Effect of Non-Normality on Tolerance Limits
Which Control Percentages in Both Tails of Normal Distribution

J. N. K. Rao*, K. Subrahmaniam*

University of Manitoba, Canada

and

D. B. Owen **

Southern Methodist University
Dallas, Texas

Effect of non-normality on Owen's tolerance limits, which control percentages \( p/2 \) in each of the tails of the normal distribution, is investigated. The limits are insensitive to departures from normality for \( p \geq 0.20 \), but the effect of non-normality is increasingly felt as \( p \) decreases. Factors for tolerance limits in the non-normal case, which require at least a rough knowledge of parental skewness \( g_i \) and kurtosis \( g_k \), are given.

Key Words
Factors for Tolerance Limits
Effect of Non-Normality
Robustness
Control of Percentages
Tails of Normal Distribution
Cornish-Fisher Method
Parental Skewness and Kurtosis

1. Introduction

Let \( x_1, x_2, \cdots, x_n \) be a random sample of size \( n \) drawn from a normal distribution with unknown mean \( \mu \) and unknown variance \( \sigma^2 \). Owen (1964, 1965) has given formulae for computing the factors \( k_1 \) and \( k_2 \) from which the tolerance limits \( \bar{x} - k_1 s \) and \( \bar{x} + k_2 s \) may be obtained. These limits specify with probability \( P \) that no more than proportion \( p_1( < \frac{1}{2}) \) of the normal distribution is below \( \bar{x} - k_1 s \) and no more than proportion \( p_2( > \frac{1}{2}) \) is above \( \bar{x} + k_2 s \), where \( \bar{x} = \frac{1}{n} \sum x_i \) and \( s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 \). These limits differ from the previous ones in the literature (e.g., Wald and Wolfowitz, 1946) which only guarantee that the proportion between \( \bar{x} - k_1 s \) and \( \bar{x} + k_2 s \) is at least \( 1 - p \) with probability \( P \). Moreover, Owen's formulae are free from approximations which characterize the previous results in the literature.

* Supported by research grants from the National Research Council of Canada.
Received Dec. 1970; revised May 1971.
As in many cases the assumption of normality may not hold or may hold only approximately, it is important to investigate how the above tolerance limits are affected by deviations from normality. To this end we follow the approach of Gayen (1949) and assume that the frequency function of \( \xi = (X - \mu)/\sigma \) is

\[
f(\xi) = G'(\xi) - \frac{\lambda_3}{4} G^{(4)}(\xi) + \frac{\lambda_4}{24} G^{(6)}(\xi) + \frac{\lambda_5}{72} G^{(8)}(\xi)
\]

where \( E(X) = \mu, V(X) = \sigma^2, G^{(\nu)}(\xi) \) is the \((\nu - 1)\)th derivation of the standardized normal density

\[
G'(\xi) = \left(2\pi\right)^{-\frac{1}{2}} \exp\left(-\frac{\xi^2}{2}\right)
\]

and \( \lambda_3 = \beta_3, \lambda_4 = \beta_4 - 3 \) are respectively the standardized third and fourth cumulants of \( X \). Since the effects of the higher order terms depending on \( \lambda_5, \lambda_6, \lambda_7, \lambda_8, \cdots \) are assumed to be negligible, the population (1) is only moderately non-normal. Moreover, too high values of \( \lambda_3 \) and \( \lambda_4 \) cannot be permitted as they will make \( f(\xi) \) negative at one or both tails and will lead to subsidiary modes. In fact, to ensure that (1) is unimodal and positive definite, \( \lambda_4 \) should be roughly between 0 and 2.4 and \( \lambda_3 \leq 0.2 \) (Barton and Dennis, 1952).

In Section 2, we give the necessary formulae for studying the effect of non-normality on the tolerance limits of Owen. An empirical investigation is undertaken in Section 3; however, it is confined to the important special case of \( p_1 = p_2 = p/2 \) and \( k_1 = k_2 = k \). Finally, in Section 4, we tabulate the values of the constant \( k \) which determine the tolerance limits \( \bar{x} - \xi_s, \bar{x} + \xi_s \) such that no more than proportion \( p/2 \) of the population (1) lies in each tail with probability \( P \geq 0.90 \). These limits require at least a rough knowledge of \( \lambda_3 \) and \( \lambda_4 \).

2. Formulae

We first obtain an expression for the probability \( P \) that no more than proportion \( p_i \) is below \( \bar{x} - \xi_s \) and no more than the proportion \( p_2 \) is above \( \bar{x} + \xi_s \) when sampling from population (1). Now

\[
P = \Pr(X \leq \bar{x} - \xi_s) \leq p_1 \quad \text{and} \quad \Pr(X \geq \bar{x} + \xi_s) \leq p_2.
\]

Let

\[
t_1 = K_p \sqrt{n}, \quad t_2 = -K_{1-p} \sqrt{n}, \quad \delta_1 = -K_{p} \sqrt{n}, \quad \delta_2 = -K_{1-p} \sqrt{n}
\]

and

\[
Z_1 = \frac{W}{\sqrt{f}} - \delta_1, \quad Z_2 = Y^2
\]

where

\[
\frac{\bar{x} - \mu}{\sigma} = \frac{Z_1}{\sqrt{n}}, \quad \frac{s^4}{\sigma^4} = \frac{Z_2}{f}, \quad f = n - 1
\]

and \( K_p \) is such that

\[
\int_{-K_p} f(\xi) \, d\xi = 1 - p_i, \quad i = 1, 2.
\]
Then (3) reduces to

\[ P = \Pr_{W,Y}(W \leq t_1 \text{ and } (W - t_2)Y \geq (\delta_1 - \delta_2)(n - 1)^{\frac{1}{2}}) \].

An approximate value of \( K_{pi} \) for a given \( p \), can be obtained by the Cornish–Fisher method (see Fisher and Cornish, 1960).

Let

\[ A_1 = t_1 f^{-1}, \quad A_2 = t_2 f^{-1} \quad \text{and} \quad R = (\delta_1 - \delta_2)/(A_1 - A_2). \]

Using the joint distribution of \( Z_1 \) and \( Z_2 \) given by Gayen (1949), we get

\[ P = P_0 + \lambda_3 P_{\lambda_3} + \lambda_4 P_{\lambda_4} + \lambda_5^2 P_{\lambda_5} \]

where

\[ P_0 = Q_i(A_1, \delta_1 ; R, \infty) - Q_i(A_2, \delta_2 ; R, \infty) \]

\[ P_{\lambda_3} = \left[ 3n/(2\pi)^{\frac{3}{2}} \Gamma\left((f/2)2^{(f-2)/2}\right)^{-1} \left( A_1^2 + 3 \right) D_i^*(A_1, \delta_1) - (A_2^2 + 3) D_i^*(A_2, \delta_2) \right] \]

\[ + 2A_2\delta_2 D_i^*(A_2, \delta_2) - 2A_1\delta_1 D_i^*(A_1, \delta_1) - (\delta_1^2 - 3n + 2) D_i^*(A_1, \delta_1) \]

\[ + (\delta_2^2 - 3n + 2) D_i^*(A_2, \delta_2) \]

(12)

and the expressions for \( P_{\lambda_4} \) and \( P_{\lambda_5} \) (which are somewhat lengthy) are given in Rao et al. (1970), where

\[ Q_i(A_1, \delta_1 ; R, \infty) = \left[ (2\pi)^{-1} \Gamma(f/2)2^{(f-2)/2} \right]^{-1} \int_{-\infty}^{\infty} G(A, y - \delta_i) y^{f-1} G'(y) dy \]

(13)

and

\[ D_i^*(A_1, \delta_1) = \frac{1}{2\pi A_1^2} G(A, R - \delta_1) R G'(R) \]

\[ + s \Gamma(s/2) 2^{s/2} A_1^{-s} \left[ Q_i(A_1, \delta_1 ; R, \infty) - Q_{i+1}(A_1, \delta_1 ; R, \infty) \right]. \]

Owen (1965) has given explicit formulae for computing (13) and the reader is referred to Rao et al. (1970) for details. For the special case \( p_1 = p_2 \) and \( k_1 = k_2 \), we have \( t_1 = -t_2 \) and \( A_1 = -A_2 \).

3. EMPIRICAL INVESTIGATION

We now investigate the effect of non-normality on Owen’s (1964) tolerance limits. To this end, we evaluate \( P \), for selected values of \( p_1 = p_2 = p/2 \), \( (\lambda_3, \lambda_4) \), and \( n \), by substituting Owen’s factors \( k_1 = k_2 = k^* \) (corresponding to \( P = 0.90 \)) in place of \( k_1 = k_2 = k \). The values of \( (\lambda_3, \lambda_4) \) chosen are \((-0.5, 0.0), (-0.5, 1.0), (-0.25, 0.0), \) and \((0.5, 0.0) \) with \( p = 0.20 \), \( 0.10 \), and \( 0.05 \). Although more extensive tables have been prepared by the authors, for different combinations of \( (\lambda_3, \lambda_4) \) and larger number of values of \( n \), we present in Table I only a selected subset of the results. We have also obtained the contribution made by each of the terms \( P_0 \), \( P_{\lambda_3} \), \( P_{\lambda_4} \), and \( P_{\lambda_5} \). As is to be expected, the value of \( P_0 \) is the predominant factor, with the other terms contributing corrections of various magnitudes. Unfortunately, it is not possible to separate the first component \( P_0 \) from the effects of \( (\lambda_3, \lambda_4) \) as has been done in other situations by various authors. [See, for example, Gayen]
Table I

Values of $P$ for two-sided tolerance limits in the non-normal case; corresponding normal value of $P$ is 0.90.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$[-.5]$</th>
<th>$[0]$</th>
<th>$[-.5]$</th>
<th>$[1]$</th>
<th>$[-.5]$</th>
<th>$[2]$</th>
<th>$[0]$</th>
<th>$[-.5]$</th>
<th>$[0]$</th>
<th>$[-.5]$</th>
<th>$[1]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>.888</td>
<td>.916</td>
<td>.898</td>
<td>.926</td>
<td>.907</td>
<td>.953</td>
<td>.591</td>
<td>.918</td>
<td>.972</td>
<td>.099</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.889</td>
<td>.888</td>
<td>.897</td>
<td>.898</td>
<td>.909</td>
<td>.904</td>
<td>.997</td>
<td>.897</td>
<td>.880</td>
<td>.086</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.886</td>
<td>.887</td>
<td>.896</td>
<td>.907</td>
<td>.901</td>
<td>.906</td>
<td>.986</td>
<td>.897</td>
<td>.885</td>
<td>.086</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>.879</td>
<td>.886</td>
<td>.894</td>
<td>.901</td>
<td>.907</td>
<td>.919</td>
<td>.994</td>
<td>.901</td>
<td>.985</td>
<td>.084</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>.868</td>
<td>.868</td>
<td>.869</td>
<td>.871</td>
<td>.871</td>
<td>.871</td>
<td>.865</td>
<td>.877</td>
<td>.868</td>
<td>.080</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>.856</td>
<td>.841</td>
<td>.868</td>
<td>.892</td>
<td>.912</td>
<td>.898</td>
<td>.877</td>
<td>.897</td>
<td>.886</td>
<td>.078</td>
<td></td>
</tr>
</tbody>
</table>

Table II

Values of $k$ for two-sided tolerance limits in the non-normal case, $(P \geq 0.90)$; parameters of non-normality ($\lambda_3, \lambda_4$).

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$0.0$</th>
<th>$0.5$</th>
<th>$1.0$</th>
<th>$1.5$</th>
<th>$2.0$</th>
<th>$2.5$</th>
<th>$3.0$</th>
<th>$3.5$</th>
<th>$4.0$</th>
<th>$4.5$</th>
<th>$5.0$</th>
<th>$5.5$</th>
<th>$6.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>5.70</td>
<td>5.07</td>
<td>4.46</td>
<td>4.13</td>
<td>4.70</td>
<td>4.70</td>
<td>4.70</td>
<td>4.70</td>
<td>4.70</td>
<td>4.70</td>
<td>4.70</td>
<td>4.70</td>
<td>4.70</td>
</tr>
<tr>
<td>8</td>
<td>2.22</td>
<td>2.01</td>
<td>1.22</td>
<td>1.22</td>
<td>1.22</td>
<td>1.22</td>
<td>1.22</td>
<td>1.22</td>
<td>1.22</td>
<td>1.22</td>
<td>1.22</td>
<td>1.22</td>
<td>1.22</td>
</tr>
<tr>
<td>15</td>
<td>2.33</td>
<td>2.65</td>
<td>2.33</td>
<td>2.33</td>
<td>2.33</td>
<td>2.33</td>
<td>2.33</td>
<td>2.33</td>
<td>2.33</td>
<td>2.33</td>
<td>2.33</td>
<td>2.33</td>
<td>2.33</td>
</tr>
<tr>
<td>30</td>
<td>2.70</td>
<td>2.32</td>
<td>2.37</td>
<td>2.37</td>
<td>2.37</td>
<td>2.37</td>
<td>2.37</td>
<td>2.37</td>
<td>2.37</td>
<td>2.37</td>
<td>2.37</td>
<td>2.37</td>
<td>2.37</td>
</tr>
<tr>
<td>50</td>
<td>2.55</td>
<td>2.15</td>
<td>2.31</td>
<td>2.20</td>
<td>2.12</td>
<td>2.08</td>
<td>2.04</td>
<td>2.01</td>
<td>1.98</td>
<td>1.95</td>
<td>1.92</td>
<td>1.89</td>
<td>1.86</td>
</tr>
</tbody>
</table>

The effect of $\lambda_4$ is to compensate for the effect of $\lambda_3$, although the correction afforded by $\lambda_4$ tends to increase $P$. A larger value of $p$ is associated with a dampened effect of the departure from normality—the value of $P$ being con-

(1949), Subrahmaniam (1966). In summary, one can see readily that the effects of the non-normality are (i) increasingly felt as the value of $p$ is decreased; (ii) dependent on the relative magnitudes and the signs of the parameters $\lambda_3, \lambda_4$; and (iii) more pronounced as the sample size $n$ increases.
sistent large for \( p = .20 \) and falling off as \( p \) decreases through 0.10 to 0.05. Thus the values of \( P \) for \( p = 0.20 \) are the most robust to departures from normality.

### 4. Two Sided Tolerance Limits

In Table 2 we present the values of \( k \) for \( p = 0.10 \) and 0.05 and selected \( N \), \( (X_3, X_4) \) which in the non-normal case would yield the value of \( P = 0.90 \) or more. Since an increase of \( k \) would correspondingly increase \( P \), it was not difficult to determine \( k \). A computer program for evaluating \( k \) corresponding to any desired combinations of \( \lambda_3, \lambda_4, n, p \) and \( P \) is available.

### 5. Acknowledgment

The authors wish to thank Mr. Arnold Yudell and Miss Lynn Folliott for their help with the programming.

### References