THE ROBUSTNESS TO NON-NORMALITY OF SIGNIFICANCE LEVELS OF THE t AND F TESTS

by

TSAU-YI HO

Thesis submitted to the Graduate Faculty of
the Virginia Polytechnic Institute
in candidacy for the degree of

MASTER OF SCIENCE

in

STATISTICS

November, 1965
Blacksburg, Virginia

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

I. INTRODUCTION

In applied statistics, we test hypotheses about unknown population parameters using a sample of observations. When the population variance is unknown, the function used to test the population mean μ is t, defined by $t = \frac{\sqrt{n}\,(\bar{x}-\mu)}{s}$, where n is the sample size, $\bar{x}$ is the sample mean and s is the sample standard deviation. Under the same condition (unknown variance), the test function used to test the null hypothesis that two population means are equal ($\mu_1=\mu_2$) is u,

defined by $u = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(N_1-1)s_1^2+(N_2-1)s_2^2}{N_1+N_2-2}\left(\frac{1}{N_1}+\frac{1}{N_2}\right)}}$, where $\bar{x}_1$ and $\bar{x}_2$ are sample means for the first sample and second sample respectively, s_1^2 and s_2^2 are their sample variances, and N_1 and N_2 are their sample sizes. For testing the homogeneity of k groups of means in a one-way classification in the analysis of variance, the test function w is defined by

$$w = \frac{\sum_{i=1}^{k} n_i(\bar{x}_{i.}-\bar{x}_{..})^2/(k-1)}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_{i.})^2/(N-k)}$$

where x_{ij} is the observation in the i-th row (group) and j-th column, $\bar{x}_{i.}$ is the i-th row (group) sample mean, $\bar{x}_{..}$ is the grand mean, n_i is the sample size for the i-th group, and $N = \sum_{i=1}^{k} n_i$ is the total number of observations. In the case of two sets of observations, the equality of the variances may be tested by using the test statistic $v = \frac{s_1^2}{s_2^2}$, where s_1^2 and s_2^2 are the respective sample variances.

Under some assumptions, the distributions of the t and u statistics have "Student's" t-distribution, while w and v have F distributions. We will call the t and u tests the one sample t-test and the two sample t-test respectively, and call the w and v tests the test of analysis of variance, and the test for the equality of two variances, respectively.

The underlying assumptions which give rise to the t and F distributions are as follows:

(1) The experimental errors are normally distributed.

(2) Error variances are homogeneous.

(3) The experimental errors are uncorrelated.

In practice we can never be sure that these assumptions all hold, and often there is good reason to suspect that at least one of them fails. The failure of an assumption will affect both the significance level and the power of the t and F tests. For instance, an experimenter may think that he is testing at the 5% level, whereas he actually may be testing at the 8% level.

When the normality assumption is relaxed, the distribution of the t and u statistics will not be the

"Student's" t-distribution nor will they have the same distribution. The distributions of the w and v statistics will not be the F distribution nor will they have the same distribution. We will discuss these two cases respectively in Chapter IV and Chapter V.

Since in this study we always use the Edgeworth series in representing a density function, we give a simple description of this series in the Appendix.

The whole study is concerned with the question: in the case of the t-tests, the test of analysis of variance, or the test for the equality of two variances, is the test size α sensitive to changes in the form of distribution? A statistical test which is insensitive to departures from the underlying assumptions is called "robust", a term introduced by Box (4). Studies of robustness have been carried out by many writers. In this study, we review several of the most interesting and most important works on the subject of robustness to non-normality, and discuss their results. We also tabulate the true significance levels of the "Student's" t-test, the test of analysis of variance and the test for the equality of two variances for some specified alternative distribution forms. The discrepancy between the true significance levels and the normal theory significance levels is one useful measure of the effects of non-normality on these tests. Therefore, this paper should be of some aid to anyone who uses the "Student's" t or F tests in experimental situations in which there is serious doubt concerning the underlying assumptions of normality.

II. REVIEW OF THE LITERATURE

Many papers have been written on the subject of effects of non-normality on the t-test and the F-test. The earliest paper was written in 1929 by Rider (23), who investigated the distribution of "Student's" ratio computed from samples of size four independently drawn from a rectangular distribution. He compared the distribution thus obtained with the classical distribution based on normal theory. Using the same technique, Pearson and Adyanthaya (21) considered sampling from specified non-normal populations with regard to effects of skewness and kurtosis of the parent populations on the derived distribution of "Student's" ratio. These experimental results showed that the t-test is more sensitive to changes in skewness than kurtosis.

In 1935 Bartlett (1) was the first author to investigate this problem theoretically. He obtained an approximate density function for t, assuming the parent population to be represented by the first three terms of an Edgeworth series. He observed that skewness in the parent population exerts a much greater effect on the level of significance than does kurtosis, which confirms results of Pearson's (21) experimental investigation. Bartlett also showed that for moderate departures from normality, the t-test may still be used with confidence.

Geary (16) obtained an expression for the density of t in samples of any size drawn from a slightly asymmetrical population specified by the first two terms of the Edgeworth series.

Gayen (13) considered the effects of both the skewness λ_3 and the kurtosis λ_4 parameters by using the first four terms of the Edgeworth series. His approximation is superior to these of previous investigations, and as a consequence his results are adopted by many authors in their explanations of the effects of non-normality on the t-test.

In his unpublished doctoral dissertation, Bradley (5) made the first analytical study of the effects of non-normality on the t- and F-tests for a general class of populations with specified density functions. He also obtained correction factors for the two-sample t-statistic and for the F-statistic in the case of high levels of significance. Part of his doctoral dissertation was later published in 1952 (6) (7).

In 1950 Gayen (15) extended his previous work (13) to the two-sample t-test. He derived the distribution of the test statistic used for testing equality of two means when each of the populations is specified by an Edgeworth series. His study showed that the significance level is seriously affected when the difference of the skewness λ_3 in the two populations is not small. In cases where the samples are taken from the same population, non-normality seems to have little effect.

Some empirical work on departures from normality has been carried out with the aid of electronic computers. Using an IBM 650 computer, Boneau (3) computed a large number of values of "Student's" ratio based on random samples drawn from distributions having specified density functions. Levels of significance as determined from the empirical frequency distribution compared favorably with those based on the normal assumption.

The above paragraphs have dealt primarily with the distribution of "Student's" t-statistic under non-normal distribution theory. We turn now to the case of the variance ratio F for testing (a) equality of a set of means in the one-way classification analysis of variance under the fixed model, and (b) equality of two population variances using random samples drawn from each.

Early work in this area was done by Pearson (22) by way of sampling experiments. The effect of non-normality on the frequency distribution of the variance ratio F was first investigated theoretically by Geary (17), who based his work on the assumption of large samples, and considered the effect of kurtosis only. He gave an approximate formula for the probability of w.

In 1950 Gayen (14) derived the density functions of w and v for non-normal populations specified by the first four terms of the Edgeworth series. The results of his study are widely quoted. There are, however, a few mathematical errors in this paper, which were pointed out by Srivastava (24) and Tiku (25).

The F-test used for testing equality of a set of means in the fixed model, one-way classification analysis of variance and the t-tests possess the property of "robustness" to non-normality. This property is not shared, however, by the test for equality of two variances. Box (4) showed that the sensitivity to non-normality is even greater when the number of variances to be compared exceeds two. This means that Bartlett's test for equality of several variances is very sensitive to non-normality.

Tiku (25) in 1964 investigated the effect of non-normality on the F-test in the analysis of variance for the case in which the error distribution from group to group is not identical.

III. A SERIES REPRESENTATION OF THE CUMULATIVE DISTRIBUTION FUNCTION OF t FOR A SINGLE SAMPLE FROM A CLASS OF NON-NORMAL POPULATIONS

3.1 Cumulative distribution

Consider drawing samples from a population with density function f(x) which satisfies the following conditions:

(a) $f(x)>0$ for all $-\infty<x<\infty$.

(b) $f(x)$ is a continuous function of x for $-\infty<x<\infty$.

(c) $\frac{d^k f(x)}{dx^k}$ is a continuous function of x for $-\infty<x<\infty$ and $k=1,2,\ldots$

Bradley (5) (6) derived the cumulative distribution of t for samples from this type of population , namely,

$$[3.1] \quad G_N(t)=1-\Big[\frac{n^{n/2}}{n!}\Gamma(n)\binom{n-1}{0}t^{-n}C_{1N}(0;0,\ldots,0)D_{1N}(0;0\ldots,0) + \ldots\ldots + \frac{n^{(n+k)/2}k!}{(n+k)!}\Gamma(n)\binom{n+k-1}{k}t^{-(n+k)} \sum_{\substack{\text{comb.}\\ a_o+\ldots+a_n=k}} C_{1N}(k;a_o,a_1\ldots,a_1,\ldots,a_n)D_{1N}(k;a_o,\ldots,a_n)+\ldots\Big]$$

for any positive value of t, and

$$[3.2] \quad G_N(t)=\frac{n^{n/2}}{n!}\Gamma(n)\binom{n-1}{0}|t|^{-n}C_{1N}(0;0,\ldots,0)D^*_{1N}(0;0,\ldots,0) + \ldots + \frac{n^{\frac{n+k}{2}}k!}{(n+k)!}\Gamma(n)\binom{n+k-1}{k}|t|^{-(n+k)} \sum_{\substack{\text{comb.}\\ a_o+\ldots+a_n=k}} C_{1N}(k;a_o,a_1,\ldots,a_n)D^*_{1N}(k;a_o,a_1,\ldots,a_n)+\ldots$$

for any negative value of t,

where n=N-1 represents the degrees of freedom, N is sample size, and C_{1N}, D_{1N} and D^{*}_{1N} will be defined later. It happens that C_{1N} is independent of the particular population and need be evaluated only once for each set of values $a_o+a_1+a_2+\ldots+a_n=k$; we will give the values of C_{1N} in the following section.

3.2 The values of C_{1N}

Bradley (5) showed that C_{1N} vanishes for all odd values of k, he calculated C_{1N} for k=0,2,4 and 6. They are sufficient for obtaining the first four terms of the expansions of $G_N(t)$. We list the values of C_{1N} for k=0,2,4 and 6 below, and omit the detailed calculation procedure, for which the reader may refer to reference (5).

$S_{n-1}=\frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}$, N = sample size, n=N-1 degrees of freedom.

k=0

$C_{1N}(0;0,\ldots\ldots\ldots\ldots\ldots\ldots,0)=S_{n-1}$

k=2

$C_{1N}(2;\ 2,\ 0,\ldots\ldots\ldots\ldots\ldots,0)=S_{n-1}$

$C_{1N}(2;\ 1,\ 1,\ 0,\ldots\ldots\ldots\ldots,0)=S_{n-1}(-\frac{1}{2})$

k=4

$C_{1N}(4;\ 4,\ 0,\ldots\ldots\ldots\ldots\ldots,0)=\frac{S_{n-1}}{N(n+2)}\cdot 3n$

$C_{1N}(4;\ 3,\ 1,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{S_{n-1}}{N(n+2)}(-3n)$

$$C_{1N}(4;\ 2,\ 2,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{S_{n-1}}{N(n+2)}\ \left(\frac{n^2+2}{2}\right)$$

$$C_{1N}(4;\ 2,\ 1,\ 1,\ 0,\ldots\ldots\ldots,0)=\frac{S_{n-1}}{N(n+2)}\ \left(-\ \frac{n^2-3n+2}{2}\right)$$

$$C_{1N}(4;\ 1,\ 1,\ 1,\ 1,\ 0,\ldots\ldots,0)=\frac{S_{n-1}}{N(n+2)}\ \left(\frac{n^2-3n+2}{8}\right)$$

<u>k=6</u>

$$C_{1N}(6;\ 6,\ 0,\ldots\ldots\ldots\ldots\ldots,0)=\frac{S_{n-1}}{N^2(n+2)(n+4)}\ 15n^2$$

$$C_{1N}(6;\ 5,\ 1,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{S_{n-1}}{N^2(n+2)(n+4)}\ (-15n^2)$$

$$C_{1N}(6;\ 4,\ 2,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{S_{n-1}}{N^2(n+2)(n+4)}\ 3n(n^2+4)$$

$$C_{1N}(6;\ 4,\ 1,\ 1,\ 0,\ldots\ldots\ldots,0)=\frac{S_{n-1}}{N^2(n+2)(n+4)}\ \left[-\frac{3n}{2}\ (n^2-5n+4)\right]$$

$$C_{1N}(6;\ 3,\ 3,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{S_{n-1}}{N^2(n+2)(n+4)}\ \left[-\frac{3}{2}\ (3n^2+2)\right]$$

$$C_{1N}(6;\ 3,\ 2,\ 1,\ 0,\ldots\ldots\ldots,0)=\frac{S_{n-1}}{N^2(n+2)(n+4)}\ [-3(n^3-3n^2+4n-2)]$$

$$C_{1N}(6;\ 3,\ 1,\ 1,\ 1,\ 0,\ldots\ldots,0)=\frac{S_{n-1}}{N^2(n+2)(n+4)}\ \left[\frac{3n^3-11n^2+12n-4}{2}\right]$$

$$C_{1N}(6;\ 2,\ 2,\ 2,\ 0,\ldots\ldots\ldots,0)=\frac{S_{n-1}}{N^2(n+2)(n+4)}\ \left[\frac{n^4-n^3+6n^2-14n+8}{6}\right]$$

$$C_{1N}(6;\ 2,\ 2,\ 1,\ 1,\ 0,\ldots\ldots,0)=\frac{S_{n-1}}{N^2(n+2)(n+4)}\left[-\frac{n^4-7n^3+24n^2-38n+20}{4}\right]$$

$$C_{1N}(6;\ 2,\ 1,\ 1,\ 1,\ 1,\ 0,\ldots,0)=\frac{S_{n-1}}{N^2(n+2)(n+4)}$$

$$\left[\frac{n^4-10n^3+35n^2-50n+24}{8}\right].$$

$$C_{1N}(6;\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 0,\ \ldots,0)=\frac{S_{n-1}}{N^2(n+2)(n+4)}$$

$$\left[-\frac{n^4-10n^3+35n^2-50n+24}{48}\right]$$

3.3 Evaluation of D_{1N} and D^*_{1N}

Bradley gave the definitions

$$D_{1N}=\frac{N^{\frac{N+K}{2}}}{a_o!a_1!\ldots\ldots a_n!}\int_0^\infty f^{a_o}(x)\ f^{a_1}(x)\ldots\ldots\ldots f^{a_n}(x)\ x^{n+k}\ dx$$

where $f^{a_i}(x)$ is the a_i-th derivative of $f(x)$ and

$$D^*_{1N}=\frac{N^{\frac{N+K}{2}}}{a_o!a_1!\ldots\ldots\ldots a_n!}\int_0^\infty f^{a_o}(-x)\ f^{a_1}(-x)\ldots f^{a_n}(-x)\ x^{n+k}\ dx.$$

If the density $f(x)$ is symmetric, then it is clear that D_{1N} is equal to D^*_{1N}.

3.4 The application of the series expansion to a particular non-normal population.

In this section we shall consider the application of the above series expansion of the cumulative distribution function of t to a particular non-normal population. We will then compare the significance level of the t-test for the non-normal population with the significance level based on normal theory. The example we consider here (Bradley (5)) is the Cauchy distribution with density function $f(x)=\frac{1}{\pi(1+x^2)}$ $-\infty<x<\infty$. It satisfies the required conditions:

(a) $f(x)>0$ for all $-\infty<x<\infty$.

(b) $f(x)$ is a continuous function of x for $-\infty<x<\infty$.

(c) $\frac{d^k f(x)}{dx^k}$ is a continuous function of x for $-\infty<x<\infty$.

We may write this density in another form, namely

$$f(x)=\frac{1}{\pi(1+x^2)} = \pi^{-1}\,\frac{\partial \tan^{-1}x}{\partial x}.$$

Now we express the v-th derivative of $f(x)$ as $f^v(x)=\pi^{-1}v!$ $(-1)^v \sin^{v+1}y\,\sin(v+1)y$ where $\sin y=(1+x^2)^{-\frac{1}{2}}$, $\sin^{v+1}y$ its $(v+1)$th derivative, and $\cos y=\frac{x}{(1+x^2)^{\frac{1}{2}}}$. The coefficients D_{1N} can be calculated by the following formula (since the Cauchy distribution is symmetric, we observe $D_{1N}=D^*_{1N}$).

$$D_{1N}=\frac{N^{\frac{(N+K)}{2}}}{a_o!a_1!\cdots\cdots a_n!}\int_0^\infty f^{a_o}(x)f^{a_1}(x)\ldots f^{a_n}(x)x^{n+k}\,dx$$

where

$$f^{a_o}(x)f^{a_1}(x)\ldots\ldots f^{a_n}(x)=a_o!a_1!\ldots a_n!\pi^{-N}(-1)^k\sin^{N+k}y$$
$$\sin(a_o+1)y\ldots\sin(a_n+1)y.$$

By the relation $x=\cot y$ and the fact that the Jacobian $\frac{dx}{dy}=\frac{-1}{\sin^2 y}$ we find

$$D_{1N}=N^{\frac{N+K}{2}}\pi^{-N}(-1)^k\int_0^{2\pi}\cos^{n+k}y\,\sin(a_o+1)y\,\sin(a_1+1)\ldots$$
$$\sin(a_n+1)y\,\csc y\,dy$$

We further note that when k=0, $a_o + a_1 + a_2 + \ldots + a_n = 0$

$$D_{1N} = N^{\frac{N}{2}} \pi^{-N} \int_0^{\pi/2} \cos^n y \sin^n y \, dy$$

and by using the Beta function

$$\beta(x,y) = 2\int_0^{\pi/2} (\sin t)^{2x-1} (\cos t)^{2y-1} \, dt,$$

we obtain

$$D_{1N}(0;\ 0,\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots,0) = \frac{N^{\frac{N}{2}}}{2\pi^N} \beta\left(\frac{n+1}{2},\frac{n+1}{2}\right).$$

When k=2 and when $a_o=2$ with a_1, a_2, ..., a_n all zero, we find $D_{1N}(2;\ 2,\ 0,\ldots\ldots\ldots\ldots\ldots\ldots,0) = N^{\frac{N+2}{2}} \pi^{-N}$

$$\int_0^{\pi/2} \cos^{n+2} y \sin^{n-1} y \sin 3y \, dy.$$

Upon writing $\sin 3y = 3\sin y - 4\sin^3 y$ this at once becomes

$$D_{1N}(2;\ 2,\ 0,\ldots\ldots\ldots\ldots\ldots\ldots,0) = \frac{N^{\frac{N+2}{2}}}{\pi^N} \int_0^{\pi/2} (3\cos^{n+2} y \sin^n y - 4\cos^{n+2} y \sin^{n+2} y)\, dy$$

$$= \frac{N^{\frac{N+2}{2}}}{\pi^N} \left[\frac{3}{2}\, \beta\left(\frac{n+3}{2},\frac{n+1}{2}\right) - 2\, \beta\left(\frac{n+3}{2},\frac{n+3}{2}\right)\right].$$

For the calculation of other terms see the Appendix of Bradley (5). Results are summarized below for the remaining case of k=2 and for the required cases of k=4 and k=6.

k=2

$$D_{1N}(2;\ 1,\ 1,\ 0,\ldots\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+2}{2}}}{\pi^{N}}\ [2\beta(\tfrac{n+5}{2},\tfrac{n+1}{2})]$$

k=4

$$D_{1N}(4;\ 4,\ 0,\ \ldots\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+4}{2}}}{\pi^{N}}\ [\tfrac{5}{2}\quad(\tfrac{n+5}{2},\tfrac{n+1}{2})-10$$

$$\beta(\tfrac{n+5}{2},\tfrac{n+3}{2})\ +\ 8\beta\ (\tfrac{n+5}{2},\tfrac{n+5}{2})]$$

$$D_{1N}(4;\ 3,\ 1,\ 0,\ldots\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+4}{2}}}{\pi^{N}}\ [8\beta(\tfrac{n+9}{2},\tfrac{n+1}{2})-4$$

$$\beta(\tfrac{n+7}{2},\tfrac{n+1}{2})]$$

$$D_{1N}(4;\ 2,\ 2,\ 0,\ldots\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+4}{2}}}{\pi^{N}}\ [\tfrac{9}{2}\ \beta(\tfrac{n+5}{2},\tfrac{n+1}{2})-12$$

$$\beta(\tfrac{n+5}{2},\tfrac{n+3}{2})\ +\ 8\beta\ (\tfrac{n+5}{2},\tfrac{n+5}{2})]$$

$$D_{1N}(4;\ 2,\ 1,\ 1,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+4}{2}}}{\pi^{N}}\ [6\beta(\tfrac{n+7}{2},\tfrac{n+1}{2})-8$$

$$\beta(\tfrac{n+7}{2},\tfrac{n+3}{2})]$$

$$D_{1N}(4;\ 1,\ 1,\ 1,\ 1,\ 0,\ldots\ldots\ldots,0)=\frac{N^{\frac{N+4}{2}}}{\pi^{N}}\ [8\beta(\tfrac{n+9}{2},\tfrac{n+1}{2})]$$

k=6

$$D_{1N}(6;\ 6,\ 0,\ldots\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+4}{2}}}{\pi^{N}}\ [32\beta(\tfrac{n+13}{2},\tfrac{n+1}{2})-40$$

$$\beta(\tfrac{n+11}{2},\tfrac{n+1}{2})\ +\ 12\beta(\tfrac{n+9}{2},\tfrac{n+1}{2})-\tfrac{1}{2}\beta(\tfrac{n+7}{2},\tfrac{n+1}{2})]$$

$$D_{1N}(6;\ 5,\ 1,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+6}{2}}}{\pi^N}\ [32\beta(\tfrac{n+13}{2},\tfrac{n+1}{2})-32\beta(\tfrac{n+11}{2},\tfrac{n+1}{2})+6\beta(\tfrac{n+9}{2},\tfrac{n+1}{2})]$$

$$D_{1N}(6;\ 4,\ 2,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+6}{2}}}{\pi^N}\ [\tfrac{15}{2}\ \beta(\tfrac{n+7}{2},\tfrac{n+1}{2})-40\beta(\tfrac{n+7}{2},\tfrac{n+3}{2})+64\beta(\tfrac{n+7}{2},\tfrac{n+5}{2})-32\beta(\tfrac{n+7}{2},\tfrac{n+7}{2})]$$

$$D_{1N}(6;\ 4,\ 1,\ 1,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+6}{2}}}{\pi^N}\ [10\beta(\tfrac{n+9}{2},\tfrac{n+1}{2})-40\beta(\tfrac{n+9}{2},\tfrac{n+3}{2})+32\beta(\tfrac{n+9}{2},\tfrac{n+5}{2})]$$

$$D_{1N}(6;\ 3,\ 3,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+6}{2}}}{\pi^N}\ [32\beta(\tfrac{n+13}{2},\tfrac{n+1}{2})-32\beta(\tfrac{n+11}{2},\tfrac{n+1}{2})+8\beta(\tfrac{n+9}{2},\tfrac{n+1}{2})]$$

$$D_{1N}(6;\ 3,\ 2,\ 1,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+6}{2}}}{\pi^N}\ [24\beta(\tfrac{n+11}{2},\tfrac{n+1}{2})-12\beta(\tfrac{n+9}{2},\tfrac{n+1}{2})-32\beta(\tfrac{n+11}{2},\tfrac{n+3}{2})+16\beta(\tfrac{n+9}{2},\tfrac{n+3}{2})]$$

$$D_{1N}(6;\ 3,\ 1,\ 1,\ 1,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+6}{2}}}{\pi^N}\ [32\beta(\tfrac{n+13}{2},\tfrac{n+1}{2})-16\beta(\tfrac{n+11}{2},\tfrac{n+1}{2})]$$

$$D_{1N}(6;\ 2,\ 2,\ 2,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+6}{2}}}{\pi^N}\ [\tfrac{27}{2}\beta(\tfrac{n+7}{2},\tfrac{n+1}{2})-54\beta(\tfrac{n+7}{2},\tfrac{n+3}{2})+72\beta(\tfrac{n+7}{2},\tfrac{n+5}{2})-32\beta(\tfrac{n+7}{2},\tfrac{n+7}{2})]$$

$$D_{1N}(6;\ 2,\ 2,\ 1,\ 1,\ 0,\ldots\ldots\ldots\ldots,0)=\frac{N^{\frac{N+6}{2}}}{\pi^N}\ [18\beta(\tfrac{n+9}{2},\tfrac{n+1}{2})-48\,\beta(\tfrac{n+9}{2},\tfrac{n+3}{2})\ +\ 32\beta(\tfrac{n+9}{2},\tfrac{n+5}{2})]$$

$$D_{1N}(6;\ 2,\ 1,\ 1,\ 1,\ 1,\ 0,\ldots\ldots\ldots,0)=\frac{N^{\frac{N+6}{2}}}{\pi^N}\ [24\beta(\tfrac{n+11}{2},\tfrac{n+1}{2})-32\,\beta(\tfrac{n+11}{2},\tfrac{n+3}{2})]$$

$$D_{1N}(6;\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 0,\ldots\ldots,0)=\frac{N^{\frac{N+6}{2}}}{\pi^N}\ [32\beta(\tfrac{n+13}{2},\tfrac{n+1}{2})].$$

We now have to calculate the values of

$$\sum_{\substack{\text{comb.}\\ a_o+..+a_n=k}} C_{1N}(k;a_o,a_1,\ldots,a_n)D_{1N}(k;a_o,a_1,\ldots,a_n) \text{ and}$$

$$\sum_{\substack{\text{comb.}\\ a_o+..+a_n=k}} C_{1N}(k;a_o,a_1,\ldots,a_n)D^*_{1N}(k;a_o,a_1,\ldots,a_n) \text{ in order to}$$

calculate the cumulative distribution of t defined in [3.1] and [3.2] for samples from a Cauchy distribution.

When <u>k=0</u>

$$\sum_{\substack{\text{comb.}\\ a_o+..+a_n=k}} C_{1N}(k;a_o,a_1,\ldots,a_n)D_{1N}(k;a_o,a_1,\ldots,a_n)=C_{1N}(0;0,\ldots,0)$$

$$D_{1N}(0;0,\ldots0)=\frac{S_{n-1}\,N^{\frac{n+1}{2}}}{2\pi^N}\ \beta(\tfrac{n+1}{2},\tfrac{n+1}{2}).$$

When <u>k=2</u> we obtain

$$\sum_{\substack{\text{comb.}\\ a_o+..+a_n=k}} C_{1N}(k;a_o,a_1,\ldots,a_n)D_{1N}(k;a_o,a_1,\ldots,a_n)$$

$$=[C_{1N}(2;2,0,\ldots,0)D_{1N}(2;2,0,\ldots,0)+C_{1N}(2;1,1,0,\ldots0)$$

$$D_{1N}(2;1,1,0,\ldots,0)]$$

$$=\frac{S_{n-1}\,N^{\frac{N+2}{2}}}{\pi^N}\left[\frac{3}{2}\beta\left(\frac{n+3}{2},\frac{n+1}{2}\right)-2\beta\left(\frac{n+3}{2},\frac{n+3}{2}\right)-\beta\left(\frac{n+5}{2},\frac{n+1}{2}\right)\right].$$

By using the following relationships

$$\beta\left(\frac{n+3}{2},\frac{n+1}{2}\right)=\frac{1}{2}\,\beta\left(\frac{n+1}{2},\frac{n+1}{2}\right)$$

$$\beta\left(\frac{n+3}{2},\frac{n+3}{2}\right)=\frac{n+1}{4(n+2)}\,\beta\left(\frac{n+1}{2},\frac{n+1}{2}\right)$$

$$\beta\left(\frac{n+5}{2},\frac{n+1}{2}\right)=\frac{n+3}{4(n+2)}\,\beta\left(\frac{n+1}{2},\frac{n+1}{2}\right)$$

the result for k=2 may be written

$$\sum_{\substack{\text{comb.}\\ a_o+a_1+\ldots a_n=k}} C_{1N}(k;a_o,a_1,\ldots,a_n)D_{1N}(k;a_o,a_1,\ldots a_n)$$

$$=\frac{S_{n-1}\,N^{\frac{n+1}{2}}\,\beta\left(\frac{n+1}{2},\frac{n+1}{2}\right)}{4\pi^N}\left(\frac{n+1}{n+2}\right).$$

Similar reductions in the cases k=4 and k=6 yield, for k=4,

$$\sum_{\substack{\text{comb.}\\ a_oa_1+\ldots+a_n=k}} C_{1N}(k;a_o,\ldots a_n)D_{1N}(k;a_o,a_1,\ldots,a_n)$$

$$=\frac{S_{n-1}\,N^{\frac{n+1}{2}}}{48\pi^N}\,\beta\left(\frac{n+1}{2},\frac{n+1}{2}\right)\frac{3(n+1)(-2n^3-5n^2+29n+14)}{(n+2)^2(n+4)}$$

and for k=6

$$\sum_{\substack{\text{comb.}\\ a_o+a_1+\ldots+a_n=k}} C_{1N}(k;a_o,a_1,\ldots a_n)D_{1N}(k;a_o,\ldots,a_n)=\frac{S_{n-1}\,N^{\frac{n+1}{2}}}{1440\pi^N}$$

$$\beta\left(\frac{n+1}{2},\frac{n+1}{2}\right)\frac{15(n+1)(48+3274n+3927n^2-1460n^3-375n^4-14n^5)}{(n+2)^2(n+4)^2(n+6)}.$$

We have thus obtained all of the functions required for calculating the cumulative distribution function of t defined in formula [3.1] for samples from the Cauchy distribution. Placing these values into formula [3.1] we obtain

$$G_{1N}=1-\frac{N^{\frac{N}{2}}n^{\frac{n-2}{2}}}{\pi^{(N+1)/2}}\frac{\beta(\frac{n+1}{2},\frac{n+1}{2})}{\Gamma(\frac{n}{2})t^n}\Big[1+\frac{n^2(n+1)}{2(n+2)^2t^2}+\frac{n^3(n+1)(14+29n-5n^2-2n^3)}{2^2 2!(n+2)^2(n+4)^2t^4}$$

$$+\frac{n^4(n+1)(48+3274n+3927n^2-1460n^3-375n^4-14n^5)}{2^3 3!(n+2)^2(n+4)^2(n+6)^2t^6}+\text{--------}\Big]$$

for positive values of t.

As we have said for the Cauchy distribution $D_{1N}=D^*_{1N}$, we note that

$$\sum_{\substack{\text{comb.}\\ a_o+a_1+..+a_n=k}} C_{1N}(k;a_o,...a_n)D_{1N}(k;a_o,a_1,...a_n)$$

$$=\sum_{\substack{\text{comb.}\\ a_o+..+a_n=k}} C_{1N}(k;a_o,..a_n)D^*_{1N}(k;a_o,a_1,..a_n).$$

Thus for negative values of t,

$$G_N(t)=\frac{N^{\frac{N}{2}}n^{\frac{n-2}{2}}}{\pi^{(N+1)/2}}\frac{\beta(\frac{n+1}{2},\frac{n+1}{2})}{\Gamma(\frac{n}{2})t^n}\Big[1+\frac{n^2(n+1)}{2(n+2)^2t^2}+\frac{n^3(n+1)(14+29n-5n^2-2n^3)}{2^2 2!(n+2)^2(n+4)^2t^4}$$

$$+\frac{n^4(n+1)(48+3274n+3927n^2-1460n^3-375n^4-14n^5}{2^3 3!(n+2)^2(n+4)^2(n+6)^2t^6}+\text{-----------}\Big].$$

Combining the above results for negative and positive values of t, we obtain the following four terms of the required expansion

$$P(|t|>t_o)=\frac{2N^{\frac{N}{2}}n^{\frac{n-2}{2}}}{\pi^{\frac{N+1}{2}}}\;\frac{\beta(\frac{n+1}{2},\frac{n+1}{2})}{\Gamma(\frac{n}{2})\,t_o^{\,n}}\;[1+\frac{n^2(n+1)}{2(n+2)^2t_o^{\,2}}$$

$$+\frac{n^3(n+1)(14+29n-5n^2-2n^3)}{2^2 2!(n+2)^2(n+4)^2t_o^{\,4}}$$

$$+\frac{n^4(n+1)(48+3274n+3927n^2-1460n^3-375n^4-14n^5)}{2^3 3!(n+2)^2(n+4)^2(n+6)^2t_o^{\,6}}$$

$$+\text{-------}]\qquad\text{------------}\quad[3.3]$$

These probabilities have been calculated and are shown in Table 1 for n = 1, 2, 3, and 4 and for values of t_o corresponding to given significance levels obtained from "Student's" t-distribution. We denote by α the significance level of the two-sided "Student's" t-test, and denote by $P_C=P(|t|>t_o)$ the true level of significance in the case of samples from a Cauchy distribution using only the first four terms in [3.3].

Table 1 True significance levels (P_C) of the two-sided t-test for samples from a Cauchy distribution

Degrees of freedom n=N-1	t_o	α	P_C
1			
	1.000	0.5	0.47
	1.376	0.4	0.32

Degrees of freedom n=N-1	t_o	α	P_C
1			
	1.963	0.3	0.21
	3.708	0.2	0.13
	6.314	0.1	0.06
	12.076	0.05	0.03
	63.657	0.01	0.006
2			
	0.816	0.5	
	1.061	0.4	0.57
	1.386	0.3	0.27
	1.886	0.2	0.13
	2.920	0.1	0.05
	4.303	0.05	0.023
	9.925	0.01	0.004
3			
	0.765	0.5	
	0.978	0.4	
	1.250	0.3	0.3
	2.353	0.1	0.05
	3.182	0.05	0.002
	5.841	0.01	0.003
4			
	0.741	0.5	

Degrees of freedom $n=N-1$	t_o	α	P_C
4			
	0.941	0.4	
	1.190	0.3	
	1.533	0.2	
	2.132	0.1	0.05
	2.776	0.05	0.019
	4.604	0.01	0.002

The series given in equation [3.3] converges very slowly for large values of N and small values of t_o. It is for this reason that some values are omitted from the table. For example, using the first four terms only, one obtains, when $t_o=0.816$ and $n=2$, $P_C=1.47$, while when $t_o=0.765$ and $n=3$, $P_C=-9.561$.

IV. THE DISTRIBUTION OF t FOR SAMPLES FROM A NON-NORMAL POPULATION REPRESENTED BY THE EDGEWORTH SERIES

It happens frequently that we do not know the density function of a random variable X, but we do know the moments of its distribution. In such a case we cannot use the technique described in Chapter III to obtain the distribution of t. In this chapter we try another approach to solve this problem.

4.1 One-sample t-test

A. The Distribution of t

Let the density of the parent population be specified by the first four terms of the Edgeworth series

$$f(x)=\Phi(x)-\frac{\lambda_3}{3!}\,\Phi^3(x)+\frac{\lambda_4}{4!}\,\Phi^4(x)+\frac{10\lambda_3^2}{6!}\,\Phi^6(x)\ldots \quad [4.1]$$

where $\Phi(x)=\frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}}$, $\Phi^v(x)$ is its v-th derivative and

$\lambda_3=\sqrt{\beta_1}$, $\lambda_4=\beta_2-3$, where $\beta_1=\mu_3^2/\mu_2^3$ and $\beta_2=\mu_4/\mu_2^2$ are defined in terms of central moments $\mu_r=E(x-\mu)^r$ about the mean μ. The parameters λ_3 and λ_4 are sometimes called coefficients of the "skewness" and "kurtosis" respectively.

Gayen (13) derived the t-distribution for a single sample of size N randomly drawn from any population which can be represented by the first four terms of the Edgeworth series as specified by formula [4.1]. He gave the following form for the derived density function:

$$f(t)=f_o(t)+\lambda_3 f_{\lambda_3}(t)-\lambda_4 f_{\lambda_4}(t)+\lambda_3^2 f_{\lambda_3^2}(t)$$

where

$$f_{\lambda_3}(t)=\frac{[3(N-1)t-(2N-1)t^3]}{6(N-1)\ \sqrt{2N\pi}\,(1+\frac{t^2}{N-1})^{\frac{N+3}{2}}}$$

$$f_{\lambda_4}(t)=\frac{\Gamma(\frac{N+2}{2})}{24N\sqrt{(N-1)\pi}\,\Gamma(\frac{N+3}{2})}\quad\frac{3(N-1)-6(N+1)t^2+(N+1)t^4}{(1+\frac{t^2}{N-1})^{\frac{N+4}{2}}}$$

$$f_{\lambda_3^2}(t)=\frac{\Gamma(\frac{N+2}{2})\left\{\begin{array}{l}3(N-1)^2(2N+11)-9(N-1)(N+3)(2N-1)t^2\\ \quad -3(N+1)(N+3)(2N+13)t^4+(N+1)(N+3)(2N+5)t^6\end{array}\right\}}{144N(N-1)^{3/2}\sqrt{\pi}\ \Gamma(\frac{N+5}{2})\,(1+\frac{t^2}{N-1})^{\frac{N+6}{2}}}$$

and where $f_o(t)$ is the normal-theory density of t for a sample of size N. The terms $f_{\lambda_3}(t)$, $f_{\lambda_4}(t)$ and $f_{\lambda_3^2}(t)$ are the corrective terms due to the parameters λ_3 and λ_4.

B. Tail area probability

We are interested in the tail area of these corrective terms and so consider the intergrals

$$\int_{-\infty}^{t_o} f(t)\,dt \quad \text{and} \quad \int_{t_o}^{\infty} f(t)\,dt\ .$$

Integrals of the above form for "Student's" density $f_o(t)$

have been tabulated extensively. For example, a table accurate to four decimal places is provided by Owen (20).

The integral $f_{\lambda_3}(t)$ is antisymmetric. It is positive for the lower tail and negative for the upper tail. For two-sided tests it cancels itself out. The integrals $f_{\lambda_4}(t)$ and $f_{\lambda_3^2}(t)$ are symmetric. These integrals may be tabulated for given values of N and t_o, either directly from their algebraic expressions or more conveniently by using the Incomplete Beta function table of Pearson. For the sake of completeness we shall give here expressions for $F_{\lambda_3}(t_o)$, $F_{\lambda_4}(t_o)$ and $F_{\lambda_3^2}(t_o)$. Thus, for the lower tail, the corrected probability integrals $F(t_o)$ will be given by

$$F(t_o)=F_o(t_o)+\lambda_3 F_{\lambda_3}(t_o)-\lambda_4 F_{\lambda_4}(t_o)+\lambda_3^2 F_{\lambda_3^2}(t_o) \qquad [4.2]$$

in which $F_o(t_o)$ has been tabulated extensively, and where

$$F_{\lambda_3}(t_o)=\int_{-\infty}^{-t_o} f_{\lambda_3}(t)\,dt=-\int_{t_o}^{\infty} f_{\lambda_3}(t)\,dt=\frac{1}{6\sqrt{2N\Pi}}\,\frac{[1+\frac{2N-1}{N-1}t_o^2]}{(1+\frac{t_o^2}{N-1})^{\frac{N+1}{2}}}$$

$$=\left(\frac{2N-1}{6\sqrt{2N\Pi}}\right) I_{x_o}\left(\frac{N-1}{2}, 1\right)-\left(\frac{N-1}{3\sqrt{2N\Pi}}\right) I_{x_o}\left(\frac{N+1}{2}, 1\right) \qquad [4.3]$$

$$F_{\lambda_4}(t_o)=\int_{-\infty}^{-t_o} f_{\lambda_4}(t_o)\,dt=\int_{t_o}^{\infty} f_{\lambda_4}(t)\,dt$$

$$= \frac{1}{12(N-1)} \frac{\Gamma(\frac{N}{2})}{\sqrt{\pi(N-1)}\,\Gamma(\frac{N-1}{2})} \frac{t_o^3 - \frac{3t_o(N-1)}{(N+1)}}{(1+\frac{t_0^2}{N-1})^{\frac{N+2}{2}}}$$

$$= \frac{N-1}{24} I_{x_o}(\frac{N-1}{2},\frac{1}{2}) - \frac{(N-1)(N-2)}{12N} I_{x_o}(\frac{N+1}{2},\frac{1}{2}) + \frac{(N+4)(N-1)}{24N} I_{x_o}(\frac{N+3}{2},\frac{1}{2})$$

[4.4]

and

$$F_{\lambda_3^2}(t_o) = \int_{-\infty}^{-t_o} f_{\lambda_3^2}(t)\,dt = \int_{t_o}^{\infty} f_{\lambda_3^2}(t)\,dt = \frac{2N+5}{36(N-1)^2} \frac{\Gamma(\frac{N}{2})}{\sqrt{\pi(N-1)}\,\Gamma(\frac{N-1}{2})}$$

$$\times \frac{t_o^5 + \frac{2t_o^3(2N-7)(N-1)}{(N+1)(2N+5)} - \frac{3(2N+11)(N-1)^2 t_o}{(N+1)(N+3)(2N+5)}}{(1+\frac{t_o^2}{N-1})^{\frac{N+4}{2}}}$$

$$= [\frac{(N-1)(2N+5)}{72} I_{x_o}(\frac{N-1}{2},\frac{1}{2}) - \frac{(N-1)(2N^2+5N+8)}{24N} I_{x_o}(\frac{N-1}{2},\frac{1}{2})$$

$$+ \frac{(N-1)(2N^2+5N+12)}{24N} I_{x_o}(\frac{N+3}{2},\frac{1}{2}) - \frac{(N-1)(2N^2+5N+12)}{72N} I_{x_o}(\frac{N+5}{2},\frac{1}{2})]$$

[4.5]

where t_o is any typical value of t, its transformation is given by $x_o = \frac{1}{1-\frac{t_o^2}{N-1}}$, and $I_{x_o}(p,q)$ is the Incomplete Beta function as defined by Pearson.

In the following paragraph we will give an example to illustrate how to calculate these integrals for given values of the sample size N and critical point t_o.

Let us consider a sample size of five (N=5) and a significance level of 2.5%. The critical t_o obtained from the usual t table is 2.7764.

Now $x_o = \dfrac{1}{1-\dfrac{t_o^2}{N-1}} = 0.3416$ and from Pearson's Incomplete Beta Function Table we find

$$I_{x_o}\left(\frac{N-1}{2},\frac{1}{2}\right)=I_{0.3416}(2,\ 0.5)=0.025$$

$$I_{x_o}\left(\frac{N+1}{2},\frac{1}{2}\right)=I_{0.3416}(2,\ \ 1)=0.1167$$

$$I_{x_o}\left(\frac{N+1}{2},1\right)=I_{0.3416}(3,\ \ 1)=0.0399$$

$$I_{x_o}\left(\frac{N+3}{2},\frac{1}{2}\right)=I_{0.3416}(4,\ 0.5)=0.0044$$

$$I_{x_o}\left(\frac{N+5}{2},\frac{1}{2}\right)=I_{0.3416}(5,\ 0.5)=0.0013\ .$$

Using formulas [4.3], [4.4] and [4.5] we obtain $F_{\lambda_3}(t_o)$ $=0.0217$, $F_{\lambda_4}(t_o)=0.0028$, $F_{\lambda_3^2}(t_o)=0.0131$, while we already know that $F_o(t_o)=0.0250$. Knowing these integrals, we are able to calculate approximately the true significance levels of the t-test for samples of size five, at a significance level of 2.5%, and for non-normal populations specified by various values of the parameters λ_3 and λ_4. These significance levels are compared in Table 2 with those of "Student's" distribution, which is given by $\lambda_3=\lambda_4=0$. For a two-sided test in the case of a sample size of 5 at 2.5% nominal value, we note that

$$P(|t|\geq t_o)=1-\int_{-t_o}^{t_o} f(t)\,dt=2[F_o(t_o)-\lambda_4 F_{\lambda_4}(t_o)+\lambda_3^2 F_{\lambda_3^2}(t_o)] \qquad [4.6]$$

The function $F_{\lambda_3^2}(t_o)$ does not appear in formula (4.6), because it is cancelled out in two-sided tests.

Table 2 True significance levels for the two-sided t-test (sample of size 5) for various values of λ_3 and λ_4 when $\alpha=0.05$

λ_4 \ λ_3^2	0.00	0.04	0.09	0.16	0.25
0.0	0.0500	0.0510	0.0524	0.0542	0.0566
0.5	0.0472	0.0482	0.0496	0.0514	0.0534
1.0	0.0444	0.0454	0.0468	0.0486	0.0510
1.5	0.0416	0.0426	0.0440	0.0458	0.0482
2.0	0.0384	0.0398	0.0412	0.0430	0.0454

In the case of a one-sided test, the tail-area will differ for the upper tail and the lower tail. We note that

$$P(t<-t_o)=F_o(t_o)+\lambda_3^2 F_{\lambda_3^2}(t_o)-\lambda_4 F_{\lambda_4}(t_o)+\lambda_3^2 F_{\lambda_3^2}(t_o)$$

for the lower tail-area, whereas

$$P(t>t_o)=F_o(t_o)-\lambda_3 F_{\lambda_3}(t_o)-\lambda_4 F_{\lambda_4}(t_o)+\lambda_3^2 F_{\lambda_3^2}(t_o)$$

for the upper tail-area.

In Table 3 we show these tail-areas.

Table 3 True significance levels for the one-sided t-test (sample of size 5) for various values of λ_3 and λ_4 when $\alpha=0.025$

λ_4	λ_3	-0.2	0.0	0.2	0.3	0.4	0.5
	λ_3^2	0.04	0.0	0.04	0.09	0.16	0.25
0.0	Upper	0.0298	0.0250	0.0212	0.0197	0.0184	0.0175
	Lower	0.0212	0.0250	0.0298	0.0327	0.0358	0.0391
0.5	Upper	0.0284	0.0236	0.0198	0.0183	0.0170	0.0161
	Lower	0.0198	0.0236	0.0284	0.0313	0.0344	0.0377
1.0	Upper	0.0270	0.0222	0.0184	0.0169	0.0156	0.0147
	Lower	0.0184	0.0222	0.0270	0.0299	0.0330	0.0363
1.5	Upper	0.0256	0.0208	0.0170	0.0155	0.0142	0.0133
	Lower	0.0170	0.0208	0.0256	0.0285	0.0316	0.0349
2.0	Upper	0.0242	0.0192	0.0156	0.0141	0.0128	0.0119
	Lower	0.0156	0.0192	0.0242	0.0271	0.0302	0.0335

It can be seen from both Table 2 and Table 3 that the effect of skewness λ_3 is rather serious while that of kurtosis λ_4 is not. In the case of the one-sided tests, we find from

Table 3 that different proportions of the area fall in the two tails when the sampled population is asymmetric.

As the sample size increases, the level of significance in the case of a non-normal population with given skewness and kurtosis approaches that for a normal population. This may be observed by comparing the corresponding entries in Table 4, which follows, and Table 2.

Table 4 True significance levels for the two-sided t-test (sample of size 10) for various values of λ_3 and λ_4 when α=0.05

λ_4 \ λ_3^2	0.0	0.04	0.09	0.16	0.25
0.0	0.0500	0.0508	0.0516	0.0530	0.0548
0.5	0.0486	0.0494	0.0502	0.0516	0.0534
1.0	0.0472	0.0480	0.0488	0.0502	0.0520
1.5	0.0458	0.0466	0.0474	0.0488	0.0506
2.0	0.0444	0.0452	0.0460	0.0474	0.0492

From Table 4 we observe, for samples of size ten, that the effect of non-normality on the significance level is not very serious for the ranges of skewness and kurtosis parameters studied.

4.2 u-test (two sample t-test)

Let us assume that random samples are independently drawn from each of two populations with parameters λ'_3, λ'_4 and λ''_3, λ''_4, respectively, whose densities can be expressed by the first four terms of the Edgeworth series

$$f(x)=\phi(x)-\frac{\lambda_3}{3!}\phi^3(x)+\frac{\lambda_4}{4!}\phi^4(x)+\frac{\lambda_3^2}{72}\phi^6(x) .$$

Gayen (15) derived the density function of the u statistic defined in the introduction as follows:

$$f(u)=\frac{1}{\sqrt{n}\ \beta(\frac{1}{2},\frac{n}{2})}\frac{1}{(1+\frac{u^2}{n})^{\frac{n+1}{2}}}+\frac{(n_{30})u^3-3(n_{21})u}{3!n\sqrt{2\Pi(n+1)}\ (1+\frac{u^2}{n})^{\frac{n+4}{2}}}$$

$$+\frac{(n_{40})u^4-6(n_{31})u^2+3(n_{22})}{4!n^{\frac{5}{2}}\beta(\frac{1}{2},\frac{n+4}{2})(1+\frac{u^2}{n})^{\frac{n+5}{2}}}+\frac{(n_{60})u^6-15(n_{51})u^4+45(n_{42})u^2-15n(n_{33})}{72n^{\frac{5}{2}}\beta(\frac{1}{2},\frac{n+4}{2})(1+\frac{u^2}{n})^{\frac{n+7}{2}}}$$

--[4.7]

where $n=N_1+N_2-2$ is the number of degrees of freedom, and

$$(n_{30})=\frac{1}{2}(\lambda'_3+\lambda''_3)\ \frac{(N_1-N_2)}{\sqrt{N_1N_2}}\ (n+2)$$

$$+\frac{1}{2}(\lambda'_3-\lambda''_3)\ \left[\frac{(N_2-N_1)^2}{\sqrt{N_1N_2}}-\frac{2\sqrt{N_1N_2}(2n+1)}{(n+2)}\right]$$

$$(n_{21})=\frac{1}{2}(\lambda'_3+\lambda''_3)\ \frac{N_2-N_1}{\sqrt{N_1N_2}}\ (n+2)$$

$$+\frac{1}{2}(\lambda'_3-\lambda''_3)\left[\frac{(N_2-N_1)^2}{\sqrt{N_1N_2}}-\frac{2n\sqrt{N_1N_2}}{(n+2)}\right]$$

$$(n_{40})=\frac{1}{2}(\lambda_4'+\lambda_4'')\left[\frac{n(N_2-N_1)^2}{\sqrt{N_1N_2}}-\frac{2n^2}{n+3}\right]$$

$$+\frac{1}{2}(\lambda_4'-\lambda_4'')\ \frac{N_2-N_1}{n+2}\left[\frac{n(N_2-N_1)^2}{N_1N_2}-\frac{6n^2}{n+3}\right]$$

$$(n_{31})=\frac{1}{2}(\lambda_4'+\lambda_4'')\left[\frac{n(N_2-N_1)^2}{N_1N_2}-\frac{2n^2}{n+3}\right]$$

$$+\frac{1}{2}(\lambda_4'-\lambda_4'')\ \frac{N_2-N_1}{n+2}\left[\frac{n(N_2-N_1)^2}{N_1N_2}-\frac{2n(n-2)}{n+3}\right]$$

$$(n_{22})=\frac{1}{2}(\lambda_4'+\lambda_4'')\left[\frac{(N_2-N_1)^2}{N_1N_2}\ \frac{n^2}{n+2}-\frac{2n^3}{(n+2)(n+3)}\right]+\frac{1}{2}(\lambda_4'-\lambda_4'')\ \frac{N_2-N_1}{n+2}$$

$$\times\left[\frac{(N_2-N_1)^2}{N_1N_2}\ \frac{n^2}{(n+2)}+\frac{2n^2(n+4)}{(n+2)(n+3)}\right]$$

$$(n_{60})=\frac{1}{2}(\lambda_3'+\lambda_3'')^2\left[\frac{(N_2-N_1)^2(n-4)}{N_1N_2}+\frac{12n}{n+3}\right]$$

$$+\frac{1}{2}(\lambda_3'^2-\lambda_3''^2)\ \frac{N_2-N_1}{n+2}\left[\frac{(N_2-N_1)^2(n-4)}{N_1N_2}-\frac{4(n^2-4n+6)}{n+3}\right]+\frac{1}{4}(\lambda_3'-\lambda_3'')^2$$

$$\times\ \frac{(N_2-N_1)^2}{N_1N_2}\ \frac{[(n-4)(n+2)^2(n+3)-12N_1N_2n(n+1)]}{(n+2)^2(n+3)}$$

$$+\ \frac{4n[3(n+1)(n+2)^2+2N_1N_2(2n^2+3n+4)]}{(n+1)(n+2)^2(n+3)}$$

$$(n_{51})=\frac{1}{4}(\lambda_3'+\lambda_3'')^2\left[\frac{(N_2-N_1)^2(5n-8)}{5N_1N_2}+\frac{36n}{5(n+3)}\right]+\frac{1}{2}(\lambda_3'^2-\lambda_3''^2)$$

$$\times\frac{N_2-N_1}{n+2}\left[\frac{(N_2-N_1)^2(5n-8)}{5N_1N_2}-\frac{4(3n^2-4n+18)}{5(n+3)}\right]+\frac{1}{4}(\lambda_3'-\lambda_3'')^2$$

$$\times\frac{(N_2-N_1)^2}{5N_1N_2}\frac{[(5n-8)(n+3)(n+2)^2-4N_1N_2(n+1)(11n+12)]}{(n+2)^2(n+3)}$$

$$+\frac{4n[9(n+1)(n+2)^2+2N_1N_2n(2n-1)]}{5(n+1)(n+2)^2(n+3)}$$

$$(n_{42})=\frac{1}{4}(\lambda_3'+\lambda_3'')^2\left[\frac{(N_2-N_1)^2 n(5n+4)}{5N_1N_2(n+2)}+\frac{12n^2}{5(n+2)(n+3)}\right]$$

$$+\frac{1}{2}(\lambda'^2_3-\lambda''^2_3)\frac{N_2-N_1}{n+2}\left[\frac{(N_2-N_1)^2 n(5n+4)}{5N_1N_2(n+2)}-\frac{4n(n^2+2)}{5(n+2)(n+3)}\right]$$

$$+\frac{1}{4}(\lambda_3'-\lambda_3'')^2\ \Big\{\frac{(N_2-N_1)^2}{5N_1N_2}$$

$$\frac{[n(5n+4)(n+2)^2(n+3)-4nN_1N_2(n+1)(7n+16)]}{(n+2)^2(n+3)}$$

$$+\frac{4n^2[3(n+1)(n+2)^2-2N_1N_2(2n^2+9n+8)]}{5(n+1)(n+2)^3(n+3)}\Big\}$$

$$(n_{33})=\frac{1}{4}(\lambda_3'+\lambda_3'')^2\left[\frac{(N_2-N_1)^2 n^2(5n+16)}{5N_1N_2(n+2)(n+4)}-\frac{12n^3}{5(n+2)(n+3)(n+4)}\right]$$

$$+\frac{1}{2}(\lambda'^2_3-\lambda''^2_3)\frac{N_2-N_1}{n+2}\left[\frac{(N_2-N_1)^2 n^2(5n+16)}{5N_1N_2(n+2)(n+4)}+\frac{4n^2(n^2+8n+18)}{5(n+2)(n+3)(n+4)}\right]$$

$$+\frac{1}{4}(\lambda_3'-\lambda_3'')^2 \frac{(N_2-N_1)^2}{5N_1N_2} \left\{\frac{n^2(n+2)^2(n+3)(5n+16)-12N_1N_2n^2(n+1)(n+4)}{(n+2)^3(n+3)(n+4)}\right.$$

$$\left.-\frac{4n^3[3(n+1)(n+2)^2+2N_1N_2(n+4)(2n+1)]}{5(n+1)(n+2)^3(n+3)(n+4)}\right\}.$$

In the particular cases of samples of equal size drawn from the same or different populations, and samples of unequal size from the same population, the distribution of the u-statistic may be easily deduced from formula [4.7]. The distribution in the case of equal-sized samples from two different populations is studied in the following paragraph.

In the case for which $N_1=N_2=N$, $n=2N-2$, the density function of u will be given by

$$f(u)=\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})\,(1+\frac{u^2}{n})^{\frac{n+1}{2}}}+\frac{1}{2}(\lambda_3'-\lambda_3'')\frac{1}{6n\sqrt{2\pi(n+2)}}\frac{[3nu-(2n+1)u^3]}{(1+\frac{u^2}{n})^{\frac{n+4}{2}}}$$

$$-\frac{1}{2}(\lambda_4'+\lambda_4'')\frac{\Gamma(\frac{n+5}{2})}{12(n+3)\sqrt{n\pi}\,\Gamma(\frac{n+4}{2})}\frac{(u^4-6u^2+\frac{3u}{n+2})}{(1+\frac{u^2}{n})^{\frac{n+5}{2}}}$$

$$+\frac{1}{4}(\lambda_3'+\lambda_3'')^2\frac{\Gamma(\frac{n+3}{2})}{12n\sqrt{n\pi}\,\Gamma(\frac{n+4}{2})}\frac{(u^6-9u^4+\frac{9nu^2}{n+2}+\frac{3u^2}{(n+2)(n+4)})}{(1+\frac{u^2}{n})^{\frac{n+7}{2}}}$$

$$+\frac{1}{4}(\lambda_3'-\lambda_3'')^2 \quad \frac{\Gamma(\frac{n+1}{2})}{144n\sqrt{n\pi}\,\Gamma(\frac{n+4}{2})} \left\{ \frac{(2n^2+9n+10)u^6-3(2n^2+1711+18)u^4}{(1+\frac{u^2}{n})^{\frac{n+7}{2}}} \right.$$

$$\left. -\frac{\frac{9n(2n^2+3n+2)u^2}{n+2}-\frac{3(2n^2+15n+10)u^2}{(n+2)(n+4)}}{(1+\frac{u^2}{n})^{\frac{n+7}{2}}} \right\}$$

$$=f_o(u)+\frac{1}{2}(\lambda_3'-\lambda_3'')f_3(u)-\frac{1}{2}(\lambda_4'+\lambda_4'')f_4(u)+\frac{1}{4}(\lambda_3'+\lambda_3'')^2 f_5(u)$$

$$+\frac{1}{4}(\lambda_3'-\lambda_3'')^2 f_6(u) \qquad [4.8]$$

where $f_o(u)$ is the normal-theory density function of u. The expressions $f_i(u)$ remaining in formula (4.8) are called the corrective factors due to the population values of the λ-coefficients appearing with them.

The integrals $\int_{-\infty}^{-u_o} f_3(u)du$ and $\int_{u_o}^{\infty} f_3(u)du$ are equal in magnitude but opposite in sign, the former being positive. All other such corrective factors have symmetric integrals. Considering the negative tail of the distribution, we write

$$P(u<-u_o)=\int_{-\infty}^{-u_o} f(u)\,du=F_o(u_o)+\frac{1}{2}(\lambda_3'-\lambda_3'')F_3(u_o)-\frac{1}{2}(\lambda_4'+\lambda_4'')F_4(u_o)$$

$$+\frac{1}{4}(\lambda_3'+\lambda_3'')^2 F_5(u_o)+\frac{1}{4}(\lambda_3'-\lambda_3'')^2 F_6(u_o) \qquad [4.9]$$

where $F_o(u_o)=\int_{-\infty}^{-u_o} f_o(u)\ du=\frac{1}{2}\ I_{x_o}(\frac{n}{2},\frac{1}{2})$

is the normal-theory cumulative u distribution.

$$F_3(u_o)=\int_{-\infty}^{-u_o} f_3(u)\ du=\frac{1}{6\sqrt{2\Pi(n+2}}\ \frac{[1+\frac{(2n+1)u_o^2}{n}]}{(1+\frac{u_o^2}{n})^{\frac{n+2}{2}}}$$

$$=(\frac{2n+1}{6\sqrt{2\Pi(n+2)}})\ I_{x_o}(\frac{n}{2},1)-(\frac{n}{3\sqrt{2\Pi(n+2)}})\ I_{x_o}(\frac{n+2}{2},1)$$

$$F_4(u_o)=\int_{-\infty}^{-u_o} f_4(u)\ du=\frac{1}{12(n+2)^2 n^{\frac{1}{2}}\beta(\frac{n+2}{2},\frac{1}{2})}\ \frac{(n+2)u_o^3-3nu_o}{(1+\frac{u_o^2}{n})^{\frac{n+3}{2}}}$$

$$=\frac{n}{8(n+2)(n+3)}\ [I_{x_o}(\frac{n}{2},\frac{5}{2})-\ 2I_{x_o}(\frac{n+2}{2},\frac{3}{2})$$

$$+\ I_{x_o}(\frac{n+4}{2},\frac{1}{2})]$$

$$F_5(u_o)=\int_{-\infty}^{-u_o} f_5(u)\ du=\frac{1}{6(n+2)^2(n+4)n^{3/2}\beta(\frac{n+2}{2},\frac{1}{2})}$$

$$\frac{[u_o^5(n+2)(n+4)-4u_o^3 n(n+4)-3n^2 u_o]}{(1+\frac{u_o^2}{n})^{\frac{n+5}{2}}}$$

$$= \frac{n}{4(n+2)(n+3)(n+5)} [5\ I_{x_o}(\tfrac{n}{2},\tfrac{7}{2}) - 9\ I_{x_o}(\tfrac{n+2}{2},\tfrac{5}{2})$$

$$+ 3\ I_{x_o}(\tfrac{n+4}{2},\tfrac{3}{2}) + I_{x_o}(\tfrac{n+6}{2},\tfrac{1}{2})]$$

$$F_6(u_o) = \int_{-\infty}^{-u_o} f_6(u)\ du$$

$$= \frac{[(2n+5)(n+4)(n+2)^2 u_o^5 + 2(2n+1)(n+4)(n-2)nu_0^3 - 3n^2 u_o(2n^2+15n+10)]}{36(n+2)^2(n+4)n^{5/2}\beta(\frac{n}{2},\frac{1}{2})(1+\frac{u_o^2}{n})^{\frac{n+5}{2}}}$$

$$= \{5(2n^2+9n+10)I_{x_o}(\tfrac{n}{2},\tfrac{7}{2}) - 3(2n^2+17n+18)I_{x_o}(\tfrac{n+2}{2},\tfrac{5}{2})$$

$$- 3(2n^2+3n+2)I_{x_o}(\tfrac{n+4}{2},\tfrac{3}{2}) + (2n^2+15n+10)I_{x_o}(\tfrac{n+6}{2},\tfrac{1}{2})\}$$

$$\times \frac{n}{24(n+1)(n+2)(n+3)(n+5)}$$

where $I_{x_o}(p,q)$ is the Incomplete Beta function, and $x_o = \frac{1}{1+\frac{u_o^2}{n}}$. For the positive tail, we write

$$P(u>u_o) = \int_{u_o}^{\infty} f(u)\ du = F_o(u_o) - \tfrac{1}{2}(\lambda_3' - \lambda_3'')F_3(u_o) - \tfrac{1}{2}(\lambda_4' + \lambda_4'')F_4(u_o)$$

$$+ \tfrac{1}{4}(\lambda_3' + \lambda_3'')^2 F_5(u_o) + \tfrac{1}{4}(\lambda_3' - \lambda_3'')^2 F_6(u_o) \qquad [4.10]$$

In the two-sided case the corrective term $F_3(u_o)$ cancels, and we write $P(|u|\geq u_o)=2F_o(u_o)-(\lambda_4'+\lambda_4'')F_4(u_o)+\frac{1}{2}(\lambda_3'+\lambda_3'')^2F_5(u_o)$

$$+\frac{1}{2}(\lambda_3'-\lambda_3'')^2F_6(u_o)\,. \qquad [4.11]$$

Thus, we can calculate approximately the true significance level for samples of equal size from two different populations the density function of which can be approximated by the first four terms of Edgeworth series. The values of $F_3(u_o)$, $F_4(u_o)$, $F_5(u_o)$ and $F_6(u_o)$ at the 2.5% nominal level of significance for n = 2, 4, 6, 8, 12, 20, 30, 40, 60, 120 and ∞ has been tabulated by Gayen (15) in Table 5 of his paper. His table is wrong, however, in column 5 and column 6. His column 5 and column 6 should be interchanged. We will show in the following Table 5 the corrective values of $F_3(u_o)$, $F_4(u_o)$, $F_5(u_o)$ and $F_6(u_o)$ at the 2.5% nominal level of significance for n = 2, 4, 6, 8, 12, 20, 24, 30, and 40.

Table 5 Correction factors due to population skewness and kurtosis for approximating the true significance level for the u-test at the 2.5% nominal level of significance for samples of equal size

n degrees of freedom	u_o	Correction Factors			
		$F_3(u_o)$	$F_4(u_o)$	$F_5(u_o)$	$F_6(u_o)$
2	4.3027	0.0149	0.0024	0.0042	0.0094
4	2.7764	0.0199	0.0024	0.0027	0.0113
6	2.4469	0.0206	0.0019	0.0014	0.0103
8	2.3060	0.0202	0.0015	0.0007	0.0090
12	2.1788	0.0188	0.0010	0.0002	0.0071
20	2.0860	0.0161	0.0006	0.0000	0.0049
24	2.0639	0.0151	0.0005	0.0000	0.0042
30	2.0423	0.0139	0.0003	0.0000	0.0035
40	2.0211	0.0123	0.0003	0.0000	0.0027

Using the correction factors shown in Table 5 and the appropriate formulas shown in [4.9], [4.10] and [4.11], we can approximate the true significance levels in the case of equal-sized samples for one-sided or two-sided tests. In Table 6 we show the approximate true significance levels in the case of samples of size 5 from each of two different mesokurtic populations ($\lambda_4' = \lambda_4'' = 0$). A nominal significance level of 2.5% was assumed for the case of one-sided test and a 5% level of significance for the two-sided test. In Table 7 we show such significance levels for similar samples but from two symmetric populations ($\lambda_3' = \lambda_3'' = 0$).

Table 6 True significance levels for the one-sided u-test (α=0.025) and the two-sided u-test (α=0.05) for samples of size 5 from two different mesokurtic ($\lambda_4'=\lambda_4''=0$) populations

λ_3'' \ λ_3'		-0.2	0.0	0.2	0.4	0.5
-0.2	Upper tail	0.0250	0.0231	0.0214	0.0197	0.0190
	Lower tail	0.0250	0.0271	0.0294	0.0319	0.0332
	Both tails	0.0500	0.0502	0.0508	0.0516	0.0522
0.0	Upper tail	0.0271	0.0250	0.0231	0.0214	0.0205
	Lower tail	0.0231	0.0250	0.0271	0.0294	0.0307
	Both tails	0.0502	0.0500	0.0502	0.0508	0.0512
0.2	Upper tail	0.0294	0.0271	0.0250	0.0232	0.0223
	Lower tail	0.0214	0.0231	0.0250	0.0272	0.0283
	Both tails	0.0508	0.0502	0.0500	0.0504	0.0506
0.4	Upper tail	0.0319	0.0294	0.0272	0.0251	0.0241
	Lower tail	0.0197	0.0214	0.0232	0.0251	0.0261
	Both tails	0.0516	0.0508	0.0504	0.0502	0.0502
0.5	Upper tail	0.0332	0.0307	0.0283	0.0261	0.0252
	Lower tail	0.0190	0.0205	0.0223	0.0241	0.0252
	Both tails	0.0522	0.0512	0.0506	0.0502	0.0504

Table 7 True significance levels for the two-sided u-test (α=0.05) for samples of size 5 from two symmetric ($\lambda_3'=\lambda_3''=0$) distributions

λ_4'' \ λ_4'	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5
-1.5	0.0545	0.0538	0.0530	0.0523	0.0515	0.0508	0.0500
-1.0	0.0538	0.0530	0.0523	0.0515	0.0508	0.0500	0.0492
-0.5	0.0530	0.0523	0.0515	0.0508	0.0500	0.0492	0.0485
0.0	0.0523	0.0515	0.0508	0.0500	0.0492	0.0485	0.0478
0.5	0.0515	0.0508	0.0500	0.0492	0.0485	0.0478	0.0470
1.0	0.0508	0.0500	0.0492	0.0485	0.0478	0.0470	0.0463
1.5	0.0500	0.0492	0.0485	0.0478	0.0470	0.0463	0.0455

The effect of non-normality in this test is seen not to be very serious. Using Table 6 we observe that the significance level will be slightly afftected if the difference in the skewness λ_3 of the sampled populations is not small. For the case in which samples are taken from the symmetrical populations or from the same population with moderate skewness, the effect of non-normality is small, and the normal theory u-test appears to be valid.

V. VARIANCE-RATIO TEST

The variance-ratio is commonly used for testing the equality of a set of means in the fixed model, one-way classification analysis of variance. It is also used for testing the equality of two variances.

5.1 The analysis of variance test for the case of non-normal errors

In the one-way classification analysis of variance, we assume the following model

$$x_{ij}=m+B_i+e_{ij}$$

where x_{ij} is the observation in the i-th row and j-th column $i=1, 2, 3,.....,k$ and $j=1, 2, 3,.....n_i$, m is the overall effect, B_i is the row effect and e_{ij} is the experimental error.

If the e_{ij} are normally and independently distributed, the ratio of "between treatment" and "within treatment" mean squares

$$w=\frac{\sum_{i=1}^{k} n_i(x_{i\circ}-\bar{x}_{\circ\circ})^2/(k-1)}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_{i\cdot})^2/(N-k)}$$

is distributed as F with k-1 and N-k degrees of freedom. Here $\bar{x}_{i\circ}$ is the sample mean of observations receiving the i-th treatment, $\bar{x}_{\circ\circ}$ is the grand mean, n_1 is the sample size for the i-th treatment, and $N = \sum_{i=1}^{k} n_i$ is the total number of observations.

The analysis of variance test w is based on several assumptions; for example, see Eisenhart (12). One important assumption is normality. That is, the experimental errors e_{ij} are normally distributed. In the following paragraphs we will study the robustness to non-normality of these errors in the analysis of variance test.

Consider k groups of observations x_{ij} ($i=1, 2,\ldots,k$; $j=1, 2,\ldots,n_i$) drawn from a population the density function of which may be approximated by the first four terms of the Edgeworth series

$$f(x)=\Phi(x)-\frac{\lambda_3}{3!}\Phi^3(x)+\frac{\lambda_4}{4!}\Phi^4(x)+\frac{\lambda_3^2}{72}\Phi^6(x)$$

Gayen (14) derived the density function of w as

$$f(w)=P(\tfrac{N_1}{2},\tfrac{N_2}{2};w)-\lambda_4(v_{22})[P(\tfrac{N_1+4}{2},\tfrac{N_2}{2};w)-2P(\tfrac{N_1+2}{2},\tfrac{N_2+2}{2};w)$$

$$+P(\tfrac{N_1}{2},\tfrac{N_2+4}{2},w)]+\lambda_3^2[(v_{31})P(\tfrac{N_1+6}{2},\tfrac{N_2}{2};w)-3(v_{32})P(\tfrac{N_1+4}{2},\tfrac{N_2+2}{2};w)$$

$$+3(v_{33})P(\tfrac{N_1+2}{2},\tfrac{N_2+4}{2};w)-(v_{34})P(\tfrac{N_1}{2},\tfrac{N_2+6}{2};w)]$$

$$= f_o(w)-\lambda_4 f_{\lambda_4}(w) + \lambda_3^2 f_{\lambda_3^2}(w) \qquad [5.1]$$

where $f_o(w)=P(\frac{N_1}{2},\frac{N_2}{2};w)=\frac{1}{\beta(\frac{N_1}{2},\frac{N_2}{2})}\;\frac{w^{\frac{N_1}{2}-1}}{(1+\frac{N_1}{N_2}w)^{\frac{N_1+N_2}{2}}}$

is the F density and the other terms $f_{\lambda_4}(w)$ and $f_{\lambda_3^2}(w)$ may be called correction functions due to population parameters λ_3 and λ_4. The remaining symbols are defined as follows:

$$V_{22}=\frac{[2N_1N_2+(k^2-k'^2)(N_1+N_2+2)]}{8(N_1+N_2+1)(N_1+N_2+2)}$$

$$V_{31}=\frac{\{2N_1N_2[N_1(N_1+N_2+7)-(N_2-22)]+(k^2-k'^2)(N_1+N_2+2)(4N_1-5N_2+16)\}}{24(N_1+N_2+1)(N_1+N_2+2)(N_1+N_2+4)}$$

$$V_{32}=\frac{\{2N_1N_2[N_1(N_1+N_2+7)-(N_2-10)]+(k^2-k'^2)(N_1+N_2+2)(4N_1-5N_2+4)\}}{24(N_1+N_2+1)(N_1+N_2+2)(N_1+N_2+4)}$$

$$V_{33}=\frac{\{2N_1N_2[N_1(N_1+N_2+7)-(N_2+2)]+(k^2-k'^2)(N_1+N_2+2)(4N_1-5N_2-8)\}}{24(N_1+N_2+1)(N_1+N_2+2)(N_1+N_2+4)}$$

$$V_{34}=\frac{\{2N_1N_2[N_1(N_1+N_2+7)-(N_2+14)]+(k^2-k'^2)(N_1+N_2+2)(4N_1-5N_2-20)\}}{24(N_1+N_2+1)(N_1+N_2+2)(N_1+N_2+4)}$$

where $k'^2=N\sum_{i=1}^{k}(\frac{1}{n_i})$ [$k'^2=k^2$ if n_i's are equal], $\Delta k^2=k^2-k'^2$.

We are now in a position to evaluate the tail-area of the above density function of w. Consider the following function

$$F(w_o)=\int_{w_o}^{\infty}f(w)\,dw=F_o(w_o)+\lambda_4[P'_{\lambda_4}(w_o)+(\Delta k^2)P''_{\lambda_4}(w_o)]$$

$$+\lambda_3^2[P'_{\lambda_3^2}(w_o)+(\Delta k^2)P_{\lambda_3^2}(w_o)] \qquad [5.2]$$

where $F_o(w_o)$ is the tail-area of the F-distribution, and the other terms are correction functions due to population parameters λ_3 and λ_4. The functions $P'_{\lambda_4}(w_o)$, $P''_{\lambda_4}(w_o)$, $P'_{\lambda_3^2}(w_o)$ and $P_{\lambda_3^2}(w_o)$ have been tabulated for values of $N_1 = 1, 6, 8, 12, 24$ and $N_2 = 1, 6, 8, 12, 20, 24, 30, 40, 60, 120$ and ∞ by Gayen (14) at the 5% significance level of the normal-theory w. We give these values in Table 10 at the end of this chapter.

In the following paragraph we apply formula [5.2] and the values shown in Table 10 to tabulate the approximate true significance levels of the test of analysis of variance in the case of observations drawn from non-normal populations. For simplification, we consider only the case of equal numbers of observations in each group. The case we consider is five groups of five observations each. Thus the term Δk^2 is formula [5.2] is zero, thus simplifying our computation. The results are shown in Table 8.

Table 8 True signification levels for the variance-ratio test (w) with four and twenty degrees of freedom when α=0.05

λ_4 \ λ_3	0.00	0.10	0.25	0.50
-1.5	0.0536	0.0537	0.0538	0.0541
-1.0	0.0524	0.0525	0.0526	0.0529
-0.5	0.0512	0.0513	0.0514	0.0517
0.0	0.0500	0.0501	0.0502	0.0505
0.5	0.0488	0.0489	0.0490	0.0493
1.0	0.0476	0.0477	0.0478	0.0481
1.5	0.0464	0.0465	0.0466	0.0469

It can be seen that non-normality of the error terms exerts only a slight effect on level of significance in the analysis of variance test. This means that the F-test for the one-way classification analysis of variance is robust to the normality assumption. From Table 8 we also note that while the w-test is practically unaffected by lack of symmetry, it is slightly affected if the distribution of experimental errors is either very flat or very peaked as indicated by λ_4.

It is well known that the t-test is a particular case of the analysis of variance test procedure. Therefore, for the case k=2, the conclusion is the same as in our previous discussion of the u-test (two sample t-test).

5.2 The test for the equality of two variances

A. Distribution of the variance-ratio v for two independent samples

Here we assume that $x_1, x_2, \ldots\ldots, x_{n_1}$ and $y_1, y_2 \ldots y_{n_2}$ are two samples drawn from two different populations with parameters λ_3', λ_4' and λ_3'', λ_4'', respectively, and that their population density functions are specified by the first four terms of the Edgeworth series.

Let

$$v = \frac{s_1^2}{s_2^2} = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})/(n_1 - 1)}{\sum_{j=1}^{n_2} (y_j - \bar{y})/(n_2 - 1)}$$

Gayen (14) found the distribution of v to be given by

$$f(v) = P\left(\frac{n_1}{2}, \frac{n_2}{2}; v\right) + [N_{21} P\left(\frac{n_1+4}{2}, \frac{n_2}{2}; v\right) - 2N_{22} P\left(\frac{n_1+2}{2}, \frac{n_2+2}{2}; v\right)$$

$$+ N_{23} P\left(\frac{n_1}{2}, \frac{n_2+4}{2}; v\right) - [N_{31} P\left(\frac{n_1+6}{2}, \frac{n_2}{2}; v\right) - 3N_{32})\left(\frac{n_1+4}{2}, \frac{n_2+2}{2}; v\right)$$

$$+ 3N_{33} P\left(\frac{n_1+2}{2}, \frac{n_2+4}{2}; v\right) - N_{34} P\left(\frac{n_1}{2}, \frac{n_2+6}{2}; v\right)$$

$$= f_o(v) + \frac{1}{2}(\lambda_4'' + \lambda_4') f_{\overline{\lambda}_4}(v) + \frac{1}{2}(\lambda_4'' - \lambda_4') f_{\underline{\lambda}_4}(v) - \frac{1}{2}(\lambda''^2_3 + \lambda'^2_3) f_{\lambda_3^2}(v)$$

$$-\frac{1}{2}(\lambda''^2_3 - \lambda'^2_3) \; f_{\underline{\lambda}_3^2}(v) \qquad [5.3]$$

where $f_o(v)=P(\frac{n_1}{2},\frac{n_2}{2};v)=\frac{1}{\beta(\frac{n_1}{2},\frac{n_2}{2})}\ \frac{v^{\frac{n_1}{2}-1}}{(1+\frac{n_1}{n_2}v)^{\frac{n_1+n_2}{2}}}$ as the F

density and

$$N_{21}=\frac{n_1n_2\{\frac{1}{2}(\lambda_4''+\lambda_4')[n_1n_2(n_1+n_2+2)-2(n_1-n_2)] -\frac{1}{2}(\lambda_4''-\lambda_4')[n_1n_2(n_1-n_2+4)+2(n_1+n_2)]\}}{8(n_1+1)(n_2+1)(n_1+n_2)(n_1+n_2+2)}$$

$$N_{22}=\frac{n_1n_2\{\frac{1}{2}(\lambda_4''+\lambda_4')[n_1n_2(n_1+n_2+2)]-\frac{1}{2}(\lambda_4''-\lambda_4')[n_1n_2(n_1-n_2)]\}}{8(n_1+1)(n_2+1)(n_1+n_2)(n_1+n_2+2)}$$

$$N_{23}=\frac{n_1n_2\{\frac{1}{2}(\lambda_4''+\lambda_4')[n_1n_2(n_1+n_2+2)+2(n_1-n_2)] -\frac{1}{2}(\lambda_4''-\lambda_4')[n_1n_2(n_1-n_2-4)-2(n_1+n_2)]\}}{8(n_1+1)(n_2+1)(n_1+n_2)(n_1+n_2+2)}$$

$$N_{31}=\frac{n_1n_2\{\frac{1}{2}(\lambda''^2_3+\lambda'^2_3)[(n_1+1)(n_2+2)(n_1+4)(n_2-1)-(n_1-1)(n_2+1)(n_2-2)(n_2-4)] +\frac{1}{2}(\lambda''^2_3-\lambda'^2_3)[2(n_1n_2-1)(n_1+2)(n_1+4)+(n_1+n_2)(n_1-1)(n_2+1)(n_2-n_1-6)]\}}{12(n_1+1)(n_2+1)(n_1+n_2)(n_1+n_2+2)(n_1+n_2+4)}$$

$$N_{32}=\frac{n_1n_2\{\frac{1}{2}(\lambda''^2_3+\lambda'^2_3)[n_1(n_1+1)(n_2+2)(n_2-1)-n_2(n_2+1)(n_2-2)(n_1-1)] +\frac{1}{2}(\lambda''^2_3-\lambda'^2_3)[n_1(n_1+1)(n_1+2)(n_2-1)+n_2(n_2+1)(n_2-2)(n_1-1)]\}}{12(n_1+1)(n_2+1)(n_1+n_2)(n_1+n_2+2)(n_1+n_2+4)}$$

$$N_{33}=\frac{\begin{array}{l}n_1n_2\{\frac{1}{2}(\lambda''^2_3+\lambda'^2_3)[n_1(n_1+1)(n_1-2)(n_2-1)-n_2(n_2+1)(n_2+2)(n_1-1)\\ +\frac{1}{2}(\lambda''^2_3-\lambda'^2_3)[n_1(n_1+1)(n_2-2)(n_2-1)+n_2(n_2+1)(n_2+2)(n_1-1)]\}\end{array}}{12(n_1+1)(n_2+1)(n_1+n_2)(n_1+n_2+2)(n_1+n_2+4)}$$

$$N_{34}=\frac{\begin{array}{l}n_1n_2\{\frac{1}{2}(\lambda''^2_3+\lambda'^2_3)[(n_1+1)(n_1-2)(n_1-4)(n_2-1)-(n_1-1)(n_2+1)(n_2+2)(n_2+4)]\\ +\frac{1}{2}(\lambda''^2_3-\lambda'^2_3)[2(n_1n_2-1)(n_2+2)(n_2+4)+(n_1+n_2)(n_1+1)(n_2-1)(n_1-n_2+6)]\}\end{array}}{12(n_1+1)(n_2+1)(n_1+n_2)(n_1+n_2+2)(n_1+n_2+4)}$$

B. The Tail-area

We are interested in the tail-area; and, therefore, we consider the function

$$F(v_o)=\int_{v_o}^{\infty} f(v)dv=F_o(v_o)+\frac{1}{2}(\lambda''_4+\lambda'_4)F_{\bar{\lambda}_4}(v_o)+\frac{1}{2}(\lambda''_4-\lambda'_4)F_{\underline{\lambda}_4}(v_o)$$

$$-\frac{1}{2}(\lambda''^2_3+\lambda'^2_3)F_{\bar{\lambda}^2_3}(v_o)-\frac{1}{2}(\lambda''^2_3-\lambda'^2_3)F_{\underline{\lambda}^2_3}(v_o) \qquad [5.4]$$

The functions $F_\lambda(v_o)$ are tabulated in Table 10 for N_1=1, 2, 3, 4, 5, 6, 8, 12, 24 and N_2=1, 2, 3, 4, 5, 6, 8, 12, 20, 24, 30, 40, 60, 120 and ∞. Using formula [5.4] and Table 10 we can calculate approximatley the true significance levels for the F-test of the equality of two variances in the case of non-normal errors. In Table 9 we show these true levels of significance for the case n_1=5, n_2=21, $\lambda'_3=\lambda''_3$ and $\lambda'_4=\lambda''_4$. The comparison leads to an F-test with four and twenty degrees of freedom. In the normal the significance limits are the same at the 5% level as for the analysis of variance test discussed in Section 5.1.

Table 9 True significance levels for the variance-ratio test (v) with four and twenty degrees of freedom when $\alpha=0.05$

λ_4 \ λ_3^2	0.00	0.10	0.25	0.50
-1.5	0.0112	0.0107	0.0099	0.0087
-1.0	0.0241	0.0236	0.0228	0.0216
-0.5	0.0371	0.0366	0.0358	0.0346
0.0	0.0500	0.0495	0.0487	0.0475
0.5	0.0630	0.0625	0.0617	0.0605
1.0	0.0759	0.0754	0.0746	0.0734
1.5	0.0889	0.0884	0.0876	0.0864

As will be seen by comparison with Table 8, the effect of non-normality on this test is much more serious than in the analysis of variance test. Box (4) showed that the sensitivity to non-normality is even greater when the number of variances to be compared exceeds two. The discrepancy does not decrease if the sample size is increased or if the samples contain an equal number of observations. It can be noted that it is the parameter λ_4 which affects the robustness of the F test for equality of two variances, whereas λ_3 has little effect.

Table 10 Corrective functions for determining the tail probabilities of (i)w and (ii)v in terms of their upper 5% normal-theory significance levels

N_2	(Test for equality of a set of means)				(Test for equality of two variances)			
	$P'_{\lambda_4}(w_o)$	$P''_{\lambda_4}(w_o)$	$P'_{\lambda_3^2}(w_o)$	$P''_{\lambda_3^2}(w_o)$	$F_{\bar{\lambda}_4}(v_o)$	$F_{\underline{\lambda}_4}(v_o)$	$F_{\bar{\lambda}_3^2}(v_o)$	$F_{\underline{\lambda}_3^2}(v_o)$
				$N_1=1$				
1	-0.0027	-0.0055	0.0054	0.0055	0.0021	-0.0042	-0.0000	0.0000
2	-0.0048	-0.0060	0.0084	0.0047	0.0066	-0.0077	0.0022	-0.0022
3	-0.0052	-0.0052	0.0073	0.0037	0.0096	-0.0091	0.0034	-0.0034
4	-0.0048	-0.0042	0.0055	0.0031	0.0110	-0.0093	0.0037	-0.0037
5	-0.0043	-0.0034	0.0039	0.0028	0.0116	-0.0089	0.0035	-0.0035
6	-0.0038	-0.0028	0.0028	0.0025	0.0117	-0.0083	0.0032	-0.0032
8	-0.0030	-0.0020	0.0015	0.0021	0.0114	-0.0070	0.0025	-0.0025
12	-0.0020	-0.0012	0.0005	0.0016	0.0103	-0.0048	0.0016	-0.0016
20	-0.0011	-0.0006	0.0001	0.0011	0.0086	-0.0021	0.0008	-0.0008
24	-0.0009	-0.0005	0.0000	0.0010	0.0080	-0.0013	0.0006	-0.0006
30	-0.0007	-0.0004	0.0000	0.0008	0.0074	-0.0004	0.0004	-0.0004
40	-0.0005	-0.0003	0.0000	0.0006	0.0067	0.0006	0.0002	-0.0002
60	-0.0003	-0.0002	0.0000	0.0004	0.0059	0.0016	0.0001	-0.0001
120	-0.0002	-0.0001	0.0000	0.0001	0.0050	0.0028	0.0001	-0.0001
∞	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
				$N_1=2$				
1	-0.0023	-0.0029	0.0053	0.0031	0.0021	-0.0042	-0.0017	-0.0017
2	-0.0043	-0.0032	0.0091	0.0026	0.0071	-0.0079	0.0024	-0.0024
3	-0.0050	-0.0029	0.0094	0.0019	0.0112	-0.0095	0.0042	-0.0037
4	-0.0052	-0.0026	0.0082	0.0015	0.0142	-0.0099	0.0052	-0.0042
5	-0.0050	-0.0023	0.0066	0.0012	0.0159	-0.0095	0.0055	-0.0039
6	-0.0047	-0.0020	0.0053	0.0010	0.0170	-0.0087	0.0055	-0.0035
8	-0.0041	-0.0015	0.0033	0.0009	0.0179	-0.0068	0.0052	-0.0024
12	-0.0030	-0.0010	0.0013	0.0007	0.0180	-0.0033	0.0046	-0.0005
20	-0.0019	-0.0006	0.0002	0.0005	0.0170	0.0012	0.0040	0.0014
24	-0.0016	-0.0005	0.0000	0.0005	0.0167	0.0026	0.0039	0.0019
30	-0.0013	-0.0004	-0.0001	0.0004	0.0161	0.0042	0.0038	0.0024
40	-0.0010	-0.0003	-0.0001	0.0003	0.0154	0.0060	0.0060	0.0029
60	-0.0006	-0.0002	-0.0001	0.0002	0.0146	0.0079	0.0038	0.0034
120	-0.0003	-0.0001	-0.0001	0.0001	0.0136	0.0101	0.0039	0.0038
∞	0.0000	0.0000	0.0000	0.0000	0.0124	0.0124	0.0042	0.0042

Table 10 (continued)

N_2	(Test for equality of a set of means)				(Test for equality of two variances)			
	$P'_{\lambda_3}(w)$	$P''_{\lambda_4}(w)$	$P'_{\lambda_3^2}(2)$	$P''_{\lambda_3^2}(w)$	$F_{\bar{\lambda}_4}(v_o)$	$F_{\underline{\lambda}_4}(v_o)$	$F_{\bar{\lambda}_3^2}(v_o)$	$F_{\underline{\lambda}_3^2}(v_o)$
				$N_1=3$				
1	-0.0020	-0.0020	0.0051	0.0022	0.0022	-0.0041	-0.0018	-0.0018
2	-0.0037	-0.0022	0.0091	0.0019	0.0074	-0.0080	0.0024	-0.0025
3	-0.0046	-0.0020	0.0102	0.0014	0.0102	-0.0099	0.0046	-0.0040
4	-0.0049	-0.0018	0.0095	0.0010	0.0154	-0.0105	0.0058	-0.0046
5	-0.0049	-0.0016	0.0082	0.0008	0.0178	-0.0101	0.0063	-0.0046
6	-0.0047	-0.0014	0.0069	0.0006	0.0194	-0.0093	0.0064	-0.0042
8	-0.0042	-0.0012	0.0047	0.0005	0.0213	-0.0073	0.0062	-0.0031
12	-0.0034	-0.0008	0.0022	0.0003	0.0225	-0.0031	0.0056	-0.0010
20	-0.0023	-0.0005	0.0005	0.0003	0.0225	0.0026	0.0050	0.0015
24	-0.0019	-0.0004	0.0002	0.0002	0.0222	0.0046	0.0049	0.0022
30	-0.0016	-0.0003	0.0000	0.0002	0.0218	0.0067	0.0048	0.0029
40	-0.0012	-0.0002	-0.0002	0.0002	0.0213	0.0092	0.0048	0.0036
60	-0.0008	-0.0002	-0.0002	0.0001	0.0206	0.0119	0.0049	0.0043
120	-0.0004	-0.0001	-0.0002	0.0001	0.0196	0.0151	0.0051	0.0050
∞	0.0000	0.0000	0.0000	0.0000	0.0185	0.0185	0.0056	0.0056
				$N_1=4$				
1	-0.0017	-0.0015	0.0050	0.0017	0.0023	-0.0040	-0.0016	-0.0016
2	-0.0032	-0.0016	0.0090	0.0015	0.0075	-0.0080	0.0025	-0.0025
3	-0.0041	-0.0015	0.0105	0.0011	0.0123	-0.0102	0.0047	-0.0042
4	-0.0044	-0.0014	0.0102	0.0008	0.0160	-0.0110	0.0060	-0.0050
5	-0.0045	-0.0012	0.0093	0.0006	0.0187	-0.0108	0.0066	-0.0051
6	-0.0045	-0.0011	0.0081	0.0004	0.0206	-0.0101	0.0068	-0.0049
8	-0.0041	-0.0009	0.0059	0.0003	0.0231	-0.0081	0.0067	-0.0039
12	-0.0034	-0.0006	0.0031	0.0002	0.0251	-0.0037	0.0060	-0.0018
20	-0.0024	-0.0004	0.0010	0.0001	0.0259	0.0029	0.0051	0.0008
24	-0.0021	-0.0003	0.0005	0.0001	0.0258	0.0052	0.0049	0.0016
30	-0.0017	-0.0003	0.0002	0.0001	0.0256	0.0078	0.0048	0.0024
40	-0.0013	-0.0002	-0.0001	0.0001	0.0252	0.0108	0.0048	0.0033
60	-0.0009	-0.0001	-0.0002	0.0001	0.0247	0.0143	0.0049	0.0041
120	-0.0005	-0.0001	-0.0002	0.0000	0.0239	0.0183	0.0051	0.0049
∞	0.0000	0.0000	0.0000	0.0000	0.0228	0.0228	0.0056	0.0056

Table 10 (continued)

N_2	(Test for equality of a set of means)				(Test for equality of two variances)			
	$P'_{\lambda_4}(w)$	$P''_{\lambda_4}(w)$	$P'_{\lambda_3^2}(w)$	$P''_{\lambda_3^2}(w)$	$F_{\bar{\lambda}_4}(v_o)$	$F_{\underline{\lambda}_4}(v_o)$	$F_{\bar{\lambda}_3^2}(v_o)$	$F_{\underline{\lambda}_3^2}(v_o)$
				$N_1=5$				
1	-0.0015	-0.0012	0.0048	0.0014	0.0024	-0.0039	-0.0013	-0.0013
2	-0.0029	-0.0013	0.0089	0.0012	0.0076	-0.0080	0.0025	-0.0025
3	-0.0037	-0.0012	0.0107	0.0009	0.0124	-0.0104	0.0048	-0.0044
4	-0.0040	-0.0011	0.0107	0.0007	0.0163	-0.0114	0.0061	-0.0053
5	-0.0042	-0.0010	0.0100	0.0005	0.0191	-0.0114	0.0068	-0.0056
6	-0.0042	-0.0009	0.0089	0.0004	0.0213	-0.0109	0.0071	-0.0054
8	-0.0040	-0.0007	0.0069	0.0003	0.0241	-0.0090	0.0069	-0.0046
12	-0.0034	-0.0005	0.0040	0.0001	0.0268	-0.0046	0.0061	-0.0026
20	-0.0025	-0.0003	0.0014	0.0001	0.0282	0.0026	0.0050	0.0000
24	-0.0021	-0.0003	0.0009	0.0001	0.0283	0.0051	0.0047	0.0008
30	-0.0018	-0.0002	0.0004	0.0001	0.0282	0.0081	0.0046	0.0017
40	-0.0014	-0.0002	0.0001	0.0001	0.0280	0.0116	0.0044	0.0026
60	-0.0010	-0.0001	-0.0001	0.0000	0.0276	0.0156	0.0044	0.0035
120	-0.0005	-0.0001	-0.0002	0.0000	0.0269	0.0204	0.0047	0.0044
∞	0.0000	0.0000	0.0000	0.0000	0.0259	0.0259	0.0052	0.0052
				$N_1=6$				
1	-0.0014	-0.0012	0.0048	0.0011	0.0024	-0.0038	-0.0011	-0.0011
2	-0.0026	-0.0011	0.0088	0.0010	0.0076	-0.0080	0.0025	-0.0025
3	-0.0033	-0.0010	0.0107	0.0008	0.0125	-0.0106	0.0048	-0.0045
4	-0.0037	-0.0009	0.0110	0.0006	0.0164	-0.0118	0.0062	-0.0055
5	-0.0039	-0.0008	0.0104	0.0005	0.0194	-0.0120	0.0069	-0.0059
6	-0.0039	-0.0008	0.0096	0.0004	0.0217	-0.0116	0.0072	-0.0059
8	-0.0038	-0.0006	0.0076	0.0002	0.0248	-0.0099	0.0071	-0.0052
12	-0.0033	-0.0005	0.0047	0.0001	0.0279	-0.0055	0.0062	-0.0034
20	-0.0025	-0.0003	0.0019	0.0001	0.0297	0.0019	0.0049	-0.0007
24	-0.0022	-0.0002	0.0013	0.0001	0.0300	0.0047	0.0046	0.0001
30	-0.0018	-0.0002	0.0007	0.0000	0.0301	0.0079	0.0043	0.0010
40	-0.0014	-0.0001	0.0003	0.0000	0.0301	0.0118	0.0040	0.0019
60	-0.0010	-0.0001	0.0000	0.0000	0.0298	0.0164	0.0039	0.0029
120	-0.0005	-0.0001	0.0000	0.0000	0.0292	0.0218	0.0041	0.0038
∞	0.0000	0.0000	0.0000	0.0000	0.0283	0.0283	0.0045	0.0045

Table 10 (continued)

N_2	(Test for equality of a set of means)				(Test for equality of two variances)			
	$P'_{\lambda_4}(w)$	$P''_{\lambda_4}(w)$	$P'_{\lambda_3^2}(w)$	$P'_{\lambda_3^2}(w)$	$F_{\bar{\lambda}_4}(v_o)$	$F_{\underline{\lambda}_4}(v_o)$	$F_{\bar{\lambda}_3^2}(v_o)$	$F_{\underline{\lambda}_3^2}(v_o)$
$N_1=8$								
1	-0.0011	-0.0008	0.0046	0.0009	0.0026	-0.0037	-0.0008	-0.0008
2	-0.0021	-0.0008	0.0087	0.0008	0.0077	-0.0080	0.0025	-0.0025
3	-0.0028	-0.0008	0.0108	0.0007	0.0125	-0.0108	0.0048	-0.0046
4	-0.0031	-0.0007	0.0113	0.0005	0.0165	-0.0124	0.0063	-0.0058
5	-0.0033	-0.0006	0.0111	0.0004	0.0195	-0.0129	0.0071	-0.0064
6	-0.0034	-0.0006	0.0104	0.0003	0.0220	-0.0128	0.0075	-0.0065
8	-0.0034	-0.0005	0.0088	0.0002	0.0254	-0.0116	0.0074	-0.0061
12	-0.0030	-0.0004	0.0059	0.0001	0.0291	-0.0075	0.0064	-0.0045
20	-0.0024	-0.0002	0.0028	0.0000	0.0317	0.0001	0.0047	-0.0020
24	-0.0021	-0.0002	0.0020	0.0000	0.0322	0.0031	0.0042	-0.0012
30	-0.0018	-0.0002	0.0013	0.0000	0.0326	0.0068	0.0037	-0.0003
40	-0.0015	-0.0001	0.0007	0.0000	0.0328	0.0112	0.0033	0.0007
60	-0.0010	-0.0001	0.0002	0.0000	0.0328	0.0167	0.0030	0.0016
120	-0.0006	0.0000	-0.0001	0.0000	0.0324	0.0234	0.0029	0.0025
∞	0.0000	0.0000	0.0000	0.0000	0.0316	0.0316	0.0033	0.0033
$N_1=12$								
1	-0.0008	-0.0005	0.0045	0.0006	0.0027	-0.0036	-0.0005	-0.0005
2	-0.0016	-0.0005	0.0084	0.0006	0.0077	-0.0079	0.0025	-0.0025
3	-0.0021	-0.0005	0.0107	0.0005	0.0125	-0.0112	0.0049	-0.0047
4	-0.0024	-0.0005	0.0116	0.0004	0.0164	-0.0131	0.0065	-0.0062
5	-0.0026	-0.0004	0.0117	0.0003	0.0195	-0.0142	0.0073	-0.0069
6	-0.0027	-0.0004	0.0014	0.0002	0.0220	-0.0145	0.0078	-0.0073
8	-0.0027	-0.0003	0.0101	0.0002	0.0257	-0.0140	0.0078	-0.0072
12	-0.0026	-0.0002	0.0075	0.0001	0.0299	-0.0109	0.0069	-0.0060
20	-0.0022	-0.0002	0.0043	0.0000	0.0335	-0.0036	0.0048	-0.0037
24	-0.0020	-0.0001	0.0033	0.0000	0.0343	-0.0005	0.0041	-0.0028
30	-0.0017	-0.0001	0.0023	0.0000	0.0350	0.0035	0.0033	-0.0020
40	-0.0014	-0.0001	0.0014	0.0000	0.0355	0.0087	0.0025	-0.0010
60	-0.0011	-0.0001	0.0007	0.0000	0.0359	0.0153	0.0018	-0.0001
120	-0.0006	0.0000	0.0001	0.0000	0.0359	0.0240	0.0013	0.0007
∞	0.0000	0.0000	0.0000	0.0000	0.0354	0.0354	0.0013	0.0013

Table 10 (continued)

N_2	(Test for equality of a set of means)				(Test for equality of two variances)			
	$P'_{\lambda_4}(w)$	$P''_{\lambda_4}(w)$	$P'_{\lambda_3^2}(w)$	$P''_{\lambda_3^2}(w)$	$F_{\bar{\lambda}_4}(v_o)$	$F_{\underline{\lambda}_4}(v_o)$	$F_{\bar{\lambda}_3^2}(v_o)$	$F_{\underline{\lambda}_3^2}(v_o)$
$N_1=24$								
1	-0.0005	-0.0003	0.0043	0.0003	0.0029	-0.0034	-0.0000	0.0000
2	-0.0009	-0.0003	0.0081	0.0003	0.0078	-0.0079	0.0026	-0.0026
3	-0.0012	-0.0002	0.0104	0.0003	0.0124	-0.0116	0.0049	-0.0049
4	-0.0014	-0.0002	0.0117	0.0002	0.0161	-0.0142	0.0066	-0.0065
5	-0.0015	-0.0002	0.0122	0.0002	0.0192	-0.0159	0.0077	-0.0075
6	-0.0016	-0.0002	0.0122	0.0001	0.0217	-0.0170	0.0082	-0.0081
8	-0.0017	-0.0002	0.0116	0.0001	0.0254	-0.0178	0.0086	-0.0084
12	-0.0017	-0.0001	0.0097	0.0001	0.0300	-0.0169	0.0079	-0.0078
20	-0.0016	-0.0001	0.0067	0.0000	0.0344	-0.0117	0.0058	-0.0058
24	-0.0015	-0.0001	0.0056	0.0000	0.0355	-0.0090	0.0049	-0.0050
30	-0.0014	-0.0001	0.0044	0.0000	0.0367	-0.0051	0.0038	-0.0041
40	-0.0012	0.0000	0.0032	0.0000	0.0378	0.0004	0.0025	-0.0032
60	-0.0009	0.0000	0.0019	0.0000	0.0387	0.0084	0.0012	-0.0022
120	-0.0006	0.0000	0.0007	0.0000	0.0394	0.0204	-0.0003	-0.0014
∞	0.0000	0.0000	0.0000	0.0000	0.0396	0.0396	-0.0013	-0.0013

The approximate true probabilities are given by:

(i) $P(w_o)=0.05+\lambda_4\{P'_{\lambda_4}(w_o)+\Delta k^2 P''_{\lambda_4}(w_o)\}+\lambda_3^2\{P'_{\lambda_3^2}(w_o)+\Delta k^2 P_{\lambda_3^2}(w_o)\}$

(ii) $P(v_o)=0.05+\frac{1}{2}(\lambda''_4+\lambda'_4)P_{\bar{\lambda}_4}(v_o)-\frac{1}{2}(\lambda''^2_3+\lambda'^2_3)P_{\bar{\lambda}_3^2}(v_o)$

$+\frac{1}{2}(\lambda''_4-\lambda'_4)P_{\underline{\lambda}_4}(v_o)-\frac{1}{2}(\lambda''^2_3-\lambda'^2_3)\; P_{\underline{\lambda}_3^2}(v_o)$

VI. EMPIRICAL STUDY OF THE EFFECT OF NON-NORMALITY ON THE DISTRIBUTIONS OF u AND w STATISTICS

A number of studies have empirically examined the effect of departures from the underlying assumptions in the case of the t, u, w and v tests. Some of these are Pearson (21), (22), Norton (19), Hack (18), and Boneau (3). Among these works Norton and Boneau's studies are the most comprehensive and significant. We will give a brief summary of their studies here.

6.1 The Norton Study

To investigate the effect of non-normality upon the distribution of the w test, Norton constructed "card populations" of 10,000 cases each. He used these "card populations" to represent the parent population from which samples were taken. Six distributions which are most frequently met in educational and psychological research were investigated. Figure 1 presents by histograms a representation of the distribution of each of these populations. Population 1, except for a finite range and lack of complete continuity, is essentially a normal distribution, and was included as a check on the sampling procedure employed. From each of these populations independently, Norton selected 3,000 sets of k random samples of n cases each. Each set thus corresponded to a hypothetical simple randomized experiment with k treatments and n cases in each treatment group. For each set (or experiment) the ratio of the mean squares for "between-treatment" and "within-treatment" was computed. An empirical distribution of 3,000 w's was thus obtained for each of the six populations. The discrepancies, in the critical upper-tail area between the empirical distribution thus obtained and the normal-theory w distribution are described by the data in Table 11.

Table 11 True significance levels for the variance ratio test (w) in empirical distributions when $\alpha=0.1$, 0.05, 0.025, 0.01

Number of population	Type of population	k	n	d.f.	α points d.f.=2,6 d.f.=3,16	0.1 3.46 2.46	0.05 5.14 3.24	0.025 7.26 4.08	0.01 10.92 5.29
1	Normal	4	5	3,16		0.0998	0.0561	0.0291	0.0144
2	Leptokurtic	3	3	2,6		0.1293	0.0783	0.0463	0.0276
2	Leptokurtic	4	5	3,16		0.1126	0.0656	0.0376	0.0163
3	Rectangular	3	3	2,6		0.1154	0.0607	0.0324	0.0177
4	Mod. Skew	4	5	3,16		0.1028	0.0515	0.0275	0.0132
5	Ext. Skew	3	3	2,6		0.0967	0.0477	0.0207	0.0080
5	Ext. Skew	4	5	3,16		0.1016	0.0476	0.0233	0.0010
6	J-Shape	3	3	2,6		0.0943	0.0480	0.0257	0.0010

From Table 11 we can see that the significance level of the test of analysis of variance is almost unaffected by the

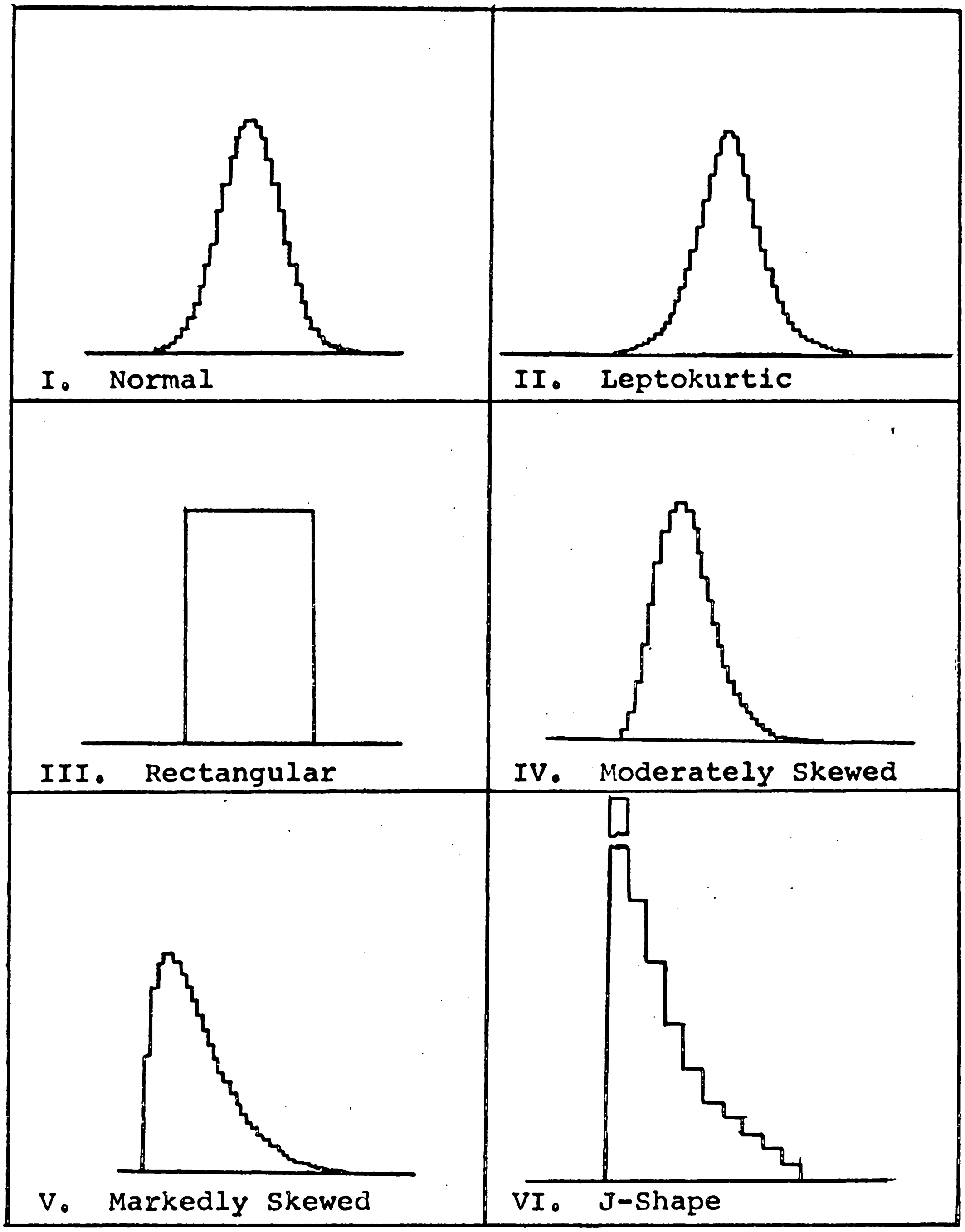

Figure 1. Histograms of populations for which empirical F-distributions were obtained in the Norton Study

skewness, and is only slightly affected by the kurtosis of the parent population. This result agrees with the theoretical investigation made by Gayen (14).

Compared with other tests, [see for example (4)] the test of analysis of variance is relatively insensitive to the form of the distribution of the parent population; this can be seen from Table 11. For example, in the case of $\alpha = 0.05$, the significance level is 0.0783 for a leptokurtic population as one extreme discrepancy and 0.0476 for an extremely skewed distribution as another.

The data in the first row of Table 11 provided a check on the sampling procedure employed in this study. As will be noted, the empirical distribution of w for samples drawn from Population 1 contained a larger proportion of upper-tail area than the theoretical. This may imply that the discrepancies reported in the remainder of Table 11 are due in part to the method of sampling, rather than to lack of normality alone. In this event, the effect of non-normality on the test of analysis of variance would actually be smaller than those reported in Table 11.

6.2 The Boneau Study

To investigate the effects of non-normality on the u-test Boneau used an IBM 650 electronic computer to compute a large number of u values. Each of these u's was based upon samples drawn at random from a population having a specified density. He then constructed a frequency distribution of the u's thus obtained. The populations he selected for this study were the translated exponential (J-Shaped with a skew to the right) having a density function of $y=e^{-(x+1)}$ $-1<x<\infty$ and the uniform distribution. These distributions represent extremes of skewness and flatness for comparison with the normal.

In the following paragraphs we will study the effects of samples from the translated exponential and uniform distributions on the significance levels of the u-test.

In Figure 2 we compare the theoretical u distribution and the empirical distribution obtained from two samples of size 5 from the translated exponential distribution. The fit is fairly close, but the proportion in the tails seems less for the empirical distribtuion than for the theoretical. By count, the significance level is 3.1% for the empirical distribution of u's from the translated exponential population, compared to the nominal 5% value. If

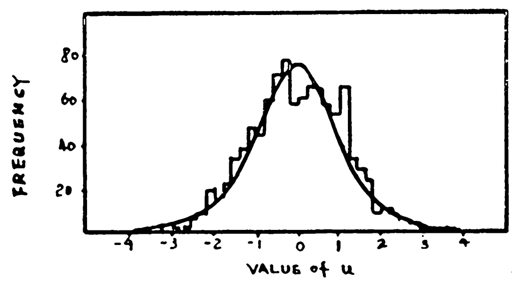

Figure 2. Empirical distribution of u's from E(0,1)5-E(0,1)E* and the theoretical distribution with 8 d.f.
* E stands for translated exponential distribution, the first number in the parenthesis is the mean of the distribution, while the second number is the variance; the number following the parenthesis is sample size.

both sample sizes are raised to 15, the corresponding percentage of u's obtained is 4%. If both samples are of size 5 from the same rectangular distribution, the fit of the theoretical u-distribution to the empirical u-distribution thus obtained is better than the previous case of samples from the translated exponential distribution. The result is as shown in Figure 3. The percentage of the obtained u's which exceed the nominal 5% values is 5.1%. For the case in which the sample sizes are both 15, the fit is equally good, with 5.0% of the cases falling outside the nominal 5% bound.

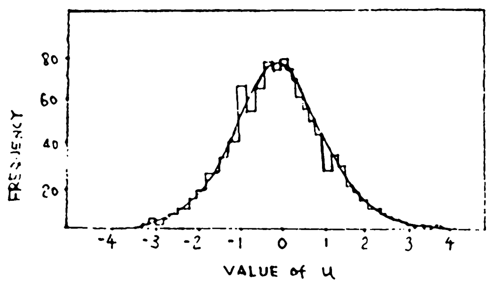

Figure 3. Empirical distribution of u's from R(0,1)5-R(0,1)5* and the theoretical distribution with 8 d.f.
* R stands for rectangular distribution, the first number in the parenthesis is the mean of the distribution, while the second number is the variance, and the number following the parenthesis is the sample size.

All Boneau's experimental results on the subject of the robustness of two-sample u-test to non-normality are summarized in Table 12.

Table 12 True significance levels for the u-test when $\alpha=0.05$ and 0.01

Symmetric distribution	Obtained significance at	
	0.05 level	0.01 level
N(0,1)5-N(0,1)5	0.053	0.009
E(0,1)5-E(0,1)5	0.031	0.003
E(0,1)15-E(0,1)15	0.040	0.004
R(0,1)5-R(0,1)5	0.051	0.01
R(0,1)15-R(0,1)15	0.050	0.015

Asymmetric distribution	Obtained significance at			
	0.05 level		0.01 level	
	Total	Larger tail	Total	Larger tail
E(0,1)5-R(0,1)5	0.064	0.05	0.033	0.025
E(0,1)15-R(0,1)15	0.056	0.039	0.016	0.012

Explanation: The letters N, E. and R refer to the population from which the sample was drawn, N for normal, E for the translated exponential, R for rectangular. The first number in the parenthesis is the mean of the distribution, in all cases zero, while the second number is the variance, the number following the parenthesis is the sample size.

The data in the first row of Table 12 provided a check on the sampling procedure employed in this study.

From Table 12 we can see that the u-test is slightly sensitive to changes in skewness, and insensitive to changes in kurtosis of the sampled population. When the sampled population is symmetric, the u-test is remarkedly robust, such as samples taken from a rectangular distribution. This result agrees with Gayen's (15) theoretical study.

All results and discussion in Boneau's study were limited to the two-tailed u-test, with some exceptions. The conclusions he reached, however, can be applied directly to the one-tailed test as well. The exceptions involve those distributions which are intrinsically asymmetric (see Table 12). In these distributions a preponderance of the sample u's fall in one tail.

VII. DISCUSSION AND CONCLUSIONS

The results of this study may prove useful to any one who contemplates using the t- or F-tests in experimental situations in which there is serious doubt about the underlying assumption of normality.

The theoretical and empirical investigations discussed in this paper show that the t-test is slightly affected by lack of symmetry, but in general the effect of non-normality is not serious. For the u-test, if the two sample sizes are equal or nearly so, and the assumed underlying populations are of the same shape or nearly so, the non-normality has little effect on its test significance level. When the sample size is increased, the effect of non-normality will be even smaller. Based on these results, we can conclude that unless the departure from normality is so extreme that it can be easily detected by mere inspection, the use of the ordinary t-test and its associated table will result in probability statements which are relatively accurate.

In the F-test of the analysis of variance, the distribution of the F-statistic is practically unaffected by lack of symmetry, but is slightly affected if the underlying population is roughly symmetrical but either very flat or very peaked. Since most non-normal distributions met in practice are non-normal primarily because of lack of symmetry rather than because of lack of the "normal" degree of peakedness, the departure from normality will probably have no appreciable effect on the validity of the F-test of analysis of variance, and the probability read from the F table may be used as close approximations to the true probabilities.

The situation is quite different for the F-test on the equality of two variances, this test is very sensitive to non-normality and will under some conditions yield significant result even when the variances are equal.

VIII. APPENDIX

THE EDGEWORTH SERIES

Let $Y_1, Y_2, \ldots\ldots, Y_n$ be a sequence of n independent random variables with finite mean μ and variance σ^2. If we denote by $F_n(x)$ and $f_n(x)$ the distribution function and density function of the standardized variable (8.1)

$x = \dfrac{\sum_{i=1}^{n} (Y_i - \mu_i)}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}$ then $F_n(x)$ and $f_n(x)$ can be expressed by

their respective series as follows:

$$F_n(x) = \Phi(x) - \frac{1}{3!}\frac{\lambda_3'}{n^{1/2}}\Phi^3(x) + \frac{1}{4!}\frac{\lambda_4'}{n}\Phi^4(x) + \frac{10}{6!}\frac{\lambda_3'^2}{n}\Phi^6(x)$$
$$- \frac{1}{5!}\frac{\lambda_5'}{n^{3/2}}\Phi^5(x) - \frac{35}{7!}\frac{\lambda_3'\lambda_4'}{n^{3/2}}\Phi^7(x) - \frac{280}{9!}\frac{\lambda_3'^3}{n^{3/2}}\Phi^9(x) \tag{8.2}$$

$$f_n(x) = \phi(x) - \frac{1}{3!}\frac{\lambda_3'}{n^{1/2}}\phi^3(x) + \frac{1}{4!}\frac{\lambda_4'}{n}\phi^4(x) + \frac{10}{6!}\frac{\lambda_3'}{n}\phi^6(x)$$
$$- \frac{1}{5!}\frac{\lambda_5'}{n^{3/2}}\phi^5(x) - \frac{35}{7!}\frac{\lambda_3'\lambda_4'}{n^{3/2}}\phi^7(x) - \frac{280}{9!}\frac{\lambda_3'^2}{n^{3/2}}\phi^9(x) \tag{8.3}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\phi(t)\,dt$, $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, $\Phi^v(x)$, $\phi^v(x)$ are v-th derivatives, and λ_v' is the cumulant of the variable

$\frac{y_i-\mu_i}{\sigma^i}$. If we use the relationship $\lambda_v = \frac{\lambda'_v}{n^{v/2}}$, where λ_v is the v-th order cumulant of the variable y, we obtain the expansions of $F_n(x)$ and $f_n(x)$ to order of n^{-1} as follows:

$$(8.4) \qquad F_n(x) = \Phi(x) - \frac{\lambda_3}{3!}\Phi^3(x) + \frac{\lambda_4}{4!}\Phi^4(x) + \frac{10\lambda_3^2}{6!}\Phi^6(x)$$

$$(8.5) \qquad f_n(x) = \phi(x) - \frac{\lambda_3}{3!}\phi^3(x) + \frac{\lambda_4}{4!}\phi^4(x) + \frac{10\lambda_3^2}{6!}\phi^6(x)$$

The convergence and asymptotic properties of these series have been studied by Cramer (8) (9). In practical applications it is of little value to know the convergence properties of these expansions. What we really want to know is whether a small number of terms, usually not more than two or four, suffice to give a good approximation to $F_n(x)$ and $f_n(x)$. This leads us to be concerned about the asymptotic properties of the expansions (8.4) (8.5). In the following paragraph we will give the conditions under which the series (8.4) (8.5) give an asymptotic expansion of $F_n(x)$ and $f_n(x)$.

Here we only consider the case that the component random variables Y_v of (8.1) have the same distribution function with mean μ and variance σ^2. The series (8.4) gives an asymptotic expansion of $F_n(x)$ if the following conditions are satisfied; (a) the absolute moment $B_k = E(|x|^k)$ of order

$k \geq 3$ is finite; (b) the component distributions are absolutely continuous. This condition is not satisfied for discrete distributions, but Wilks (26) said "... as a matter of fact, even if $F_n(x)$ has no p.d.f. (that is, $F_n(x)$ may be a discrete c.d.f.), [8.4] is still a useful approximation.....".

In addition to the above two conditions, if the derivative $F_n'(x)$ is of bounded variation in the whole domain $-\infty<x<\infty$, then the series (8.5) gives an asymptotic expansion of $f_n(x)$.

Conditions for the validity of the asymptotic expansion for sums of non-identically distributed random variables are somewhat more restrictive. These conditions have been proved by Cramer (9).

For large values of x, the expansion (8.5) will sometimes yield small negative values for $f_n(x)$. This fact is one objection to using the Edgeworth series in representing the density function. For this point, Barton and Dennis (1) have pointed out that when λ_4 lies roughly between 0 and 2.4 and $\lambda_3^2 \leq 0.2$, the series (8.5) gives a positive definite, unimodal function.

In the following paragraph we will give a numerical example and use χ^2 test of goodness of fit to test how good the Edgeworth series (8.5) fits an actual set of observations. (This example is quoted from Cramer (5) p. 441.)

Table 13 The distribution of the breadth of n=12000 beans

class number i	Breadth of bean mm	Observed frequency n_i	Expected frequency np_i		
			Normal	First app.	Second app.
1	6.70-6.95	32	67.6	17.5	26.6
2	6.95-7.20	103	132.2	98.3	90.4
3	7.20-7.45	239	309.8	291.5	277.2
4	7.45-7.70	624	617.3	648.9	636.8
5	7.70-7.95	1187	1045.7	1142.2	1141.1
6	7.95-8.20	1650	1505.8	1630.4	1639.9
7	8.20-8.45	1883	1842.3	1918.1	1931.6
8	8.45-8.70	1930	1919.9	1892.4	1906.2
9	8.70-8.95	1638	1697.9	1587.3	1599.5
10	8.95-9.20	1130	1277.3	1158.8	1163.5
11	9.20-9.45	737	817.0	752.4	745.1
12	9.45-9.70	427	444.2	441.9	427.3
13	9.70-9.95	221	205.3	235.6	223.8
14	9.95-10.25	110	80.7	112.7	109.1
15	10.25-10.50	57	27.0	47.5	49.7
16	10.50-10.75	32	10.0	24.5	32.2
Total		12000	12000	12000	12000
$\bar{y}=8.512$ $s=0.6163$ $g_1=-0.2878$ $g_2=0.1953$			$\chi^2=196.5$ (13 d.f.)	$\chi^2=34.3$ (12 d.f.)	$\chi^2=14.9$ (11 d.f.)

The table shows the distribution of the breadths of n=12000 beans; it gives also the expected frequencies calculated by the following formulas, and the corresponding values of χ^2 (tests of goodness of fit)

(a) normal . . . . $\phi(x)=\frac{1}{\sqrt{2\Pi}}e^{-\frac{x^2}{2}}$

(b) first app. . . . . $\phi(x)-\frac{g_1}{3!}\,\phi^3(x)$

(c) second app. . . . . . $\phi(x)-\frac{g_1}{3!}\,\phi^3(x)+\frac{g_2}{4!}\,\phi^4(x)+\frac{10g_1^2}{6!}\,\phi^6(x)$

where g_1 and g_2, the estimates of λ_3 and λ_4, respectively, are calculated from the grouped sample, using Sheppard's corrections. The estimated standard deviation s in the table, also calculated from the grouped sample, uses Sheppard's correction $s^2=\frac{1}{n}\sum_{i=1}^{n} n_i(y_i-\bar{y})^2 - \frac{h^2}{12}$, where $y_i=y_1-(i-1)h$, h is the class interval and y_i is the i-th class expected value.

From the χ^2 values in Table 13, we observe that, in the first two cases, the deviations of the sample from the hypothetical distributions are highly significant, while in the third case the agreement is satisfactory.

Figure 4 shows by means of a histogram the agreement between the sample of 12,000 breadths of beans and the hypothetical density function of the corresponding population as approximated by the Edgeworth expansion (8.5).

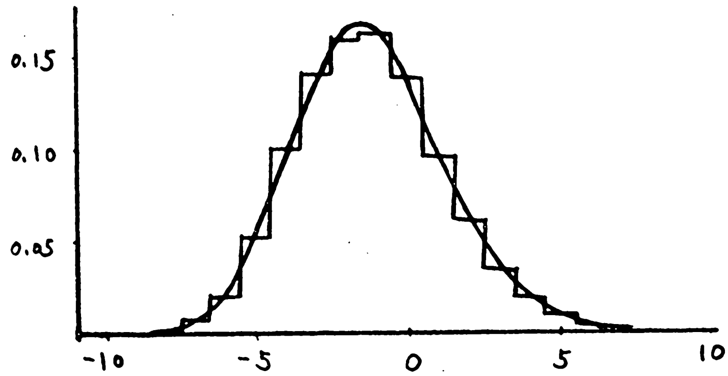

Figure 4 Histogram for the breadths of 12000 beans and frequency curve according to Edgeworth's series.

IX. ACKNOWLEDGEMENTS

The author wished to express his sincere appreciation to his thesis advisor, Dr. Richard G. Krutchkoff, for suggesting the problem, and for his patience and continued encouragement throughout this study.

The author also expresses his sincere appreciation to Dr. D. Jensen for his time in reading the manuscript and for his constructive criticism and suggestions which added greatly to this study.

X. BIBLIOGRAPHY

1. Bartlett, M. S. "The effect of non-normality on the t-distribution", Proc. Cambridge Philo. Soc. 31 (1935), p. 223-231.

2. Barton, D. E. and Dennis, K. E. "The conditions under which Gram-Charlier and Edgeworth series are positive definite and unimodal", Biometrika 39 (1952), p. 425-427.

3. Boneau, C. A. "The effects of violations of assumptions underlying the t-test", Psy. Bull. 57 (1960), p. 49-64.

4. Box, G. E. P. "Non-normality and tests on variance", Biometrika 40 (1953), p. 318-335.

5. Bradley, R. A. "Sampling distribution for non-normal universes", unpublished Ph.D. thesis, Univ. of N.C.

6. Bradley, R. A. "The distribution of the t and F statistics for a class of non-normal populations", Va. Jour. Sci. 3 (1952), p. 1-34.

7. Bradley, R. A. "Corrections for non-normality in the use of the two-sample t and F tests at high significance level", Ann. Math. Stat. 23 (1952), p. 103-113.

8. Cramer, H. "On the composition of elementary errors", Skand. Aktuar. 11 (1928), p. 13-74

9. Cramer, H. "Random variable and probability distributions", Cambridge University Press, No. 36.

10. Cramer, H. "Mathematical methods of statistics", Princeton University Press.

11. Edgeworth, F. Y. "The law of errors", Trans. Camb. Phil. Soc. 20 (1905), p. 36-65.

12. Eisenhart, C. "The assumptions underlying the analysis of variance", Biometrics 3 (1947), p. 1-21.

13. Gayen, A. K. "The distribution of Student's t in random samples of any size drawn from non-normal universe", Biometrika 36 (1949), p. 353-363.

14. Gayen, A. K. "The distribution of the variance ratio in random samples of any size drawn from non-normal universe", Biometrika 37 (1950), p. 236-255.

15. Gayen, A. K. "Significance of difference between the means of two non-normal samples", Biometrika 37 (1950), p. 399-408.

16. Geary, R. C. "The distribution of Student's ratio for non-normal samples", Jour. Roy. Stat. Soc., series B, 3 (1936), p. 178-184.

17. Geary, R. C. "Testing for normality", Biometrika 34 (1947), p. 209.

18. Hack, H. R. B. "An empirical investigation into the distribution of the F-ratio in samples from two non-normal populations", Biometrika 44 (1958), p. 260-265.

19. Norton, D. W. "An empirical investigation of some effects of non-normality and heterogeneity on the F-distribution", unpublished Ph.D. Thesis, State Univ. of Iowa, 1952.

20. Owen, D. B. "The power of Student's t-test", Jour. Amer. Stat. Assoc. 60 (1965), p. 320-330.

21. Pearson, E. S. and Adyanthaya, N. K. "The distribution of frequency constants in small samples from non-normal symmetrical and skew populations", Biometrika 21 (1929), p. 259-286.

22. Pearson, E. S. "The analysis of variance in the case of non-normal variation", Biometrika 23 (1931), p. 114-133.

23. Rider, P. R. "On the distribution of the ratio of mean to standard deviation", Biometrika 21 (1929), p. 124-143.

24. Srivastava, A. B. L. "Effect of non-normality on the power of the analysis of variance test", Biometrika 46 (1959), p. 114-122.

25. Tiku, M. L. "Approximating the general non-normal variance-ratio sampling distribution", Biometrika 51 (1964), p. 83-95.

26. Wilks, S. S. "Mathematical statistics", John Wiley & Sons, Inc., New York, 1961.

The vita has been removed from the scanned document

ABSTRACT

A statistical test is called "robust" if it is insensitive to departures from the underlying assumptions, this term was introduced by Box (4).

Theoretical study made by Gayen (13), (14) and (15) showed that "Student's" t-test and the closely related F-test of analysis of variance are insensitive to departures from normality. But the F-test on the equality of two variances is very sensitive to such departures.

Empirical studies made by Norton (19) and Boneau (3) agree with Gayen's theoretical conclusions. Norton studied the effect of non-normality on the F-test of analysis of variance, and showed that the form of the sampled population had very little effect on this test. For example, for the case of three groups of sample sizes 3, for the 5% level, the percentages exceeding the theoretical limits were 7.83 and 4.77% respectively for sampling from a leptokurtic and an extremely skewed population. Such property of robustness to non-normality on the F-test of analysis of variance is also possessed by t-test. Boneau's empirical study on the effect of non-normality on the two-sample t-test showed that for two samples of size 5, the significance level is respectively 3.1 and 5.1% for the empirical distribution of t's from the exponential and uniform distribution compare to

the nominal 5% value. The discrepancy is decreased when sample size is increased.