The Journal of National Chiao Tung University
Vol. 2, December 1976
pp. 111-122

最佳連續抽樣法之討論
Some Optimal Continuous Sampling Plans

戴久永 Chiu-Yung Tai

Department of Transportation Management, N. C. T. U

(Received October 19, 1976)

ABSTRACT — Beattie (1962) proposed a continuous acceptance sampling plan for attributes inspection which seems to appear well-suited and applicable to the situations when production is continuous, sampling test is destructive, a constant sampling rate is required, and discrimination between two quality levels is required with a small sample size. Prairie and Zimmer (1973) presented brief underlying mathematics and provided tables and graphs of the ARL for a variety of plans to aid in the design and evaluation of Beattie's continuous acceptance sampling procedure. However, their approach suffers the restriction that the parameters are integers. In this paper, a method for deriving the non-integer parameters is proposed, we begin by discussing the previously published method. Then we describe our new approach. Finally, some tables and graphs are provided for determining a plan along with examples of their use.

I. Introduction

Consider, for instance, the situation that a continuous flow of manufactured products are submitted for the final inspection, each item is classified as defective or non-defective. It is desired to accept or reject the product based on the number of defectives observed, where the sampling test may be destructive, the number of items available for testing is relatively small, and the sampling rate is required to keep constant. It is desired to discriminate between the prescribed Acceptable Quality Level (AQL) p_1 and Rejectable Quality Level (RQL) p_2, such that the procedure should accept (reject) at p_1 (p_2) with a high probability. A procedure that appear to be well suited for this type of situation is one developed by Beattie (1962) which based on cumulative sums.

Beattie's CSP is to set up a cusum chart which has an accept zone and a reject zone for the cusum. A small sample of size n is selected at regular intervals of time from the process and the number of defectives y in the sample is recorded. A reference value k is chosen such that the cusum $S_m=\Sigma(y_j-k)$ decreases if $p=p_1$ and increases if $p=p_2$, this allows the decision lines ($S_m=h$ and $S_m=h+h^*$) on the accept and reject zones to remain horizontal. The cusum S_m is computed and plotted according to the rules below. An example is shown in Fig. 2 (see later).

The rules for plotting are as follows:

(1) Start the cumulation at zero.

(2) Accept product as long as $S_m<h$. When $S_m<0$, return the cumulation to zero.

(3) Reject product when $S_m \geq h$, restart cumulation at $S_m=h+h^*$ and continue rejecting product until $S_m<h$. When $S_m>h+h^*$, return cumulation to $h+h^*$.

(4) When h is crossed or reached from above, accept product and restart cumulation at zero.

In the above description, values of n, k, h and h* can be found so that product will be accepted $100(1-\alpha)\%$ of the time when $p=p_1$, i.e., $P(a|p_1)=1-\alpha$ and rejected $100\beta\%$ of the time when $p=p_2$, i.e., $P(a|p_2)=\beta$.

When the sampling rate is the same in both zones, the probability of acceptance (the proportion of product accepted) for a given quality p is

$$P(a|p)=\frac{L(0,p)}{L(0,p)+L^*(h+h^*,p)}$$

Where $L(0,p)$ is the ARL in the accept zone and $L^*(h+h^*,p)$ is the ARL in the reject zone. By ARL we mean the average number of samples taken before the cusum reaches or across the decision line.

A common method of evaluating the suitability of a sampling plan to a particular situation is the OC curve. If we have a large collection of sampling plans and corresponding OC curves, and can choose a curve that describes our required degree of discrimination, (i.e. satisfies the specified α and β) we can then choose the appropriate plan.

To obtain the OC curve for a given Beattie's plan, it is necessary to find solutions for $L(0,p)$ and $L^*(h+h^*,p)$ for various values of n, k, h and h*. The values of $L(0,p)$ and $L^*(h+h^*,p)$ can be computed by the following two equations derived by Ewan and Kemp (1960) which may be solved iteratively or by standard matrix inversion methods.

In the accept zone, let $L(z,p)$ denote the ARL for the Beattie's plan which starts at a point z units above the 0 boundary.

$$L(z,p)=1+\sum_{y=1}^{h-1} L(y,p)f(y+k-z)+L(0,p)\sum_{x=0}^{k-z} f(x) \qquad (2.1)$$

$z=0,1,2,\ldots h-1$. In the reject zone, let $L^*(z,p)$ denote the ARL which the first cumulation starts at a point z units below the h+h* boundary. Then

$$L^*(z,p)=1+\sum_{y=1}^{h^*-1} L^*(y,p)f(z-y+k)+L^*(h+h^*,p)\sum_{x=z}^{\infty} f(k+x) \qquad (2.2)$$

$z=h+1,h+2,\ldots h+h^*$, in each equation $f(x)$ is the probability of having exactly x defectives in the sample; $f(x)$ is taken here as the binomial or Poisson density functions.

Prairie and Zimmer (1973) presented brief underlying mathematics to Beattie's acceptance sampling plan, and provided graphs and tables for determining a plan under the binomial assumption and the Poisson approximation. However, the parameters of the plans are restricted to be positive integers.

Based on an empirical and an intuitive argument, Ewan and Kemp (1960) conjectured that the optimal choice of k is np_0 where n is the sample size, and p_0 is a value between p_1 and p_2. As long as k is restricted to be an integer, large sample size will be required for small fraction defective. In order to minimize the amount of sampling needed to satisfy the desired requirement and still allow k to be equal to np_0 without the necessity for a large sample size, a fraction value of k is proposed. In this paper, we describe the Beattie's procedure with non-integer parameters, and present tables and graphs for determining a plan to discriminate between the prescribed AQL and RQL. Under the binomial assumption and the Poisson approximation.

II. The Non-Integer Parameters Approach

In order to illustrate the procedure for the non-integer situation, consider the following hypothetical situation. Suppose a plan is defined by n=5, k=2/3, h=2, h*=1, the results from 15 samples are given below.

Sample	1	2	3	4	5	6	7	8
d	1	0	2	1	1	0	1	0
d-k	1/3	-2/3	4/3	1/3	1/3	-2/3	1/3	-2/3
S_m	1/3	-1/3→0	4/3	5/3	2→3	7/3	8/3	2→ 0
Sample	9	10	11	12	13	14	15	
d	0	2	0	1	1	0	1	
d-k	-2/3	4/3	-2/3	1/3	1/3	-2/3	1/3	
S_m	-2/3→0	4/3	2/3	1	4/3	2/3	1	

Figure 1. The numerical cumulative sum tabulation.

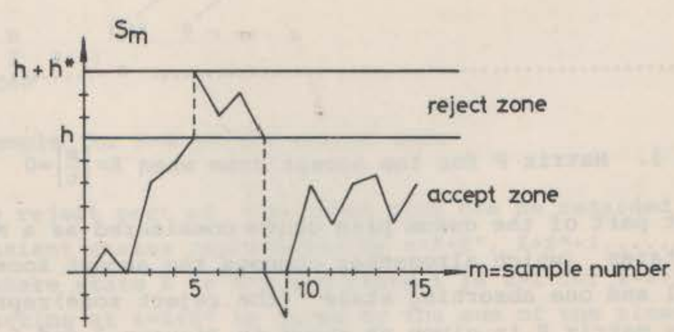

Figure 2. Form of cusum chart for the acceptance of the continuous production process.

With this plan, accept or refect decisions are made regularly on the basis of the sample at hand together with preceding ones.

III. ARL Derived with a Markov Chain

Beattie's continuous sampling plan with sample size n is determined by the parameters (k,h,h*). Let $k=\frac{s}{t}$, both s and t are integers, s is not a multiple of t. The variable (y-k) will have (n+1) possible values; i.e., $y-k=\frac{yt-s}{t}$, y=0, 1,2,...n. We define the symbol z to be the number of steps of size $\frac{1}{t}$ above 0 represented by $\sum_{1}^{m}(y_j-k)$. That is, $z=0,1,2,\ldots\frac{sh}{k},\frac{sh}{k}+1,\frac{sh}{k}+2,\ldots\frac{s(h+h^*)}{k}$ is equivalent to $\Sigma(y_j-k)=0,\ \frac{1}{t},\ \frac{2}{t},\ldots\ h,\ h+\frac{1}{t},\ldots h+h^*$. We assume h to be a multiple of k although it is not necessary. For convenience, we define $Z=\frac{sh}{k}$ and $Z^*=\frac{sh^*}{k}$. Both Z and Z* are integers. (They are the sizes of matrices for the accept zone and the reject zone.)

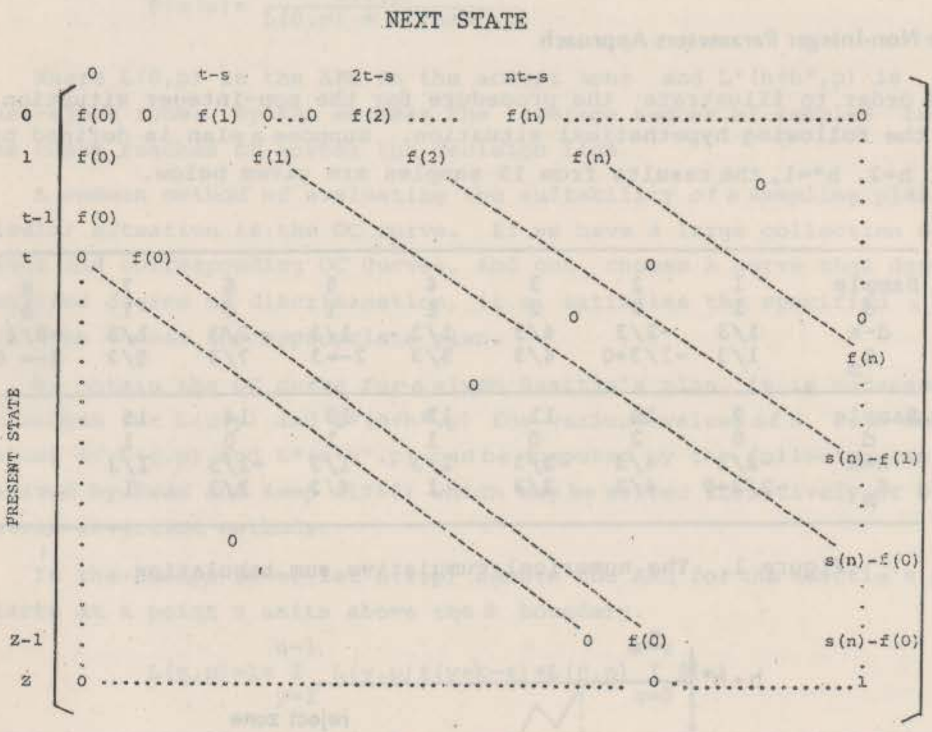

Figure 3. Matrix P for the accept zone when $K=\left[\frac{s}{t}\right]=0$

The accept part of the cusum plan can be considered as a Markov chain with Z transient states (which altogether compose the accept zone) represented by z=0,1,2,...Z-1 and one absorbing state (the reject zone) represented by $z \geq Z$. The transition matrix P is given as shown in Figures 3 and 5 where state i is z=i for i=0,1,2,...,Z-1 and state Z is $z \geq Z$. The matrix of transient states Q, is formed by removing the last row and column from P. The fundamental matrix is $N=(I-Q)^{-1}$, (see example in Figure 4), and the ARL having started at z =0 is given by the sum of the elements of the first row of N. (The sums of the other rows give the ARLs having started at the other levels.)

Let n=3, k=2/3, h=3 we have the matrix for accept as follows

$$P = \begin{bmatrix} f(0) & f(1) & 0 & 0 & f(2) & 0 & 0 & f(3) & 0 & 0 \\ f(0) & 0 & f(1) & 0 & 0 & f(2) & 0 & 0 & f(3) & 0 \\ f(0) & 0 & 0 & f(1) & 0 & 0 & f(2) & 0 & 0 & f(3) \\ 0 & f(0) & 0 & 0 & f(1) & 0 & 0 & f(2) & 0 & f(3) \\ 0 & 0 & f(0) & 0 & 0 & f(1) & 0 & 0 & f(2) & f(3) \\ 0 & 0 & 0 & f(0) & 0 & 0 & f(1) & 0 & 0 & s(3)-s(1) \\ 0 & 0 & 0 & 0 & f(0) & 0 & 0 & f(1) & 0 & s(3)-s(1) \\ 0 & 0 & 0 & 0 & 0 & f(0) & 0 & 0 & f(1) & s(3)-s(1) \\ 0 & 0 & 0 & 0 & 0 & 0 & f(0) & 0 & 0 & s(3)-s(0) \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$I-Q = \begin{bmatrix} 1-f(0) & -f(1) & 0 & 0 & -f(2) & 0 & 0 & -f(3) & 0 \\ -f(0) & 1 & -f(1) & 0 & 0 & -f(2) & 0 & 0 & -f(3) \\ -f(0) & 0 & 1 & -f(1) & 0 & 0 & -f(2) & 0 & 0 \\ 0 & -f(0) & 0 & 1 & -f(1) & 0 & 0 & -f(2) & 0 \\ 0 & 0 & -f(0) & 0 & 1 & -f(1) & 0 & 0 & -f(2) \\ 0 & 0 & 0 & -f(0) & 0 & 1 & -f(1) & 0 & 0 \\ 0 & 0 & 0 & 0 & -f(0) & 0 & 1 & -f(1) & 0 \\ 0 & 0 & 0 & 0 & 0 & -f(0) & 0 & 1 & -f(1) \\ 0 & 0 & 0 & 0 & 0 & 0 & -f(0) & 0 & 1 \end{bmatrix}$$

where $s(n) = \sum_{i=0}^{n} f(i)$

Figure 4. Example for K=0 in the accept zone

Similarly, the reject part of the cusum plan can be regarded as a Markov chain with Z^* transient states represented by $z=Z+Z^*$, $Z+Z^*-1$, ... $Z+1$ and one absorbing state, where state Z is $z \leq Z$ and state i is $z=i$ for $i=Z+1$, $Z+2$, ... $Z+Z^*$. The ARL starting at $z=Z+Z^*$ is given by the sum of the elements of the last row of N.

Further details of Markov chain theory can be found in Finite Markov Chain by Kemeny and Snell (1959) and Stochastic Processes by Parzen (1962).

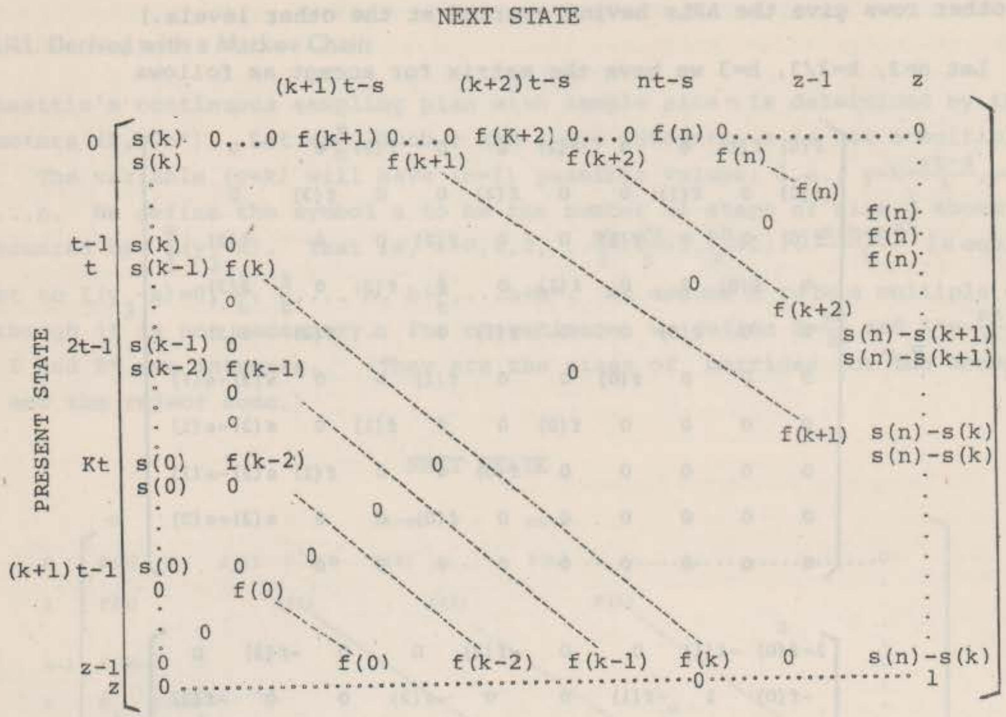

Figure 5. Matrix p for the accept zone when $K \geq 1$.

From now on, let L*(h+h*,p) denote the ARL in the reject zone for integer values k, h, and h* and L*(Z+Z*,p) for non-integer value of k, h, and h*.

IV. Computations

For given n, k, h and h* we can use the matrices (I-Q) derived in the preceding section to solve numerically for L(0,p) and L*(Z+Z*,p) to satisfy the specified α and β values over a suitable range of p by using computer. Results of the computations are similar to the Figures 9 through 12, where the figures shown are curves of ARL versus product quality for the different value of n,k, h and h*. In the Figures 9-10, curves of L(0,p) and L*(Z+Z*,p) are given for the binomial case and in Figures 11-12 for the Poisson case. Note that the binomial would generally be the distribution to be used; however, when the Poisson provides a good approximation to the binomial, it is more convenient to use the Poisson distribution. In TABLE 2 (see later) some specific plans are given for various levels of quality. These plans are such that $P(a|p_1 \text{ (or } m_1))=0.95$, $P(a|p_2(\text{or } m_2))=0.10$ and $L(0,p_2(\text{or } m_2))$ is less than 6.

V. Illustrations and the Design of a Sampling Plan

(1) Suppose that a plan is selected with a uniform sampling rate and n=5, k=3/8, h=1,h*=1. Using Figures 9 and 10, the OC curve can be obtained $P_u(a|p)$ is shown in Figure 6 as a function of quality for some selected values of p. If the sampling rates r_a and r_r are not taken to be equal, a significantly different OC curve will result. Figure 7 illustrates the OC curve for the same plan as just described but with r_a=0.1 and r_r=1.0. Some values of $P_v(a|p)$ as a function of quality as shown in Figure 6.

P	0.01	0.02	0.05	0.10	0.18	0.20
$P_u(a\|p)$	0.989	0.957	0.765	0.389	0.090	0.060
$P_v(a\|p)$	0.999	0.996	0.970	0.864	0.497	0.390

Figure 6. Tabulation for comparison of $P_u(a|p)$ and $P_v(a|p)$

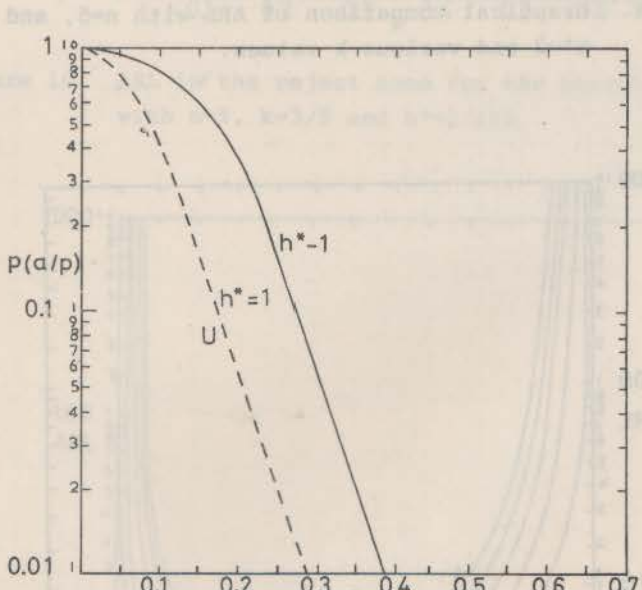

Figure 7. Graphical comparison of OC curves for $P_u(a|p)$ and $P_v(a|p)$ with n=5, k=3/8, h=1 and h*=1.

A plan with r_a=0.1, r_r=1.0, and the same protection at p=0.18, i.e., P(a|0.18) =0.10 can be obtained by changing h* and thus L*(Z+Z*, 0.18). For instance, the same protection at p=0.18 is provided by the original plan results from using h*=2 instead of 1.

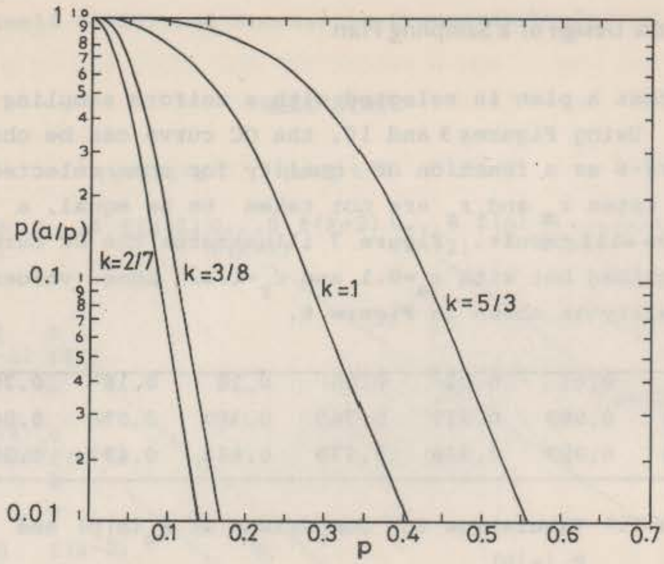

Figure 8. Graphical comparison of ARL with n=5, and h=1, h*=2 and various k values.

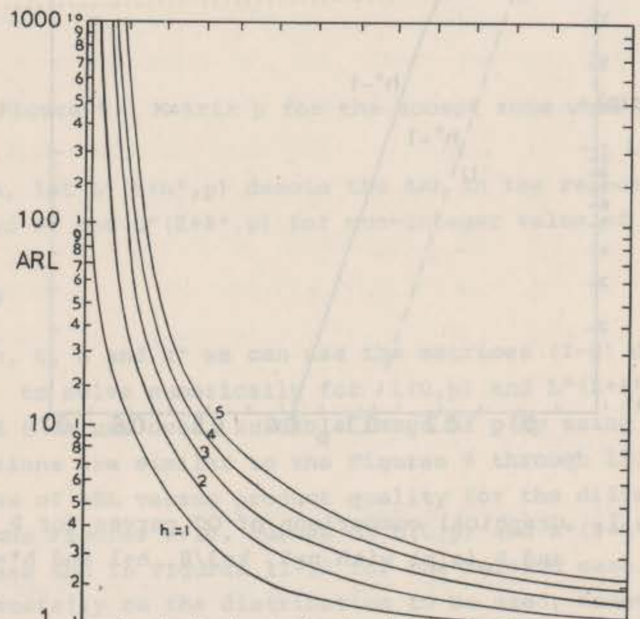

Figure 9. ARL in the accept zone for the binomial sampling with n=5, k=3/8 and h*=1(1)5.

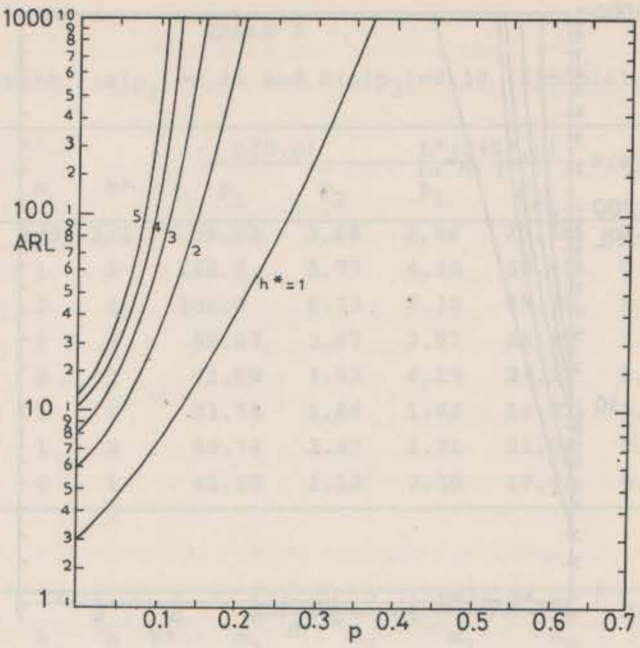

Figure 10. ARL in the reject zone for the binomial sampling with n=5, k=3/8 and h*=1(1)5.

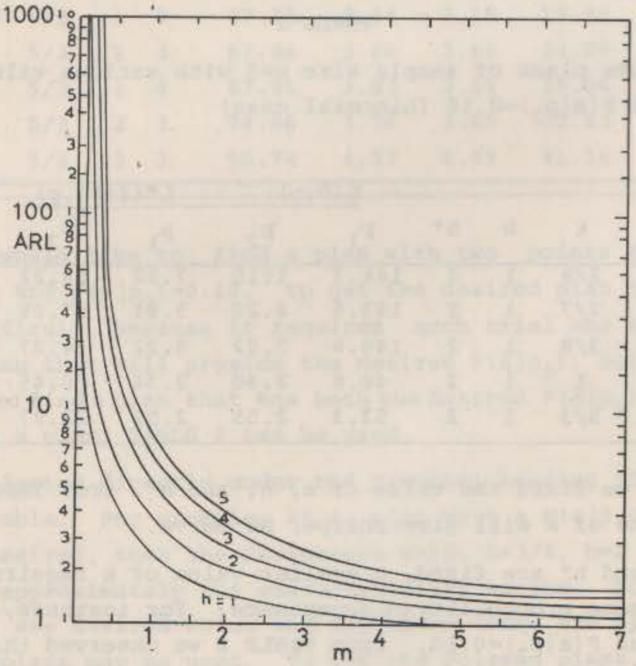

Figure 11. ARL in the accept zone for the Poisson sampling with k=2/7, h=1(1)5.

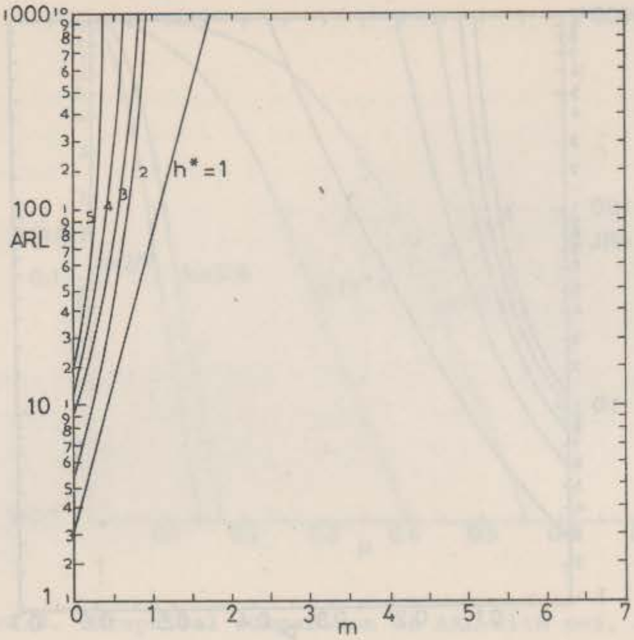

Figure 12. ARL in the reject zone for the Poisson sampling with k=2/7, h*=1(1)5.

TABLE 1

Comparison of some plans of sample size n=5 with various values of k such that $P(a|p_1)=0.95$ and $P(a|p_2)=0.10$ (binomial case)

					L(0,p)		L*(Z+Z*,p)		$P(a\|p_1)$	$P(a\|p_2)$
p_1	p_2	k	h	h*	p_1	p_2	p_1	p_2		
0.01	0.07	2/9	1	2	146.3	1.10	7.22	68.22	0.93	0.10
0.01	0.09	2/7	1	2	193.5	8.22	5.81	57.08	0.96	0.09
0.015	0.11	3/8	1	2	140.9	7.02	5.22	44.27	0.95	0.11
0.05	0.26	1	1	2	48.8	2.60	2.38	20.45	0.95	0.10
0.13	0.43	5/3	1	2	52.3	2.55	2.51	22.97	0.95	0.10

(2) Suppose we fixed the value of n, h, and h*, from TABLE 1 or Figure 8, the smaller value of k will give sharper OC curve.

(3) When h and h* are fixed, a smaller value of k requires smaller sample size to get the same probability of acceptance. For instance, let h=1 and h*=1, $P(a|p_1)=0.95$ and $P(a|p_2)=0.10$, from TABLE 2 we observed that when k=2/9, n=5 while k=3/8, n=10. This is one of the advantages to use a non-integer value k. Because if the sample size required is small, we can save inspection costs.

TABLE 2

Some specific plans with $P(a|p_1)=0.95$ and $P(a|p_2)=0.10$ (Binomial case)

						L(0,p)		L*(Z+Z*,p)		$P(a\|p_1)$	$P(a\|p_2)$
p_1	p_2	n	k	h	h*	p_1	p_2	p_1	p_2		
0.02	0.17	5	2/9	4/3	2/3	55.32	3.68	2.86	25.35	0.94	0.10
0.01	0.11	5	2/9	1	1	146.3	5.77	4.30	38.40	0.96	0.10
0.02	0.12	5	2/9	2	1	200.0	6.73	6.18	48.39	0.97	0.11
0.01	0.09	10	3/8	1	1	80.47	3.67	2.97	25.96	0.96	0.10
0.02	0.10	10	3/8	2	1	92.59	4.53	4.13	34.54	0.95	0.11
0.07	0.27	10	5/3	1	1	33.34	1.69	1.82	14.37	0.95	0.11
0.06	0.23	10	5/3	1	2	50.74	2.47	2.31	21.01	0.95	0.10
0.10	0.29	10	5/3	2	1	43.10	2.12	2.36	19.44	0.95	0.11

(Poisson case)

						L(0,p)		L*(Z+Z*,p)		$P(a\|m_1)$	$P(a\|m_2)$
m_1	m_2	R	k	h	h*	m_1	m_2	m_1	m_2		
0.07	0.72	0.10	2/7	1	1	102.00	4.71	3.52	31.39	0.96	0.10
0.11	0.94	0.12	3/8	1	1	66.29	3.74	2.90	25.76	0.95	0.10
0.20	1.46	0.19	3/8	2	1	88.58	4.51	3.92	35.14	0.95	0.10
0.27	1.14	0.24	3/8	3	1	98.63	5.24	4.88	43.35	0.95	0.10
0.57	2.46	0.23	5/3	1	2	47.71	2.44	2.18	19.58	0.95	0.10
0.53	2.13	0.25	5/3	1	3	57.86	3.00	2.68	24.09	0.95	0.10
0.50	1.92	0.26	5/3	1	4	67.61	3.83	3.16	28.56	0.95	0.10
0.80	2.29	0.35	5/3	2	3	74.96	3.76	3.65	32.93	0.95	0.10
1.00	2.40	0.42	5/3	3	3	90.74	4.57	4.59	41.16	0.95	0.10

(4) Suppose we would like to find a plan with two points fixed on the OC curve, $P(a|p_1)=0.95$ and $P(a|p_2)=0.10$. To get the desired plan through use of the figures is difficult, because it requires much trial and error. It may be easy to find a plan that will provide the desired $P(a|p_2)$, but a lot of juggling may be needed to find a plan that has both the desired $P(a|p_1)$ and $P(a|p_2)$. To aid in selecting a plan, TABLE 2 can be used.

A plan may be selected directly under the binomial heading if the required p_1 and p_2 are available. For example, if a plan with a $P(a|0.01)=0.95$ and $P(a|0.09)=0.10$ is desired, then the parameters n=10, k=3/8, h=2 and h*=1 will define a plan with approximately the characteristics wanted. For situations where a p_1 and a p_2 are desired which are not given under the binomial plans, one of the Poisson plans may be used. To use the Poisson plans, first calculate the ratio $R=p_1/p_2$ and then looking at TABLE 2 that has a value of R close to the one computed and use the plan given on that line. For instance, suppose we would like to find a plan with $p_1=0.01$ and $p_2=0.1$, R=0.1 and the plan given for R=0.1 is h=1, h*=1 and k=2/7. For such a plan $m_1=0.07$ and $m_2=0.72$. From the relation m=np, we have n=0.07/0.01=7, and a plan n=7 and the parameter values given above will have the required p_1 and p_2.

(5) Assume that a plan is desired with P(a|0.005)=0.95, and that an ARL in the reject zone L*(Z+Z*,0.005) of approximately 5 would be acceptable. Then P(a|0.005)=L(0,0.005)/L((0,0.005)+L*(Z+Z*,0.005))=L(0,0.005)/(L(0,0.005)+5)=0.95, L(0,0.005)=95. From Figure 14, it is seen that a plan with k=2/7, h=2 and hence m=0.15 could be used. Using the relation m=np, the necessary sample size with m=0.15 and p=0.005 is n=30. From Figure 12 it is seen that L*(Z+Z*, 0.005) is approximately 5 for m=0.15 and h*=1. Thus a plan with k=2/7, h=2, h*=1 and n=30 gives the desired characteristics.

References

1. Beattie, D. W., "A continuous acceptance sampling procedure based upon a cumulative sum chart for the number of defectives.", Applied Statistics 11, 137-147 (1962).
2. Ewan, W. D. & Kemp, K. W., "Sampling inspection of continuous processes with no autocorrelation between successive results." Biometrika 47, 363-380 (1960).
3. Kemeny, J. G., Snell, J. L., "Finite Markov Chain." Van Nostrand, Princeton, N. J., (1959).
4. Parzen, E., "Stochastic Processes." Holden-Day, S. F., (1962).
5. Prairie, R. R. & Zimmer, W. J., "Graphs, tables, and discussion to aid in the design and evaluation of the acceptance sampling procedure based on cumulative sums." Jol. of Quality Tech. 5, No. 2, 58-66, (1973).