SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NY 14853-7501

TECHNICAL REPORT NO. 863

September 1989
(Revised May 1991)

PREDICTION LIMITS ON OUTGOING QUALITY
FOR CONTINUOUS SAMPLING PLANS

by

Lisa M. McShane¹ and
Bruce W. Turnbull

This research was supported in part by Grant GM 28364 for the U.S. National Institutes of Health, and in part by a fellowship from the U.S. Army Research Office through the Mathematical Sciences Institute at Cornell University.

¹Present address: National Institutes of Health, Federal Building - Room 7C08, 7550 Wisconsin Avenue, Bethesda, MD 20892.

PROBABILITY LIMITS ON OUTGOING QUALITY
FOR CONTINUOUS SAMPLING PLANS

Lisa M. McShane¹ and Bruce W. Turnbull
School of Operations Research and Industrial Engineering
Cornell University
Ithaca, NY 14853

ABSTRACT

We investigate the performance of Dodge's (1943) continuous sampling plan 1 (CSP-1) when production run lengths are short or moderate or when the input process is not i.i.d. Bernoulli. For finite run lengths, methods are presented for computing upper and lower percentage points (probability limits) for the distributions of the outgoing quality and fraction inspected when the input process is Markov. Both rectifying and nonrectifying inspection are considered. Particular attention is paid to the special case of i.i.d. Bernoulli input. Comparisons are made with long run maximum averages under various input process assumptions, including Dodge's AOQL for the i.i.d. case, an AOQL for Markov input, and unrestricted AOQL (UAOQL) values. These limits are proposed as new performance measures, intended to supplement the tables in MIL-STD-1235B for use in selecting CSP-1 plans. We also discuss the problem of interval estimation of the incoming and outgoing quality from CSP-1 inspection data.

KEY WORDS: Continuous sampling plan (CSP-1), Average outgoing quality,
Inspection by attributes, probability limits, Markov chain,
MIL-STD-1235B, Markov dependent production process,
Estimation

¹Present address: National Institutes of Health
Federal Building - Room 7C08
7550 Wisconsin Avenue
Bethesda, MD 20892

1. INTRODUCTION

Various quality control sampling plans have been proposed for monitoring continuous production processes. In this paper, we consider the particular plan developed by Dodge (1943), termed CSP-1 (Continuous Sampling Plan-1). This plan has been incorporated into MIL-STD-1235B (1981) and has found many applications. Various modifications and elaborations of CSP-1 have been proposed (for a review, see Banzhaf and Brugger, 1970), yet the plan remains popular perhaps because of its simplicity. In this paper we present new performance measures for use in examining the robustness of the standard measures of effectiveness of the plan when production runs are of short or moderate length or when assumptions made concerning the input process do not hold.

Our own motivation for studying this class of sampling plans developed when consulted on its application to the quality monitoring of laboratory analyses of river water samples taken in the vicinity of an electricity generating power plant. A laboratory was contracted to analyze water samples from the river to monitor the river ichthyoplankton population. Each water sample was analyzed by a technician who would count and sort ichthyoplankton in the sample according to species and lifestage. This counting procedure was tedious and subject to error, so a CSP-1 inspection plan was used to tag certain of the technician's analyses for reanalysis by an expert. In this setting, an inaccurate analysis is considered a "defective". An adaptive plan such as CSP-1 was thought appropriate to provide an incentive for a low technician error rate because reanalysis costs were borne partly by the contracting laboratory. The water samples were typically analyzed in batches of a few hundred samples. In such a situation, it would be of interest to the laboratory manager to know, on a per batch basis, what inspection cost might be anticipated and how many inaccurate analyses might escape detection by the CSP-1 monitoring. The usual performance measures are for long run averages and do not take into account any measure of variability. Selection of the appropriate sampling plan depends on operational and cost constraints, but improved performance measures will better assist in this choice.

The plan CSP-1 alternates between two modes of operation, namely (1) screening inspection and (2) sampling inspection. The process starts in mode 1 in which 100% of the items are inspected until a prespecified "clearance number", i , of consecutive nondefective (or "conforming") items have been

observed. The inspection process then proceeds to mode 2 in which only a prespecified “sampling fraction”, f , of the items are inspected. When a defective is found, inspection reverts to mode 1.

Sampling at rate f in mode 2 can be accomplished in one of three different ways — systematic, random, or probability sampling (Sackrowitz, 1975). We will only consider the case of probability sampling in which each item is inspected with probability f . In most situations the differences are not important (Yang 1983, p.63); in Section 2, we indicate how results can be obtained for systematic and block random sampling.

Inspection may be either rectifying or nonrectifying. With rectifying inspection, defectives found during inspection are repaired or replaced with nondefective items. In nonrectifying inspection, defectives are simply discarded.

The outgoing quality (OQ_N) is the proportion of defective items which escape detection in a finite production run of length N . We define the Average Outgoing Quality $AOQ_N = E[OQ_N]$. The primary measure of effectiveness for continuous sampling plans has been the long run Average Outgoing Quality, AOQ_∞ or simply AOQ . For rectifying (R) inspection the “denominator” of the outgoing quality is the total number of items in the production run, whereas for nonrectifying (NR) inspection the denominator is the number of items in the production run less the number found defective during inspection. When we need to make the distinction, we will use the terms OQ_N^R , AOQ_N^R , AOQ^R and OQ_N^{NR} , AOQ_N^{NR} , AOQ^{NR} . Another important quantity of interest is the fraction of items inspected, denoted FI_N , and its expectation $AFI_N = E[FI_N]$. The long run average fraction inspected is $AFI = AFI_\infty$. Typically the AOQ and AFI have been used to aid in the selection of a plan; in the case of CSP-1, this involves choosing a clearance number i and a sampling fraction f .

Dodge (1943) and most subsequent authors assumed that inspection was perfect and that the incoming items came from a process that was “in statistical control”. This implies that the incoming items form an i.i.d. (independent, identically distributed) Bernoulli sequence such that each has some constant probability p of being defective. For this situation, the long run AOQ and AFI are given by (Dodge, 1943):

$$AOQ^R(p) = \frac{p(1-f)q^i}{f + (1-f)q^i} \quad AOQ^{NR}(p) = \frac{p(1-f)q^{i-1}}{f + (1-f)q^{i-1}} \quad (1)$$

$$AFI(p) = \frac{f}{f + (1-f)q^i} \quad (2)$$

where $q = 1 - p$. The maximum value of AOQ, or Average Outgoing Quality Limit, is defined as $AOQL^R = \max\{AOQ^R(p); 0 \leq p \leq 1\}$ for rectifying inspection, and $AOQL^{NR}$ is defined similarly. Although other measures of effectiveness (e.g., the AEDL by Hillier, 1964) and other approaches (Anscombe, 1958) have been proposed, the AOQL remains the primary criterion for evaluating continuous sampling plans.

Several authors, e.g., Hillier (1964, p.166), have noted that the AOQ and AFI, being long run averages, are not satisfactory measures of performance for short or moderate production runs of length N . First, they may differ from the exact means, AOQ_N , AFI_N ; second, they do not reflect the variability inherent in finite run lengths. For the case of rectifying inspection and an i.i.d. Bernoulli process input, Lasater (1970) used a Markov chain model of CSP-1 to derive the mean, variance, and “Chebyshev-type” probability limits for OQ_N . Blackwell (1977) and Yang (1983) derived approximation formulae for AFI_N and AOQ_N^R under i.i.d Bernoulli input.

Another way that the formulae (1), (2) may be deceiving is their reliance on the i.i.d. Bernoulli input assumption. This assumption has been criticized by Hillier (1964) and others as being too simplistic. Lasater (1970) showed that under certain extra-binomial variation models, the expressions (1), (2) for AOQ and AFI remain valid with p being replaced by its expectation. Lieberman (1953) derived an expression, assuming rectifying inspection, for an “unrestricted” AOQ (“UAOQL”) attained under a “least favorable” condition in which the production process is dependent on the inspection process. Endres (1969) derived the analogous expression for nonrectifying inspection. Expressions are given by

$$UAOQL^R = \frac{f^{-1} - 1}{i + f^{-1}} \quad (3a) \quad \text{and} \quad UAOQL^{NR} = \frac{f^{-1} - 1}{i + f^{-1} - 1} \quad (3b)$$

Several authors, including Sackrowitz (1975), criticized the UAOQL as unrealistic because it is attained only under pathological cases. Sackrowitz (1975) proposed another upper bound excluding the pathological cases in which the production process is dependent on the inspection process. We denote

this bound by $AOQL_S$. Assuming rectifying inspection, the formula is given by

$$AOQL_S = \frac{(f^{-1} - 1)[1 - (1 - f)^{\lambda^*}]}{\lambda^* + i} \quad (4)$$

where λ^* is the positive integer which maximizes the quantity $(\lambda + i)^{-1}[1 - (1 - f)^\lambda]$.

Rajarshi and Kumar (1983) and Kumar and Rajarshi (1987) proposed a model “intermediate” between the i.i.d and unrestricted models where the quality fluctuates between two levels according to a two-state time homogeneous Markov chain. Using the theory of success runs, the authors obtained expressions for the AOQ and AOQL under rectifying inspection, where the AOQL is defined to be the maximum value of the AOQ over all parameters of the Markov chain. Kumar and Rajarshi (1987), Rykova (1973), and Bidjarano and Wasmuth (1976) describe manufacturing processes in which this correlated input model is appropriate. The i.i.d. Bernoulli model is a special case.

In the next section we present a Markov chain formulation of the CSP-1 inspection process under the more general Kumar/Rajarshi incoming quality model. We use it to develop asymptotic results including expressions for the AFI, AOQ, and AOQL under both rectifying and nonrectifying inspection. In Section 3, we use our formulation to derive expressions for the mean, variance, and approximate percentage points for the distributions of the FI_N , OQ_N^R , and OQ_N^{NR} . These percentage points will be termed “probability limits”.

In Sections 4 and 5, we restrict attention to the special case of the i.i.d. Bernoulli process input. Although somewhat unrealistic, this model serves as a convenient “starting point for the analytical solution of some of the problems raised in continuous sampling” (Blackwell 1977, p.259) and is the one considered by most authors. In Section 4, we show via examples how even under the i.i.d. Bernoulli process input, the frequently used AOQL can differ significantly from $AOQL_N = \max\{AOQ_N(p), 0 \leq p \leq 1\}$ and even more significantly from a maximum upper probability limit that takes into account the variability of the OQ for finite production runs. In Section 5, we show how probability limits can be used to obtain confidence intervals and point estimates of the incoming quality p based on inspection data. In Section 6, we examine how AOQL values are affected if the incoming stream of items does not form an i.i.d. Bernoulli process, i.e., is not in “statistical control”. This is done by comparison with UAOQL values and with values obtained under the Kumar/Rajarshi model. Some concluding remarks are contained in Section 7.

2. MARKOV CHAIN FORMULATION OF CSP-1 AND ASYMPTOTIC RESULTS

In this section we describe the CSP-1 monitoring process using a discrete time Markov chain. Its formulation differs slightly from those used by Lieberman and Solomon (1955), Hillier (1964), Roberts (1965), Lasater (1970), Blackwell (1977), and others. We need a state space that is somewhat larger than those used by previous authors in order to treat all of the quantities of interest to us, and the transition probability matrix reflects the fact that we are using the more general Kumar/Rajarshi (1983, 1987) correlated input model.

We will adopt the convention that time m ($1 \leq m \leq N$) is the time at which the m 'th item in the production process is presented. Our model has $i+5$ states, where i is the clearance number. For $1 \leq j \leq i+1$, the j 'th state, labelled S_j , denotes the situation where the current item is inspected in mode 1 (screening inspection) and it completes a sequence of $j-1$ consecutive nondefectives since the last defective was found. The process is said to be in state $i+2$, labelled (I, D) , or $i+3$, labelled (I, ND) , if the current item is sampled in mode 2 and found to be, respectively, defective or nondefective. Finally, the process is in state $i+4$ (NI, ND) , or in state $i+5$ (NI, D) , if the current item is not inspected and is either nondefective or defective, respectively. Of course, the latter fact will be unknown at the time of operation of the plan.

Under the Kumar/Rajarshi (1987) correlated input model we have,

$$\Pr[(m+1)\text{st item is defective} | m\text{'th item nondefective}] = a \quad (5)$$

$$\Pr[(m+1)\text{st item is defective} | m\text{'th item defective}] = 1 - b$$

for parameters $0 \leq a, b \leq 1$. Then the transition matrix P is given in Figure 1. The entries are one-step transition probabilities for CSP-1 with clearance number i , sampling fraction f , and assuming the Kumar/Rajarshi (a, b) input model. For the usual i.i.d. Bernoulli(p) model we simply substitute $a = 1 - b = p$. The representation does assume probability sampling in mode 2. By substantially increasing the state space, either systematic or block random sampling can also be represented by a Markov chain model. However, the analytical details and computations are more cumbersome; treatment of the probability sampling situation captures all of the essential ideas. In fact, for our application to quality control of laboratory analyses, probability sampling is preferred in any case in order to avoid observer bias.

Figure 1

Transition Matrix (P) for Markov chain representation of CSP-1

State	S1	S2	S3	. . .	Si	Si+1	I,D	I,ND	NI,ND	NI,D
S1	1-b	b	0	. . .	0	0	0	0	0	0
S2	a	0	1-a	. . .	0	0	0	0	0	0
S3	a	0	0	. . .	0	0	0	0	0	0
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
Si	a	0	0	. . .	0	1-a	0	0	0	0
Si+1	0	0	0	. . .	0	0	fa	f(1-a)	(1-f)(1-a)	(1-f)a
I,D	1-b	b	0	. . .	0	0	0	0	0	0
I,ND	0	0	0	. . .	0	0	fa	f(1-a)	(1-f)(1-a)	(1-f)a
NI,ND	0	0	0	. . .	0	0	fa	f(1-a)	(1-f)(1-a)	(1-f)a
NI,D	0	0	0	. . .	0	0	f(1-b)	fb	(1-f)b	(1-f)(1-b)

Excluding the trivial cases when a or b equal 0 or 1, the Markov chain is finite, irreducible, and aperiodic. Thus if we let π_j^m denote the probability that the process is in state j ($1 \leq j \leq i+5$) at time m ($m \geq 1$), then there exists a vector of limiting probabilities $\pi = (\pi_1, \pi_2, \dots, \pi_{i+5})$ where $\pi_j = \lim_{m \rightarrow \infty} \pi_j^m$. Solving the steady state equations $\pi = \pi P$ we obtain

$$\begin{aligned}
\pi_1 &= af[1 - b(1 - a)^{i-1}]/D \\
\pi_j &= (1 - a)^{j-2}abf/D \quad 2 \leq j \leq i+1 \\
\pi_{i+2} &= (1 - a)^{i-1}abf/D = \pi_{i+1} \\
\pi_{i+3} &= bf(1 - a)^{i-1}[b + f(1 - a - b)]/D \\
\pi_{i+4} &= b(1 - f)(1 - a)^{i-1}[b + f(1 - a - b)]/D \\
\pi_{i+5} &= ab(1 - f)(1 - a)^{i-1}/D
\end{aligned} \tag{6}$$

where $D = (a + b)[f + (1 - f)b(1 - a)^{i-1}]$.

We immediately have the following limiting formulae for rectifying and nonrectifying inspection.

$$AOQ^R(a, b) = \pi_{i+5} = \frac{a}{a + b} \{1 - [1 + b(1 - a)^{i-1}(\frac{1}{f} - 1)]^{-1}\}, \tag{7a}$$

$$AOQL^R = \{[i(\frac{1}{f} - 1)^{-1}(1 - \frac{1}{f})^{1-i}]^{1/2} + 1\}^{-2}. \tag{7b}$$

Here $AOQL^R$ is defined as $\max\{AOQ^R(a, b); 0 \leq a, b \leq 1\}$, and the values of a and b which give the $AOQL^R$ are

$$a^* = \frac{1}{i} \quad b^* = [i(\frac{1}{f} - 1)(1 - \frac{1}{i})^{i-1}]^{-1/2}. \quad (7c)$$

For nonrectifying inspection,

$$AOQ^{NR}(a, b) = \frac{\pi_{i+5}}{1 - (\pi_1 + \pi_{i+2})} = \frac{a(1-f)(1-a)^{i-1}}{f + (a+b)(1-f)(1-a)^{i-1}} \quad (8a)$$

and its maximum ($AOQL^{NR}$) is given by

$$AOQL^{NR} = \{i(\frac{1}{f} - 1)^{-1}(1 - \frac{1}{i})^{1-i} + 1\}^{-1}. \quad (8b)$$

The $AOQL^{NR}$ value is only attained in the limiting case:

$$a^* = \frac{1}{i} \quad b^* = 0. \quad (8c)$$

We can also obtain an expression for the long run average fraction inspected given by

$$AFI(a, b) = 1 - \pi_{i+4} - \pi_{i+5} = \frac{(a+b)f + bf(1-f)(1-a)^{i-1}(a+b-1)}{(a+b)[f + (1-f)b(1-a)^{i-1}]}. \quad (9)$$

The formulae for $AOQ^R(a, b)$ and $AFI(a, b)$ agree with those given by Rajarshi and Kumar (1983). They did not consider $AOQ^{NR}(a, b)$. For the i.i.d. Bernoulli(p) model, we substitute $p = a = 1 - b$. In that special case, the steady state probabilities (6) become those found by Roberts (1965), and the expressions in (7a), (8a), and (9) reduce to the familiar results (1) and (2).

We shall also be interested in $\hat{p}_N$, the proportion of defective items found among those inspected during a production run of length N . The quantity $\hat{p}_N$ is suggested in MIL-STD-1235B, to be used as an estimate of the incoming quality. As $N \rightarrow \infty$, $\hat{p}_N$ converges almost surely to

$$\hat{p}_\infty = \frac{\pi_1 + \pi_{i+2}}{1 - (\pi_{i+4} + \pi_{i+5})}, \quad (10)$$

where the $\{\pi_j\}$ are given by (6). Note that the expression (10) is not equal to $p = a/(a+b)$, the long run proportion of incoming defectives, unless $a = 1 - b = p$, the i.i.d. Bernoulli(p) situation (in which case $\hat{p}_N$ is the maximum likelihood estimate of p). In general, therefore, $\hat{p}_N$ is an inconsistent estimator.

3. PROBABILITY LIMITS AND VARIABILITY FOR FINITE PRODUCTION RUNS

As mentioned in Section 1, the AOQ and AFI, being long run averages, may not be satisfactory measures of performance of CSP-1 for finite production runs. In the Appendix A it is shown how to

derive exact expressions for the mean and variance of FI_N and OQ_N^R .

For runs of length N , the means of FI_N and OQ_N^R are given exactly by

$$AFI_N = 1 - \sum_{m=1}^N [\pi_{i+4}^m + \pi_{i+5}^m]/N \quad (11a)$$

$$AOQ_N^R = \sum_{m=1}^N \pi_{i+5}^m/N \quad (11b)$$

where the $\{\pi_j^m\}$ are defined in Section 2. To start inspection in mode 1 (100% inspection), the process starts in state 1 immediately prior to initiation (time 0). Thus

$$\pi_j^m = P_{1j}^m \quad (1 \leq k \leq i+5, m \geq 1) \quad (12)$$

where the notation P_{ij}^m represents the (i, j) th element of the m -step transition matrix given by $\underline{P}^m$, the m 'th power of $\underline{P}$. The exact expression for $\text{Var}[OQ_N^R]$ is given by

$$N^2 \text{Var}[OQ_N^R] = \sum_{m=1}^N \pi_{i+5}^m - \left(\sum_{m=1}^N \pi_{i+5}^m \right)^2 + 2 \sum_{n=1}^{N-1} \sum_{m=n+1}^N P_{i+5, i+5}^{m-n} \pi_{i+5}^n, \quad (13)$$

and the exact expression for $\text{Var}[FI_N]$ is given by Equation (A1) in Appendix A.

For the special case of i.i.d. Bernoulli(p) input, the exact mean and variance of FI_N can be more easily computed in terms of those of OQ_N^R via the relations

$$AFI_N = 1 - AOQ_N^R/p, \quad (14a)$$

$$\text{Var}[FI_N] = \text{Var}[OQ_N^R]/p^2 - (1 - p)AOQ_N^R/(Np^2). \quad (14b)$$

In the context of the i.i.d. Bernoulli(p) input model, Lasater (1970) obtained the expressions (11b) and (13). Yang (1983) used the relation (14a).

Exact expressions for the mean and variance of OQ_N^{NR} for nonrectifying inspection and for the MIL-STD-1235B estimator $\hat{p}_N$ are more difficult because they are defined as ratios where the denominator is no longer N , but random. However, we can use the Taylor series expansion technique (the “delta method”, Rao 1965, p. 321). In Appendix B, we derive approximate means and variances of OQ_N^{NR} and $\hat{p}_N$. A simulation study for the i.i.d. Bernoulli(p) case found these approximations to be excellent ones. Details are given in McShane (1989). The formulae for the exact means and variances of OQ_N^R and FI_N and for the approximate means and variances of OQ_N^{NR} and $\hat{p}_N$ are complicated but can easily be evaluated using a computer. The formulae require finding elements of successive powers of the transition matrix $\underline{P}$. However, the amount of computation can be reduced by exploiting the

sparseness of the transition matrix. A computer program is listed in McShane (1989) and a copy of it is available from the first author.

The derivations of means and variances of OQ_N^R , OQ_N^{NR} , FI_N , and $\hat{p}_N$, as given in Appendices A and B, use the fact that these random quantities can be expressed as continuous functions of the occupation times of the various states in the Markov chain. By appealing to a theorem of Romanovsky (1970, p. 214) concerning the asymptotic multivariate normality of the limiting joint distribution of these occupation times, we are able to conclude that the distributions of OQ_N^R , OQ_N^{NR} , FI_N , and $\hat{p}_N$ can be approximated by normal distributions with means and variances as given earlier in this section. This suggests an approximate γ percentage point, or probability limit, $OQ_N^R(\gamma)$, for OQ_N^R of the form

$$OQ_N^R(\gamma) = E[OQ_N^R] + z_\gamma \sqrt{\text{Var}[OQ_N^R]}$$

where z_γ is the 100γ percentile of a standard normal distribution, i.e., $\Pr[Z < z_\gamma] = \gamma$. Similarly, probability limits can be constructed for OQ_N^{NR} , FI_N , and $\hat{p}_N$ using means and variances derived in this section and the appropriate entries from the standard normal tables. Upper and lower probability limits are useful in quantifying the variability present when considering moderate production run lengths. In fact, a continuity correction improves the approximation to normality. This is described in the next section.

Figure 2 displays means and probability limits for OQ_N^R for a CSP-1 plan with $i = 8$, $f = 1/7$, $N = 500$, in the i.i.d. Bernoulli(p) case, plotted as a function of the incoming quality p . These were the parameters used in the river water sample analyses plan described in Section 1. It can be seen that AOQ_N^R and probability limits are quite different from the limiting AOQ value.

[Insert Figure 2 about here.]

4. SUPPLEMENTARY TABLES FOR MIL-STD-1235B

For the next two sections we restrict our attention to the special case of i.i.d. Bernoulli(p) input in which $p = a = 1 - b$. The MIL-STD-1235B describes a collection of continuous sampling plans including CSP-1 along with corresponding AOQL values. The AOQL values assume the i.i.d. input model. In this section we show how these AOQL values, based on long run mean behaviour, can differ

substantially from the maximum mean for finite production runs, $AOQL_N$, and from maximum upper probability limits.

Maximum means are defined as $AOQL_N^R = \max_{0 \leq p \leq 1} AOQ_N^R$, and similarly for $AOQL_N^{NR}$. Maximum upper probability limits for the outgoing quality under rectifying inspection and nonrectifying inspection, respectively, are defined as

$$OQL_N^R(\gamma) = \max_{0 \leq p \leq 1} \left\{ AOQ_N^R + z_\gamma \sqrt{\text{Var}[OQ_N^R]} + \frac{1}{2N} \right\} \quad (15a)$$

$$OQL_N^{NR}(\gamma) = \max_{0 \leq p \leq 1} \left\{ AOQ_N^{NR} + z_\gamma \sqrt{\text{Var}[OQ_N^{NR}]} + \frac{1}{2 \left(N - \sum_{m=1}^N [\pi_1^m + \pi_{i+2}^m] \right)} \right\} \quad (15b)$$

where π_1^m and π_{i+2}^m are given by (12). Because the random variables OQ_N^R and OQ_N^{NR} are discrete it was found that the inclusion of a continuity correction, one half the reciprocal of the number of outputted items in the run, improved the normality approximation. In a large simulation study, McShane (1989) studied the accuracy of the normal approximation and compared the probability limits (15a, b) to uncorrected limits, and to limits based on logit and arcsine transformations. Despite its ad hoc nature, this simple continuity correction worked best overall in the many cases considered.

[Insert Table 1 about here.]

The maxima in equations (15a, b) were found using numerical search techniques (Cooper and Steinberg 1970, p. 147). A computer program for performing these calculations is listed in McShane (1989), and a copy of it is available from the first author.

Table 1 gives a listing of maximum means, $AOQL_N^{NR}$, and maximum upper probability limits, $OQL_N^{NR}(\gamma)$ for $\gamma = 90\%$, 95% , and 99% for a production run of length $N = 500$. An extensive set of similar tables for both rectifying and nonrectifying inspection, a variety of production run lengths, and a variety of CSP-1 and CSP-F plans selected from Tables 2-A, 3-A-11, and 3-A-12 in MIL-STD-1235B are given in McShane (1989). (CSP-F plans are CSP-1 plans applied to finite production runs with nonrectifying inspection. MIL-STD-1235B supplies tables with $AOQL_N^{NR}$ for these plans.) Table 1 and the tables given by McShane (1989) are intended to serve as an adjunct to the tables in MIL-STD-1235B. Since interest here is in finite production runs, these tables are provided mostly for plans with small to moderate i values (clearance numbers). They were constructed to have a format similar to the

format of the MIL-STD-1235B tables so the AQL and AOQL indexing was retained. We now demonstrate the use of the tables with two examples.

Example 1: Suppose that CSP-1 will be used with nonrectifying inspection on a production run of length $N=500$, and it is desired that the outgoing quality not exceed 5.00%. Using the traditional AOQL criterion, it would be acceptable to select any of the plans in the column corresponding to $AOQL_{\infty} = 4.94\%$ in Table 1, i.e., $(i=6, f=1/2)$, $(i=10, f=1/3)$, . . . , $(i=33, f=1/25)$. $AOQL_{500}$ as well as $AOQL_{\infty}$ for each of these plans is just under 5.00%. However, if we stipulate that with 95% probability the actual outgoing quality (not just the expected outgoing quality) should not exceed 5.00%, we need to look at the maximum upper 95% probability limits. We see that the maximum upper 95% probability limits for the plans given above range from 6.99% for $(i=10, f=1/3)$ to 7.87% for $(i=27, f=1/15)$. None of these would meet the 5.00% criterion. In order to have the maximum 95% upper probability limit not exceed 5.00% we would need to select from amongst the plans in the column corresponding to $AOQL_{\infty} = 2.90\%$, i.e., $(i=10, f=1/2)$, $(i=16, f=1/3)$, . . . , $(i=57, f=1/25)$. The maximum upper probability limits for those plans range from 4.50% for $(i=10, f=1/2)$ to 4.83% for $(i=38, f=1/10)$. Plans in the columns corresponding to $AOQL_{\infty} = 1.22\%$ and 1.90% would also satisfy the 5.00% criterion, but would require more inspection (and therefore greater cost).

Example 2: Suppose that the CSP-F plan with parameters $(i=32, f=1/4)$ is selected from MIL-STD-1235B Table 3-A-12 and will be used with rectifying inspection on a production run of length $N=2000$. We wish to know an upper bound on the outgoing quality for the batch of 2000 items such that with 99% probability that upper bound is not exceeded. This will correspond to the maximum 99% upper probability limit. We first note that the $AOQL_N$ of 1.90% given for Table 3-A-12 in MIL-STD-1235B is computed assuming nonrectifying inspection and therefore is not correct for this situation in which rectifying inspection will be used. This paper provides methods to obtain both the correct $AOQL_N$ values and maximum upper probability limits. The $AOQL_N$ is obtained by maximizing Equation (11b) with respect to the incoming quality, p . The maximum upper 99% probability limit is given by

Equation (15a). Carrying out these calculations (McShane, 1989), we obtain $AOQL_{2000} = 1.79\%$ and a maximum upper 99% probability limit of 2.87%.

5. INTERVAL ESTIMATES FOR INCOMING AND OUTGOING QUALITY

The primary purpose of the CSP-1 plan is for the screening of incoming product; however, estimates of incoming quality based on the inspection data may also be of interest. Probability limits for $\hat{p}_N$ can be inverted to form confidence intervals for the incoming quality p . Let $A(p, \alpha)$ represent an interval of values for $\hat{p}_N$ with endpoints at the $\alpha/2$ and $1-(\alpha/2)$ probability limits, so that $\Pr_p\{\hat{p}_N \in A(p, \alpha)\} = 1-\alpha$ ($0 < \alpha < 1$). Then $CI = \{p: \hat{p}_N \in A(p, \alpha)\}$ is a $100(1-\alpha)$ percent confidence interval for p based on the observed value of $\hat{p}_N$. Point estimates of p are also of interest. From the results of Section 3 and the simulations in McShane (1989), it can be seen that the MIL-STD-1235B estimator $\hat{p}_N$ tends to underestimate p . An estimator with less negative bias is the estimator p^* which is given by the solution to the equation $E_p[\hat{p}_N] = \hat{p}_{obs}$ where $\hat{p}_{obs}$ is the observed value of $\hat{p}_N$.

As an example consider the operation of CSP-1 with $i = 8$, $f = 1/7$, and $N = 500$. This was the scheme adopted for monitoring the quality of laboratory analyses of river water samples described in Section 1. Table 2 gives the 2-sided 95% probability limits and means for $\hat{p}_N$ for selected values of incoming quality p . Figure 3 displays the bias of $\hat{p}_N$ as a function of p . Now suppose that in the run of $N = 500$ items, 90 items had been inspected of which 6 were found to be defective. Thus $\hat{p}_N = 6/90 = 6.67\%$. Using the inversion method described above and Table 2, we obtain a point estimate $p^* = 7.00\%$. By interpolating in the second and third columns of that table, we obtain a 95% confidence interval for p of (2.94%, 12.27%).

[Insert Table 2 about here.]

[Insert Figure 3 about here.]

Confidence intervals for the incoming quality also allow for the construction of data-based prediction and tolerance intervals for the outgoing quality, OQ_N . (We omit the superscript R or NR ; the theory applies to both rectifying and non-rectifying inspection.) Let $CI = (p_L, p_U)$ be a $100(1-\alpha)$ percent confidence interval for p , as constructed above. Let $I = N \cdot (FI_N)$ denote the

number of items inspected and M the number of items passed. Thus $M = N$ for rectifying inspection; for non-rectifying inspection M is equal to N decreased by the number of observed defectives. Note that conditional on I and M , the number of passed undetected defectives, $M \cdot (OQ_N)$, is binomially distributed $B(N-I, p)$. Finally, define

$$H = \frac{1}{M} \max_{p \in CI} [(N-I)p + z_\gamma \sqrt{(N-I)p(1-p)}].$$

Using the normal approximation to the binomial:

$$\begin{aligned} P[P[OQ_N \leq H \mid I, M] \geq \gamma] &\geq P[p \in CI] \\ &= 1 - \alpha. \end{aligned}$$

Thus H is a γ -content tolerance limit for OQ_N at confidence level $1-\alpha$ (Guttman 1970, p. 4).

Alternatively a prediction interval for OQ_N for a future production run can be constructed.

Defining

$$K = \max_{p \in CI} \{OQ_N(\gamma)\}$$

we have that

$$\begin{aligned} P[OQ_N \leq K] &\geq P[OQ_N \leq K \mid p \in CI]P[p \in CI] \\ &\geq P[OQ_N \leq OQ_N(\gamma)]P[p \in CI] \\ &= \gamma(1-\alpha), \end{aligned}$$

and thus K is an upper β -prediction limit for OQ_N , also called a β -expectation tolerance limit where $\beta = \gamma(1-\alpha)$, cf. Guttman (1970, p. 5).

Suppose in the preceding example we wished to compute a one-sided upper prediction limit for the outgoing quality and rectifying inspection is used. From Figure 2, we note that the upper 99% probability limits are increasing on the range of the confidence interval (2.94%, 12.27%). Therefore, the maximum upper prediction limit occurs at the upper endpoint, $p = 12.27\%$. The value of the upper probability limit there is approximately 11.58% (Figure 2). Thus a 94% ($= 95\% \times 99\%$) prediction limit is $K = 11.58\%$.

6. AOQL MEASURES WHEN INPUT PROCESS IS NOT IN “STATISTICAL CONTROL”

In Section 4 we showed the dangers of ignoring variability when considering criteria based on the AOQ for finite production runs. We discussed in Section 1 how the validity of the assumption of an i.i.d. Bernoulli input process (“statistical control”) poses another threat to the applicability of the AOQL values given in MIL-STD-1235B. If no assumptions are made at all, then we have Lieberman’s (1953) UAOQL bound for rectifying inspection given by (3a) and Endres’ (1969) UAOQL bound for nonrectifying inspection given by (3b). From our earlier argument, these bounds are too extreme and a compromise can be found by generalizing the i.i.d. Bernoulli model to that of Kumar and Rajarshi (1983) with expressions for $AOQL^R$ and $AOQL^{NR}$ given by (7b) and (8b). For the rectifying case, another compromise is the Sackrowitz upper bound on the AOQ given by (4).

For selected CSP-1 plans, Tables 3 and 4 compare the “usual” AOQL values obtained under the i.i.d. Bernoulli input assumption with upper probability limits for the OQ_N for production runs of length 500 and 1500, with the AOQL under the more general Kumar/Rajarshi model, and with Lieberman’s (1953) UAOQL bound and Sackrowitz’s (1975) upper bound for the rectifying inspection case, and Endres’ (1969) UAOQL bound for the nonrectifying inspection case. It can be seen from these examples that relaxing the usual statistical control assumptions about the input process can have a large effect on the AOQ and AOQL values.

[Insert Table 3 about here.]

[Insert Table 4 about here.]

For example, if rectifying inspection is used with $i=8$ and $f=1/7$, we see from Table 4 that the usual AOQL would be stated as 10.03%. Under the Kumar Rajarshi model, the $AOQL = 12.38\%$, achieved when $a = 12.5\%$ and $b = 23.03\%$, or $p = a/(a + b) = 35.18\%$. For more extreme situations, we could report Sackrowitz’s $AOQL_S = 26.57\%$ or Lieberman’s $UAOQL^R = 40.00\%$. If we were to consider upper probability limits for the i.i.d. case for production runs of length 500 and 1500, we would give upper bounds for the outgoing quality as 13.73% or 12.20%, respectively. Upper probability limits for OQ_N for the Kumar/Rajarshi model could also be derived using the methods described in Section 3 . (It might be reasonable to amend Equation (12) by assuming that the input

process initially starts in the stationary state.) Tables analogous to Table 1 could be constructed; however, we believe that such tables would be of limited value because the OQ probability limits would be unacceptably high for most applications. Thus they are omitted.

7. CONCLUSION

It is clear that care should be taken in interpreting the AOQ and AOQL, which are the usual measures of effectiveness of CSP-1 plans. Even if the input process is in statistical control, these long run average measures can be very deceiving for finite production runs because the AOQ and AOQL may differ from their finite run counterparts (AOQ_N and $AOQL_N$), and they do not take into account any measure of variability. The deficiencies of the AOQ and AOQL are also evident when we consider input process models which deviate from the i.i.d. Bernoulli model.

8. ACKNOWLEDGMENTS

We thank the Editor, Associate Editor, and two referees for their helpful comments and suggestions for improving this paper. This research was supported in part by Grant GM 28364 for the U.S. National Institutes of Health, and in part by a fellowship from the U.S. Army Research Office through the Mathematical Sciences Institute at Cornell University.

APPENDIX A

Here we present the method for computing the means and variances of OQ_N^R and FI_N .

Define a sequence of indicator functions as follows:

$$I_{(j)}^m = \begin{cases} 1 & \text{if in state } j \text{ at time } m \\ 0 & \text{otherwise.} \end{cases}$$

Write FI_N as

$$FI_N = 1 - \sum_{m=1}^N [I_{(i+4)}^m + I_{(i+5)}^m] / N.$$

Taking expected values of both sides yields

$$E[FI_N] = 1 - \sum_{m=1}^N [\pi_{i+4}^m + \pi_{i+5}^m] / N.$$

Thus we obtain Equation (11a). To compute the variance, write

$$\begin{aligned} \text{Var}[FI_N] &= E[(1 - FI_N)^2] - [E(1 - FI_N)]^2 \\ &= E\left\{\sum_{m=1}^N \sum_{n=1}^N [I_{(i+4)}^m + I_{(i+5)}^m][I_{(i+4)}^n + I_{(i+5)}^n] / N^2\right\} - \left(E\left\{\sum_{m=1}^N [I_{(i+4)}^m + I_{(i+5)}^m] / N\right\}\right)^2. \end{aligned}$$

So,

$$\begin{aligned} N^2 \text{Var}[FI_N] &= \sum_{m=1}^N [\pi_{i+4}^m + \pi_{i+5}^m] - \left(\sum_{m=1}^N [\pi_{i+4}^m + \pi_{i+5}^m]\right)^2 \\ &\quad + 2 \sum_{n=1}^{N-1} \sum_{m=n+1}^N \left\{ (P_{i+5, i+5}^{m-n} + P_{i+5, i+4}^{m-n}) \pi_{i+5}^n + (P_{i+4, i+4}^{m-n} + P_{i+4, i+5}^{m-n}) \pi_{i+4}^n \right\}. \end{aligned} \quad (A1)$$

Equations (11b) and (13) are obtained similarly, recognizing that OQ_N^R can be written as $OQ_N^R = \sum_{m=1}^N I_{(i+5)}^m / N$.

APPENDIX B

Here we present the method for computing the approximate means and variances of OQ_N^{NR} and $\hat{p}_N$.

The MIL-STD-1235B estimator of the incoming quality, $\hat{p}_N$, can be written as

$$\hat{p}_N = \left[\sum_{m=1}^N [I_{(1)}^m + I_{(i+2)}^m] \right] / \left[N - \sum_{m=1}^N [I_{(i+4)}^m + I_{(i+5)}^m] \right].$$

This has the form of a ratio estimator X/Y so we have that, approximately,

$$E\left[\frac{X}{Y}\right] \approx \frac{\mu_X}{\mu_Y} - \frac{\sigma_{XY}^2}{\mu_Y^2} + \frac{\mu_X \sigma_Y^2}{\mu_Y^3} \quad (B1)$$

$$\text{Var}\left[\frac{X}{Y}\right] \approx \frac{\sigma_X^2}{\mu_Y^2} + \frac{\mu_X^2 \sigma_Y^2}{\mu_Y^4} - \frac{2\mu_X \sigma_{XY}^2}{\mu_Y^3} \quad (B2)$$

where $\mu_X = E[X]$, $\mu_Y = E[Y]$, $\sigma_X^2 = \text{Var}[X]$, $\sigma_Y^2 = \text{Var}[Y]$, and $\sigma_{XY}^2 = \text{Cov}[X, Y] = E[XY] - E[X] \cdot E[Y]$.

Here $Y = N \cdot \text{FI}_N$, so

$$E[Y] = N \cdot E[\text{FI}_N] \quad \text{and} \quad \text{Var}[Y] = N^2 \cdot \text{Var}[\text{FI}_N].$$

Applying the same methods presented in Appendix A to derive the mean and variance of FI_N , we replace states $i+4$ and $i+5$ by 1 and $i+2$, and we have

$$\begin{aligned} E[X] &= \sum_{m=1}^N [\pi_1^m + \pi_{i+2}^m] \quad \text{and} \\ \text{Var}[X] &= \sum_{m=1}^N [\pi_1^m + \pi_{i+2}^m] - \left(\sum_{m=1}^N [\pi_1^m + \pi_{i+2}^m] \right)^2 \\ &\quad + 2 \sum_{n=1}^{N-1} \sum_{m=n+1}^N \left\{ (P_{i+2, i+2}^{m-n} + P_{i+2, 1}^{m-n}) \pi_{i+2}^n + (P_{1, 1}^{m-n} + P_{1, i+2}^{m-n}) \pi_1^n \right\}. \\ E[XY] &= N \cdot E \left\{ \sum_{m=1}^N [I_{(1)}^m + \pi_{(i+2)}^m] \right\} - E \left\{ \sum_{m=1}^N \sum_{n=1}^N [I_{(i+4)}^m + I_{(i+5)}^m] [I_{(1)}^n + I_{(i+2)}^n] \right\} \\ &= N \sum_{m=1}^N [\pi_1^m + \pi_{i+2}^m] \\ &\quad - \sum_{n=1}^{N-1} \sum_{m=n+1}^N \left\{ (P_{1, i+4}^{m-n} + P_{1, i+5}^{m-n}) \pi_1^n + (P_{i+2, i+4}^{m-n} + P_{i+2, i+5}^{m-n}) \pi_{i+2}^n \right\} \\ &\quad - \sum_{n=1}^{N-1} \sum_{m=n+1}^N \left\{ (P_{i+4, 1}^{m-n} + P_{i+4, i+2}^{m-n}) \pi_{i+4}^n + (P_{i+5, 1}^{m-n} + P_{i+5, i+2}^{m-n}) \pi_{i+5}^n \right\}. \end{aligned}$$

These quantities are then substituted into Equations (B1) and (B2) to obtain $E(\hat{p}_N)$ and $\text{Var}(\hat{p}_N)$.

The same ratio estimator technique can be used to derive an approximate mean and variance for OQ_N^{NR} by expressing OQ_N^{NR} in terms of indicator functions as

$$\text{OQ}_N^{NR} = \sum_{m=1}^N I_{(i+5)}^m \Bigg/ \left[N - \sum_{m=1}^N [I_{(1)}^m + I_{(i+2)}^m] \right]$$

and proceeding as above. The approximate mean and variance of OQ_N^{NR} are given by the right hand sides of (B1) and (B2), with

$$\mathbb{E}[X] = N \cdot \text{AOQ}_N^R, \quad \text{Var}[X] = N^2 \text{Var}[\text{OQ}_N^R],$$

$$\mathbb{E}[Y] = N - \sum_{m=1}^N [\pi_1^m + \pi_{i+2}^m],$$

$$\begin{aligned} \text{Var}[Y] = & \sum_{m=1}^N [\pi_1^m + \pi_{i+2}^m] - \left(\sum_{m=1}^N [\pi_1^m + \pi_{i+2}^m] \right)^2 \\ & + 2 \sum_{n=1}^{N-1} \sum_{m=n+1}^N \left\{ (P_{i+2, i+2}^{m-n} + P_{i+2, 1}^{m-n}) \pi_{i+2}^n + (P_{1, 1}^{m-n} + P_{1, i+2}^{m-n}) \pi_1^n \right\}, \quad \text{and} \end{aligned}$$

$$\sigma_{XY}^2 = \mathbb{E}[XY] - \mu_X \mu_Y$$

where

$$\begin{aligned} \mathbb{E}[XY] = & N^2 \text{AOQ}_N^R - \sum_{n=1}^{N-1} \sum_{m=n+1}^N (P_{i+5, 1}^{m-n} + P_{i+5, i+2}^{m-n}) \pi_{i+5}^n \\ & - \sum_{n=1}^{N-1} \sum_{m=n+1}^N (P_{1, i+5}^{m-n} \pi_1^n + P_{i+2, i+5}^{m-n} \pi_{i+2}^n). \end{aligned}$$

REFERENCES

- Anscombe, F. J. (1958), "Rectifying inspection of a continuous output," *Journal of the American Statistical Association*, 53, 702-719.
- Banzhaf, R. A. and Brugger, R. M. (1970), "Review of standards and specifications: MIL-STD-1235(ORD), Single- and multi-level continuous sampling procedures and tables for inspection by attributes," *Journal of Quality Technology*, 2, 41-53.
- Bidjarano, M. and Wasmuth, C. W. (1976), *The development of a quality control simulator*, M.Eng. Thesis, Cornell University, Ithaca, New York.
- Blackwell, M. T. R. (1977), "The effects of short production runs on CSP-1," *Technometrics*, 19, 259-263.
- Cooper, L. and Steinberg, D. (1970), *Introduction to Methods of Optimization*, Philadelphia: W. B. Saunders Company.
- Dodge, H. F. (1943), "A sampling inspection plan for continuous production," *The Annals of Mathematical Statistics*, 14, 264-279.
- Endres, A. C. (1969), "The computation of the unrestricted AOQL when defective material is removed but not replaced," *Journal of the American Statistical Association*, 64, 665-668.
- Guttman, I. (1970), *Statistical Tolerance Regions: Classical and Bayesian*, London: Charles Griffin & Co.
- Hillier, F. S. (1964), "New criteria for selecting continuous sampling plans," *Technometrics*, 6, 161-178.
- Kumar, V. S. S. and Rajarshi, M. B. (1987), "Continuous sampling plans for Markov dependent production processes," *Naval Research Logistics Quarterly*, 34, 629-644.
- Lasater, H. A. (1970), *On the robustness of a class of continuous sampling plans under certain types of process models*, Ph.D. Dissertation, Rutgers University, New Brunswick, New Jersey.

- Lieberman, G. (1953), "A note on Dodge's continuous inspection plan," *Annals of Mathematical Statistics*, 24, 480-484.
- Lieberman, G. and Solomon, H. (1955), "Multi-level continuous sampling plans," *The Annals of Mathematical Statistics*, 26, 686-704.
- McShane, L. M. (1989), *Statistical quality control procedures for monitoring laboratory analyses*, Ph.D. Dissertation, Cornell University, Ithaca, New York.
- MIL-STD-1235B (1981), *Single- and multi-level continuous sampling procedures and tables for inspection by attributes*, Superintendent of Documents, Government Printing Office, Washington, D.C.
- Rajarshi, M. B. and Kumar, V. S. S. (1983), "Dodge's CSP-1 for Markov dependent process using probability sampling," *Journal of the Indian Statistical Association*, 21, 99-111.
- Rao, C. R. (1965), *Linear Statistical Inference and Its Applications*, New York: John Wiley and Sons, Inc.
- Roberts, S. W. (1965), "States of Markov chains for evaluating continuous sampling plans," *Transactions of 17th Annual All Day Conference on Quality Control*, Metropolitan Section ASQC and Rutgers University. New Brunswick, New Jersey, 106-111.
- Romanovsky, V. I. (1970), *Discrete Markov Chains*, Groningen (Netherlands): Wolters-Noordhoff Publishing.
- Rykova, L. V. (1973), "Sample checking on the basis of an alternative criterion with allowance for the probability of occurrence of a series of defective articles," *Engineering Cybernetics*, 11, 243-251. (Translated from the Russian.)
- Sackrowitz, H. (1975), "A note on unrestricted AOQLs," *Journal of Quality Technology*, 7, 77-86.
- Yang, G. L. (1983), "A renewal-process approach to continuous sampling plans," *Technometrics*, 25, 59-67.