Technical Report No. 93

October, 1966

INVESTIGATION OF REJECTION RULES FOR OUTLIERS IN SMALL SAMPLES FROM THE NORMAL DISTRIBUTION. IV: ESTIMATION IN THE CASE WHERE BOTH PARAMETERS ARE UNKNOWN.

by

Irwin Guttman and Dennis Smith*

*Smith currently at HRB-Singer, Inc., State College, Pa.

This work was partially supported by the National Science Foundation and the Wisconsin Alumni Research Foundation, October 1966.
0. Summary

We assume that we have a sample \((y_1, \ldots, y_n)\) of \(n\) independent observations, hopefully all from \(N(\mu, \sigma^2)\), but where perhaps one of the observations may be spurious, arising from either \(N(\mu + \sigma, \sigma^2)\) or \(N(\mu, (1+b) \sigma^2)\). However, unlike the assumptions of the previous reports, we assume that we are concerned with either

(i) estimating \(\mu\) when \(\sigma^2\) is unknown, or

(ii) estimating \(\sigma^2\) when \(\mu\) is unknown.
1. Introduction

This report, an extension of the work done in the previous reports in this series, considers the estimation problem when both parameters of the parent normal distribution are unknown. Our investigation here is limited to the case where \( n = 3 \). This is because of the inherent difficulties of evaluating the premiums and protections for large sample sizes. These difficulties are discussed in section 5.

In section 2 we discuss the case of estimating \( \mu \) when \( \sigma^2 \) is unknown \( (n=3) \) and obtain the formulas for the premiums and protections of the A, W, and S-Rules.

In section 3 we discuss the case of estimating \( \sigma^2 \) when \( \mu \) is unknown \( (n=3) \) and obtain the formulas for the premiums and protections of the A, W, and S-Rules.

Numerical results are given in section 4.
2. Estimation of $\mu$ When $\sigma^2$ Unknown (n=3)

Suppose a sample of three observations $(y_1, y_2, y_3)$ is taken, hopefully from a normal distribution $N(\mu, \sigma^2)$, but where perhaps one observation is spurious, being from $N(\mu + \alpha \sigma, \sigma^2)$ or $N(\mu, (1+b) \sigma^2)$. A rejection rule may, of course, be applied to try to offset any bias in the resultant estimation of $\mu$.

From Report # 90, we see that these rules use estimators whose values are dependent upon the rejection boundary

$$\max (|z_{(1)}|, |z_{(3)}|) = Cv$$

where

$$v^2 = \frac{1}{2} \sum_1^3 (y_i - \bar{y})^2 = \frac{1}{2} \sum_1^3 z_i^2,$$

as defined in section 5.1 of Report # 90.

For a sample of three, the ordered residuals are $z_{(1)} < z_{(2)} < z_{(3)}$, and since $z_{(2)} = -z_{(1)} - z_{(3)}$, we have that

$$v^2 = z_{(1)}^2 + z_{(1)} z_{(3)} + z_{(3)}^2$$

Let us now assume for the moment that $|z_{(1)}| > |z_{(3)}|$ . Thus, our interest is in the boundary $|z_{(1)}| = Cv$. Since we have assumed that $|z_{(1)}| > |z_{(3)}|$, i.e., $-z_{(1)} > z_{(3)}$, it follows that $|z_{(1)}| > v$.

From (3.1) of Report # 91, we see that the region of definition of $(z_{(1)}, z_{(3)})$ is

$$\{(z_{(1)}, z_{(3)}) \mid \frac{-z_{(1)}}{2} < z_{(3)} < -2z_{(1)} \}.$$  \hspace{1cm} (2.1)

Thus, if $|z_{(1)}| > |z_{(3)}|$, then $v < |z_{(1)}| < \frac{2}{3^{\frac{3}{2}}} v$, and likewise, if
If $|z(3)| > |z(1)|$, then $v < |z(3)| < \frac{2}{3^{1/2}} v$. Hence, if $C < 1$, a rejection will always be made, and if $C > \frac{2}{3^{1/2}}$, a rejection will never be made.

Let us return to the assumption that $|z(1)| > |z(3)|$, and consider again the rejection boundary $|z(1)| = C v$. Solving for $z(3)$ in terms of $z(1)$, we have that

$$z(3) = z(1) \left[ -\frac{1}{2} \pm \frac{(4-3C^2)^{1/2}}{2C} \right].$$

This, coupled with the region of definition (2.1) and the assumption that $|z(1)| > |z(3)|$, implies that the rejection region for the observation $y(1)$ corresponding to $z(1)$ is

$$-z(1) \cdot \max \left[ 1, \frac{1}{2}, \frac{1}{2} - \frac{(4-3C^2)^{1/2}}{2C} \right] < z(3) < -z(1) \cdot \min \left[ 1, \frac{1}{2} + \frac{(4-3C^2)^{1/2}}{2C} \right]$$

and since $\frac{(4-3C^2)^{1/2}}{2C} > 0$, the rejection region is

$$\frac{-z(1)}{2} < z(3) < -Rz(1)$$

or, equivalently,

$$-2z(3) < z(1) < \frac{-z(3)}{R}$$

where

$$R = \begin{cases} \frac{1}{2} + \frac{(4-3C^2)^{1/2}}{2C} & \text{if } C \geq 1 \\ 1 & \text{if } C < 1 \end{cases}$$

(2.2)

Similarly, we find that when $|z(3)| > |z(1)|$, the rejection region is

$$\frac{-z(1)}{R} < z(3) < -2z(1).$$

Since we have written the rejection region in terms of $z(1)$ and $z(3)$, we need not be concerned with the distribution of $v$. Let us now define, for a
given $R$, the events (or regions)

\[ T_0 = \{ (z(1), z(3)) \mid 0 < z(3) < \infty, -\frac{z(3)}{R} < z(1) < -\frac{z(3)}{R} \} \]

(2.3)

\[ T_1 = \{ (z(1), z(3)) \mid 0 < z(3) < \infty, -2z(3) < z(1) < -z(3) \} \]

(2.4)

\[ T_3 = \{ (z(1), z(3)) \mid 0 < z(3) < \infty, -\frac{z(3)}{R} < z(3) < -2z(3) \} \]

(2.5)

Hence, using the definitions of $T_0$, $T_1$, and $T_3$, we see from sections 5.2, 5.3, and 5.4 of Report # 90 that the estimators $\hat{u}_A$, $\hat{u}_W$, and $\hat{u}_S$, may be written as $\hat{u}_A = \bar{y} + A$, $\hat{u}_W = \bar{y} + W$, and $\hat{u}_S = \bar{y} + S$, where

\[
A = \begin{cases} 
0 & \text{if } (z(1), z(3)) \in T_0 \\
-\frac{z(1)}{2} & \text{if } (z(1), z(3)) \in T_1 \\
-\frac{z(3)}{2} & \text{if } (z(1), z(3)) \in T_3 
\end{cases}
\]

(2.6)

\[
W = \begin{cases} 
0 & \text{if } (z(1), z(3)) \in T_0 \\
-\frac{z(3) + 2z(1)}{3} & \text{if } (z(1), z(3)) \in T_1 \\
\frac{z(1) + 2z(3)}{3} & \text{if } (z(1), z(3)) \in T_3 
\end{cases}
\]

(2.7)
and

\[
S = \begin{cases} 
0 & \text{if } (z_{(1)}, z_{(3)}) \in T_0 \\
\frac{CV + z_{(1)}}{3} & \text{if } (z_{(1)}, z_{(3)}) \in T_1 \\
\frac{CV - z_{(3)}}{3} & \text{if } (z_{(1)}, z_{(3)}) \in T_3 
\end{cases}
\]  

(2.8)

With the estimators written in this form, we shall proceed to consider premiums and protections.

2.1. Premiums (n=3)—Estimation of \( \mu \) When \( \sigma^2 \) Unknown

For a given rejection constant \( C \), the premiums are of the same functional form as in the case when \( \sigma^2 \) is known. In fact, the analysis of section 2.1 of Report # 91 carries through, except that the functions \( A \), \( W \), and \( S \) are now different and given by (2.6), (2.7), and (2.8), respectively, reflecting, of course, the fact that \( \sigma^2 \) is no longer assumed known. Hence, with this modification, we still need to evaluate the expectations \( E(A^2) \), \( E(W^2) \), and \( E(S^2) \), and substitute the values of these quantities into (2.4), (2.6), and (2.8) of Report # 91, respectively, to obtain the premiums "charged" by the three rules under consideration.

Now, it is easily verified that the density of \( (z_{(1)}, z_{(3)}) \) is

\[
\frac{1}{\sigma^2} f \left( \frac{z_{(1)}}{\sigma}, \frac{z_{(3)}}{\sigma} \right), \text{ where the density } f \text{ is given by (3.1) of Report # 91.}
\]

Thus, using the symmetry of the null case, we have for the A-Rule that

\[
\text{Premium} = \frac{3}{\sigma^2} E(A^2)
\]
\[
\frac{3}{2} \int_{0}^{\infty} \int_{-2z(3)}^{\infty} \frac{z(1)^2}{2} f \left( \frac{z(1)}{\sigma}, \frac{z(3)}{\sigma} \right) \, dz(1) \, dz(3)
\]

Applying the transformation

\[
z(1) = \sigma x(1)
\]

\[
z(3) = \sigma x(3)
\]

we have that

\[
\text{Premium} = \frac{3}{2} \int_{0}^{\infty} \int_{-2x(3)}^{\infty} x(1)^2 f \left( x(1), x(3) \right) \, dx(1) \, dx(3)
\]

which is the same form as (3.2) of Report # 91, except for the region of integration. The expressions for the premiums under the W-Rule and S-Rule are of the same integral form as expressions (3.3) and (3.4) of Report # 91, respectively, but with the limits of integration determined by region \( T_1 \), (2.4), instead of the region \( G_1 \), defined in section 3.1 of Report # 91.

The value of \( R \), and hence of \( C \), corresponding to a given premium and rule may be determined by iteration. Table 4.1 lists the values of \( C \) corresponding to premiums of 5%, 4%, 3%, 2%, 1%, and .5%. 

2.2. Protections (n=3)---Estimation of $\mu$ When $\sigma^2$ Unknown, Biased Mean Case

The determination of protection when a sample of three observations contains two from $N(\mu, \sigma^2)$ and one from $N(\mu + a\sigma, \sigma^2)$, $\sigma^2$ unknown, is similar to the case where $\sigma^2$ is known. By using the same type of procedure as in section 2.1, we see that the protections afforded by the A-Rule, W-Rule, and S-Rule may be evaluated using expressions (3.6), (3.8), and (3.9) of Report # 91, respectively, with the region of integration determined by $T_1$, (2.4), replacing the region of integration determined by $G_1$, and the region of integration determined by $T_3$, (2.5), replacing the region of integration determined by $G_2$ (see section 3.1 of Report # 91). Numerical results are presented in section 4.

2.3. Protections (n=3)---Estimation of $\mu$ When $\sigma^2$ Unknown, Biased Variance Case

Again, the analysis of Report # 91 may be used to compute protections when a sample of three contains two observations from $N(\mu, \sigma^2)$ and one from $N(\mu, (1+b) \sigma^2)$, and $\sigma^2$ is unknown. Protections afforded by the A-Rule, W-Rule, and S-Rule may be obtained from (3.11), (3.12), and (3.13) of Report # 91, respectively, with the region $H_1$ now defined as

$$H_1 = \{ (z(1), z(3), \bar{y}) \mid (z(1), z(3)) \in T_1, -\infty < \bar{y} < \infty \}.$$  

Numerical results are presented in section 4.
3. Estimation of $\sigma^2$ When $\mu$ Unknown ($n=3$)

Let us now consider the estimation of $\sigma^2$ based on a sample of three, hopefully all from $N(\mu, \sigma^2)$, where we wish to guard against a spurious observation from $N(\mu + a\sigma, \sigma^2)$ or $N(\mu, (1+b)\sigma^2)$, and where $\mu$ is unknown.

From expressions (5.2.3), (5.3.3), and (5.4.3) of Report # 90 we see that for a sample of size three, the rejection rules for estimating $\sigma^2$ have boundaries of the form $\text{max}(z_{(1)}^2, z_{(3)}^2) = Kv^2$, which may be written as $\text{max}|z_{(1)}, |z_{(3)}| = \frac{1}{K}v$. Thus, the rejection boundary is equivalent to that of the rule for estimating $\nu$, with $K = C^2$.

Hence, each estimator may be written in the form

$$
\hat{\sigma}^2 = \begin{cases} 
Dv^2 & \text{if } (z_{(1)}, z_{(3)}) \in T_0 \\
Dg_1(z) & \text{if } (z_{(1)}, z_{(3)}) \in T_1 \\
Dg_3(z) & \text{if } (z_{(1)}, z_{(3)}) \in T_3 
\end{cases}
$$

(3.1)

Using the notation of section 5.1 of Report # 90 we have for the $A$-Rule that

$$
g_1(z) = v_{(1)}^2 \\
g_3(z) = v_{(3)}^2
$$

(3.2)

Similarly, for the $W$-Rule,

$$
g_1(z) = \text{max} (v_{(1, 1)}^2, v_{(3, 1)}^2) \\
g_3(z) = \text{max} (v_{(1, 3)}^2, v_{(2, 3)}^2)
$$

(3.3)

Since $v_{(2, 1)}^2 = v_{(3, 1)}^2$ and $v_{(1, 3)}^2 = v_{(2, 3)}^2$, we may write for the $W$-Rule,

$$
g_1(z) = v_{(3, 1)}^2 \\
g_3(z) = v_{(1, 3)}^2
$$

(3.4)
For the S-Rule,
\[ g_1(z) = \frac{1}{2} \left( v_{(1)}^2 + Kv^2 \right) \]
\[ g_3(z) = \frac{1}{2} \left( v_{(3)}^2 + Kv^2 \right) \]  
(3.5)

Using these expressions, we may calculate premiums and protections.

3.1. **Premiums (n=3)---Estimation of \( \sigma^2 \) When \( \mu \) Unknown**

In the null case when all three observations are from \( N(\mu, \sigma^2) \), we have that \( v^2 \sim \frac{\sigma^2}{2} \chi_2^2 \). Hence,

\[ E(v^2) = \sigma^2 \]
\[ V(v^2) = \sigma^4 \]

and

\[ E(v^4) = 2\sigma^4 \]

Now, by definition, we have that for a rejection rule using (unbiased) estimator \( \hat{\sigma}^2 \),

\[ \text{Premium} = \frac{V(\hat{\sigma}^2) - V(v^2)}{V(v^2)} \]

which we may express as

\[ \text{Premium} = E\left( \frac{\hat{\sigma}^4}{\sigma^4} \right) - 2 \]  
(3.6)

For a given constant \( C \) (or \( K \)), we may calculate the \( R \) of expression (2.2), and by following the procedure of section 2.1 of Report #92 we may determine the appropriate constant \( D \) by solving equation (3.7) for \( D \). This
where \( f(z(1), z(3)) \) is given by expression (3.1) of Report #91. The expressions for \( E(\hat{\sigma}_w^4) \) and \( E(\hat{\sigma}_d^4) \) may, of course, be written in integral form similar to (3.10).

The values of D, C, and K corresponding to premiums of 5%, 4%, 3%, 2%, 1%, and 0.5% are listed in section 4 for each of the three rejection rules considered.

3.2. Protections(n=3)—Estimation of \( \sigma^2 \) When \( \mu \) Unknown, Biased Mean Case

If a spurious observation from \( N(\mu + a\sigma, \sigma^2) \) is present in the sample of three, we see that \( v^2 \sim \frac{\sigma^2}{2} \chi_2^2, \frac{2a^2}{3} \), where \( \chi_2^2, \frac{2a^2}{3} \) denotes a chi-square variate with 2 degrees of freedom and noncentrality parameter \( \frac{2a^2}{3} \). Hence,

\[
E(v^2) = \sigma^2(1 + a^2) \frac{2}{3}
\]

\[
V(v^2) = \sigma^4(1 + 2a^2) \frac{2}{3}
\]

\[
E(v^4) = \sigma^4(2 + 4a^2) + \frac{a^4}{9}
\]

Now, for a rejection rule using estimator \( \hat{\sigma}^2 \),

\[
\text{Protection} = \frac{E(v^2 - \sigma^2)^2 - E(\hat{\sigma}^2 - \sigma^2)^2}{E(v^2 - \sigma^2)^2} \quad (3.11)
\]

Since we have that

\[
E(v^2 - \sigma^2)^2 = \sigma^4(1 + 2a^2) + \frac{a^4}{9} \quad (3.12)
\]
we see we need only calculate \( E(\hat{\sigma}^2 - \sigma^2)^2 \) in order to calculate the protection. To do this, we may use the same type procedure as in section 3.1, and write

\[
E(\hat{\sigma}^2 - \sigma^2)^2 = E(\hat{Dv}^2 - \sigma^2)^2 \\
+ \text{Prob}[(z_{(1)}, z_{(3)}) \in T_1].
\]

\[
E[(D_{g_1}(z) - \sigma^2)^2 - (Dv^2 - \sigma^2)^2 | (z_{(1)}, z_{(3)}) \in T_1]
\]

\[
+ \text{Prob}[(z_{(1)}, z_{(3)}) \in T_3].
\]

\[
E[(D_{g_3}(z) - \sigma^2)^2 - (Dv^2 - \sigma^2)^2 | (z_{(1)}, z_{(3)}) \in T_3]
\]

(3.13)

where \( T_1 \) and \( T_3 \) are defined by (2.3) and (2.4), respectively. Using the fact that the density of \((z_{(1)}, z_{(3)})\) is \( \frac{1}{\sigma^2} f(\frac{z_{(1)}}{\sigma}, \frac{z_{(3)}}{\sigma}; a) \), we may evaluate (3.13), using the density \( f(z_{(1)}, z_{(3)}; a) \) given by expression (3.5) of Report # 91, and then together with (3.12) we may substitute into (3.11) to obtain the protection.

Section 4 gives the protections corresponding to premiums of 5% and 1%, for various biases "a".
3.3. Protections (n=3)---Estimation of $\sigma^2$ When $\mu$ Unknown, Biased Variance Case

If, in the sample of three, there is one spurious observation from $N(\mu, (1+b) \sigma^2)$, then $v^2 \sim \sigma^2 \left[ \chi_1^2 + (1+\frac{2b}{3}) \chi_1^1 \right]$. Hence,

\[
E(v^2) = \sigma^2(1+\frac{b}{3})
\]
\[
V(v^2) = \sigma^4(1+\frac{2b}{3} + \frac{4b^2}{9})
\]
\[
E(v^4) = \sigma^4(2+\frac{4b}{3} + \frac{b^2}{3})
\]

We see from the formula for protection (3.11) that we must evaluate $E(v^2 - \sigma^2)^2$ and $E(\hat{\sigma}^2 - \sigma^2)^2$. To do this we note that

\[
E(v^2 - \sigma^2)^2 = \sigma^4(1+\frac{2b}{3} + \frac{b^2}{3})
\]

(3.14)

and write $E(\hat{\sigma}^2 - \sigma^2)^2$ in the same form as (3.13), which we may evaluate using the density of $(z(1), z(3))$ which now reflects the biased variance of the spurious observation. This density is $\frac{1}{\sigma^2} g(\frac{z(1)}{\sigma}, \frac{z(3)}{\sigma}; b)$, where $g(z(1), z(3); b)$ may be obtained from the density $h(z(1), z(3), \bar{y}; b)$ given in (3.10) of Report # 91 by integrating over the range of $\bar{y}$. That is, we have that

\[
g(z(1), z(3); b) = \int_{-\infty}^{\infty} h(z(1), z(3), \bar{y}; b) \, d\bar{y}
\]

Thus, we may obtain the protection by substituting (3.13) and (3.14) into (3.11). Numerical Results are listed in the next section.

4. Tables and Graphs of Results (n=3)

The following self-explanatory tables and graphs present the premiums and protections calculated by numerical integration.
<table>
<thead>
<tr>
<th>R</th>
<th>0.98926</th>
<th>1.15470</th>
<th>1.15470</th>
<th>0.0050</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>0.93575</td>
<td>1.15468</td>
<td>1.15468</td>
<td>0.0100</td>
</tr>
<tr>
<td>R</td>
<td>0.86212</td>
<td>1.15464</td>
<td>1.15464</td>
<td>0.0200</td>
</tr>
<tr>
<td>R</td>
<td>0.80625</td>
<td>1.15456</td>
<td>1.15456</td>
<td>0.0300</td>
</tr>
<tr>
<td>R</td>
<td>0.75935</td>
<td>1.15445</td>
<td>1.15445</td>
<td>0.0400</td>
</tr>
<tr>
<td>R</td>
<td>0.71813</td>
<td>1.15431</td>
<td>1.15431</td>
<td>0.0500</td>
</tr>
</tbody>
</table>

The regression constants, c, for the estimation of H_i

**Table 4.1**
<table>
<thead>
<tr>
<th>T</th>
<th>C</th>
<th>D</th>
<th>K</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00762</td>
<td>1.3333</td>
<td>1.0250</td>
<td>1.3333</td>
<td>1.0850</td>
</tr>
<tr>
<td>1.0032</td>
<td>1.3333</td>
<td>1.0250</td>
<td>1.3333</td>
<td>1.0850</td>
</tr>
<tr>
<td>1.01549</td>
<td>1.3333</td>
<td>1.0250</td>
<td>1.3333</td>
<td>1.0850</td>
</tr>
<tr>
<td>1.0319</td>
<td>1.3333</td>
<td>1.0250</td>
<td>1.3333</td>
<td>1.0850</td>
</tr>
<tr>
<td>1.0500</td>
<td>1.3333</td>
<td>1.0250</td>
<td>1.3333</td>
<td>1.0850</td>
</tr>
</tbody>
</table>

**S-Rule**

**A-Rule**

**Premium constants for the estimation of Q.**

**Table 4.2**
<table>
<thead>
<tr>
<th>Protection for a 1% premium</th>
<th>Protection for a 5% premium</th>
</tr>
</thead>
<tbody>
<tr>
<td>w(β + 0.0) is present in a sample of size three.</td>
<td></td>
</tr>
<tr>
<td>Corresponding to premiums of 5% and 1% when a spurious observation is present.</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3
<table>
<thead>
<tr>
<th>1%</th>
<th>S-Rule</th>
<th>W-Rule</th>
<th>A-Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
</tbody>
</table>

**N(π, (1+β) 0.5) is present in a sample of size three.**

Estimation of π.

Proportions corresponding to premiums of 5% and 1% when a spurious observation from Table 4.4.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.064</td>
<td>0.044</td>
<td>0.044</td>
<td>0.06</td>
<td>0.036</td>
<td>0.036</td>
<td>0.036</td>
<td>0.036</td>
<td>0.036</td>
</tr>
<tr>
<td>0.053</td>
<td>0.044</td>
<td>0.044</td>
<td>0.06</td>
<td>0.027</td>
<td>0.027</td>
<td>0.027</td>
<td>0.027</td>
<td>0.027</td>
</tr>
<tr>
<td>0.044</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.019</td>
<td>0.019</td>
<td>0.019</td>
<td>0.019</td>
<td>0.019</td>
</tr>
<tr>
<td>0.036</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td>0.027</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
</tr>
<tr>
<td>0.019</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td>0.010</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.008</td>
<td>0.008</td>
<td>0.008</td>
<td>0.008</td>
<td>0.008</td>
</tr>
<tr>
<td>0.007</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Table 4.5

Protection for a 1% Premium

Protection for a 5% Premium

N(μ + α, σ) is present in a sample of size three. (Estimation of σ.)

Protection corresponding to premium of 5% when a spurious observation from
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.46</td>
<td>600**-</td>
<td>600**-</td>
<td>2.28</td>
<td>0.20</td>
<td>0.20</td>
<td>1.4</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.41</td>
<td>900**-</td>
<td>900**-</td>
<td>2.22</td>
<td>0.09</td>
<td>0.09</td>
<td>1.2</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>0.36</td>
<td>900**-</td>
<td>900**-</td>
<td>2.21</td>
<td>0.03</td>
<td>0.03</td>
<td>1.0</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>0.30</td>
<td>400**-</td>
<td>400**-</td>
<td>1.94</td>
<td>0.03</td>
<td>0.03</td>
<td>1.0</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>0.24</td>
<td>400**-</td>
<td>400**-</td>
<td>1.69</td>
<td>0.04</td>
<td>0.04</td>
<td>8.0</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>0.16</td>
<td>800**-</td>
<td>800**-</td>
<td>1.33</td>
<td>0.03</td>
<td>0.03</td>
<td>1.2</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>0.08</td>
<td>400**-</td>
<td>400**-</td>
<td>0.81</td>
<td>0.03</td>
<td>0.03</td>
<td>0.4</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>0.03</td>
<td>300**-</td>
<td>300**-</td>
<td>0.44</td>
<td>0.05</td>
<td>0.05</td>
<td>2.9</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.08</td>
<td>200**-</td>
<td>200**-</td>
<td>0.44</td>
<td>0.05</td>
<td>0.05</td>
<td>1.2</td>
<td>0.05</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Note: (p) is present in a sample of size three. (Estimation of $O_2$)
5. Discussion of Results

As can be seen from the preceding tables and graphs, the protection afforded by the W-Rule and the A-Rule is extremely small for even moderate biases. Thus, we must echo Anscombe (1960) and say that these two rejection rules are "utterly useless and absurd" for a sample size of three. We must, however, add a note of optimism in regard to the S-Rule, which performs extremely well with respect to the A and W-Rules, both for estimating \( \mu \) and for estimating \( \sigma^2 \). Of course, even the protections afforded by the S-Rule are relatively low when compared with those tabulated in Reports # 91 and 92. This is to be expected, however, since both parameters \( \mu \) and \( \sigma^2 \) are now assumed unknown.

As mentioned previously, we have omitted sample sizes \( n>3 \) because analytic calculations are abhorrent and time consuming, while consideration of Monte Carlo procedures get bogged down because of the form of the rejection boundaries. For example, to extend the Monte Carlo procedure derived in Report # 91 for estimating \( \mu \), we find that what is involved is random sampling of \( (z_1, \ldots, z_{n-1}) \) subject to conditions of the form

\[
\frac{z_{n-1}^2}{n-1} \geq \frac{2c^2}{n-1} (z_1^2 + \ldots + z_{n-1}^2 + z_1z_2 + \ldots + z_{n-1}z_{n-1}).
\]

The products \( z_iz_j \) add considerable complications to the problem, for essentially we find we need to sample in a cone in \((n-1)\)-space, which is no easy task.
We must admit that we found it an unsurmountable task, but perhaps someone else will not.

Certainly a welcome next step would be the investigation of the rejection rules for larger sample sizes, and we would suggest that work on the S-Rule be the primary task.

(1) Actually, the skeleton of a Monte Carlo procedure existed, but it was abandoned when it became apparent that it was completely inefficient insofar as requirements for computer time were concerned.
Bibliography


