

AD-A009 866

A PREDICTION INTERVAL APPROACH TO DEFINING VARIABLES
SAMPLING PLANS FOR FINITE LOTS REQUIRED TO BE OF HIGH
QUALITY: SINGLE SAMPLING FOR GAUSSIAN PROCESSES

Kenneth W. Fertig, et al

Rocketdyne

Prepared for:

Office of Naval Research

1 February 1975

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER <i>AD-ACC 9 866</i>
4. TITLE (and Subtitle) Prediction Interval Approach to Defining Variables Sampling Plans for Finite Lots Required to be of High Quality: Single Sampling for Gaussian Processes		5. TYPE OF REPORT & PERIOD COVERED Technical Report 2/1/75
6. AUTHOR(s) Kenneth W. Fertig Nancy R. Mann		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Rocketdyne, Division of Rockwell International 6633 Canoga Avenue, Canoga Park, CA 91304		8. CONTRACT OR GRANT NUMBER(s) N00014-73-C-0474
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlington, Va 22217		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-321
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE 2/1/75
		13. NUMBER OF PAGES 23
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for Public Release; Distribution Unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Variables Sampling Plans Finite Lots Prediction Intervals Gaussian Processes		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) See Title page		

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. Department of Commerce
Springfield, VA 22151

PRICES SUBJECT TO CHANGE

149094

A Prediction Interval Approach to Defining Variables Sampling Plans for Finite Lots Required to be of High Quality: Single Sampling for Gaussian Processes

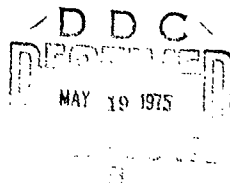
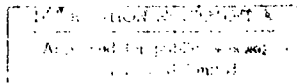
by

Kenneth W. Fertig and Nancy R. Mann

ABSTRACT

A prediction-interval approach and an assumption of a Gaussian manufacturing process are used to derive a variables sampling plan applicable to finite lots required to be of very high quality. Values tabulated for calculating acceptance regions are such that, with high probability, accepted lots will have zero defects. Further, tables are given for selecting sample size for a specified combination of lot size and acceptable quality level. Comparisons show that, for a fixed risk level, substantial savings in required sample size can be effected over those specified by comparable hypergeometric sampling plans. Lots ranging in size from 5 to 100 are considered.

Bounds for the probability of accepting a lot with a fixed number of defects and coming from a prescribed manufacturing process are derived and prove to be very tight. Not only are these bounds useful in defining the required sample size in the sampling plan context, but the method of derivation has application to other areas of sampling, e.g. from truncated or stratified distributions.



The research presented herein was performed at Rocketdyne, a Division of Rockwell International, under the sponsorship of the Office of Naval Research, Contract number N00014-73-C-0474, Task NR-042-321.

1. INTRODUCTION

In the following, consideration is given to sampling from finite lots for which the underlying process distribution is Gaussian with unknown mean and variance. It is assumed that accepted lots are required to be of very high quality, that is, with high probability, accepted lots should have zero defects. An item will be said to be defective if its measured parameter, x , is less than some specified value L (greater than a specified U).

Suppose one were to make no attempt to use the knowledge concerning the distributional form of the process that generated the lot. Acceptance or rejection of the lot would then be based solely on the attributes data, i.e. whether or not the observed x 's were less than L (or exceeded U). A required sample size would probably be determined for a specified combination of lot size, rejectable quality level and acceptable quality level by use of the hypergeometric distribution, perhaps from tables of Lieberman and Owen [8]. Use of binomial sampling plans which inherently make no provision for the finite size of the lot, can lead to sample-size requirements that are unnecessarily large, perhaps even exceeding the lot size.

Here it is assumed that we record the variables data, the sample measurements, $x_1, \dots, x_n$, of the normal variate that determines whether or not an item is defective. One would expect that the required sample size associated with a specified combination of lot size and acceptable quality level based on the additional information $x_1, \dots, x_n$ and an assumption of normality for the process distribution would be less than required

solely on the basis of whether or not the x 's are less than L (greater than U). What is desired, therefore, is a small-lot variables sampling plan (one that makes use of sample measured values and the distributional form of the process) that can be applied to lots required to be of very high quality. In the following, only a lower specification limit L will be discussed. The theory developed can be applied to upper specification limits by appropriate changes in notation and sign.

The small-lot Gaussian sampling plans of Fertig and Mann [3], the theory for which is described in [5], are based on a decision-theoretic approach and an economic loss function. Use of the loss function requires that rejected lots be screened and defective items be replaced by nondefectives. Because of this requirement, these sampling plans cannot be applied, for example, to situations of destructive testing. The sampling plans to be described in the sequel, however, do not depend on the utilization of a loss function and can be used even if the sampling situation involves destructive testing. Moreover, these plans are based on what will generally be considered a more classical approach involving none of the Bayesian methods required in the decision theoretic approach, e.g. the use of "optimal" priors on the process parameters to compute the posterior expectation of the number of defects remaining in the lot after sampling.

The new sampling plans are "Double Zero" sampling plans in the sense described by Ellis [2], that is, lots are required to have zero defects with high probability and detection of a defect in a sample requires rejection of the corresponding lot. Ellis [2] maintains that such plans, which are used currently at Pratt and Whitney Aircraft, "are especially effective in the metal working industry." The new plans differ from Ellis'

in that these are based on actual variables data while his are not and here there is no consideration given to Ellis' "gray areas" of marginal quality.

2. BASIS OF THE SAMPLING PLANS

A sampling plan can be described in terms of its acceptance criteria and the manner of specifying the sample size as a function of lot size and given requirements on quality level. Here, the approach used in defining the acceptance criteria is based on the theory of classical prediction intervals. Thus, for a sample of size n from a lot of size N , the sample variables data $x_1, \dots, x_n$, are used to predict whether or not the $N-n$ unsampled items in the lot are all nondefective. The actual prediction interval is computed on the basis of the sampling distribution of

$$R_{n,N-n} = (\bar{X} - Y_{1,N-n})/S \quad (2.1)$$

where $Y_{1,N-n}$ is the smallest observable value of the random variate X in the unsampled portion of the lot and $\bar{X}$ and S are the random variates with realizations $\bar{x}$ and s , respectively, calculated from $x_1, \dots, x_n$. If we let $\theta_{\gamma,n,N-n}$ be the 100th percentile of $R_{n,N-n}$, then $(\bar{X} - \theta_{\gamma,n,N-n} S, +\infty)$ is the 100 percent lower prediction interval for $Y_{1,N-n}$.

Methodology for relatively easy and inexpensive computer calculation of $\theta_{\gamma,n,N-n}$ was developed by Fertig and Mann [6]. The procedure depends on the numerical integration of a single integral (shown in Section 5), in contrast to the earlier approach of Hahn [7] which involved a multivariate t -distribution and numerical integration of a double integral.

Care must be exercised in the interpretation of a prediction interval. Specifically, in this case it means that if pairs of samples of sizes n and $N-n$ are created repeatedly by the manufacturing process, the 100 γ percent random prediction interval $(\bar{X} - \theta_{\gamma,n,N-n} S, +\infty)$ with $\bar{X}$ and S

computed from the first sample of size n of the pair, will contain $Y_{1,N-n}$, the smallest observable variate in the corresponding second sample of the pair, 100 γ percent of the time. If the prediction interval contains the smallest observation $y_{1,N-n}$, then of course, it contains all the observations in the second sample, the unsampled portion of the lot in the present context.

2.1 Acceptance Criteria

For the sampling-plan tables presented in this paper, we have set $\gamma = 0.90$. The value of $\theta_{Y,n,N-n}$ for each n and N considered is given in Table 1. The basis for choosing the sample size n will be discussed in section 2.2. Once the sample $x_1, \dots, x_n$ has been observed and $\bar{x}$ and s computed, the following specific acceptance criteria for the lot can be applied.

Accept the lot if

$$x_i > L \quad \text{for } i=1, \dots, n \quad (2.2)$$

and

$$\bar{X} - \theta_{Y,n,N-n} s > L. \quad (2.3)$$

These criteria follow directly from the consideration that only those lots that have no defective items in them are acceptable. The first condition (2.2) specifies that there must be no defectives in the sample and the second condition specifies that the sample must predict with a high degree of confidence (90 percent) that there are no defectives in the remaining portion of the lot.

2.2 Selection of Sample Size

In the classical operating-characteristics approach to indexing sampling plans, the acceptance criteria and sample size are chosen so that the producer's

TABLE 1. FINITE LOT VARIABLES SAMPLING PLAN ACCEPTANCE FACTORS,
 θ , n , $N-n$, FOR $\gamma = 0.90$ AND ACCEPTANCE QUALITY
 LEVEL PERCENTILES, AQLP, FOR $1 - \alpha = 0.95$

Lot Size	Sample Size	θ (0.90)	AQLP (0.95)	Lot Size	Sample Size	θ (0.90)	AQLP (0.95)	Lot Size	Sample Size	θ (0.90)	AQLP (0.95)
5	2	7.594	14.949	15	2	13.320	26.143	25	2	15.353	30.123
	3	3.066	5.460		3	5.491	9.590		3	6.256	10.905
	4	1.831	3.156		4	4.104	6.732		4	4.682	7.652
6	2	8.730	17.166	6	2	3.174	4.844	6	2	3.675	5.573
	3	3.615	6.387		3	2.935	4.374		3	3.438	5.087
	4	2.484	4.183		4	2.745	4.014		4	3.267	4.737
7	2	9.619	18.905	7	2	2.580	3.711	7	2	3.134	4.468
	3	4.009	7.058		3	2.422	3.432		3	3.025	4.250
	4	2.873	4.789		4	2.257	3.155		4	2.931	4.066
8	2	10.345	20.323	8	2	13.592	26.676	8	2	2.847	3.905
	3	4.316	7.580		3	5.594	9.768		3	2.770	3.761
	4	3.150	5.223		4	4.184	6.859		4	2.696	3.627
9	2	11.376	22.535	9	2	5.590	5.644	9	2	2.625	3.501
	3	4.565	8.006		3	3.246	4.949		3	2.553	3.378
	4	3.364	5.560		4	3.010	4.480		4	2.478	3.255
10	2	12.476	24.720	10	2	2.826	4.126	10	2	2.401	3.142
	3	4.775	8.364		3	2.669	3.834		3	2.326	3.019
	4	3.538	5.834		4	2.525	3.574		4	2.251	2.896
11	2	13.632	27.105	11	2	3.844	5.326	11	2	2.176	2.773
	3	4.954	8.672		3	3.689	4.951		3	2.101	2.650
	4	3.684	6.065		4	3.545	4.573		4	2.026	2.527
12	2	14.847	29.486	12	2	4.256	6.974	12	2	1.951	2.404
	3	5.112	9.041		3	4.111	6.596		3	1.876	2.281
	4	3.809	6.264		4	3.966	6.218		4	1.801	2.158
13	2	16.122	31.867	13	2	4.872	8.475	13	2	1.726	2.035
	3	5.385	9.590		3	4.727	8.097		3	1.651	1.912
	4	4.041	6.713		4	4.582	7.719		4	1.576	1.789
14	2	17.467	34.250	14	2	5.696	10.883	14	2	1.501	1.666
	3	5.952	10.006		3	5.551	9.505		3	1.426	1.543
	4	4.598	7.129		4	5.406	9.127		4	1.351	1.420
15	2	18.852	36.633	15	2	6.510	13.266	15	2	1.276	1.297
	3	6.207	11.175		3	6.365	11.887		3	1.201	1.174
	4	4.833	8.298		4	6.220	11.509		4	1.126	1.051
16	2	20.287	39.016	16	2	7.624	15.705	16	2	1.151	1.026
	3	6.458	12.300		3	7.479	14.326		3	1.076	0.903
	4	5.084	9.423		4	7.334	13.947		4	1.001	0.780
17	2	21.772	41.399	17	2	8.828	18.901	17	2	1.026	0.971
	3	6.703	13.684		3	8.683	17.522		3	0.951	0.848
	4	5.329	10.807		4	8.538	16.143		4	0.876	0.725
18	2	23.222	43.776	18	2	10.232	23.855	18	2	0.901	0.896
	3	7.078	15.069		3	10.087	22.476		3	0.826	0.773
	4	5.704	12.192		4	9.942	21.097		4	0.751	0.650
19	2	24.722	46.155	19	2	11.686	26.901	19	2	0.776	0.821
	3	7.478	16.354		3	11.541	25.522		3	0.701	0.698
	4	6.104	13.477		4	11.396	24.143		4	0.626	0.575
20	2	26.172	48.534	20	2	13.180	31.855	20	2	0.751	0.896
	3	7.933	17.558		3	13.035	30.476		3	0.676	0.773
	4	6.559	14.681		4	12.890	29.097		4	0.601	0.650
21	2	27.672	50.911	21	2	15.024	36.766	21	2	0.776	0.971
	3	8.188	19.062		3	14.879	35.387		3	0.701	0.848
	4	6.814	16.185		4	14.734	34.008		4	0.626	0.725
22	2	29.222	53.288	22	2	16.868	44.479	22	2	0.801	0.996
	3	8.703	20.232		3	16.723	43.100		3	0.726	0.873
	4	7.329	17.355		4	16.578	41.721		4	0.651	0.750
23	2	30.772	55.665	23	2	18.912	51.941	23	2	0.826	1.096
	3	9.188	21.676		3	18.767	50.562		3	0.751	0.973
	4	7.814	18.799		4	18.622	49.183		4	0.676	0.850
24	2	32.322	58.042	24	2	21.056	62.101	24	2	0.851	1.196
	3	9.594	23.120		3	20.911	60.722		3	0.776	1.073
	4	8.220	20.243		4	20.766	59.343		4	0.701	0.950
25	2	33.872	60.419	25	2	23.190	71.921	25	2	0.876	1.296
	3	10.006	24.194		3	23.045	70.542		3	0.801	1.173
	4	8.632	21.317		4	22.900	69.163		4	0.726	1.050

and consumer's risks satisfy certain constraints. Specifically, the producer's risk, the probability of rejecting lots at the acceptable quality level (AQL) must be less than or equal to α and the consumer's risk, the probability of accepting lots at the rejectable quality level (RQL) must be less than or equal to $1-\gamma$. For each sample size, the acceptance region given in (2.1) is designed to give the consumer lots of high quality. It is granted that (2.2) and (2.3) do not lead to the classical protection criteria at the RQL mentioned above, but they do lend themselves to a consumer's-risk interpretation. Specifically, only those lots that the user predicts not to be of rejectable quality (one or more defects) are accepted. Here, the consumer is "risking" a faulty prediction at the $\frac{100\gamma}{\gamma}$ th confidence level.

Taking this as a proper consumer's risk control, we are free to choose the sample size to control the producer's risk. Here we can closely follow the classical interpretation, in that we require that lots with zero defects will be accepted at least 100 $(1-\alpha)$ percent of the time. In order to compute this probability, however, we must specify the quality of the process the lot comes from as measured by p , the average fraction defective. Letting $\Phi(\cdot)$ be the cumulative standard normal distribution, μ and σ the mean and standard deviation of the manufacturing process, we have that $p = \Phi((1-\mu)/\sigma)$. If we let k be the number of defects in the lot, we have that the sample size, n , must be such that:

$$P \{ \text{accepting lot} \mid k=0, p \} \geq 1-\alpha \quad (2.4)$$

If the number of defects in the lot were not specified as a condition in (2.4) and only (2.3) were used as an acceptance criterion, then the

probability of accepting the lot could be computed directly as a function of p using the non central t -distribution. However, this is not the case. Furthermore, an exact expression for the left hand side of (2.4) is not available in general. Fortunately, upper and lower bounds for the left hand side of (2.4) are derivable. These bounds prove to be quite tight for all values of n and N of practical importance in the sampling-plan application. A theorem giving expressions for the bounds is proved in Section 4 for general k . If we let $\underline{P}$ be the lower bound and require that n be such that

$$\underline{P} \geq 1 - \alpha, \quad (2.5)$$

then (2.4) is satisfied. This is exactly what has been done to generate the values in Table 1. Specifically, Table 1 gives to each N and n tabulated, the acceptance parameter $\theta_{Y,n,N-n}$ for $Y=0.90$ and the acceptable quality level percentile (AQLP) defined as

$$z_p = (L - \mu) / \sigma \quad (2.6)$$

with $p = \Phi(z_p)$ for $1 - \alpha = 0.95$.

It is interesting to note that the upper bound $\bar{P}$ for the probability of acceptance is, for many combinations of N and n shown in Table 1, equal to $\underline{P}$. It is for these combinations of values of n and N that (2.4) is an exact equality. This occurs for all n less than $N/2$ up to $N=25$, for all n less than $N/3$ up to $N=50$, for all n less than $N/4$ up to $N=80$ and for all $n \leq 22$ for $N=90$ and 100 . In all cases the difference is less than 1% between $\bar{P}$ and $\underline{P}$ for $n \leq N/2$, and less than 0.08% for $n \leq N/3$. Values for $1 - \alpha = 0.90$ as well as all values of $\bar{P}$ are given in Fertig and Mann [5].

Before we proceed with a proof of the theorem referred to above, we give an example of use of the tables.

3. EXAMPLE OF USE OF THE TABLES

Consider a lot of size 50. Suppose one wishes to reject the lot even if it has only one defect. The binomial plan, which ignores the finite size of the lot, would require a sample of size 114 in order to reject a lot ninety percent of the time that is $2\% = 100 (1/50) \%$ defective. The hypergeometric plan would require a sample of size 45 to insure a 90% rejection rate. Since it is required to have such high quality in the accepted lot, one might wish to use a plan based on the assumption that the manufacturing process producing the lot is very good. If variables data are being recorded, and it can be assumed that these data are normally distributed, then the procedure presented herein will apply. Specifically, referring to Table 1, we find that a sample size of 9, for example, will guarantee a 95% acceptance rate if the manufacturing process mean is at least 5.147σ above the lower specification limit. We would accept the lot if the corresponding sample, $x_1, \dots, x_9$, had no defects and, from Table 1, its mean was at least 3.632 sample standard deviations above L. From this latter criterion we see why the process mean must be better than 5σ above L in order to guarantee a 95% acceptance rate.

At this point, it might be appropriate to discuss a criticism of the above procedure, namely expecting the normality assumption to hold at the extreme tails of the process distribution (5σ from the mean). In fact, however, normality in the tails is not as necessary as one might first think to maintain accuracy in the calculation of the probability of acceptance. The normality assumption in the tails is used only in converting from the AQLP, $(1-p)/\sigma$, to the process fraction defective, p . In fact, since p will tend to be very small, it can be off by many orders of magnitude because of a

lack of normality in the tails and still the probability of acceptance will remain relatively unaffected. This can be seen by inspecting the binomial probabilities in (4.4) and (4.5). The important parameter in the calculation of the probability of acceptance is not the fraction defective per se, but rather the acceptable quality level percentile $(L-\mu)/\sigma$ which is directly proportional to the noncentrality parameter in (4.4). It is this parameter which must be controlled by the manufacturer as well as the process average defective in order to guarantee a high probability of acceptance. It is for this reason that the tables are given in terms of AQLP rather than the AQL. Even if the quality is not as good as the AQLP, all that occurs is a higher rejection rate, essentially what is occurring with the hypergeometric plan wherein practically the entire lot is being screened. On the other hand, if the process quality is as good as the AQLP, a substantial savings in sample size has been effected.

Finally, we should remark that requiring very high manufacturing quality levels is not unreasonable in most cases in which lots would be unacceptable if they contained even a single defect.

4. THE BOUNDS ON THE PROBABILITY OF ACCEPTANCE OF A LOT

In this section we prove the following theorem which gives expressions for bounds on the probability of acceptance of a lot of size N on the basis of (2.2) and (2.3) when we specify fraction defective p in the process, number of defectives k in the lot and confidence level γ that an accepted lot has zero defectives. We note that this theorem may well be of interest outside the sampling-plan context. It applies generally to truncated-sampling applications and stratified-sampling problems.

Theorem

Let $T(t|v, \delta)$ be the cumulative non-central t -distribution with v degrees of freedom and noncentrality parameter δ . Let z_p be the $100p$ th percentile of the standard normal distribution. Let θ be the variables acceptance factor defined by (2.3) for a lot of size N and sample of size n . Consider lots that have k defects and come from a process which produces fraction p defective. The probability of accepting such lots satisfies:

$$\underline{P} \leq P \{ \text{accepting lot } |k, p\} \leq \overline{P} \quad (4.1)$$

where

$$\overline{P} = \min (1, P^*) \quad (4.2)$$

$$\underline{P} = P^* - D, \quad (4.3)$$

with

$$P^* = \frac{H(0; k, n, N)}{B(0; p, n)} \left[1 - T(\sqrt{n} \theta | n-1, -\sqrt{n} z_p) \right] \quad (4.4)$$

and

$$D = \frac{H(0; k, n, N)}{B(0; p, n)} \sum_{\ell=1}^m \frac{[1 - T(\sqrt{n-\ell} \theta_{\ell}^* | n-\ell-1, -\sqrt{n-\ell} z_p)] B(\ell; p, n)}{B(0; p, n-\ell)}, \quad (4.5)$$

and

$$H(\ell; k, n, N) = \frac{\binom{k}{\ell} \binom{N-k}{n-\ell}}{\binom{N}{n}}, \quad (4.6)$$

$$B(\ell; p, n) = \binom{n}{\ell} p^{\ell} (1-p)^{n-\ell}, \quad (4.7)$$

$$\theta_{\ell}^* = \left(n \theta \frac{(n-\ell-1)/(n-\ell)}{(n-\ell)(n-1)-n\ell\theta^2} \right)^{1/2} \quad (4.8)$$

$$m = \max \left\{ \text{all integers} < \frac{n(n-1)}{n-1+n\theta^2} \right\}. \quad (4.9)$$

The values of θ in (4.4) and of θ_{ℓ}^* in (4.5), as can be seen from (4.1), depend upon the specified value of the confidence level γ as well as n and $N-n$. Also, as will be seen subsequently, summation on ℓ in (4.5) should be taken as zero if $m=0$. For this case $\underline{p}=\overline{p}$ and, therefore, the probability of accepting the lot can be computed exactly.

To prove (4.1), we first note that since the events $\ell=0$ and $\ell \neq 0$ are disjoint all-inclusive events, then

$$\begin{aligned} P \{ \text{accepting lot} \mid k, p \} & \quad (4.10) \\ &= P \{ \text{accepting lot} \mid \ell=0, k, p \} P \{ \ell=0 \mid k, p \} \\ &+ P \{ \text{accepting lot} \mid \ell \neq 0, k, p \} P \{ \ell \neq 0 \mid k, p \} \end{aligned}$$

Since we reject the lot if $\ell \geq 1$, the second term on the right in (4.10) is zero. When k is given, the probability distribution of ℓ is hypergeometric. Therefore,

$$\begin{aligned} P \{ \ell=0 \mid k, p \} &= \frac{\binom{k}{\ell} \binom{N-n}{k-\ell}}{\binom{N}{n}} \\ &\equiv H(0; k, n, N) \end{aligned}$$

When $\ell=0$, the decision whether or not to accept the lot rests on the value of the random variable $L^* = \bar{X}-OS$. Thus, we see that

$$P \{ \text{accepting lot} | k, p \} \quad (4.11)$$

$$= P \{ L^* > L | \ell=0, k, p \} H(0; k, n, N)$$

Finding the bounds in (4.1) now reduces to finding bounds for the probability that L^* is greater than L , given that it is based on a sample with no defects coming from a lot with k defects produced by a process with p fraction defective. Since we assume the items in the lot are stochastically independent observations from the process and since we are considering a series of lots all from the same process, sampling from the lot and requiring the sample to have zero defects is the same as sampling from the process and still requiring the sample to have zero defects. That is

$$P \{ L^* > L | \ell=0, k, p \} = P \{ L^* > L | \ell=0, p \}$$

where the expression on the right can be computed assuming a sample of size n is taken from a process of quality P with the restriction that no defects are in the sample. But since the events $\ell=\ell$ for $\ell=0, \dots, n$ are disjoint all-inclusive event for this procedure, we see that

$$P \{ L^* > L | \ell=0, p \} = \frac{P \{ L^* > L | p \}}{P \{ \ell=0 | p \}} \quad (4.12)$$

$$\sum_{\ell=1}^n \frac{P \{ L^* > L | \ell=\ell, p \} P \{ \ell=\ell | p \}}{P \{ \ell=0 | p \}}$$

In order to proceed, we need the following three lemmas.

Lemma 1 Consider a sample of size n_1 from a process of quality p . Let $L_1^* = \bar{X}_1 - OS_1$ where $\bar{X}_1$ and S_1 are the observable sample mean and standard deviation $\left(\bar{X}_1 = \sum_{i=1}^{n_1} X_i / n_1, S_1^2 = \sum_{i=1}^{n_1} (X_i - \bar{X}_1)^2 / (n_1 - u) \right)$,

Then

$$P \{ L_1^* > L | p \} = 1 - T \left(\sqrt{n_1} \ 0 | n_1 - 1, \sqrt{n_1} \ z_p \right)$$

Lemma 2 $P \{L_1^* > L | p\} < P \{L_1^* > L | \ell_1 = 0, p\}$

where ℓ_1 is the number of defects in the sample of size n_1

Lemma 3 $P \{L_1^* > L | \ell_1 = \ell_1, p\} < P \{L_1^{*'} > L | x_1, \dots, x_{\ell_1} = L, \ell_1' = 0, p\}$

where $L_1^{*'}$ is formed by replacing all observations in the sample that are less than L with the value L itself. (Thus, ℓ_1' , the number of defects remaining, is zero.)

The latter two lemmas follow easily if one notes that L_1^* is an increasing function of each of the sample observations. Thus restricting these observations to be above certain limits increases the probability of observing an L_1^* greater than L .

The first lemma follows from the definition of a non-central t-variate. Specifically,

$$\begin{aligned} P \{L_1^* > L | p\} &= P \{ \bar{X}_1 - \theta s_1 > L | p \} \\ &= P \left(\frac{\sqrt{n_1} \frac{\bar{X}_1 - \mu}{\sigma} - \sqrt{n_1} \frac{L - \mu}{\sigma}}{s_1 / \sigma} > \sqrt{n_1} \theta \right) \end{aligned}$$

Since $\frac{L - \mu}{\sigma} = z_p$, $\sqrt{n} (\bar{X} - \mu) / \sigma$ is $N(0, 1)$, and s_1 / σ is the square root of an independent chi square variate divided by its degrees of freedom ($n_1 - 1$), the lemma follows directly by recalling that $t = (z + \delta) / \chi_v^2 / \sqrt{v}$ is a non-central t variate if z is standard normal and χ_v^2 is an independent chi square with degrees of freedom v .

Since the distribution of $\hat{p}$, given the sample comes from a process of quality p , is binomial with parameters p and n , we see from Lemma 1 that (4.12) implies

$$P\{L^* > L | \ell=0, p\} \leq (1 - T(\sqrt{n}\theta | n-1, -\sqrt{n}z_p)) / B(0; n, p) \quad (4.13)$$

From (4.11) and (4.13) we see we can bound the probability of accepting the lot from above by the expression for P^* given in (4.4). This establishes (4.2).

In order to establish a lower bound on the probability of accepting the lot, one could use Lemma 2 directly in (4.11). A sharper bound, however, will in general be given by considering the summation in (4.12). Let L^* be based on an observable sample with ℓ defectives, $X_1, \dots, X_\ell < L$ and $n-\ell$ nondefectives $X_{\ell+1}, \dots, X_n \geq L$. Let $\bar{X}_2 = \sum_{i \geq \ell+1} X_i / (n-\ell)$, $S_2^2 = \frac{1}{(n-\ell-1)} \sum_{i \geq \ell+1} (X_i - \bar{X}_2)^2$.

We now seek to compute $P\{L^* > L | X_1, \dots, X_\ell = L, \ell' = 0, p\}$ defined in Lemma 3. After some algebra, we find that using Lemma 3 gives

$$P\{\bar{X} - \theta S > L | \ell = \ell, p\} \leq P\left\{\left(\frac{n-\ell}{n} - \frac{\theta^2 \ell}{n-1}\right) (\bar{X}_2 - L)^2 > \frac{n-\ell-1}{n-\ell} \frac{n}{n-1} \theta^2 S_2^2 \mid \ell_2 = 0, p\right\}$$

This probability can be non-zero only if $\frac{n-\ell}{n} - \frac{\theta^2 \ell}{n-1} > 0$. Thus, ℓ must be strictly less than $n(n-1)/(n-1+\theta^2)$. This is the value given for m in (4.9). Since the number of defects, ℓ_2 , in the portion of the sample to be used to compute $\bar{X}_2$ is zero, we have $\bar{X}_2 \geq L$. Therefore, we see that for $\ell \leq m$,

$$P\{\bar{X} - \theta S > L | \ell = \ell, p\} \leq P\left\{\frac{1}{\ell} (\bar{X}_2 - L) / S_2 > n\theta \frac{(n-\ell-1)/(n-\ell)}{(n-\ell)(n-1) - n\ell\theta^2} \mid \ell_2 = 0, p\right\}^{1/2}$$

Since $\bar{X}_2$ and S_2 are based on an observable sample of size $n_2 = n-\ell$ that has zero defects, we can apply Lemma 2 and a relation similar to the first half of this proof (in which the expression for P^* is derived) to find that

$$P\{\bar{X} - \bar{O}S > L | \ell = \ell, p\} \leq \frac{P\{\bar{X}_2 - \bar{O}_\ell^* S_2 > L | p\}}{P\{\ell_2 = 0 | p\}}$$

Here, the probability in the numerator on the right is now computed with no restriction on the number of defects in the sample of size $n_2 = n - \ell$. Therefore, we can apply Lemma 1 to get for each $\ell = 1, \dots, m$

$$P\{\bar{X} - \bar{O}S > L | \ell = \ell, p\} \leq \frac{(1 - T(\sqrt{n - \ell} \bar{O}_\ell^* | n - \ell - 1, -\sqrt{n - \ell} z_p))}{B(0; p, n - \ell)} \quad (4.14)$$

and

$$P\{\bar{X} - \bar{O}S > L | \ell = \ell, p\} = 0 \quad (4.15)$$

for $\ell = m+1, \dots, n$.

From (4.15) and the definition of m , we see that the sum in (4.14) should be taken as zero if $m=0$. For $m \geq 1$, we see from (4.11), (4.12), (4.14), and (4.15), that D is as given in (4.5). This completes the proof of the theorem.

5. CALCULATION OF TABLE ENTRIES

The distribution of the random variate $R_{n,N-n} = (\bar{X} - Y_{1,N-n}) / S$, given by (2.1), which provides one of the criteria for acceptance or rejection of a lot of size N on the basis of a sample of size n , is considered by Fertig and Mann [6]. They show that

$$P(R_{n,N-n} < 0) = \frac{N-n}{\sqrt{2\pi}} \int_{-\infty}^{\infty} [1 - \Phi(z)]^{N-n-1} \exp(-z^2/2) T(\sqrt{N} \theta | n-1, -\sqrt{n} z) dz, \quad (4.1)$$

where $T(\cdot | \nu, \delta)$ indicates the cumulative noncentral t-distribution function with degrees of freedom and noncentrality parameter δ and $\Phi(\cdot)$ indicates the cumulative standard normal distribution function with mean zero and variance one.

For each combination of N and n given in Table 1 $P(R_{n,N-n} < 0) = .90$ was solved iteratively for θ by use of Muller's [10] method, a multipoint iteration scheme which determines a next guess by using inverse parabolic interpolation based on the previous two iterates and their midpoint. A first guess for the value of θ , was provided by the approximation suggested in Mann, Schafer and Singpurwalla [9] and investigated by Fertig and Mann [6]. Sixty-four point Hermite Gauss quadrature was used to evaluate the integral given by equation (4.1), the function $\Phi(\cdot)$ having been determined from $\text{erf}(x) = 2\Phi(x\sqrt{2}) - 1$, $x \geq 0$, with the error function $\text{erf}(\cdot)$ computed using a double precision IBM subroutine and the noncentral t-distribution function computed using an algorithm of B.E. Cooper [1]. Where the combination of values of n and $N-n$ agree with those of Hahn [7], computed using the multivariate t-distribution, our computed values agree with his to the number of significant figures given.

The acceptable quality level percentile AQLP defined as $z_p = -(u-L)/\sigma$, with p equal to AQL, the acceptable quality level, and related to z_p through $p = \Phi(z_p)$, was determined iteratively by requiring equality in (2.5) with $\underline{p}$ given by (4.3). Muller's [10] method, discussed above, was used in this iterative determination.

REFERENCES

- [1] Cooper, B.E., "The Integral of a Non-Central t-Distribution," Journal of the Royal Statistical Society, Ser. C. 17, No. 2 (1968), 193-4.
- [2] Ellis, E.W., "Double Zero Attribute Sampling Plans," Annual Technical Conference Transactions, American Society for Quality Control (1966), 340-7.
- [3] Fertig, K.W. and Mann, N.R., "Finite-Lots Variables Sampling Plans Applicable When the Underlying Manufacturing Process is Normal," Research Report RR72-03, Rocketdyne, Canoga Park, Calif., 1972 (submitted for publication).
- [4] Fertig, K.W. and Mann, N.R., "A Decision-Theoretic Approach to Defining Variables Sampling Plans for Finite Lots: Single Sampling for Exponential and Gaussian Processes," Journal of the American Statistical Association, 69 (September 1974), 665-71.
- [5] Fertig, K.W. and Mann, N.R., "A Prediction-Interval Approach to Defining Variables Sampling Plans for Finite Lots Required to be of High Quality: Single Sampling for Gaussian Processes," Research Report RR74-07, Rocketdyne, Canoga Park, Calif. 1974.
- [6] Fertig, K.W. and Mann, N.R., "The Determination of One-Sided Prediction Intervals for Normal and Lognormal Distributions, with Tables," in Reliability and Fault-Tree Analysis, Eds. R. Barlow and P. Chatterjee, SIAM Series in Applied Mathematics, 1975.
- [7] Hahn, G.J., "Factors for Calculating Two-Sided Prediction Intervals for Samples from a Normal Distribution," Journal of the American Statistical Association, 64 (September 1969), 878-88.
- [8] Lieberman, G.J. and Owen, D.B., Tables of the Hypergeometric Probability Distribution; Stanford University Press, 1961.
- [9] Mann, N.R., Schafer, R.E. and Singpurwalla, N.D., Methods for Statistical Analysis of Reliability and Life Data; John Wiley, 1974.
- [10] Muller, D.F., "A Method for Solving Algebraic Equations Using an Automatic Computer," Mathematical Tables and Computations, 10 (1956), 208-15.