Biom. J. 25 - 1983 - 4, 327-343

Forschungszentrum für Tierproduktion, AdL der DDR, Dummerstorf-Rostock

Robustness of Statistical Methods
II. Methods for the one-Sample Problem

G. HERRENDÖRFER, D. RASCH and K. D. FEIGE

Abstract

This paper presents a new exact method of the investigation of the robustness in the one-sample- case of the *u*-, *t*- and x^2-statistics for discrete alternatives to the underlying normal distribution. First results for sample sizes up to n=230 are given.

Key words: Robustness, *t*-test, confidence estimation, exact method, discrete alternatives.

1. Introduction

The first part of this work (RASCH and HERRENDÖRFER, 1981) proposed quantitative definitions of the robustness of statistical procedures and the simulation method for investigating this property was described together with an appropriate experimental design for obtaining results with the desired precision. The extensive literature dealing with this range of problems was reviewed.

This part will first collate the most important results given in the literature regarding the exact investigation of the robustness of the *t*-test and the corresponding interval estimation. Finally, an approach is given for calculating the robustness of the confidence estimation of an expectation and, quite analogously, for tests of μ and σ^2. Basically this approach can also be applied to other statistical problems.

Let $\underline{y}$ be $N(\mu; \sigma^2)$-distributed, and let $\underline{\mathbf{y}}' = (\underline{y}_1, \ldots, \underline{y}_n)$ be a random sample. Then

$$(1) \qquad d_\alpha = \left\langle \bar{\underline{y}} - t\left(n-1;\, 1-\frac{\alpha}{2}\right)\frac{\underline{s}}{\sqrt{n}},\ \bar{\underline{y}} + t\left(n-1;\, 1-\frac{\alpha}{2}\right)\frac{\underline{s}}{\sqrt{n}} \right\rangle$$

is a $(1 - \alpha)$-confidence interval for μ if the $\underline{y}_i$ are distributed as $\underline{y}$. In other words

$$P_N(\mu \in \underline{d}_\alpha) = 1 - \alpha\ .$$

If, on the other hand, $\underline{y}$ has some other distribution denoted, say, by g, then

$$(2) \qquad P_g\,(\mu \in \underline{d}_\alpha) = 1 - \alpha(n, g)\ .$$

In the case of discrete distributions with a finite support $\alpha(n, g)$ can be calculated exactly. Initial results in this connection will be presented.

The exact (analytical) calculation of the confidence coefficient of an interval estimation

or the power function of a test has hitherto been possible in only a few cases and for very small n when the distribution is not normal. GRONOW (1953) investigated two formulae for the approximative t-test for comparing two means. He used a method described by DAVID and JOHNSON (1951) to calculate approximately the distribution of a ratio of two random variables when the denominator is always non-negative. This involved using the first four moments. He also assumed that the initial distribution can be described by symmetric GRAM-CHARLIER series of type A with certain coefficients. This was used as a basis to approximate the t-test-statistics. Exact calculation of the power was not possible. No error estimation was made in the paper.

For samples of the same size, γ_2 (measure of kurtosis) does not affect the power of the t-test significantly. This also applies to the inhomogeneity of variances. Theoretical studies on robustness of the t-test are given in BARTLETT (1935), GEARY (1936, 1947) and GAYEN (1949). They used EDGEWORTH series of the initial distributions up to the second term. GAYEN (1949) also considered the fourth moment of the distribution in his work. He examined the first kind risk of the test.

Apart from the fundamental experimental work done by PEARSON and ADYANTHAYA (1929) on the power of the t-test for non-normal distributions, the publication by GHURYE (1949) was the first in which an attempt was made to assess theoretically the influence of nonnormality. He studied the effect of skewness on the power function. SRIVASTAVA (1958) characterized the deviation of the initial distribution from the normal distribution by the first four moments (standardized third and fourth cumulants).

He was also able to obtain only approximate results. LADERMAN (1939) gave an exact result. He investigated the distribution of $\underline{t}$ for n = 2 and continuous distribution (one-sample case) and used geometrical methods. PERLO (1933) derived the distribution of $\underline{t}$ for $n = 3$ and a uniform distribution from $\underline{y}$. HYRENIUS (1950) considered the distribution of $\underline{t}$ for composite normal distributions, and TIKU (1963) used the first eight cumulants of the initial distribution. BRADLEY (1952) regarded the distribution function of $\underline{t}$ as a multiple integral of the density of the random sample; to calculate the integral he also used series development which he broke off after a few terms. In other words he could only propose an approximative solution.

RATCLIFFE (1968) tried to find the minimum size of sample to which the t-test can be applied and thus studied a question of basic importance for the use of the test. He did this by studying several practically occurring distribution types and the uniform distribution, the continuous distribution types and the uniform distribution, the continuous distributions being approximated by 72-point distributions.

The t-statistic was investigated by means of the simulation method ($N_E = 2000$). The value of n varied between $n = 3$ and $n = 80$. He came to the following conclusions:

- the effect of non-normality decreases with increasing n, and the smaller α is, the larger n must be;
- the distribution of $\underline{t}$ is only slightly affected by positive kurtosis if $n > 3$;
- the distribution of $\underline{t}$ is greatly affected by negative kurtosis for $n \leq 30$. If the kurtosis is not smaller than -1.2, i.e. if it is greater than that of the uniform distribution, it can be ignored for $n = 15$ and more;
- skewness of the initial distribution causes skewness in the distribution of $\underline{t}$. This can be expected to have a noticeable effect on the t-test up to $n = 50$. The t-test can still be affected by extreme skewness at $n = 80$.

SANSING and OWEN (1974) considered the joint distribution of $\underline{\bar{y}}$ and $\underline{s^2}$ for distributions with positive density in the R^1 and with certain integration properties. They derived the

distribution of $\underline{t}$ from these distributions. The family of generalized normal distributions was used as an example. An approximation of the density of $\underline{t}$ was derived for this example. They succeeded, under the given conditions, in stating the joint distribution of $\underline{y}$ and $\underline{s}^2$ recursively. They also took the fact that $\underline{y}$ and $\underline{s}^2$ are dependent into account, something that is often ignored in papers using density approximations.

BLUMENTHAL and GOVINDARAJULU (1977) investigated the robustness of STEIN'S two-stage *t*-test for the one-sample case. They proceeded from the assumption that the initial distribution can be represented as a mixed distribution composed of two normal distributions with different means and identical variances. LEE and GURLAND (1977) studied the *t*-test for the one-sample case using a mixed distribution comprising two normal distributions with different means and different variances. They succeeded in giving the distribution of $\underline{t}^2$ exactly as an infinite series.

They gave the real values of $\alpha(n, g)$ for $n = 5$ and $n = 10$ for the nominal α-values 0.01; 0.05 and 0.10. Considerable deviations were still evident for $n = 10$.

In the second volume of their book KENDALL and STUART (1967) presented a few asymptotic results for the *t*-test and used them to draw some conclusions regarding robustness.

Finally, we shall mention here two further very extensive simulations by POSTEN (1978, 1979) although the method is not exact. However these are probably the most extensive results obtained hitherto regarding the robustness of the *t*-test. He dealt with both the one-sample and two-sample cases. Distributions of the PEARSON system were investigated as alternatives to the normal distribution. They can be described in terms of skewness and kurtosis (i.e. by the first four moments).

The values $0 \leq \gamma_1^2 \leq 2$ and $-1.6 \leq \gamma_2 \leq 4.8$ were permissible. The author considered 87 points (γ_1, γ_2) and used $n = 5\ (5)\ 30$. The number of *t*-values generated for each point and for each *n* was 100,000 and the power function was estimated at various points. The two-sided tests are much more robust than the one sided. For $n = 30$ and above the two-sided *t*-test can be considered robust to the deviations from the normal distribution described.

The distribution of the *t*-test statistic could be found analytically for only a few distributions. Hitherto, discrete distributions such as are very often encountered in practice have not been admissible as alternative distributions to the normal distributions in recent investigation. However, for discrete distribution with a finite support the exact calculation of the power function of the known parametric procedure (for a sample of finite size) is always possible.

2. An exact method for investigating the robustness of statistical procedures for discrete distributions with a finite support.

Let

$$(3) \qquad g_k = \begin{pmatrix} y_1, y_2, \ldots, y_k \\ p_1, p_2, \ldots, p_k \end{pmatrix}$$

be a discrete distribution with *k* points in its support, and let $\underline{\mathbf{y}}' = (\underline{y}_1, \ldots, \underline{y}_n)$ be a random sample with its components distributed according to (3). The probability of the realization

$$(4) \qquad P_{\mathfrak{N}} = \binom{y_1, \ldots, y_k}{n_1, \ldots, n_k} \qquad \sum n_i = n,\ n_i \geqq 0$$

is

$$(5) \qquad P_{g_k}(\underline{\mathfrak{y}} = P_{\mathfrak{N}}) = \frac{n!}{n_1! \ldots n_k!} p_1^{n_1} \ldots p_k^{n_k} = p_{\mathfrak{N}} .$$

The distribution of $\underline{\mathbf{y}}$ is thus polynomial. The number of possible (different) samples is given by

$$(6) \qquad M = \binom{n+k-1}{k-1} .$$

Equation (1) can be used to calculate the statistic $\underline{d_\alpha}$ for each possible sample. If g_k is known, $E(\underline{y}_i) = \mu$ and $V(\underline{y}_i) = \sigma^2$ are also known, it is therefore possible to check, for each such sample, whether a correct or a false decision has been made. The M possible samples must then be arranged in some fixed order, whereupon we unite the p_R in accordance with the fixed order to obtain a vector $\boldsymbol{P}$. We shall call the corresponding resultant vector $\boldsymbol{H}$. The components of $\boldsymbol{H}$ are 0 if the correct decision has been made for the sample concerned and 1 otherwise. What is actually meant by the term "correct decision" in this context must be defined separately for each group of statistical procedures. In the case above, a correct decision would mean that the interval concerned covers the known parameter.

If $\underline{d_\alpha}$ were an α-test, the null hypothesis can at first be given as correct. In this event the assumption of the null hypotheses for the specific sample would be a correct decision.

If it is postulated that $\boldsymbol{H}_A$ (i.e. some specific value of $\boldsymbol{H}_A$) is correct, the correct decision would have been made if $\boldsymbol{H}_A$ had been accepted. The resultant vector $\boldsymbol{H}$ for a selection procedure can be set up analogously. These two vectors can be used to calculate the probability that the inference drawn by a statistical procedure is correct or false for any discrete distribution with a finite support and for any finite n for which d_α can be calculated. We have

$$(7) \qquad \boldsymbol{A = P'H}$$

If d_α is an interval estimation, we denote A by

$$\boldsymbol{A = \alpha\,(n, g_k).}$$

If d_α is a test and H_0 correct, we would also have

$$\boldsymbol{A = \alpha\,(n, g_k).}$$

On the other hand, where the alternative hypothesis is correct, we would obtain

$$\boldsymbol{A = \alpha\,(n, g_k, H_A).}$$

The second kind risk for a specific value of the alternative hypothesis and a specific distribution is the object of the calculation.

The *approach* described here can also be applied to the k-sample problem. ($k \geq 2$). In this case, however, the number of possible samples is considerably larger and the calculation can be performed on fast computers only for small n. For instance, the number of possible samples of the same size would be M^2 even for the two-sample case.

3. Example for application of the method

Let there be a distribution

$$g_3 = \begin{pmatrix} -\sqrt{3}, & 0, & \sqrt{3} \\ \frac{1}{6}, & \frac{4}{6}, & \frac{1}{6} \end{pmatrix} \tag{8}$$

This distribution has the parameters $\mu = 0$, $\sigma^2 = 1$ and $\gamma_1 = \gamma_2 = 0$. We shall calculate the interval $\underline{d}_\alpha$ using (1) from a random sample with g_3-distributed components. We shall take $\alpha = 0.05$ as our nominal α. How large is $\alpha\ (n, g_3)$ for this distribution if $n = 3$?

Table 1 shows all possible random samples and their probabilities. For instance, according to (5)

$$p_{\mathfrak{R}}^{(1)} = \frac{3!}{3!\,0!\,0!}\left(\frac{1}{6}\right)^3 \cdot \left(\frac{4}{6}\right)^0 \cdot \left(\frac{1}{6}\right)^0 = \frac{1}{6^3}\,.$$

Table 1

Possible samples and their probabilities for distribution (8) and $n = 3$

	No. 1	2	3	4	5	6	7	8	9	10
$-\sqrt{3}$	3	2	2	1	1	1	0	0	0	0
0	0	1	0	2	1	0	3	2	1	0
$\sqrt{3}$	0	0	1	0	1	2	0	1	2	3
	$\frac{1}{6^3}$	$\frac{12}{6^3}$	$\frac{3}{6^3}$	$\frac{48}{6^3}$	$\frac{24}{6^3}$	$\frac{3}{6^3}$	$\frac{64}{6^3}$	$\frac{48}{6^3}$	$\frac{12}{6^3}$	$\frac{1}{6^3}$

The means. variances, intervals and resultant vector are shown in Table 2. From this table we see that false decisions were made in only two cases. Hence,

$$\alpha(3, g_3) = P'H = \frac{2}{6^3} = \frac{1}{108} \approx 0.01\,.$$

Table 2

Means, variances, interval limits I_l and I_u and the resultant vector for distribution (8), $n = 3$ and interval (1)

	No. 1	2	3	4	5	6	7	8	9	10
$\bar{g}$	$-\sqrt{3}$	−1.15	−0.58	−0.58	0	−0.58	0	0.58	1.15	$\sqrt{3}$
s^2	0	1	4	1	3	4	0	1	1	0
I_l	$-\sqrt{3}$	−3.63	−5.54	−3.06	−4.30	−4.38	0	−1.90	−1.33	$\sqrt{3}$
I_u	$-\sqrt{3}$	1.33	4.38	1.90	4.30	4.54	0	3.06	3.63	$\sqrt{3}$
H'	1	0	0	0	0	0	0	0	0	1

Interval (1) holds for the class G_1 of normal distributions and $n \geq 2$. Let us now consider G_2, i.e. the class G_1 after extension by g_3 in (8). It must be remembered that G_2 contains only distributions for which $\gamma_1 = \gamma_2 = 0$. If we take $\varepsilon = 0.005$, (1) is not ε -robust when applied to the class G_2 of distributions conforming with the definition given in the first part of this contribution because

$$|\alpha\ (3, g_3) - 0.05| \geq 0.005.$$

It is, however, one sided 0.005-robust.

If we consider the intervals for our example we find that most boundaries are situated

outside the support of the distribution. In this case this is because the sample size is very small. The distributions of $\underline{\bar{y}}$ and $\underline{s}^2$ - are also of interest when assessing robustness. The common distribution of $\underline{\bar{y}}$ and $\underline{s}^2$ can also be given, as Table 3 shows, for all discrete distributions with a finite support if n is finite.

Table 3
Distribution of $\underline{\bar{y}}$ and $\underline{s}^2$ for g_3 according to (8) and $n = 3$

	$\bar{y}$							
s^2	$-\sqrt{3}$	-1.15	-0.58	0	0.58	1.15	$\sqrt{3}$	$\sum$
0	$\frac{1}{6^3}$	0	0	$\frac{64}{6^3}$	0	0	$\frac{1}{6^3}$	$\frac{66}{6^3}$
1	0	$\frac{12}{6^3}$	$\frac{48}{6^3}$	0	$\frac{48}{6^3}$	$\frac{12}{6^3}$	0	$\frac{120}{6^3}$
3	0	0	0	$\frac{24}{6^3}$	0	0	0	$\frac{24}{6^3}$
4	0	0	$\frac{3}{6^3}$	0	$\frac{3}{6^3}$	0	0	$\frac{6}{6^3}$
$\sum$	$\frac{1}{6^3}$	$\frac{12}{6^3}$	$\frac{51}{6^3}$	$\frac{88}{6^3}$	$\frac{51}{6^3}$	$\frac{12}{6^3}$	$\frac{1}{6^3}$	1

We now wish to calculate the power function of the two-sided t-test of the hypotheses

H_0: $\mu_0 = 0$

H_A: $\mu_0 \neq 0$

We shall do this by considering the distributions

$$g_3(\mu) = \begin{pmatrix} -\sqrt{3}+\mu, & \mu, & \sqrt{3}+\mu \\ \frac{1}{6}, & \frac{4}{6}, & \frac{1}{6} \end{pmatrix} \tag{9}$$

and representing the power function as a function of

$$T = \frac{\mu - \mu_0}{\sigma} \tag{10}$$

for = 3 and $\mu_3(g)$. We shall use Tables 1 and 2 for our calculations. Since the distribution is symmetric we can restrict our attention to $\mu \geq 0$. At the point $T = 0$ the power function has the value $1 - \beta\ (0, 3, g_3) = 1/108$. Any change in T results only in a shift of the intervals given in Table 2. If T is situated in the interval (0; 1.33), an appropriate value must be added to all interval boundaries. But the resultant vector remains $H' = (1, 0, 0, 0, 0, 0, 1, 0, 0, 1)$ for all $T \in (0; 1.33)$.

We thus obtain the value $1 - \beta\ (T, 3, g_3(\mu)) = 66/6^3 = 33/108$.

The values of the power function are given in Table 4.

In calculating the power function of the t-test we proceeded from the interval estimation and considered the *closed* interval. The test cannot be performed in its normal form because $s^2 = 0$ appears as an estimation. However, it is also quite possible to give an unambiguous decision rule by means of the so called least significant difference (or interval (1)). For normal distributions the power function of the t-test is a monotonically increasing function in R^+. The t-test has no longer this property when used for the distributions $g_3(\mu)$. The alternative hypothesis in this case is accepted for the first sample "due to the false reason". But this "false reason" vanishes at the point $\mu = \sqrt{3}$ if μ is enlarged. Hence, H_0 will be accepted for the same "displaced" sample although it is the false decision. This effect produces the drop in the power function at the point $\mu = \sqrt{3}$.

Table 4

Power function of the t-test for $g_3(\mu)$ and $n=3$, $T \in R^+$

T	$1-\beta\,(T, 3, g_3(\mu))$
0	$\frac{2}{216}$
(0; 1.33]	$\frac{66}{216}$
(1.33; $\sqrt{3}$)	$\frac{78}{216}$
[$\sqrt{3}$; $\sqrt{3}$]	$\frac{77}{216}$
[$\sqrt{3}$; 1.90]	$\frac{78}{216}$
(1.90; 3.06)	$\frac{126}{216}$
(3.06; 3.63)	$\frac{174}{216}$
(3.63; 4.30]	$\frac{186}{216}$
(4.30; 4.38)]	$\frac{210}{216}$
(4.38; 5.54]	$\frac{213}{216}$
(5.54; ∞)	1

4. First results

4.1 Discrete distributions that have been investigated

In the literature concerning the robustness of statistical procedures it is often tacitly assumed that it is sufficient to consider the first four moments of a distribution and that the further moments have no significant effect on the properties being studied. Our first considerations are designed to check the validity of this assumption. Even BRADLEY (1963, 1964, 1978, 1980a, b, c) warned against excessive trust in the robustness of the most common tests.

We consider all distributions for which the first four moments are finite, and let the variance be positive. We can divide the set of these distributions into classes $K(\mu;\sigma^2;\gamma_1;\gamma_2)$ with fixed parameters. The normal distributions are elements of the classes $K(\mu;\sigma^2;0;0)$. It is clear that, for the following investigations, it will be sufficient if we restrict our attemption to the classes $K(0; 1 ; \gamma_1; \gamma_2)$, i.e. if we consider only distributions that are standardized in respect of μ and σ^2. We first scrutinize the class $K(0; 1; 0; 0)$ rather more closely. If the above assumption were correct, we should be able to take any distribution of this class without significantly affecting the properties of the statistical method. This class naturally also contains discrete distributions.

A discrete distribution of the type defined by (3) is unambiguously defined by $2k$ - 1 parameters. The class $K(0; 1; 0; 0)$ contains no two-point distributions but a one-parametric family of three-point distributions from which we choose the particular distribution g_3 defined in (8). In addition for (8) we shall extend our investigations to

cover the distributions

$$g_5 = \begin{pmatrix} -2, & -1, & 0, & 1, & 2 \\ \frac{1}{12}, & \frac{2}{12}, & \frac{6}{12}, & \frac{2}{12}, & \frac{1}{12} \end{pmatrix}$$

and

$$g_7 = \begin{pmatrix} -3, & -2, & -1, & 0, & 1, & 2, & 3 \\ \frac{1}{180}, & \frac{9}{180}, & \frac{45}{180}, & \frac{70}{180}, & \frac{45}{180}, & \frac{9}{180}, & \frac{1}{180} \end{pmatrix}$$

all of which lie in $K(0;1;0;0)$

4.2 Estimation of a mean and comparison of a mean with a constant when the variance is known

Let the random variable $\underline{y}$ be $N(\mu;\sigma^2)$-distributed, and let σ^2 be known. Then

(11) $$\left\langle \bar{\underline{y}}_{\cdot} - \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} ; \; \bar{\underline{y}}_{\cdot} + \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right\rangle$$

is an interval estimation for the confidence coefficient 1 - α if $(y_1, y_2, ..., y_n)$ is a random sample with components distributed as $\underline{y}$. The test for the hypotheses

(12) $H_0: \mu_0 = 0$

$H_A: \mu_0 \neq 0$

can also be performed over interval (11). If σ^2 and μ are known, the transformed quantity

(13) $$\underline{u} = \frac{\underline{y} - \mu}{\sigma}$$

can be used regardless of the distribution, to investigate the interval (from "observation" $\underline{u}_i$)

(14) $$\left\langle \bar{\underline{u}}_{\cdot} - \frac{1}{\sqrt{n}} u_{1-\frac{\alpha}{2}} ; \; \bar{\underline{u}}_{\cdot} + \frac{1}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right\rangle$$

independently of μ and σ. If $\underline{y}$ is $N(\mu; \sigma^2)$-distributed, (14) is an interval estimation for 0 with the confidence coefficient 1 - α. What is the probability that this interval covers the parameter μ = 0 if $\underline{y}$ is not normally distributed? This question will be studied for the distributions g_3, of the class $K(0; 1; 0; 0)$. The results are •shown in figures 1, 2 and 3. These show that the *u*-test is robust for $\varepsilon = 0.2\alpha$ if $n \geq n_0$ (n_0 shown in the figures) if the distribution under H_0 is given by (8) but it is not robust for $\varepsilon = 0.1\alpha$ if $n < 230$. It is clear that the moments of g_3 are the same as for the normal distribution up to the 5th order. This means that the first moment differing between the two distributions is that of the 6th order.

For samples of small size there seems to be good reason to doubt the results of the investigations, based mainly on PEARSONS' distributions since no such deviations of the magnitude found in the class $K(0; 1; \gamma_1; \gamma_2)$ are suspected or investigated.

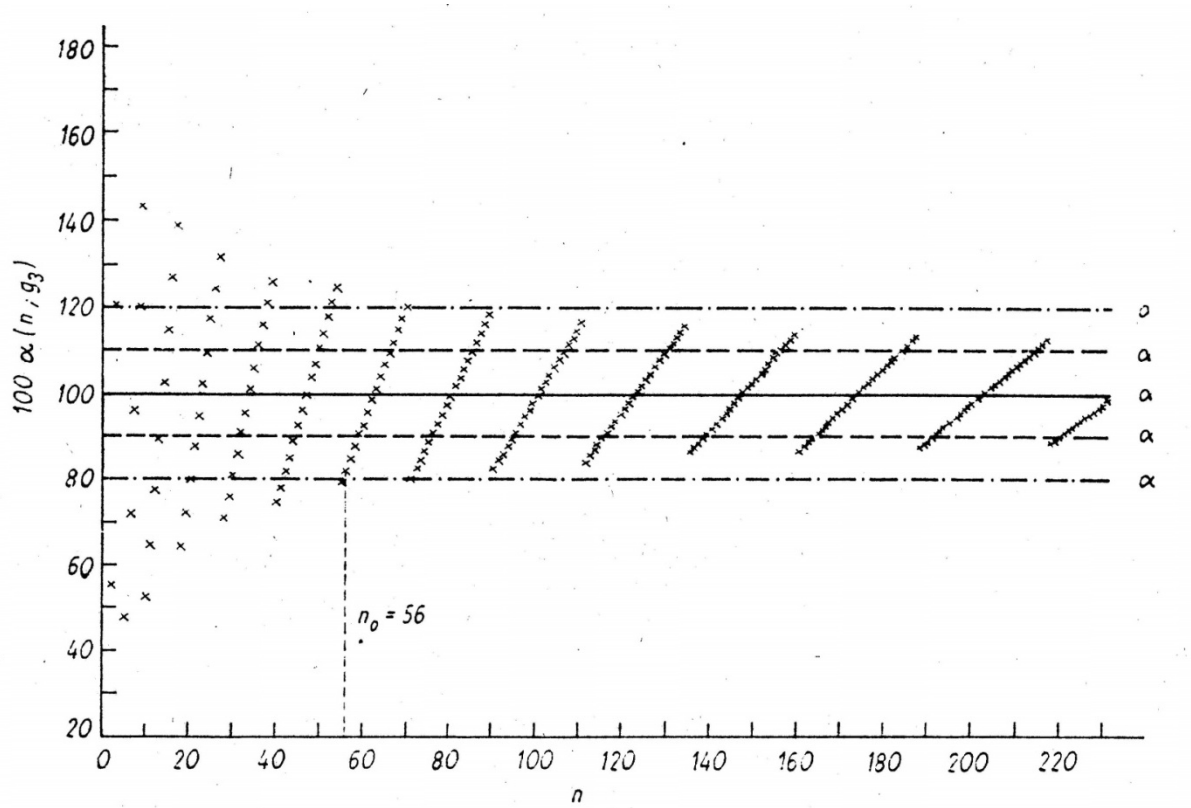

Fig. 1. $\alpha(n, g)$ for the u-test, g_3 in (8) and nominal $\alpha = 0.10$

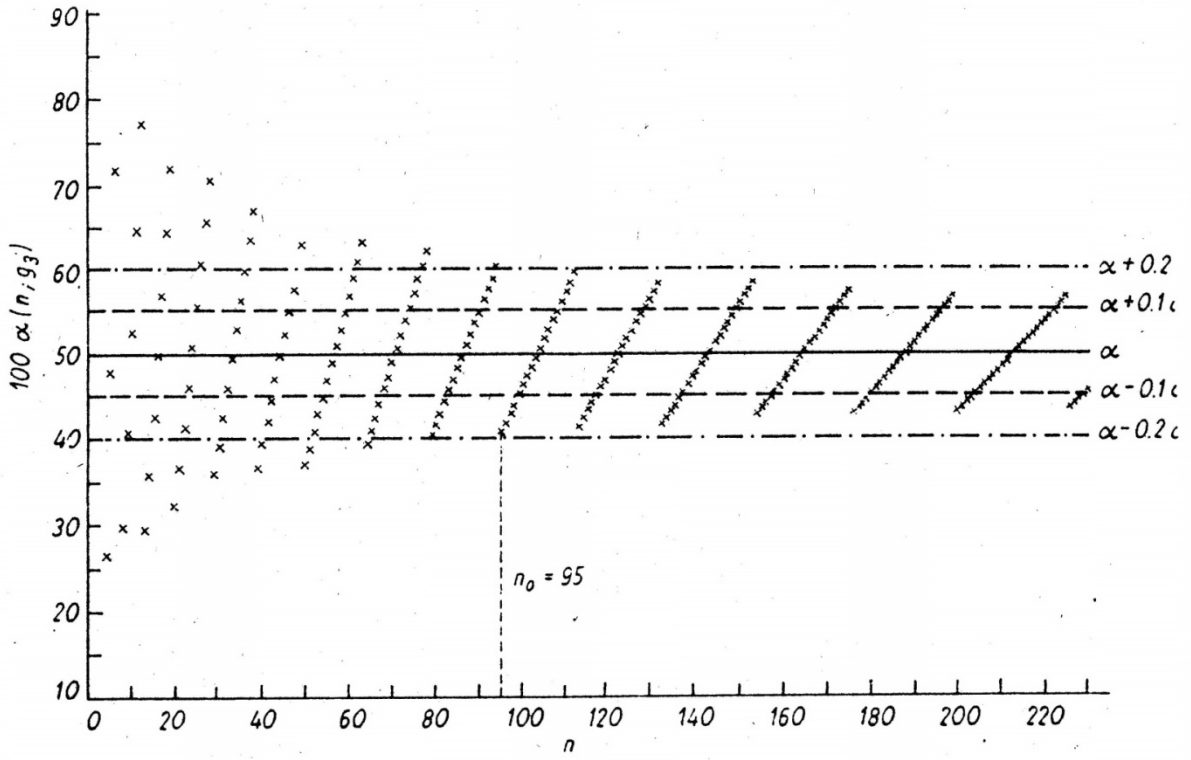

Fig. 2. $\alpha(n, g)$ for the u-test, g_3 in (8), and nominal $\alpha = 0.05$

4.3 Estimation of a mean and comparison of a mean with a constant when the variance is unknown

Let the random variable $\underline{y}$ be $N(\mu; \sigma^2)$-distributed, and let σ^2 be unknown. Equation (1) then gives an interval estimation for the confidence coefficient $1 - \alpha$ if $(\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_n)$ is a random sample comprising components that are distributed as $\underline{y}$.

Hypothesis (12) can also be tested using the interval (1). Even if the values of μ and σ^2 are unknown the transformed quantity

$$\underline{u} = \frac{\underline{y} - \mu}{\sigma}$$

can still be used regardless of the distribution (which again will be assumed to possess the first four moments) and the interval

$$(15) \qquad \left\langle \bar{\underline{u}}_{\cdot} - \frac{\underline{s}_u}{\sqrt{n}}\, t\left(n-1;\, 1-\frac{\alpha}{2}\right);\ \bar{\underline{u}}_{\cdot} + \frac{\underline{s}_u}{\sqrt{n}}\, t\left(n-1;\, 1+\frac{\alpha}{2}\right) \right\rangle$$

can be investigated independently of μ and σ^2. If y is $N(\mu; \sigma^2)$-distributed, (16) is an interval estimation for 0 with the confidence coefficient $1 - \alpha$. Here $\underline{s}_u^2$ is the estimation function of the variance of a random variable that is $N(0; 1)$-distributed. We shall again put the question of the probability with which the interval (16) covers the parameters 0 when y is not $N(\mu; \sigma^2)$-distributed, proceeding from the same distributions as we used in 4.2. The results of the calculations for the real $\alpha(n; g_k)$ are shown in figures 4, 5 and 6 for $k = 3$. The calculation of $\alpha(n; g_k)$ by means of the conventional test statistic $\underline{t}$ for g_3 is impossible because estimates $s_u = 0$ appear with a probability $p > 0$. The closed interval (16) must therefore always be used as a starting point. If $s_u = 0$, the interval contains exactly one point.

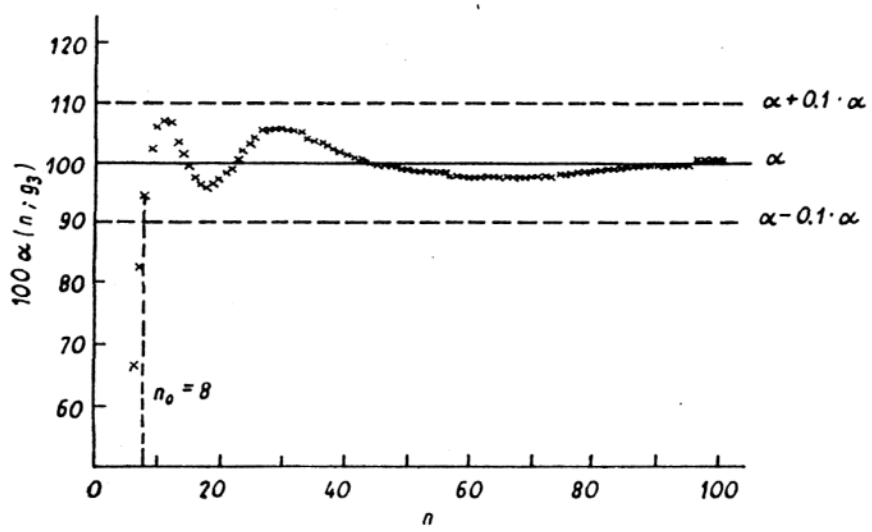

Fig. 4. $\alpha(n, g)$ for the t-test, g_3 in (8) and nominal $\alpha = 0.10$

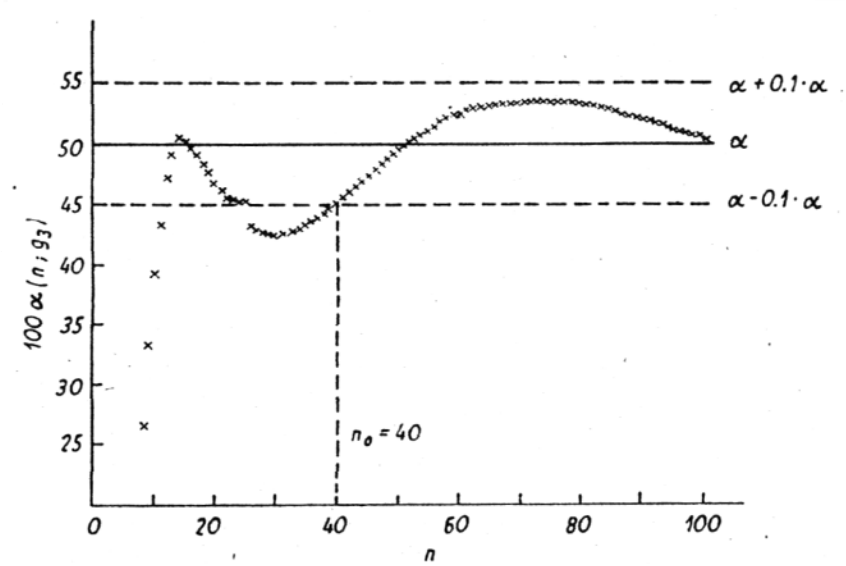

Fig. 5. $\alpha(n, g)$ for the t-test, g_3 in (8) and nominal $\alpha = 0.05$

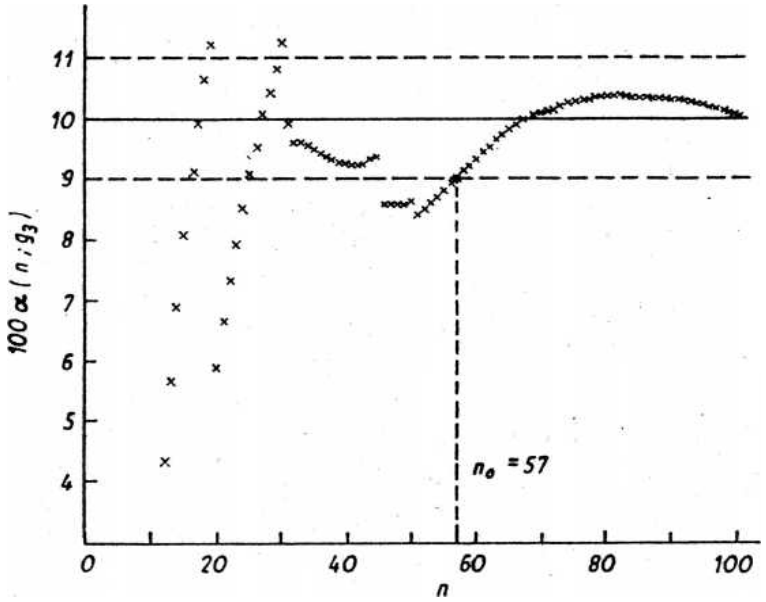

Fig. 6. $\alpha(n, g)$ for the t-test, g_3 in (8) and nominal $\alpha = 0.01$

H_0 must be accepted if this is zero and rejected otherwise. But a false decision has been taken if H_A is accepted.

The values above which the procedure can be considered robust for $\varepsilon = 0.1\alpha$ are shown in figures 4, 5 and 6. It is assumed that, for larger n, $\alpha(n; g_k)$ differs from α by less than 0.1α. These investigations have also been continued up to 230.

It must be noted that we have here considered only one three-point distribution of the one-parameter class of three-point distributions in $K(0;\ 1;\ 0;\ 0)$. The $\alpha(n; g_k))$, and thus also the values n from which the procedure may be considered robust, for other three-point distributions may be different.

4.4 Estimation of a variance and comparison of a variance with a constant

Let the random variable $\underline{y}$ be $N(\mu; \sigma^2)$-distributed, and let $(\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_n)$ be a random sample with components distributed as $\underline{y}$. Then

$$(16) \qquad \left\langle \frac{\underline{s}^2 (n-1)}{\chi^2\left(n-1;\, 1-\frac{\alpha}{2}\right)};\ \frac{\underline{s}^2 (n-1)}{\chi^2\left(n-1;\, \frac{\alpha}{2}\right)} \right\rangle$$

is an interval estimation of o^2 for the confidence-coefficient 1 - α. The hypotheses

$$(17) \qquad H_0 : \sigma^2 = \sigma_0^2$$
$$H_A : \sigma^2 \neq \sigma_0^2$$

can also be tested by means of (17). We consider the quantity (18)

$$(18) \qquad \underline{h} = \frac{\underline{y}}{\sigma_0} .$$

The interval

$$(19) \qquad \left\langle \frac{\underline{s}_h^2 (n-1)}{\chi^2\left(n-1;\, 1-\frac{\alpha}{2}\right)};\ \frac{\underline{s}_h^2 (n-1)}{\chi^2\left(n-1;\, \frac{\alpha}{2}\right)} \right\rangle$$

is an estimation of $o^2 = 1$ with a confidence coefficient 1 -α. We want to know the probability with which interval (19) will cover the parameters $\sigma^2 = 1$ if the random quantity is not normally distributed. Table 5 shows the results of the calculations for (g_3, g_5 and g_7. The values at which the χ^2-test is first robust for $\varepsilon=0.2\alpha$ are enclosed in frames. The fact that they are, on average, higher than for the *t*-test is worthy of note.

The χ^2-test appears to be somewhat less robust than the *t*-test but considerably more robust than the *u*-test. However, the $\chi^2\alpha(n; g_k)$ fluctuate about the nominal values of α. It is conspicuous that the χ^2-test in the class $K(\mu; \sigma^2; 0; 0)$ is absolutely useless for $n \leq 4$ but then rapidly (for g_5 and g_7, not for g_3) approaches the nominal values of α.

The two other test procedures didn't such major deviations for small *n*. Furthermore, the improvement from g_3 to g_5 and g_7 is conspicuous. At g_7 the *t*-test and the χ^2-test can scarcely be distinguished from each other as far as robustness within the class $K(\mu; \alpha^2; 0; 0)$ is concerned.

Further investigations will be made with respect to other cases $K(\mu; \alpha^2; \gamma_4; \gamma_2)$ with $\gamma 1 \neq 0$, $\gamma 2 \neq 0$.

Table 5

$\alpha(n; g_k)$ for the test or interval estimation of σ^2 as a function of n for $k=3.5$ and 7 in ‰

n	nominal α ‰ for g_3			nominal α ‰ for g_5			nominal α ‰ for g_7		
	100	50	10	100	50	10	100	50	10
2	556	556	500	389	333	333	352	293	293
3	444	333	306	194	174	139	149	121	97
4	284	204	199	137	81	64	88	52	36
5	172	150	132	75	58	33	57	57	15
6	130	99	88	131	43	18	87	35	9
7	88	88	59	134	61	9	97	42	7
8	97	55	40	134	64	26	112	41	11
9	177	44	27	115	80	16	99	56	8
10	139	29	18	109	60	19	87	49	10
11	109	98	12	109	66	18	89	51	8
12	98	68	8	106	63	15	88	48	8
13	116	53	40	116	62	20	101	47	10
14	109	42	29	112	67	15	98	53	10
15	121	68	20	112	57	18	99	46	10
16	102	71	16	110	59	16	107	49	10
17	114	63	11	110	54	17			
18	82	61	21	113	53	14			
19	128	57	25	109	59	15			
20	103	72	20	109	54	13			
21	101	61	14	103	54	14			
22	130	58	18	106	55	15			
23	117	48	13						
24	116	61	22						
25	88	43	16						
26	107	56	16						
27	124	48	12						
28	104	54	17						
29	109	65	19						
30	106	53	16						
31	117	56	12						
32	95	44	14						
33	95	63	19						
34	108	61	15						
35	108	49	12						
36	102	55	13						
37	94	61	16						
38	118	55	16						
39	108	57	15						
40	98	49	12						
41	107	61	13						
42	105	55	16						
43	104	51	13						

Acknowledgements

We wish to thank Mr. **RIEMANN,** Computer Centre of the WILHELM - PIECK - Universität Rostock, for performing the computations,

Zusammenfassung

Es wird eine exakte Methode zur Beurteilung der Robustheit der *u*-, *t*- und χ^2-Maßzahl im Einstichprobenproblem gegenüber diskreten Alternativen zur Normalverteilung vorgestellt. Erste Ergebnisse bis zu Stichprobenumfängen von n = 230 werden angegeben.

References

AUTORENKOLLEKTIV: Probleme der angewandten Statistik, Heft 4, Robustheit I, Herausgeber: AdL der DDR, Forschungszentrum für Tierproduktion Dummerstorf-Rostock.

BARTLETT, M. S., 1935: The effect of non-normality on the *t*-distribution. Proc. Comb. Phil. Soc. **31**, 223-231.

BLUMENTHAL, S.; GOVINDARAJULU Z., 1977: Robustness of Stein's two-stage procedure for mixtures of normal population. JASA **72**, 192-196.

BRADLEY, R. A., 1952: The distribution of the *t* and *F* statistics for a class of non-normal populations. Virginia J. Sci. **3**, 1-32.

BRADLEY, J. V., 1963: Studies in Research Methodology: IV. A sampling study of the central limit theorem and the robustness of one-sample parametric tests. AMRL Technical Documentary Report.

BRADLEY, J. V., 1964: Studies in research methodology VI: The central limit effect for a variety of populations and the robustness of *z, t* and *F*. Aerospace Medical Research Laboratories (AMRL) TR - **64** - 123.

BRADLEY, J. V., 1978: Robustness? British J. Math. Statist. Psychol. **31**, 144-152.

BRADLEY, J. V., 1980a: Nonrobustness in one-sample *Z* and *t* tests: A large scale sampling study. Bull. Psychonomic Soc. **15**, 29-32.

BRADLEY, J. V., 1980b: Nonrobustness in classical tests on means and variances. A large-scall sampling study. Bull. Psychonomic Soc. **15**, 275-278.

BRADLEY, J. V., 1980c: Nonrobustness in *Z*, *t* and *F* tests at large sample sizes. Bull. Psychonomic Soc. **15**, 333-336.

DAVID, F. N.; JOHNSON, N. L., 1951: The effect of non-normality on the power function of the *F*-test in the analysis of variance. Biometrika **38**, 43-57.

GAYEN, A. K., 1949: Student's *t* in samples of any size from non-normal universes. Biometrika **36**, 353-369.

GEARY, R. C., 1936: The distribution of "Student's" ratio for non-normal samples JRSS 3, 178-184.

GEARY, R. C., 1947: Testing for normality. Biometrika **34**, 209-242.

GHURYE, S. G., 1949: On the use of students *t*-test in a asymmetrical population. Biometrika **36**, 426-430.

GRONOW, D. G. C., 1953: Non-normality in two-sample *t*-tests. Biometrika **40**, 222-225.

HYRENIUS, H., 1950: Distribution of "Student"-Fisher's *t* in samples from compound normal functions. Biometrika **37**, 429.

KENDALL, M. G.; STUART, A., 1967: The advanced theory of statistics, inference and relationship, Volume **2**. Griffin, London.

LADERMAN, J., 1939: The distribution of "Student's" ratio for samples of two items drawn from non-normal universes. Ann. Math. Statist. **10**, 376-379.

LEE, A. F. S.; GURLAND, J., 1977: One-Bample *t*-test when sampling from a mixture of normal distributions. Ann. Statist. **5**, 803-807.

PEARSON, E. S.; ADYANTHAYA, N. K. 1929: Small samples from non-normal populations. Biometrika **21**, 259-286.

PERLO, V. P., 1933: Student's ratio for samples of 3 from a rectangular population. Biometrika **25**, 203-204.

POSTEN, H. O., 1978: The robustness of the two-sample *t*-test over the Pearson system. J. Statist. Comput. Simul. **6**, 295-311.

POSTEN, H. O., 1979: The robustness of the one-sample *t*-test over the Pearson system. J. Statist. Comput. Simul. **9**, 133-149.

RASCH, D., HERRENDÖRFER, G., 1981: Robustheit statistischer Methoden I. Begriffsbildung und Entwicklung. Rostock. Math. Kolloq. **17**, 87-104.

RATCLIFFE, J. F., 1968: The effect on the *t*-distribution of non-normality in the sampled population. Appl. Statist. **17**, 42-48.

SANSING, R. C.; OWEN, D. B., 1974: The density of the *t*-statistic for non-normal distributions. Commun. Statist. **3**, 139-155.

SRIVASTAVA, A. B. L., 1958: Effect on non-normality on the power function of *t*-test. Biometrika **45**, 421-430.

TIKU, M. L., 1963: Student's *t*-distribution in terms of Hermite and Laguerre polynomials. J. Indian Math. Soc. **27**, 91.

Manuscript received: 20.8.1981

Author's address:

Dr. sc. G. HERRENDÖRFER
Prof. Dr. D. RASCH
Dr. K. D. FEIGE

Forschungszentrum für Tierproduktion
Abt. Biomathematik und EDV
DDR - 2551 Dummerstorf, Krs. Rostock