_β^2 to obtain an asymptotic formula for λ.

To derive a large sample formula for r, consider the function

$$f(x, y) = \frac{1}{\sqrt{2\pi}} \int_{x-y}^{x+y} e^{-t^2/2}\, dt - \gamma = 0 \tag{2}$$

where for convenience $\dfrac{1}{\sqrt{N}}$ and r are replaced by x and y. It is desired to express y as a power series in x. Let y_0 be defined by $f(0,y_0) = 0$. Since $f(x,y)$ is a con-

[1] This paper reports work done in the Statistical Research Group, Division of War Research, Columbia University, under Contract OEMsr-618 with the Applied Mathematics Panel, National Defense Research Committee, Office of Scientific Research and Development. The work was first reported in an unpublished memorandum, "Computation of Factors for Tolerance Limits when the Sample is Large" (SRG No. 559, September 24, 1945). A brief account of the application of tolerance limits, including tables, will be published in *Techniques of Statistical Analysis* described in the footnote on page 217.

TABLE 1

Comparative Values of Exact and Approximate λ

β \ N	γ	50 Exact	50 Approximate	50 Difference	100 Exact	100 Approximate	100 Difference	160 Exact	160 Approximate	160 Difference
.75	.75	1.25480	1.25147	.00333	1.21808	1.21698	.00110	1.20161	1.20108	.00053
	.95	2.13774	2.13226	.00548	2.07533	2.07349	.00184	2.04728	2.04639	.00089
	.999	3.58821	3.57979	.00842	3.48401	3.48112	.00289	3.43704	3.43563	.00141
.95	.75	1.39621	1.38467	.01154	1.31050	1.30670	.00380	1.27204	1.27022	.00182
	.95	2.37866	2.35921	.01945	2.23279	2.22635	.00644	2.16728	2.16420	.00308
	.999	3.99259	3.96080	.03179	3.74835	3.73776	.01059	3.63850	3.63341	.00509
.99	.75	1.51184	1.48901	.02283	1.38251	1.37511	.00740	1.32566	1.32215	.00351
	.95	2.57565	2.53698	.03867	2.35546	2.34290	.01256	2.25865	2.25268	.00597
	.999	4.32325	4.25926	.06399	3.95429	3.93343	.02086	3.79189	3.78196	.00993

Comparative Values of Exact and Approximate λ—Continued

β \ N	γ	500 Exact	500 Approximate	500 Difference	800 Exact	800 Approximate	800 Difference	1000 Exact	1000 Approximate	1000 Difference
.75	.75	1.17733	1.17724	.00009	1.17126	1.17122	.00004	1.16891	1.16888	.00003
	.95	2.00593	2.00578	.00015	1.99559	1.99552	.00007	1.99158	1.99153	.00005
	.999	3.36769	3.36744	.00025	3.35034	3.35022	.00012	3.34361	3.34352	.00009
.95	.75	1.21501	1.21470	.00031	1.20062	1.20047	.00015	1.19502	1.19491	.00011
	.95	2.07013	2.06960	.00053	2.04562	2.04536	.00026	2.03608	2.03589	.00019
	.999	3.47547	3.47459	.00088	3.43433	3.43390	.00043	3.41831	3.41800	.00031
.99	.75	1.24268	1.24208	.00060	1.22198	1.22169	.00029	1.21395	1.21374	.00021
	.95	2.11727	2.11626	.00101	2.08201	2.08152	.00049	2.06832	2.06797	.00035
	.999	3.55462	3.55292	.00170	3.49543	3.49460	.00083	3.47244	3.47186	.00058

tinuous function of x and y, and since $\left.\dfrac{\partial f}{\partial y}\right|_{\substack{x=0\\y=y_0}} \neq 0$, the function $y(x)$ defined implicitly by (2) is continuous. Since $\dfrac{dy}{dx} = -\dfrac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} = \tanh xy$, the higher derivatives of $y(x)$ exist and are continuous and $y(x)$ permits of a finite Taylor's expansion. The coefficients of odd powers of x drop out and we obtain

$$y = y_0 + \frac{y_0}{2!}x^2 + \frac{3y_0 - 2y_0^3}{4!}x^4 + 0(x^6),$$

or returning to the original notation and retaining terms in $1/N$,

$$r \sim r_\infty\left(1 + \frac{1}{2N}\right). \tag{3}$$

If x_p is defined by $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x_p} e^{-t^2/2}\,dt = p$ we know from [3] that

$$\frac{\chi_\beta^2}{n} \sim 1 + \frac{\sqrt{2}\,x_{1-\beta}}{\sqrt{n}} + \frac{2}{3}\frac{x_{1-\beta}^2 - 1}{n}. \tag{4}$$

Proceeding formally and retaining terms in $1/N$ we obtain

$$\left(\frac{n}{\chi_\beta^2}\right)^{\frac{1}{2}} = \left(1 - \frac{x_{1-\beta}}{\sqrt{2N}} + \frac{4 + 5x_{1-\beta}^2}{12N}\right)$$

and multiplying by the expression for r given by equation (3) we find the desired expansion for λ.

$$\lambda \sim r_\infty\left(1 - \frac{x_{1-\beta}}{\sqrt{2N}} + \frac{5x_{1-\beta}^2 + 10}{12N}\right). \tag{5}$$

Recall that both r_∞ and $x_{1-\beta}$ are readily obtainable from tables of the normal curve; in fact, r_∞ is defined by

$$\frac{1}{\sqrt{2\pi}}\int_{-r_\infty}^{r_\infty} e^{-t^2/2}\,dt = \gamma \text{ and } x_{1-\beta} \text{ is defined by } \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x_{1-\beta}} e^{-t^2/2}\,dt = 1 - \beta.$$

A comparative table of approximate and exact values of λ is given in Table 1. From the table we see that for $N \geqq 800$ the error is less than 1 in the 4th significant figure, and for $N \geqq 160$ the error is less than 1 in the 3rd significant figure within the limits of β and γ considered. The approximation will be less exact for higher values of β and γ.

REFERENCES

[1] A. Wald and J. Wolfowitz, "Tolerance limits for a normal distribution," *Annals of Math. Stat.*, Vol. 17 (1946), pp. 208–215.

[2] C. M. Thompson, "Tables of percentage points of the χ^2 distribution," *Biometrika*, Vol. 32 (1941–42), pp. 188–9.

[3] Henry Goldberg and Harriet Levine, "Approximate formulas for the percentage points and normalization of t and χ^2," *Annals of Math. Stat.*, Vol. 17 (1946).

THE PROBABILITY DISTRIBUTION OF THE MEASURE OF A RANDOM LINEAR SET

By David F. Votaw, Jr.

Naval Ordnance Laboratory

1. Introduction. Consider a random sample $0_n(x_1, \cdots, x_n)$ of n values of a one-dimensional random variable x having cumulative distribution function $F(x)$. Let there be associated with each x an interval of length D centered at x