ON SOME JOINT PROBABILITIES USEFUL IN MIXED ACCEPTANCE SAMPLING

by

E. G. Schilling

and

H. F. Dodge

TECHNICAL REPORT NO. N-26

December, 1966

Research supported by the Army, Navy, Air Force and NASA under Office of Naval Research Contract No. Nonr 1014r(18) Task No.
NR 0442-241 with Rutgers - The State University. Reproduction in whole or in part is permitted for any purpose of the United States Government. Distribution of this document is unlimited.
ON SOME JOINT PROBABILITIES USEFUL IN MIXED ACCEPTANCE SAMPLING

by

E. G. Schilling

and

H. F. Dodge

1

1. INTRODUCTION

1.1. Scope

Mixed acceptance sampling plans, i.e. sampling schemes combining constituent variables and attributes results in a double sampling procedure, provide considerable advantage over more conventional plans in certain applications. This has been pointed out by Dodge [1] and Bowker and Goode [2], among others. Such plans have not, however, found application in proportion to their merit, partly because methods of computing the probabilities necessary for the evaluation of the operating characteristics of such plans for all attributes acceptance numbers have not been developed.

Of particular importance is the case involving a single specification limit, known standard deviation, assuming an underlying normal distribution of product, when the results of the variables and attributes constituents of the plan are not kept independent. At present the only detailed investigation in this area appears to be the pioneering work of Gregory and Resnikoff [3]. Savage [4] points out that Gregory and Resnikoff were able to find feasible computing techniques for the necessary probabilities for this case only when \( c = 0 \), a considerable restriction. This report includes an alternative derivation of the necessary probabilities when \( c = 0 \) and then presents a derivation of the probabilities for \( c > 0 \), together with

\[1\] The material in this report is based in part on work done in preparation of a doctoral dissertation at Rutgers - The State University.

Distribution of this document is unlimited.
a scheme for readily computing these probabilities.

1.2. The Mixed Acceptance Sampling Plan

Mixed acceptance sampling plans are of two types, "independent" and "dependent"
Independent plans are those in which the results of the variables and attributes
components are kept independent. Such plans have been discussed by Bowker and
Goode [1], Gregory and Resnikoff [3], and Schilling [5]. The values of probability
of acceptance, \( P_a \), of such plans are readily determined through the use of existing
tables and procedures. Probably of more importance, both theoretically and practi-
cally, are dependent plans, those in which the variables and attributes results of
the procedure are combined in such a way as to make them dependent. Attention
will be directed here to dependent plans involving a single specification limit,
known standard deviation, and the assumption of an underlying normal distribution
of product. For reasons of symmetry, only an upper specification limit will be
considered.

Let:

\[ n_1 = \text{the first sample size.} \]
\[ A = \text{variables acceptance limit such that if the sample mean,} \]
\[ \bar{x} \leq A \text{ the lot is accepted}^2. \]
\[ n_2 = \text{second sample size.} \]
\[ c = \text{attributes acceptance number.} \]

There are a number of procedures for carrying out a dependent mixed plan. A
typical one is that originally proposed by Bowker and Goode [1] which can be
summarized as follows:

\[^2\text{Of the several methods of specifying the variables constituent of known-standard-}
\text{deviation (c) variables plans, designation by sample size (n_1) and acceptance}
\text{limit on the sample average (A), simplifies the notation somewhat. Note that}
\text{A = U-kv for upper specification limit U and customary variables acceptance}
\text{factor k; also that A = (K_p - k) in the notation of Gregory and Resnikoff [3].} \]
1. Determine parameters of plan: $n_1$, $A$, $n_2$, $c$.

2. Take a random sample of size $n_1$ from the lot.

3. Determine the sample average $(\bar{x})$ of this first sample and,
   a. If $\bar{x} \leq A$ accept the lot and cease inspection.
   b. If $\bar{x} > A$ continue inspection as in 4.

4. Take a second sample of size $n_2$ from the lot and determine
   the number of defectives $m$ contained in the total combined
   sample of $n_1 + n_2$ and,
   a. If $m \leq c$ accept the lot and cease inspection.
   b. If $m > c$ reject the lot and cease inspection.

Clearly, since the combined sample is subjected to the attributes test, the variables and attributes aspects of the plan are not independent. The probability of acceptance under this procedure becomes:

$$P_a = P(\bar{x} \leq A) + \sum_{i=0}^{c} \sum_{j=0}^{c-1} P_i(j; n_2) P(j; n_2)$$

where

- $P(\bar{x} \leq A)$ = probability that the sample mean does not exceed the acceptance limit $A$
- $P(j; n_2)$ = probability of $j$ defectives in a sample of size $n_2$
- $P_i(j; n_2)$ = joint probability of $\bar{x} > A$ and $i$ defectives in a sample of size $n_1$.

Examination of the formula for the probability of acceptance, $P_a$, will indicate that all probabilities contained therein have been tabulated or can be calculated by elementary methods, with the exception of $P_i(j; n_2)$. The same would be true for modified forms of the above procedure, say one having two acceptance numbers, $c_1$ and $c_2$, for the 2 samples. Thus, this joint probability is a very important factor in determining the properties of mixed sampling plans generally. It provides the key, and can be used directly, for determining the operating characteristic curves of mixed plans of the general type to be considered here -- for single specification limit, and known standard deviation.
2. METHOD OF EVALUATING $P_n(i, \bar{x} > A)$

2.1. Approach

This section will present a method of evaluating $P_n(i, \bar{x} > A)$, the probability that a sample of size $n$ will have exactly $i$ defectives and a sample mean ($\bar{x}$) greater than some limit $A$. First to be considered will be the case in which $i = 0$. This will be followed by the case in which $i > 0$. The method of evaluation used was suggested from the results of Gregory and Resnikoff [3] — the solution for this probability when $i = 0$ is the same as that of their exposition $^3$.

In determining these probabilities, we will let $p_n(i, \bar{x})$ be the joint probability density of $i$ and $\bar{x}$ in a sample of $n$. Then, given a particular value $i = i_o$, integration of this density over an appropriate region for $\bar{x}$ will give $P_n(i, \bar{x} > A)$ if the function is integrable and the integral exists.

2.2. Evaluation of $P_n(0, \bar{x} > A)$

Without loss of generality, since the $z$ transformation will always be appropriate, consider random samples of $n$ from the standard normal distribution, $\mathcal{N}(0, 1)$. Choose $z_U$ so that:

$$\frac{1}{\sqrt{2\pi}} \int_{z_U}^{\infty} e^{-\frac{t^2}{2}} dt = p .$$

Thus, $z_U$ is an upper limit such that only 100$p$ percent of the population exceeds the said limit. The probability density for samples with mean $\bar{z}$ and no observations greater than $z_U$ is desired, that is, samples with mean $\bar{z}$ such that if the observations are ordered:

$^3$The Gregory and Resnikoff evaluation of a similar probability for $i > 0$ appears to be subject to question due to the particular restrictions placed on the region of integration.
Now the probability that none of the observations in a random sample from 
\( \Phi(0,1) \) is a distance further from the sample mean than the limit \( z_u \) can be found from the distribution of the deviation of \( z(n) \) from the sample mean, 
\( u_n = (z(n)^2) \), requiring that 
\( (z(n)^2) < (z_u^2) \). This distribution has been tabulated both by 
Nair [6] and Grubbs [7] as \( F_n(u) \) for \( \Phi(0,1) \) where 
\( F_n(u) = P(u_n \leq u) \) for a sample of size \( n \).

It can be shown that the extreme deviate from the sample mean, \( u_n \), and the
sample mean \( \overline{x} \) are independent. Consider the following theorem:

**Theorem 1.** Let \( X_1, X_2, \ldots, X_n \) denote a random sample from a
distribution having a p.d.f. \( f(x; \theta) \), \( \gamma < \theta < \delta \). Let
\( Y_1 = u_1(X_1, X_2, \ldots, X_n) \) be a sufficient statistic for
\( \theta \) and let the p.d.f. \( g_1(y_1; \theta), \gamma < \theta < \delta \), of \( Y_1 \) be complete.
Let \( Z = u_1(X_1, X_2, \ldots, X_n) \) be any other statistic (not a
function of \( Y_1 \) alone). If the distribution of \( Z \) does not
depend upon \( \theta \), then \( Z \) is stochastically independent of the
sufficient statistic \( Y_1 \).

Here \( Y_1 = \overline{x} \) while \( Z = u_n \). Also \( \overline{x} \) is, for every \( \sigma^2 > 0 \), a sufficient statistic
for \( \mu, -\infty < \mu < \infty \), and the density function of the sufficient statistic \( \overline{x} \) is
complete. The cumulative distribution function of \( u_n \) is

\[
F_n(u_n \leq u) = \frac{n}{\sqrt{n-1}} \int_0^u \frac{e^{-1/2 \frac{x^2}{n-1}}}{\sqrt{2\pi}} \, dx
\]

with

\[
F_2(u_n \leq u) = 2 \sqrt{\pi} \int_0^u \frac{1}{\sqrt{2\pi}} \, e^{-\frac{1}{2}x^2} \, dx
\]

---

which does not contain u and hence \( u_n = (z_{(n)} - \bar{z}) \) is stochastically independent of the sufficient statistic \( \bar{z} \).

So when \( i = 0 \), the joint density for \( z \) in the range \(-\infty < \bar{z} \leq z_U\) is

\[
p_n(0, \bar{z}) = p(\bar{z}) \cdot F_n(z_{(n)} - \bar{z}) = p(\bar{z}) \cdot F_n(z_U - \bar{z}).
\]

But \( p_n(0, \bar{z}) \) is Riemann integrable in the region \(-\infty < \bar{z} \leq z_U\) since it is the product of two continuous functions in the interval. For some value \( z_A \) such that

\(-\infty < z_A < z_U, \)

\[
P_n(0, \bar{z} > z_A) = \int_{z_A}^{z_U} \phi(0, \bar{z}) \cdot F_n(z_U - \bar{z}) \, d\bar{z} = \int_{z_A}^{z_U} \frac{\bar{z}^2}{2\pi} e^{-\frac{\bar{z}^2}{2}} \cdot F_n(z_U - \bar{z}) \, d\bar{z}
\]

and the right hand side can be evaluated by numerical integration procedures.

2.3. Evaluation of \( P_n(i, \bar{z} > z_A), i > 0 \)

Again without loss of generality, consider random samples from the standard normal distribution \( \phi(0,1) \). The probability density for samples with mean \( \bar{z} \) and \( i \) observations greater than \( z_U \) is desired, that is, samples with mean \( \bar{z} \) such that if the observations are ordered:

\(-\infty < z(1) \leq z(2) \leq \ldots \leq z(n-1) \leq z_U < z(n-i+1) \leq \ldots \leq z(n) < \infty .\)

Suppose the sample to be selected sequentially. Further, suppose that the sequence of observations giving rise to such a sample was in the order indicated, that is, the first \((n-i)\) observations less than or equal to \( z_U \) and the last \( i \) observations...
greater than $z_U$. Then the sample could be divided into two subsamples, the first 
$(n-1)$ values $\leq z_U$ and the second $(i)$ values $> z_U$. If these subsamples are regarded 
separately, the means of the two samples would be related to the overall mean of 
the grand sample as:

$$(n-1)\overline{x}_1 + (i)\overline{x}_2 = (n)\overline{x}$$

$$\overline{x}_1 = \frac{n\overline{x} - i\overline{x}_2}{n-1}$$

while clearly

$$-\infty < \overline{x}_1 < z_U$$

$$z_U < \overline{x}_2 < \infty.$$  

Now to have a grand mean $\overline{x}$ and $i$ defectives, there must be associated with 
$\overline{x}_1$, no sample value $> z_U$ and with $\overline{x}_2$ no sample value $< z_U$. But since the grand 
sample has been divided into two subsamples, each of them can be treated in 
exactly the manner of the above section for $p_n(0, \overline{x})$. So for $\overline{x}_1$:

$$p_{n-1}(0, \overline{x}_1) = p(\overline{x}_1)F_{n-1}(z_U - \overline{x}_1).$$

And treating the lower limit on the second sample as if it were an upper limit:

$$p_1(i, \overline{x}_2) = p(\overline{x}_2)F_1(z_U - \overline{x}_2)$$

since $F_n(u) = P(u_n \leq u) = P(\overline{-u} \leq \overline{u}).$

Note that in the above section it was shown that $(\overline{z})$ is independent of $\overline{x}$ 
for the sample. Since each of the subsamples is being treated as an entity in 
itself, the magnitude of $\overline{x}_1$ will not affect $F_n(u)$. This, of course, must be the 
case for the statistic $\left(\frac{\overline{x}(n) - \overline{x}}{\sigma}\right)$ to be useful in testing outliers, its principal 
application.
For a grand mean \( \bar{x} \) in the range \(-\infty < \bar{x} < \infty \) and \( \bar{x}_1 \) and \( \bar{x}_2 \) related as shown above, the joint density is such that

\[
\rho_n(i, \bar{x}) = \binom{n}{i} p_{n-1}(0, \bar{x}_1)p_1(i, \bar{x}_2)
\]

\[
= \binom{n}{i}p(\bar{x}_1)p_{n-1}(s_2-\bar{x}_1)p(\bar{x}_2)p_1(\bar{x}_2-s_2)
\]

since there are \( \binom{n}{i} \) possible combinations of the observations to give \( n-i < U \) and \( i > U \).

Now \( p_n(i, \bar{x}) \) certainly does not define a continuous surface. However, for a fixed value of \( i = i_0 \), it becomes a product of continuous functions and is therefore Riemann integrable over an appropriate finite range. Furthermore, it is dominated by an appropriate multiple of the normal density and therefore the improper integral exists. Again for any specified value \( z_A < s_A < \infty \), \( p_n(i, \bar{x}) \) must be integrated over a suitable region to obtain the desired probability, \( p_n(i, \bar{x} > s_A) \). Given \( \bar{x} \), if \( \bar{x}_2 \) is allowed to vary so that \( s_2 < \bar{x}_2 < \infty \), then \( \bar{x}_1 \) will be determined by the relationship:

\[
\frac{n\bar{x} - i\bar{x}_2}{n-1} < \bar{x}_1 < s_0.
\]

For \( p_n(i, \bar{x} > s_A) \), however, \( \bar{x} \) is never \( < s_A \), \( s_A \) being the greatest lower bound for the interval, so

\[
\frac{n\bar{s}_A - i\bar{x}_2}{n-1} < \bar{x}_1 < s_A.
\]

Hence, for \( i = i_0 > 0 \) and appropriate regions of integration for \( \bar{x}_1 \) and \( \bar{x}_2 \), the regions \( \mathbb{H}_1 \) and \( \mathbb{H}_2 \) respectively.
The probability \( P_n(i, z > z_A) \) can be evaluated by use of the formulas:

\[
P_n(0, z > z_A) = \int_{z_A}^{\infty} \frac{1}{2\pi} \exp \left( -\frac{z^2}{2} \right) F_n(z, \bar{z}) \, dz,
\]

and, for \( i = 0 > 0 \)

\[
P_n(i, z > z_A) = \left( \begin{array}{c} n \\ i \end{array} \right) \int_{z_A}^{\infty} \frac{1}{2\pi} \exp \left( -\frac{z^2}{2} \right) \left( \frac{1}{\sqrt{i(n-1)}} \right) \exp \left( -\frac{i\bar{z}^2 + (n-1)z_A^2}{2\sqrt{i(n-1)}} \right) F_n(z, \bar{z}) \, dz,
\]

\( i > 0 \)

which can be integrated numerically as will be shown.

2.4. Summary

The probability \( P_n(i, z > z_A) \) can be evaluated by use of the formulas:

\[
P_n(0, z > z_A) = \int_{z_A}^{\infty} \frac{1}{2\pi} \exp \left( -\frac{z^2}{2} \right) F_n(z, \bar{z}) \, dz,
\]

and, for \( i = 0 > 0 \)

\[
P_n(i, z > z_A) = \left( \begin{array}{c} n \\ i \end{array} \right) \int_{z_A}^{\infty} \frac{1}{2\pi} \exp \left( -\frac{z^2}{2} \right) \left( \frac{1}{\sqrt{i(n-1)}} \right) \exp \left( -\frac{i\bar{z}^2 + (n-1)z_A^2}{2\sqrt{i(n-1)}} \right) F_n(z, \bar{z}) \, dz,
\]

\( i > 0 \)

\( P_n(0, z > z_A) \) and \( P_n(i, z > z_A) \), then, represent values of the probability desired

for samples taken from a standard normal universe. Using the \( z \) transformation

for a particular \( i = i_0 \).
and conversely

$$P_n(i, \bar{\mu} > A) = P_n(i, \bar{\mu} = \frac{\bar{\mu} + \mu}{\sigma} > A = \frac{A - \mu}{\sigma})$$

3. COMPUTATION OF $P_n(i, \bar{\mu} > A)$

3.1. Scheme for Computation

The computation of

$$P_n(0, \bar{z} > z_A) = \int_{z_A}^{z_U} \frac{-n\sigma^2}{2} e^{-F_n(\bar{z})} d\bar{z}$$

can be accomplished by the application of appropriate methods of numerical integration to the integrand between the limits indicated. Values of $F_n(\bar{z})$ may be obtained by interpolation from the tables of $F_n(u)$ of both Nair [6] and Grubbs [7] to arrive at the functional values.

For a particular $i = 1$, the computation of

$$P_n(i, \bar{\mu} > z_A) = \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{(i\bar{z}_2 + (n-1)\bar{z}_1)^2}{2\sigma^2}} \frac{F_1(\bar{z}_2 - \bar{z}_1) F_{n-1}(\bar{z}_1 - \bar{z}_1) d\bar{z}_1 d\bar{z}_2}{n-1}$$

is greatly simplified by rewriting the integral as
\[ P_n(1, \bar{x} > s_A) = \frac{\sigma}{n-1} \int \frac{\sqrt{2\pi}}{2\pi} \cdot F_1(s_2 - s_U) \]

\[ \left[ \begin{array}{c}
\begin{array}{c}
\frac{(n-1)\bar{x}_1}{2} \\
\frac{(n-1)\bar{x}_2}{2}
\end{array}

\begin{array}{c}
s_U \\
\frac{\sigma^2 - s_U^2}{n-1}
\end{array}

\begin{array}{c}
\frac{\sqrt{2\pi}}{2\pi} \\
\frac{\sigma^2 - s_U^2}{n-1}
\end{array}

\end{array} \right]
\]

and noting that

\[ P_{n-1}(0, \bar{x} > \frac{ns_A - 15}{n-1}) = \int \frac{\sqrt{2\pi}}{2\pi} \cdot F_{n-1}(s_U - \bar{x}) \, ds_2 \]

when calculated for a sample of size \((n-1)\). This is exactly the expression which appears in brackets in the formula for \(P(1, \bar{x} > s_A)\) above.

So,

\[ P(1, \bar{x} > s_A) = \left( \frac{n}{n-1} \right) \int \frac{\sqrt{2\pi}}{2\pi} \cdot F_1(s_2 - s_U) \left[ P_{n-1}(0, \bar{x}_1 > \frac{ns_A - 15}{n-1}) \right] \, ds_2 \]

Treating the integrand as one function for purpose of numerical integration it is possible to interpolate for values of \(P_{n-1}(0, \bar{x}_1 > \frac{ns_A - 15}{n-1})\) once tables of this probability are available. In the same way \(F_1(s_2 - s_U)\) may be obtained from existing tables, where available, to arrive at values of the integrand.

In evaluating the integral expression for \(P_n(1, \bar{x} > s_A)\) it is worthwhile to note that the values of \(i\) for which \(F_1(s_2 - s_U)\) is calculated will generally be small. But \(F_1(u) = 1\) for all values of \(u\), since, for \(n = 1\), it represents \(P(x_1 < x_1 < u)\) for
u > 0. Further, for \( i = 2 \):

\[
F_2(u) = 2 \sqrt{2} \int_0^u \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx
\]

which is simply the error function for which standard approximations are available. These relationships simplify the computation of \( P_n(1, \bar{z} > z_1) \) still further.

5. CONCLUSION

This report presents a method for evaluating joint probabilities necessary to determine the operating characteristics of dependent mixed plans for the case of single specification limit, known standard deviation, assuming an underlying normal distribution. It also provides a technique for facilitating the computation of these joint probabilities. Such values will be of substantial aid to any comprehensive studies that may be made to determine the best composite structure and the relative merits of dependent mixed plans, or to explore the possibility of developing suitable sets of mixed plans for practical application. Tables of values for a number of small sample sizes are being developed.
6. REFERENCES


This report presents a method of evaluating certain joint probabilities necessary to determine the operating characteristics of dependent mixed (variables and attributes) sampling inspection plans for the case of single specification limit, known standard deviation, and normal distribution. It also provides a technique for facilitating the computation of these joint probabilities.
### Security Classification

<table>
<thead>
<tr>
<th>KEY WORDS</th>
<th>LINK A</th>
<th>LINK B</th>
<th>LINK C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed Sampling Plans</td>
<td>ROLE</td>
<td>WT</td>
<td>ROLE</td>
</tr>
<tr>
<td>Variables-Attributes Sampling Plans</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sampling Inspection Plans</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acceptance Sampling</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial.

6. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

7. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

8a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

8b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).

9. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

10. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include address.

13. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

   (1) "Qualified requesters may obtain copies of this report from DDC."

   (2) "Foreign announcement and dissemination of this report by DDC is not authorized."

   (3) "U.S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through "

   (4) "U.S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through "

   (5) "All distribution of this report is controlled. Qualified DDC users shall request through "

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.