APPLIED MATHEMATICS AND STATISTICS LABORATORY
STANFORD UNIVERSITY
CALIFORNIA

TABLES OF THE NON-CENTRAL t-DISTRIBUTION

DENSITY FUNCTION, CUMULATIVE DISTRIBUTION FUNCTION AND PERCENTAGE POINTS

By

GEORGE J. RESNIKOFF

and

GERALD J. LIEBERMAN

TECHNICAL REPORT NO. 32

April 1, 1957

PREPARED UNDER CONTRACT N6onr-25126
(NR-042-002)
FOR
OFFICE OF NAVAL RESEARCH

TABLES OF THE NON-CENTRAL t-DISTRIBUTION

DENSITY FUNCTION, CUMULATIVE DISTRIBUTION FUNCTION
AND PERCENTAGE POINTS

By

George J. Resnikoff

and

Gerald J. Lieberman

TECHNICAL REPORT NO. 32
APRIL 1, 1957

PREPARED UNDER CONTRACT N6onr-25126
(NR-042-002)

GERALD J. LIEBERMAN, PROJECT DIRECTOR
FOR
OFFICE OF NAVAL RESEARCH

APPLIED MATHEMATICS AND STATISTICS LABORATORY
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Stanford University Press, Stanford, California
London: Oxford University Press

Printed in the United States of America
by Stanford University Press

Library of Congress Catalog Card Number: 57-7832

Preface

The computation of the tables contained herein was undertaken to facilitate the construction of variables sampling plans in industrial acceptance sampling under Office of Naval Research contract N6onr-25126.

Because it is believed that the tables have numerous other applications to both practical and theoretical statistics, it was decided to make them generally available by publishing them in book form.

The authors wish to acknowledge their indebtedness to the following persons: Albert H. Bowker, for originally suggesting the computation of the tables and for encouragement and advice during the course of the work; Joseph Carter, for assistance in much of the programming; Gladys R. Garabedian, for invaluable assistance in preparing the text material; and last but not least, Herbert Solomon, without whose inspiration the work could never have been initiated.

The authors imply no responsibility on the part of these people for any inacccuracies that may exist in the tables.

January 25, 1957

GEORGE J. RESNIKOFF
Research Associate, Applied Mathematics and Statistics Laboratory, Stanford University

GERALD J. LIEBERMAN
Associate Professor of Statistics and Industrial Engineering, Stanford University

Contents

TABLES OF THE NON-CENTRAL t-DISTRIBUTION

TABLES OF THE NON-CENTRAL t-DISTRIBUTION: DENSITY FUNCTION, CUMULATIVE DISTRIBUTION FUNCTION AND PERCENTAGE POINTS

1. Introduction

Let z be a random variable distributed normally about zero with unit standard deviation, and let w be a random variable distributed independently of z as χ^2/f with f degrees of freedom. If t is defined by

$$t = \frac{z+\delta}{\sqrt{w}}$$

where δ is some constant, then t is said to have the non-central t-distribution with f degrees of freedom and non-centrality parameter δ.

The probability density of t is given by

$$h(f,\delta,t) = \frac{f!}{2^{\frac{f-1}{2}}\,\Gamma(\frac{f}{2})\sqrt{\pi f}}\, e^{-\frac{1}{2}\frac{f\delta^2}{f+t^2}} \left(\frac{f}{f+t^2}\right)^{\frac{f+1}{2}} Hh_f\left(\frac{-\delta t}{\sqrt{f+t^2}}\right)$$

where

$$Hh_f(y) = \int_0^\infty \frac{v^f}{f!} e^{-\frac{1}{2}(v+y)^2}\, dv.$$

The tables contained herein give values of the probability integral, of the probability density function and of the percentage points of the non-central t-statistic for selected values of the parameters f and δ.

Important tables related to the non-central t-statistic have been published previously by Johnson and Welch [4]. Their tables do not deal directly with the probability integral, nor is it possible to obtain from them values of the density function. The Johnson and Welch tables facilitate the computation of $\delta(f,t_o,\epsilon)$, that is, that value of the parameter δ for which $\Pr[t \le t_o | f,\delta] = 1-\epsilon$, for seventeen values of ϵ. By a more extended computation it is possible to obtain the percentage points corresponding to these seventeen values of ϵ. The percentage points corresponding to the special values $\epsilon = .05$ and $\epsilon = .95$ may be obtained with somewhat less labor. The tables may also be used to compute values of the probability integral as a function of δ, for fixed f and t.

Other tables of the probability integral are those of Neyman [9] and Neyman and Tokarska [10], which were computed for the purpose of obtaining the power curve of the Student t - test.

2. Description of the Tables

Three tables of the non-central t-statistic are given: the probability integral, the probability density function, and the percentage points of t.

Tabulation of the non-central t - distribution requires a table of triple entry since the distribution depends on the two parameters f and δ, where f is the number of degrees of freedom and δ is the non-centrality parameter.

The argument used throughout the tables is $x = t/\sqrt{f}$. The reason for using $t/\sqrt{f}$ as the argument instead of t itself is that the range for this argument is roughly the same whatever the values of the parameters f and δ. This enables a somewhat more compact form of tabulation than would be possible if the argument were t itself.

The ranges of the parameters f and δ are as follows: f ranges from 2 to 24 by steps of 1 and from 24 to 49 by steps of 5; and $\delta = \sqrt{f+1}\, K_p$ where K_p is the standardized normal random variable exceeded with probability p; that is,

$$\int_{K_p}^{\infty} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, dy = p.$$

The values of p which were used are $p = .2500, .1500, .1000, .0650, .0400, .0250, .0100, .0040, .0025$ and $.0010$. A table of $\sqrt{f+1}\, K_p$ for these values of f and p is given on page 3 .

These values of the parameters were chosen so as to cover the useful range adequately. In particular, the choice of $f = 2(1)24$ and $24(5)49$ appears to fill the needs for most practical applications. It is felt that for f greater than 49, a normal approximation [12] is adequate. The values for f of $24(5)49$ were chosen rather than f of $25(5)50$ because most practical applications require that $f = n-1$ where n is the sample size. Sample sizes which are multiples of 5 are prevalent in industrial statistics.

The indexing of $\delta = \sqrt{f+1}\, K_p$ as a function of K_p requires a word of explanation. General examples of the usefulness of these tables appear in the next section. However, the development of these tables was motivated by the necessity of using the non-central t - statistic for problems involving the fraction of normally distributed random variables falling above and/or below specified values. The

Table of $\sqrt{f+1}\ K_p$

K_p	0.674490	1.036433	1.281552	1.514102	1.750686	1.959964	2.326348	2.652070	2.807034	3.090232
f \ p	.2500	.1500	.1000	.0650	.0400	.0250	.0100	.0040	.0025	.0010
2	1.168251	1.795155	2.219712	2.622501	3.032277	3.394757	4.029353	4.593520	4.861925	5.352439
3	1.348980	2.072867	2.563103	3.028204	3.501372	3.919928	4.652696	5.304140	5.614068	6.180465
4	1.508205	2.317536	2.865636	3.385635	3.914653	4.382613	5.201872	5.930208	6.276718	6.909970
5	1.652156	2.538733	3.139147	3.708777	4.288287	4.800912	5.698365	6.496218	6.875800	7.569492
6	1.784532	2.742145	3.390667	4.005937	4.631880	5.185577	6.154938	7.016717	7.426713	8.175986
7	1.907745	2.931476	3.624775	4.282527	4.951688	5.543615	6.579905	7.501186	7.939490	8.740497
8	2.023469	3.109300	3.844655	4.542306	5.252058	5.879892	6.979044	7.956209	8.421101	9.270697
9	2.132924	3.277490	4.052622	4.788011	5.536156	6.197950	7.356558	8.386581	8.876620	9.772173
10	2.237029	3.437461	4.250426	5.021708	5.806369	6.500465	7.715623	8.795921	9.309878	10.249141
11	2.336501	3.590311	4.439425	5.245003	6.064554	6.789514	8.058705	9.187039	9.723850	10.704879
12	2.431907	3.736914	4.620700	5.459172	6.312188	7.066751	8.387767	9.562174	10.120904	11.141991
13	2.523710	3.877979	4.795127	5.665251	6.550467	7.333514	8.704397	9.923137	10.502959	11.562591
14	2.612288	4.014089	4.963428	5.864091	6.780378	7.590908	9.009906	10.271422	10.871595	11.968418
15	2.697959	4.145734	5.126206	6.056408	7.002744	7.839856	9.305391	10.608279	11.228135	12.360929
16	2.780992	4.273324	5.283972	6.242802	7.218264	8.081138	9.591778	10.934764	11.573697	12.741354
17	2.861618	4.397214	5.437163	6.423790	7.427532	8.315423	9.869858	11.251779	11.909236	13.110745
18	2.940033	4.517708	5.586154	6.599817	7.631064	8.543285	10.140315	11.560104	12.235576	13.470010
19	3.016410	4.635071	5.731273	6.771270	7.829306	8.765225	10.403744	11.860417	12.553437	13.819939
20	3.090900	4.749534	5.872807	6.938487	8.022651	8.981683	10.660665	12.153311	12.863445	14.161223
21	3.163637	4.861304	6.011010	7.101767	8.211446	9.193046	10.911539	12.439310	13.166156	14.494474
22	3.234739	4.970560	6.146105	7.261378	8.395995	9.399657	11.156772	12.718880	13.462061	14.820233
23	3.304311	5.077466	6.278295	7.417554	8.576575	9.601823	11.396731	12.992436	13.751601	15.138985
24	3.372449	5.182167	6.407758	7.570509	8.753430	9.799820	11.631739	13.260349	14.035169	15.451162
29	3.694333	5.676779	7.019347	8.293078	9.588903	10.735165	12.741932	14.525985	15.374757	16.925900
34	3.990335	6.131623	7.581761	8.957548	10.357198	11.595303	13.762860	15.689857	16.606636	18.282061
39	4.265848	6.554980	8.105244	9.576021	11.072311	12.395901	14.713116	16.773162	17.753240	19.544345
44	4.524615	6.952607	8.596909	10.156904	11.743959	13.147838	15.605616	17.790625	18.830155	20.729908
49	4.769363	7.328691	9.061938	10.706317	12.379220	13.859038	16.449764	18.752965	19.848726	21.851242

solution of such problems involves K_p. The particular values of p chosen coincide with the Acceptable Quality Levels (AQL's) of Military Standard 105A, Sampling Procedures and Tables for Inspection by Attributes.[1/] It should be noted that this choice does not restrict the usefulness of these tables, but rather gives a method of adequately representing the useful range of δ for the applications.

In the subsequent sections the probability integral of t will be denoted by $P(f,\delta,x) = Pr[t/\sqrt{f} \leq x \mid f,\delta]$. The probability density of t, tabled with argument $t/\sqrt{f}$, will be denoted by $P'(f,\delta,t/\sqrt{f})$. The percentage points of t will be denoted by $x(f,\delta,\epsilon)$, where x is the value such that $Pr[t/\sqrt{f} > x \mid f,\delta] = \epsilon$.

For negative values of δ the following relationships are useful:

$$P'(f, -\delta,x) = P'(f,\delta, -x),$$

$$P(f, -\delta,x) = 1-P(f,\delta, -x)$$

and

$$x(f, -\delta,\epsilon) = -x(f,\delta,1-\epsilon).$$

3. Examples of the Uses of the Probability Density Function of the Non-Central t-Statistic

a. The WAGR Sequential Test

The WAGR test is a sequential procedure for testing the null hypothesis H_o, that the proportion of a normal population exceeding a given constant U is p_o (given), against the alternative hypothesis H_1, that the proportion is p_1 (given). The name WAGR stems from the initials of several of the individuals who proposed this test: Wald, Arnold, Goldberg and Rushton. A description of the WAGR test follows.

Let $Y_1, Y_2, \ldots$ be a sequence of independent observations on a normally distributed random variable with mean μ and variance σ^2.

Let $\bar{Y}$ be the arithmetic mean of the first n of the observations. Define u_n by

$$u_n = \frac{1}{\sqrt{n}} \left(\frac{\sum_{i=1}^{n} (U-Y_i)}{\sqrt{\sum_{i=1}^{n} (Y_i-\bar{Y})^2}} \right) \qquad n=2,3,\ldots$$

1/ These same values of p will be the AQL's of a new military standard for sampling inspection by variables.

Then $t_n = \sqrt{n-1}\, u_n$ is a non-central t-statistic with parameters $f = n-1$ and $\delta = \sqrt{n}\ \frac{U-\mu}{\sigma} = \sqrt{n}\, K_p$. The ratio of the probability density function of t_n under H_1 to the probability density function of t_n under H_o is, in the present notation, given by

$$\lambda_n = \frac{P'(n-1, \sqrt{n}\, K_{p_1}, u_n)}{P'(n-1, \sqrt{n}\, K_{p_o}, u_n)} .$$

The test procedure is to observe λ_n sequentially for $n = 2,3,\ldots$. The first time that the inequality

$$\frac{\beta}{1-\alpha} \leq \lambda_n \leq \frac{1-\beta}{\alpha}, \quad 0 < \alpha + \beta < 1, \quad 0 < \alpha, \quad 0 < \beta$$

is violated, the null hypothesis is either accepted or rejected. If $\lambda_n > \frac{1-\beta}{\alpha}$, reject H_o; if $\lambda_n < \frac{\beta}{1-\alpha}$, accept H_o.

It has been shown [8] that the WAGR is a true sequential test as defined by Wald and hence that the probability of accepting H_o is approximately $1-\alpha$ when p_o is the true population proportion, and is approximately β when p_1 is the true proportion. A proof that this test reaches a decision with probability one has been given [2].

In actually carrying out the test, one initially observes Y_1 and Y_2 and computes u_2. The ratio λ_2 is not available from the present tables since it corresponds to $f = n-1 = 1$. However λ_2 may be expressed as follows:

$$\lambda_2 = \left[\frac{\varphi(z_1/u_2)}{\varphi(z_o/u_2)}\right] \left[\frac{z_1\, \Phi(z_1) + \varphi(z_1)}{z_o\, \Phi(z_o) + \varphi(z_o)}\right]$$

where

$$z_j = \frac{\sqrt{2}\, K_{p_j}\, u_2}{\sqrt{u_2^2 + 1}}, \quad j = 0,1 ,$$

$$u_2 = \frac{(2U - Y_1 - Y_2)}{|Y_1 - Y_2|} ,$$

$$\varphi(z) = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} ,$$

and

$$\Phi(z) = \int_{-\infty}^{z} \varphi(v)dv.$$

The functions $\varphi(z)$ and $\Phi(z)$ are conveniently available in a single volume of tables of the normal probability function [7].

After this first step, if the inequality is not violated, one observes Y_n (and hence u_n) for $n = 3,4,\ldots$ until the test is terminated. Each time u_n is computed, the ratio λ_n is obtained directly from the appropriate table of the non-central t-density function, if p_1 and p_o are among the ten tabulated values of p; otherwise, interpolation on the non-centrality parameter is required.

A numerical example of the application of the tables of the probability density to the WAGR test follows. Let $U = 10$, $p_o = .01$, $p_1 = .065$, $\alpha = .05$ and $\beta = .10$. Then $\frac{\beta}{1-\alpha} = .1053$ and $\frac{1-\beta}{\alpha} = 18.0000$. Suppose the first observations are $Y_1 = 8.10$ and $Y_2 = 8.70$; then

$$u_2 = \frac{[2(10) - 8.10 - 8.70]}{|8.10 - 8.70|} = 5.3333 ,$$

$$z_o = \frac{(3.2900)(5.3333)}{\sqrt{29.4441}} = 3.2337$$

and

$$z_1 = \frac{(2.1413)(5.3333)}{\sqrt{29.4441}} = 2.1046.$$

From the tables of the normal probability function, we obtain: $\varphi(z_o) = .002139$, $\varphi(z_1) = .04356$, $\Phi(z_o) = .9994$, $\Phi(z_1) = .9823$, $\varphi(z_o/u_2) = \varphi(.6063) = .3320$ and $\varphi(z_1/u_2) = \varphi(.3946) = .3691$. Hence

$$\lambda_2 = [\frac{.3691}{.3320}] [\frac{(2.1046)(.9823) + (.04356)}{(3.2337)(.9994) + (.002139)}] = .7257 .$$

Thus, the inequality is not violated at the $n=2$ level and we take another observation. Let $Y_3 = 7.20$; then

$$u_3 = \frac{1}{\sqrt{3}} (\frac{6.00}{\sqrt{1.14}}) = 3.2444.$$

From the tables of the density function of t, we obtain

$$P'(2, \sqrt{3}\, K_{.01}, 3.2444) = .1414$$

and

$$P'(2, \sqrt{3}\, K_{.065}, 3.2444) = .0976.$$

Hence

$$\lambda_3 = \frac{.0976}{.1414} = .6902$$

and the inequality still holds at the $n=3$ level. One would then take another observation and proceed as for $n=3$ above, obtaining the values of the density function from the appropriate table. This operation is continued until the inequality is violated.

b. An Application of the Probability Density Function to the Computation of Moments of a Function of the Non-Central t - Statistic

The first two moments of the non-central t - statistic are given by the expressions

$$E(t) = \frac{\sqrt{\frac{f}{2}}\,\Gamma(\frac{f-1}{2})\delta}{\Gamma(\frac{f}{2})}$$

and

$$E(t^2) = \frac{f(1+\delta^2)}{f-2} .$$

The moments of several useful functions of t are known only approximately, for large values of the parameter f. One application of the tables of the probability density to the problem of computing the moments of an arbitrary function $g(t)$ of the non-central t - statistic is by straightforward numerical integration.

The k^{th} moment of $g(t)$ can be written as

$$E[(g(t))^k] = \int_{-\infty}^{\infty} (g(t))^k\, h(f,\delta,t)dt$$

where $h(f,\delta,t)$ is the analytical expression for the non-central t - density given in the introductory section.

This may be approximated by

$$E[(g(t))^k] \sim \sum_{t_i} (g(t_i))^k \, P'(f,\delta, \frac{t_i}{\sqrt{f}})$$

where $\sum$ refers to summation of a grid over the t_i, such as the well-known trapezoidal rule, Simpson's rule, or other more elaborate integration formulas.

Particular examples of useful functions of the non-central t-statistic are two estimates of the proportion p of a normal population which lies above a given limit U. Let the population mean be μ and the population variance be σ^2; then p is expressed as

$$p = \int_{\frac{U-\mu}{\sigma}}^{\infty} \frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}} \, dv .$$

An estimate of p which is often given is the biased estimate $\tilde{p}(\overline{Y},s)$ which is obtained by replacing μ and σ by their sample estimates in the expression for p. Let $Y_1, Y_2, \ldots, Y_n$ be a sample of n observations and let

$$\overline{Y} = \frac{\sum Y_i}{n}$$

and

$$s = \sqrt{\frac{\sum (Y_i - \overline{Y})^2}{n-1}} .$$

Then

$$\tilde{p} = \int_{\frac{U-\overline{Y}}{s}}^{\infty} \frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}} \, dv .$$

Another estimate of p which has appeared recently in the literature [6] is the uniformly minimum variance unbiased estimate defined by

$$\hat{p} = \begin{cases} I_z(\frac{n-2}{2}, \frac{n-2}{2}) & \text{for } 0 < z < 1 \\ 0 & \text{for } z \leq 0 \\ 1 & \text{for } z \geq 1 \end{cases}$$

where $z = \frac{1}{2} - \frac{1}{2} \frac{U-\bar{Y}}{s} \frac{\sqrt{n}}{n-1}$ and $I_z(a,b)$ is the Incomplete Beta Function ratio with parameters a and b.

The quantity

$$\frac{\sqrt{n}\ (U-\bar{Y})}{s}$$

has the non-central t-distribution with degrees of freedom $f = n-1$ and $\delta = \sqrt{n}\ K_p$, so that both estimates are functions of the non-central t-statistic.

The two estimates $\tilde{p}$ and $\hat{p}$ are asymptotically equivalent and asymptotically efficient. The latter estimate, $\hat{p}$, is unbiased and has the smallest variance among all unbiased estimates of p. The former estimate, $\tilde{p}$, is biased so that its mean square error $E(\tilde{p}-p)^2$ is more relevant than its variance. No small sample comparison of the mean square errors of these estimates is available. It has sometimes been assumed that $\tilde{p}$ is a "best" estimate in the sense of least mean square error. Numerical integrations performed by the writers using the present tables of the probability density show that this is not the case, and in fact their relative merit depends on the value of the parameter p at which the mean square error is computed.

Suppose we wish to obtain the first moment of the estimate $\tilde{p}$ of the proportion defective of a normal population for $p = .0250$. Let $n = 17$. Then we use the expression

$$E[g(t)] \sim \sum_{t_i} g(t_i)\ P'(16, \sqrt{17}\ K_{.0250}, \frac{t_i}{\sqrt{16}})$$

where

$$g(t) = \tilde{p} = \int_{\frac{U-\bar{Y}}{s} = \frac{t}{\sqrt{n}}}^{\infty} \frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}}\ dv.$$

Note that we use $t/\sqrt{f}$ as argument for computation purposes to avoid interpolation in the tables of the non-central t-distribution. But the increment to be used in applying the numerical integration formula must be the increment in t itself. In the example below, the increment for argument $t/\sqrt{f}$ is .25, but the increment in t is $.25\sqrt{f} = 1.00$. We choose arbitrarily to use every fifth point given in the probability density tables, but the integration may be done with as fine an interval as desired. The values for $g(t_i)$ were obtained from tables of the normal probability integral. Applying the trapezoid rule, we obtain

$$\frac{.25\sqrt{16}}{2} \sum a_i g(t_i) \; P'(16, \sqrt{17}\, K_{.0250}, \frac{t_i}{\sqrt{16}}) = .0294$$

where the a_i are the appropriate coefficients for the trapezoid rule.

$\frac{t_i}{\sqrt{16}}$	$P'(16, \sqrt{17}\, K_{.0250}, \frac{t_i}{\sqrt{16}})$	$\frac{t_i}{\sqrt{17}}$	$g(t_i)$
.75	.0000	.7276	.2334
1.00	.0016	.9701	.1660
1.25	.0240	1.2127	.1126
1.50	.1044	1.4552	.0728
1.75	.2009	1.6977	.0448
2.00	.2277	1.9403	.0262
2.25	.1837	2.1828	.0145
2.50	.1195	2.4254	.0076
2.75	.0679	2.6679	.0038
3.00	.0356	2.9104	.0018
3.25	.0178	3.1530	.0008
3.50	.0087	3.3955	.0003
3.75	.0042	3.6380	.0001
4.00	.0020	3.8806	.0001
4.25	.0010	4.1231	.0000
4.50	.0005	4.3656	.0000
4.75	.0002	4.6082	.0000
5.00	.0001	4.8507	.0000
5.25	.0001	5.0932	.0000
5.50	.0000	5.3358	.0000

4. Examples of the Use of the Probability Integral and Percentage Points of the Non-Central t-Statistic

a. Sampling Inspection by Variables for Fraction Defective, with One Standard Given

A class of problems in which the non-central t-statistic plays an important role is in the computation of operating characteristic (OC) curves for sampling inspection by variables procedures. Suppose an object is classified as defective or non-defective according to whether the value of a characteristic exceeds or falls short of a fixed standard U. The percent defective, p, in a lot is defined as the percentage of the objects falling above U. The general sampling inspection problem is the formulation of a procedure which will accept the lot if p is sufficiently small. In particular, sampling inspection by variables can be used if the measured characteristic, Y, is a normally distributed random variable. (Note that this variable Y is distinct from that random variable which depends solely on whether the object is defective or non-defective.) The acceptance procedure for the case in which the mean, μ, and the standard deviation, σ, are unknown is to accept the lot if

$$\bar{Y} + ks \leq U$$

where k is a constant, and $\bar{Y}$ and s, the sample mean and sample standard deviation computed from a sample of size n, are given by:

$$\bar{Y} = \frac{\sum Y_i}{n}$$

and

$$s = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n-1}} .$$

The acceptance criterion can be rewritten as

$$\frac{\sqrt{n}\,(U-\bar{Y})}{s} \geq \sqrt{n}\,k$$

and further as

$$\left\{ \frac{\sqrt{n}\,(U-\mu)}{\sigma} - \frac{\sqrt{n}\,(\bar{Y}-\mu)}{\sigma} \right\} \Big/ \frac{s}{\sigma} \geq \sqrt{n}\,k .$$

But the expression on the left has a non-central t-distribution with $f = n-1$ and $\delta = \frac{\sqrt{n}\,(U-\mu)}{\sigma} = \sqrt{n}\,K_p$. Hence the acceptance criterion can be written as

$$t \geq \sqrt{n}\,k.$$

The OC curve for this procedure, that is, the probability of accepting the lot, can be written as

$$\Pr\left\{ t \geq \sqrt{n}\, k \right\} = 1-\Pr\left\{ t \leq \sqrt{n}\, k \right\} = 1-\Pr\left\{ \frac{t}{\sqrt{n-1}} \leq \sqrt{\frac{n}{n-1}}\, k \right\}$$

$$= 1-P(n-1, \sqrt{n}\, K_p, \sqrt{\frac{n}{n-1}}\; k) .$$

Ten points on the OC curve corresponding to the ten values of p found in the table can be obtained immediately.

For example, for $n = 10$ and $k = 1.72$, $\Pr\left\{ t \geq \sqrt{n}\, k \right\} = 1-P(9, \sqrt{10}\, K_p, 1.813)$. By interpolation in the Table of the Probability Integral for $f = 9$, the following points are obtained:

$p = .2500$	$1-P(9, \sqrt{10}\, K_{.2500}, 1.813) = .0215$
$p = .1500$	$1-P(9, \sqrt{10}\, K_{.1500}, 1.813) = .1038$
$p = .1000$	$1-P(9, \sqrt{10}\, K_{.1000}, 1.813) = .2242$
$p = .0650$	$1-P(9, \sqrt{10}\, K_{.0650}, 1.813) = .3851$
$p = .0400$	$1-P(9, \sqrt{10}\, K_{.0400}, 1.813) = .5685$
$p = .0250$	$1-P(9, \sqrt{10}\, K_{.0250}, 1.813) = .7177$
$p = .0100$	$1-P(9, \sqrt{10}\, K_{.0100}, 1.813) = .8976$
$p = .0040$	$1-P(9, \sqrt{10}\, K_{.0040}, 1.813) = .9692$
$p = .0025$	$1-P(9, \sqrt{10}\, K_{.0025}, 1.813) = .9843$
$p = .0010$	$1-P(9, \sqrt{10}\, K_{.0010}, 1.813) = .9961$

These tables can also be used to design sampling inspection plans. If two points on the OC curve, $(p_1, 1-\alpha)$ and (p_2, β), are given, the required values of n and k can be found. These values are obtained from the relationships:

$$P(n-1, \sqrt{n}\, K_{p_1}, \sqrt{\frac{n}{n-1}}\, k) = \alpha$$

$$P(n-1, \sqrt{n}\, K_{p_2}, \sqrt{\frac{n}{n-1}}\, k) = 1-\beta .$$

The values of n and k are found by trial and error. Theoretically, the correct value of f has the property that there exists a value of x for which the entry in the p_1 column equals α and simultaneously the entry in the p_2 column equals $1-\beta$; that is,

$$P(n-1, \sqrt{n}\, K_{p_1}, x) = \alpha$$

$$P(n-1, \sqrt{n}\, K_{p_2}, x) = 1-\beta.$$

However, because of the necessary discreteness in the sample size, an integral value of f possessing the above property is not usually obtainable. Therefore, the following procedure for $f \leq 24$ is established. Two consecutive values of f, say f^* and $f^* + 1$, are found such that

(1) for f^* there exists an x' for which

$$P(n-1, \sqrt{n}\, K_{p_1}, x') \geq \alpha$$

and simultaneously

$$P(n-1, \sqrt{n}\, K_{p_2}, x') \leq 1-\beta$$

and

(2) for $f^* + 1$ there exists an x'' for which

$$P(n-1, \sqrt{n}\, K_{p_1}, x'') \leq \alpha$$

and simultaneously

$$P(n-1, \sqrt{n}\, K_{p_2}, x'') \geq 1-\beta.$$

The value $f^* + 1$ is the desired f, and the appropriate value of x is found by interpolating in the p_1 column with degrees of freedom $f^* + 1$ for the given value α. The value of k is obtained by substituting this value of x in $k = \sqrt{\frac{n-1}{n}}\, x$. Note that this procedure always yields a value of β less than the desired one. This has the effect of establishing a sampling plan that is a little stricter than the one desired in the sense that lots of incoming quality p_2 will be accepted less frequently than prescribed.

For example, suppose an OC curve is to be constructed passing through the points $(p_1, 1-\alpha) = (.01, .99)$ and $(p_2, \beta) = (.15, .10)$. Thus, $\alpha = .01$ and $1-\beta = .90$. $f^* = 16$ and $f^* + 1 = 17$ have the properties described above, that is,

$$P(16, \sqrt{17}\, K_{.01}, 1.55) = .0104 \geq .01$$

$$P(16, \sqrt{17}\, K_{.15}, 1.55) = .8936 \leq .90$$

and

$$P(17, \sqrt{18}\, K_{.01}, 1.55) = .0090 \leq .01$$

$$P(17, \sqrt{18}\, K_{.15}, 1.55) = .9022 \geq .90 \; .$$

Thus, $f = 17$ is chosen. It should be noted that x' and x'' need not have the same numerical value. Because of the relative coarseness of the interval in the argument x, the value of x, corresponding to a specified value of α for f, will often be identical to the value of x corresponding to α for $f+1$. With this f and by linear interpolation in the p_1 column, the value of $x = 1.560$ provides an $\alpha = .01$. The value of k is found from $k = \sqrt{\frac{17}{18}}\, 1.560 = 1.516$. Thus, for the values of $n = 18$ and $k = 1.516$, the OC curve will pass through $(p_1, 1-\alpha) = (.01, .99)$ but will not pass through $(p_2, \beta) = (.15, .10)$ exactly. Instead, for these values of n and k, the curve will pass through $(.15, .0941)$.

For $24 \leq f \leq 49$, it is suggested that a first approximation for $n = f+1$ be obtained by the normal approximation which is given by

$$n = \frac{2(K_\alpha + K_\beta)^2 + (K_\alpha K_{p_2} + K_\beta K_{p_1})^2}{2(K_{p_1} - K_{p_2})^2}$$

where the numerical value obtained is always rounded to the next higher integer and where K_ϵ is a normal deviate exceeded with probability ϵ; that is, K_ϵ is defined by

$$\frac{1}{\sqrt{2\pi}} \int_{K_\epsilon}^{\infty} e^{-\frac{y^2}{2}}\, dy = \epsilon \; .$$

Call this value $\tilde{f}$. Then interpolate in the Tables of the Probability Integral of t for the x corresponding to p_1 and α for the two tabled values of f adjacent to $\tilde{f}$; then obtain the $(1-\beta)$ value corresponding to these x's in the p_2 column. If $\tilde{f}$ is already one of the tabulated values of f, the desired values of x and $1-\beta$ are obtained by these interpolations. Finally interpolate on f for the x corresponding to p_1 and α and for $(1-\beta)$.

For example, suppose a sampling plan is to be constructed which will have an OC curve passing through (.065, .99) and (.25, .04). From the normal approximation to f, the value of $\tilde{f} = 36$ is obtained. The two tabled values of f adjacent to 36 are 34 and 39. For f = 34, by interpolation it is found that x = 1.043 corresponds to $p_1 = .065$ and $\alpha = .01$; and for this value of x, $1-\beta = .9507$. Similarly for f = 39, x = 1.066 and $1-\beta = .9698$ are found. Then, by interpolation for $\tilde{f} = 36$ on these pairs of x and $(1-\beta)$ values, x = 1.052 and $1-\beta = .9583$ are obtained. Then $k = \sqrt{\frac{36}{37}}\,1.052 = 1.038$. Thus for the values n = 37 and k = 1.038, the OC curve will pass through (.065, .99) and (.25, .0417) which is close enough for most practical applications. It must be noted that these results are only as accurate as the method of interpolation used.

If the value of f required exceeds the range of the table the normal approximation for both n and k [12] can be used.

The solution to the problem of finding values of n and k corresponding to two given points on the OC curve can also be obtained by using the Table of the Percentage Points of t. This table can be used (and should be used) whenever $1-\alpha$ and β are values of ϵ given in this table. Theoretically, the solution is obtained by finding values of n and x such that

$$x(n-1, \sqrt{n}\,K_{p_1}, 1-\alpha) = x(n-1, \sqrt{n}\,K_{p_2}, \beta) .$$

However, again the discreteness of f does not allow an exact solution and it is necessary to proceed as above, finding two consecutive values of f, f* and f* + 1, such that for f*

$$x(n-1, \sqrt{n}\,K_{p_1}, 1-\alpha) < x(n-1, \sqrt{n}\,K_{p_2}, \beta)$$

and for f* + 1

$$x(n-1, \sqrt{n}\,K_{p_1}, 1-\alpha) > x(n-1, \sqrt{n}\,K_{p_2}, \beta).$$

The proper value of f is f* + 1 and x is the value $x(n-1, \sqrt{n}\,K_{p_1}, 1-\alpha)$ for this f. This procedure yields a slightly stricter OC curve than the one specified.

For example, if once again a sampling plan is to be constructed passing through the points $(.01, .99)$ and $(.15, .10)$. The Tables of the Percentage Points of t may be used since $\epsilon = .99$ and $\epsilon = .10$ are both tabled values. For $f^* = 16$

$$x(16, \sqrt{17}\, K_{.01}, .99) = 1.545 < x(16, \sqrt{17}\, K_{.15}, .10) = 1.566$$

whereas for $f^* + 1 = 17$

$$x(17, \sqrt{18}\, K_{.01}, .99) = 1.561 > x(17, \sqrt{18}\, K_{.15}, .10) = 1.545.$$

Thus $f = 17$ is chosen and $k = \sqrt{\frac{17}{18}}\, 1.561 = 1.517$. The difference between this solution and the one using the probability integral is due to the fact that the percentage point table yields more accurate results than does inverse linear interpolation in the Tables of the Probability Integral.

If the values of p_1 and p_2 chosen are not those found in the table, a solution can be obtained by interpolation on p or K_p.

Finally, these tables can also be used to design sampling plans when the value of n is given and a single point $(p_1, 1-\alpha)$ is specified. The solution yields the appropriate value of k and is obtained from the expression

$$P(n-1, \sqrt{n}\, K_{p_1}, \sqrt{\frac{n}{n-1}}\, k) = \alpha$$

when the Table of the Probability Integral is used. The value of x which satisfies the above expression is used to determine $k = \sqrt{\frac{n-1}{n}}\, x$.

If the Table of Percentage Points is used, the value of x corresponding to $x(n-1, \sqrt{n}\, K_{p_1}, 1-\alpha)$ yields the appropriate solution and $k = \sqrt{\frac{n-1}{n}}\, x$.

For example, if $p_1 = .01$, $\alpha = .05$, and $n = 15$, and the Table of Percentage Points is used, the value of x corresponding to

$$x(14, \sqrt{15}\, K_{.01}, .95) = 1.736$$

and

$$k = \sqrt{\frac{14}{15}}\, 1.736 = 1.677.$$

b. Sampling Inspection by Variables for Fraction Defective, with Two Standards Given

An unsolved problem is that of constructing tests of significance about the proportion defective p, with given significance level α, when p is characterized by two specifications, a lower standard L, and an upper standard U, so that

$$p = p_U + p_L = \int_{\frac{U-\mu}{\sigma}}^{\infty} \frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}} dv + \int_{-\infty}^{\frac{L-\mu}{\sigma}} \frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}} dv .$$

This is sometimes called the two-sided proportion defective. However it is possible to devise such tests of significance for p with significance level very near a given value α by the use of a procedure based on the uniformly minimum variance estimate $\hat{p}$, defined by

$$\hat{p} = \hat{p}_U + \hat{p}_L = I_{z_U}(\frac{n-2}{2}, \frac{n-2}{2}) + I_{z_L}(\frac{n-2}{2}, \frac{n-2}{2})$$

where

$$z_U = \frac{1}{2} - \frac{1}{2} \frac{U-\bar{Y}}{s} \frac{\sqrt{n}}{n-1} \quad ; \quad z_L = \frac{1}{2} - \frac{1}{2} \frac{\bar{Y}-L}{s} \frac{\sqrt{n}}{n-1} .$$

I_z is the Incomplete Beta Function Ratio.

If $z < 0$, then I_z is taken to be zero; if $z > 1$, I_z is taken to be one.

If $L = -\infty$ so that $p_L = 0$,

$$p = p_U = \int_{\frac{U-\mu}{\sigma}}^{\infty} \frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}} dv.$$

Then $\hat{p}_L = 0$ and the uniformly minimum variance unbiased estimate of the one-sided proportion defective p_U is

$$p_U = \begin{cases} I_{z_U}(\frac{n-2}{2}, \frac{n-2}{2}) & \text{for } 0 < z_U < 1 \\ 0 & \text{for } z_U \leq 0 \\ 1 & \text{for } z_U \geq 1 \end{cases}$$

If $U = +\infty$ so that $p_U = 0$, then $\hat{p}_U = 0$ and the uniformly minimum variance unbiased estimate of p_L is

$$p_L = \begin{cases} I_{z_L}(\frac{n-2}{2}, \frac{n-2}{2}) & \text{for} \quad 0 < z_L < 1 \\ 0 & \text{for} \quad z_L \le 0 \\ 1 & \text{for} \quad z_L \ge 1 \end{cases}$$

To test $p_U \le p_o$ against $p_U > 0$ the procedure is as follows:

1. Choose a significance level α and sample size n.

2. Using the Tables of the Percentage Points, solve the equation

$$x(n-1, \sqrt{n}\, K_{p_o}, 1-\alpha) = \sqrt{n-1} \quad (1-2\, \beta_{p*})$$

for β_{p*}.

3. β_{p*} is defined by $I_{\beta_{p*}}(\frac{n-2}{2}, \frac{n-2}{2}) = p*$. By interpolation in the tables of the Incomplete Beta Function determine $p*$.

4. Draw a sample of n and compute

$$z_U = \frac{1}{2} - \frac{1}{2}\, \frac{U-\bar{Y}}{s}\, \frac{\sqrt{n}}{n-1} .$$

5. From the Table of the Incomplete Beta Function, using z_U as argument determine $\hat{p}$.

6. If $\hat{p} > p*$, the hypothesis $p_U \le p_o$ is rejected with probability exactly α.

If in step 3 the critical region $\hat{p}_U > p*$ is such that $p* = 0$, then for this combination of p_o and n, the significance level α is unobtainable with the test statistic $\hat{p}_U$. This situation occurs only for the combination of very small p_o and very small n, so that $x(n-1, \sqrt{n}\, K_{p_o}, 1-\alpha) > \sqrt{n-1}$.

The extension of this procedure to the case of testing $p_L \le p_o$ involves only replacing z_U by

$$z_L = \frac{1}{2} - \frac{1}{2}\, \frac{\bar{Y}-L}{s}\, \frac{\sqrt{n}}{n-1} .$$

The operating characteristic function (probability of acceptance) for these one-sided procedures is obtained from the Table of the Probability Integral and is given by

$$1-P[n-1, \sqrt{n}\, K_p, x(n-1, \sqrt{n}\, K_{p_o}, 1-\alpha)].$$

Thus it is seen to be equivalent to the one-sided acceptance procedure given in the preceding section where the criterion is $\overline{Y} + ks \leq U$.

The relation between k and p^* for given α is

$$k = \frac{n-1}{\sqrt{n}} (1-2\beta_{p*}).$$

Unless one is interested in simultaneously testing an hypothesis about proportion defective p and obtaining an unbiased estimate of p, the foregoing procedure affords no advantages over the criterion

$$\overline{Y} + ks \leq U.$$

However in the two-sided case the uniformly minimum variance unbiased estimate provides a procedure for testing $p = p_U + p_L \leq p_o$. Computations in [11] show that the probability distribution of $\hat{p} = \hat{p}_U + \hat{p}_L$ is essentially dependent only on p and only very slightly on the partition of p into its components p_U and p_L. To this extent the following statement holds:

$$\Pr\left\{ \hat{p} \leq p^* | p \right\} \sim 1-P[n-1, \sqrt{n}\, K_p, \sqrt{n-1}\, (1-2\beta_{p*})].$$

The two-sided test procedure is as follows:

1. Choose a significance level α and sample size n.

2. Using the Table of the Percentage Points, solve the equation

$$x(n-1, \sqrt{n}\, K_{p_o}, 1-\alpha) = \sqrt{n-1}\, (1-2\beta_{p*})$$

for β_{p*}.

3. By interpolation in the Table of the Incomplete Beta Function determine p^*.

4. Draw a sample of size n and compute both

$$z_U = \frac{1}{2} - \frac{1}{2}\, \frac{U-\overline{Y}}{s}\, \frac{\sqrt{n}}{n-1}$$

$$z_L = \frac{1}{2} - \frac{1}{2}\, \frac{\overline{Y}-L}{s}\, \frac{\sqrt{n}}{n-1}.$$

5. From the Table of the Incomplete Beta Function, using z_U and z_L as

arguments, determine $\hat{p}_U$ and $\hat{p}_L$ and hence $\hat{p} = \hat{p}_U + \hat{p}_L$.

6. If $\hat{p} > p^*$ the hypothesis $p \leq p_o$ is rejected with probability very nearly equal to α.

If the critical region $\hat{p} > p^*$ is such that $p^* = 0$ the significance level α is unachievable for this combination of p_o and n.

The operating characteristic function is obtained from the Table of the Probability Integral and is given approximately by

$$1-P[n-1, \sqrt{n}\, K_p, x(n-1, \sqrt{n}\, K_{p_o}, 1-\alpha)].$$

For the two-sided test procedure steps 1, 2 and 3 and the OC curve are identical with those of the one-sided case. Therefore any test criterion p^* and OC curve for a one-sided test may be used as approximations for the two-sided test.

An example of this procedure follows. It is assumed that the hypothesis to be tested is that the proportion defective p_o is less than or equal to .01. A sample size of 10 is to be used with significance level $\alpha = .10$. To find the test criterion the equation

$$x(9, \sqrt{10}\, K_{.01}, .90) = \sqrt{9}\,(1-2\beta_{p^*})$$

is solved for β_{p^*}. From the Tables of the Percentage Points $x(9, \sqrt{10}\, K_{.01}, .90) = 1.807$ so that $\beta_{p^*} = .1988$. By entering the Tables of the Incomplete Beta Function Ratio the value $p^* = .0327$ is found.

As a result of sampling 10 items the following statistics are observed

$$\frac{U-\bar{Y}}{s} = 2.034 \quad \text{and} \quad \frac{\bar{Y}-L}{s} = 1.925.$$

From these sample values

$$z_U = .14266 \quad \text{and} \quad z_L = .16181$$

are computed. The estimates $\hat{p}_U = .0101$ and $\hat{p}_L = .0159$ are found by entering the Tables of the Incomplete Beta Function Ratio. Thus $\hat{p} = .0101 + .0159 = .0260 < p^* = .0327$ so that the hypothesis is not rejected on the basis of this sample. The OC curve for this procedure is given by $1-P(9, \sqrt{10}\, K_p, 1.807)$. Ten points on this curve obtained from the Tables of the Probability Integral are given below.

p	OC function
.0010	.9963
.0025	.9849
.0040	.9702
.0100	.9000
.0250	.7217
.0400	.5728
.0650	.3890
.1000	.2270
.1500	.1054
.2500	.0219

c. Confidence Intervals for Proportions from a Normal Population

Denote by p the proportion of a normal population with unknown mean μ and unknown standard deviation σ which lies above a fixed standard U. In variables acceptance sampling p is termed the proportion defective.

The limits p_1 and p_2 of a confidence interval for p, with confidence coefficient γ, may be found by estimating $\sqrt{n}\, K_p = \sqrt{n}\, \frac{U-\mu}{\sigma}$ by the non-central t-statistic based on a sample of n observations, $Y_1, Y_2, \ldots, Y_n$. This non-central t-statistic is

$$\frac{\sqrt{n}\,(U-\bar{Y})}{s}$$

where

$$\bar{Y} = \frac{\sum Y_i}{n}$$

and

$$s = \sqrt{\frac{\sum (Y_i-\bar{Y})^2}{n-1}} \quad .$$

Confidence limits for $K_p = \frac{U-\mu}{\sigma}$ are determined by solving the equations

$$x(n-1, \sqrt{n}\, K_{p_1}, 1-\gamma_1) = \sqrt{\frac{n}{n-1}}\; \frac{U-\bar{Y}}{s}$$

$$x(n-1, \sqrt{n}\, K_{p_2}, \gamma_2) = \sqrt{\frac{n}{n-1}}\; \frac{U-\bar{Y}}{s}$$

for K_{p_1} and K_{p_2} using the Tables of the Percentage Points of the non-central t-statistic. Here $\gamma_1 + \gamma_2 = 1-\gamma$. It is usually necessary to interpolate on K_p in these tables for this purpose. One then converts K_{p_2} and K_{p_1} into p_2 and p_1, the upper and lower limits, respectively, of the confidence interval for p with coefficient γ. The most convenient table for converting K_p to p is [5]. However, any table of the cumulative normal distribution will serve.

A numerical example of the use of the Tables of the Percentage Points to obtain confidence limits for proportion defective p follows: Suppose a sample of size $n = 20$ is drawn and on the basis of the sample mean $\bar{Y}$ and sample standard deviation s, the statistic $\frac{U-\bar{Y}}{s}$ is found to be 1.834. A confidence interval for p is to be constructed with confidence coefficient $\gamma = .90$. Let $\gamma_1 = .05$ and $\gamma_2 = .05$. Then

$$\sqrt{\frac{n}{n-1}}\ \frac{U-\bar{Y}}{s} = \sqrt{\frac{20}{19}}\ 1.834 = 1.882.$$

Solve the equations

$$x(19, \sqrt{20}\ K_{p_1},\ .95) = 1.882$$

$$x(19, \sqrt{20}\ K_{p_2},\ .05) = 1.882$$

for p_1 and p_2. Interpolation should be performed on K_p, not on p. Values of K_p can be found in the table on page 3. For this example, the following results are obtained:

$$K_{p_1} \doteq 2.429021; \quad p_1 = .0076$$

$$K_{p_2} = 1.209322; \quad p_2 = .1133.$$

Thus the confidence interval for p with confidence coefficient $\gamma = .90$ based on this sample of 20 items is

$$[.0076,\ .1133].$$

If a lower specification L is given so that

$$p = \int_{-\infty}^{\frac{L-\bar{Y}}{s}} \frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}}\, dv$$

the equations which yield confidence limits for K_p and hence for p are

$$x(n-1, \sqrt{n}\, K_{p_1}, 1-\gamma_1) = \sqrt{\frac{n}{n-1}}\, \frac{(\bar{Y}-L)}{s}$$

$$x(n-1, \sqrt{n}\, K_{p_2}, \gamma_2) = \sqrt{\frac{n}{n-1}}\, \frac{(\bar{Y}-L)}{s} .$$

d. The Power of Student's t-Test

In testing the hypothesis H_o that the mean of a normal distribution is $\mu = \mu_o$ against the alternative that $\mu > \mu_o$, Student's t-test consists of calculating

$$t = \frac{\sqrt{n}\, (\bar{Y}-\mu_o)}{s}$$

where $\bar{Y} = \Sigma Y_i/n$ and $s = \sqrt{\frac{\Sigma (Y_i-\bar{Y})^2}{n-1}}$ and rejecting the hypotheses if

$$t > t_o .$$

If the level of significance is α, the following probability statement holds,

$$P(n-1, 0, t_o/\sqrt{n-1}) = 1-\alpha .$$

The value of t_o is obtained from a Table of the Percentage Points of Student's t-distribution.

Ten other points on the power curve (probability of rejection) can be obtained as follows:

When the true mean is μ, the quantity t has a non-central t-distribution with degrees of freedom $n-1$, and non-centrality parameter $\frac{\sqrt{n}\, (\mu-\mu_o)}{\sigma}$, that is,

$$t = \left\{ \frac{\sqrt{n}\, (\bar{Y}-\mu)}{\sigma} + \frac{\sqrt{n}\, (\mu-\mu_o)}{\sigma} \right\} \Big/ \frac{s}{\sigma} .$$

The power of the test is therefore given by

$$1-P(n-1, \frac{\sqrt{n}\, (\mu-\mu_o)}{\sigma}, t_o/\sqrt{n-1}) .$$

The values $\frac{\mu-\mu_o}{\sigma}$ (the abscissa of the power curve) which correspond to a given power are found by equating $\frac{\mu-\mu_o}{\sigma} = K_p$. Thus, for the ten values of p given in the probability integral table, ten values of $\frac{\mu-\mu_o}{\sigma}$ are obtained, and one minus the power associated with these points are read out of the table. These values appear in the entry corresponding to $x = t_o/\sqrt{n-1}$.

As an example, the power of the t-test at ten points will be obtained from the Table of the Probability Integral for $f = 4 (n = 5)$. A level of significance equal to five percent is chosen. For this case, $t_o = 2.132$ and $\frac{t_o}{\sqrt{n-1}} = 1.066$.

$\frac{\mu-\mu_o}{\sigma} = K_p$	1 - Power	Power
0.6745	.6512	.3488
1.0364	.3948	.6052
1.2816	.2409	.7591
1.5141	.1329	.8671
1.7507	.0636	.9364
1.9600	.0296	.9704
2.3263	.0060	.9940
2.6521	.0012	.9988
2.8070	.0005	.9995
3.0902	.0001	.9999

e. The Coefficient of Variation

This example of the use of the Tables of the Probability Integral of the non-central t-statistic deals with the coefficient of variation,

$$V = \sigma/\mu \ ,$$

where σ is the standard deviation of a normal distribution and μ is the mean of a normal distribution. An estimate of V is provided by the sample coefficient of variation $v = \frac{s}{\bar{Y}}$, where $\bar{Y} = \sum Y_i/n$ and $s = \sqrt{\frac{\sum(Y_i-\bar{Y})^2}{n-1}}$.

The following statistic

$$\frac{\sqrt{n}}{v} = \frac{\sqrt{n}\,\bar{Y}}{s} = \left\{ \frac{\sqrt{n}\,(\bar{Y}-\mu)}{\sigma} + \frac{\sqrt{n}\,\mu}{\sigma} \right\} \Big/ \frac{s}{\sigma}$$

has a non-central t-distribution with n-1 degrees of freedom and non-centrality parameter $\delta = \frac{\sqrt{n}\,\mu}{\sigma}$. Thus

$$\Pr\left\{ v > v_o \right\} = \Pr\left\{ \frac{\sqrt{n}}{v} \le \frac{\sqrt{n}}{v_o} \right\}$$

$$= \Pr\left\{ \sqrt{\frac{n}{n-1}}\ \frac{1}{v} \le \sqrt{\frac{n}{n-1}}\ \frac{1}{v_o} \right\} = P(n-1, \frac{\sqrt{n}\,\mu}{\sigma}, \sqrt{\frac{n}{n-1}}\ \frac{1}{v_o}).$$

Suppose a test of the hypothesis that $V = V_o$ at the ϵ level of significance is to be made against the alternative that $V > V_o$. The procedure used will be to reject the hypothesis if $v > v_o$. The critical value of the test v_o can be obtained from the identity

$$x(n-1, \frac{\sqrt{n}}{V_o}, 1-\epsilon) = \sqrt{\frac{n}{n-1}}\ \frac{1}{v_o}.$$

The solution will usually require an interpolation on the non-centrality parameter. Similarly, the power of the test can be obtained from the relationship Power $= P(n-1, \frac{\sqrt{n}\,\mu}{\sigma}, \sqrt{\frac{n}{n-1}}\ \frac{1}{v_o})$. Ten points on the power curve can be obtained by choosing the abscissa values corresponding to $\frac{1}{V} = K_p$.

As an example let $V_o = 1$, $n = 9$, and $\epsilon = .05$. Using the Table of the Percentage Points the value $v_o = 0.495$ for $f = 8$ and $\epsilon = .05$ is obtained from the expression

$$x(8, 3, .95) = 2.141 = \sqrt{\frac{9}{8}}\ \frac{1}{v_o}.$$

The ten points on the power curve are as follows:

$V = \frac{1}{K_p}$	Power
1.4826	.9880
.9649	.9428
.7803	.8719
.6605	.7661
.5712	.6256
.5102	.4873
.4299	.2635
.3771	.1245
.3563	.0814
.3236	.0332

5. Computational Methods

The tables were computed on the IBM Card Programmed Computer, Model II. This machine does not readily lend itself to table look-up, so that it was decided to generate the probability density function on the machine directly, without using existing tables of the Hh_f function [1]. In any case the range of the latter tables would not have sufficed for the purpose at hand.

The Hh_f function in the non-central t-density obeys the recurrence relation

$$fHh_f(x) = Hh_{f-2}(x) - xHh_{f-1}(x).$$

In particular

$$\sqrt{2\pi}\ Hh_o(x) = \int_{-\infty}^{x} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\, dt \quad \text{and} \quad \sqrt{2\pi}\ Hh_{-1}(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

Repeated application of the recurrence formula shows that Hh_f may be expressed as

$$Hh_f(x) = P_f(x)\ Hh_o(x) + Q_f(x)\ Hh_{-1}(x)$$

where P_f and Q_f are polynomials. It can easily be demonstrated that P_f and Q_f obey the same recurrence laws, that is, that

$$fP_f(x) = P_{f-2}(x) - xP_{f-1}(x)$$

with

$$P_o(x) = 1 \quad \text{and} \quad P_{-1}(x) = 0$$

and that

$$fQ_f(x) = Q_{f-2}(x) - xQ_{f-1}(x)$$

with

$$Q_o(x) = 0 \quad \text{and} \quad Q_{-1}(x) = 1.$$

A table of the polynomials P_f and Q_f was prepared for $f = 1$ to $f = 20$. The first few are given below:

f	$P_f(x)$	$Q_f(x)$
1	1	$-x$
2	$\frac{1}{2!}(-x)$	$\frac{1}{2!}(1+x^2)$
3	$\frac{1}{3!}(2+x^2)$	$\frac{1}{3!}(-3x-x^3)$

The polynomials and exponential functions in the non-central t-density were generated on the machine by standard programs. In order to compute the normal probability integral Hh_0, a rational function approximation due to Cecil Hastings [3] was used. This approximation yields an order of accuracy consistent with the number of significant figures available on the machine.

For f greater than 20, the polynomials P_f and Q_f became too cumbersome, causing the running time on the machine to be excessive. It was decided to use an asymptotic expansion for the Hh_f function. The following expression was derived using the method of Steepest Descent:

$$Hh_f(x) = \frac{1}{f!} t^f e^{-\frac{1}{2}(t+x)^2} \sqrt{\frac{2\pi t^2}{f+t^2}} [1 - \frac{3f}{4(f+t^2)^2} + \frac{5f^2}{6(f+t^2)^3}]$$

where

$$t = \frac{-x + \sqrt{x^2 + 4f}}{2} .$$

This approximation was tested in the following ways. The Hh_{20} functions were computed both by the method described previously and by the asymptotic expansion, and the two methods were compared. Both methods were then compared with Airey's tables of the Hh_{20} function.

The final results for the density function were adjudged to be accurate to at least five decimal places. (Seven places were retained in all functions during the computations.) In order to limit the size of the printed tables the final tabulation was rounded down to the four places shown.

The probability integral was computed by summing the probability density function by numerical integration. The results were spot-checked in scores of places against the tables of Johnson and Welch. The results usually agreed to four decimal places. In rare cases they differed by no more than one or two units in the third decimal place.

The discrepancies between the values obtained from the Johnson and Welch tables and the present Table of the Probability Integral may be due to interpolation in the former tables. Johnson and Welch indicate that results obtainable from their tables may be in error to this extent.

The Tables of the Percentage Points were obtained by inverse interpolation on the probability integral using six-point Lagrangian interpolation polynomials. The percentage points were checked in numerous instances by comparison with results obtained from the Johnson and Welch tables. The percentage points are believed to be correct in the second decimal place throughout, and to differ occasionally from the true values by no more than one or two units in the third decimal place.

Bibliography

[1] British Association for the Advancement of Science, Mathematical Tables, Vol. 1, Cambridge, England: University Press (1946).

[2] David, Herbert T., and Kruskal, William H., "The WAGR sequential t-test reaches a decision with probability one," Annals of Mathematical Statistics, 27(1956), 797-805.

[3] Hastings, Jr., Cecil, Approximations for Digital Computers, Princeton, New Jersey: Princeton University Press (1955).

[4] Johnson, N. L., and Welch, B. L., "Applications of the non-central t-distribution," Biometrika, 31(1940), 362- 89.

[5] Kelley, Truman Lee, The Kelley Statistical Tables, Cambridge, Mass.: Harvard University Press (1948).

[6] Lieberman, Gerald J., and Resnikoff, George J., "Sampling plans for inspection by variables," Journal of the American Statistical Association, 50(1955), 457-516.

[7] Mathematical Tables Project, Work Projects Administration for the City of New York, F.W.A., Tables of Probability Functions, Vol. II, Washington, D. C.: Government Printing Office (1942).

[8] National Bureau of Standards, Tables to Facilitate Sequential t-Tests, Applied Mathematics, Series 7, Washington, D. C.: Government Printing Office (1951).

[9] Neyman, J., "Statistical problems in agricultural experimentation," Supplement to the Journal of the Royal Statistical Society,2(1935), 107-180.

[10] Neyman, J., and Tokarska, B., "Errors of the second kind in testing 'Student's' hypothesis," Journal of the American Statistical Association, 31(1936), 318-326.

[11] Resnikoff, George J., A New Two-Sided Acceptance Region for Sampling by Variables, Technical Report No. 8, Office of Naval Research, Contract No. N6onr-25126, Applied Mathematics and Statistics Laboratory, Stanford University, (1952).

[12] Wallis, W. Allen, "Use of variables in acceptance inspection for percent defective," Chapter 1 in Statistical Research Group, Techniques of Statistical Analysis for Scientific and Industrial Research and Production and Management Engineering, New York: McGraw-Hill Book Company, Inc., (1947), 6-93.