Anders Hald

Two Notes on Attribute Sampling Plans

Preprint
November
1975

16

Institute of Mathematical Statistics
University of Copenhagen

Anders Hald

TWO NOTES ON ATTRIBUTE SAMPLING PLANS

Preprint 1975 No. 16

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

November 1975

AN APPROXIMATION TO THE BINOMIAL OC BY MEANS OF THE POISSON OC FOR MULTIPLE SAMPLING PLANS

by

Anders Hald

Institute of Mathematical Statistics

University of Copenhagen

SUMMARY. Assuming that the Poisson OC function for a multiple sampling plan is known it is shown how an approximation to the binomial OC may be found by means of a conversion formula similar to the one valid for single sampling.

KEY WORDS. Approximation to binomial OC. Multiple sampling. Double sampling.

Let there be given a multiple sampling plan and a table of its OC function based on the Poisson distribution. The problem is to find a good approximation to the corresponding binomial OC using the values in the Poisson table.

Let c denote the acceptance number and n the sample size for a single sampling plan so that the binomial OC function is $B(c,n,p)$. For given probability of acceptance, P say, we define the OC fractile or percentage point p_P as solution to the equation $B(c,n,p_P) = P$ for $0 \leqq P \leqq 1$.

Similarly, we define the Poisson fractile λ_P as solution to the equation $G(c,n\lambda_P) = P$, where $G(c,m)$ denotes the Poisson distribution function with mean m. To find λ_P we introduce the auxiliary function $m_P(c)$ defined by the equation $G(c,m_P) = P$. This function $m_P(c) = \frac{1}{2}\chi^2_{1-P}(2c+2)$ has been extensively tabulated, also for non-integral values of c. Since $\lambda_P = m_P(c)/n$ it is easy to find the Poisson fractiles for any value of n from a table of $m_P(c)$.

For small values of p the Poisson fractile may be used as an approximation to p_P. It is well known that a better approximation is obtained from the formula

$$\hat{p}_P = m_P(c)/\{n + \tfrac{1}{2}(m_P(c) - c)\} = \lambda_P/\{1 + \tfrac{1}{2}(\lambda_P - h)\} \quad \text{for } h = c/n.$$

Numerical investigations show that the <u>relative</u> error of $\hat{p}_P$ is less than 2 % for $(c+1)/(n+1) \leqq 0.25$ under the assumption that $n \geqq 20$ and $0.01 \leqq P \leqq 0.99$. In general, the relative error is increasing with $(c+1)/(n+1)$ and decreasing with P so that the largest relative error within the domain defined above is obtained for $(c+1)/(n+1) = 0.25$ and $P = 0.01$. For other values of the parameters the relative error is considerably smaller. Hence, the approximation is sufficiently accurate for most applications in sampling inspection.

Consider now a multiple sampling plan characterized by the parameters (b,n_1) where n_1 denotes the size of the first sample and b denotes a vector consisting of the acceptance and rejection numbers and the remaining sample sizes divided by n_1. For a k-stage sampling plan b is a vector of dimension $3k-2$. With this notation the OC function for the multiple sampling plan computed from the Poisson distribution may be written as $H(b,n_1\lambda)$, say analogously to the OC function $G(c,n\lambda)$ for single sampling. Introducing the auxiliary function $v_P(b)$ defined by the equation $H(b,v_P) = P$, $0 \leqq P \leqq 1$, we get the corresponding OC fractile for the multiple sampling plan, λ_P^* say, as $\lambda_P^* = v_P(b)/n_1$.

Since the OC for any multiple sampling plan may be approximated by the OC for a single sampling plan it is to be expected that the approximate relationship between the Poisson and binomial fractiles discussed above will be valid also for multiple sampling , at least within the domain where the equivalent single sampling plan gives a good approximation to the OC for the multiple sampling plan. Numerical investigations confirm this expectation. Hence, we have the following approximation to the binomial OC fractile

$$\hat{p}_P = v_P(b)/\{n_1 + \tfrac{1}{2}(v_P(b) - n_1h)\} = \lambda_P^*/\{1 + \tfrac{1}{2}(\lambda_P^* - h)\},$$

where $h = c/n$ is found from the equivalent single sampling plan.

To find the equivalent single sampling plan we choose as usual two risks α and β, $0 < \beta < 1-\alpha < 1$, determine the corresponding Poisson fractiles $\lambda_{1-\alpha}^* = v_{1-\alpha}(b)/n_1$ and $\lambda_\beta^* = v_\beta(b)/n_1$ for the multiple sampling plan, solve the equation $m_\beta(c)/m_{1-\alpha}(c) = \lambda_\beta^*/\lambda_{1-\alpha}^*$ for c and compute $n = m_\beta(c)/\lambda_\beta^*$. Of course, h depends on the choice of α and β, but not much if the usual values of (α,β) are adhered to. Notice that $n_1h = v_\beta(b)c/m_\beta(c)$.

A table of $v_P(b)$ for multiple sampling plans combined with the conversion formula above will give the binomial OC fractiles with sufficient accuracy for most sampling inspection problems. The importance of this result is, of course, that we save the nearly impossible task of tabulating the binomial OC functions. In a large number of cases investigated the relative error has been less than 5 % if the conditions mentioned for the equivalent single sampling plan have been fulfilled.

An example taken from Mil.Std.105D for AQL = 0.04 has been given in Table 1. The table contains for single, double and multiple sampling the OC fractiles computed from the Poisson distribution, from the approximation formula and from the binomial distribution, respectively. Equivalent single sampling plans for $\alpha = 0.05$ and $\beta = 0.10$ have been used. It will be seen that the approximation formula gives satisfactory results even in the present case where the equivalent single sample size is about 20. As mentioned above the relative error is largest for $P = 0.01$ where $\hat{p} > p$.

In Table 2 we have given some values of the function $v_P(b)$ for the multiple sampling plan in Table 1 and shown how the approximation formula may be used to find the binomial OC fractiles for $n_1 = 3$, 10 and 20. It will be seen that the relative error is a rapidly decreasing function of n_1 or the equivalent n.

Acknowledgements.

I am grateful to Mr. Uffe Møller for carrying out the computations.

Table 1. Exact and approximative OC functions for single, double and multiple sampling plans.

	$n = 20$			$n_1 = n_2 = 13$			$n_1 = \cdots = n_7 = 5$		
	Ac 2			Ac 0 3			Ac −1 0 0 1 2 3 4		
	Re 3			Re 3 4			Re 2 3 3 4 4 5 5		
100P	100λ	$100\hat{p}$	100p	100λ	$100\hat{p}$	100p	100λ	$100\hat{p}$	100p
99	2.18	2.27	2.27	2.79	2.91	2.94	2.22	2.31	2.36
95	4.09	4.21	4.22	4.89	5.04	5.07	4.35	4.49	4.53
90	5.51	5.64	5.64	6.36	6.52	6.54	5.86	6.00	6.03
75	8.64	8.70	8.70	9.47	9.55	9.56	9.03	9.09	9.07
50	13.4	13.1	13.1	14.0	13.8	13.8	13.7	13.4	13.3
25	19.6	18.7	18.7	19.7	18.9	18.8	19.8	18.9	18.5
10	26.6	24.6	24.5	26.2	24.3	24.1	27.1	25.0	24.3
5	31.5	28.4	28.3	30.7	27.9	27.6	32.7	29.4	28.4
1	42.0	36.2	35.8	40.7	35.5	34.6	46.9	39.6	37.6
	c=2 n=20 h=0.1			c=2.63 n=23.6 h=0.111			c=2.12 n=20.3 h=0.105		

Table 2. Values of $v_P(b)$ and approximate and exact binomial OC fractiles for multiple sampling plans with $n_1 = n_2 = \cdots = n_7$, acceptance and rejection numbers as in Table 1 and $n_1 = 3$, 10 and 20. $c = 2.12$, $n_1 h = 0.522$.

		$n_1 = 3$		$n_1 = 10$		$n_1 = 20$	
100P	v_P	$100\hat{p}$	100p	$100\hat{p}$	100p	$100\hat{p}$	100p
99	0.1108	3.96	4.14	1.13	1.14	.560	.562
95	0.2175	7.64	7.79	2.21	2.22	1.10	1.10
90	0.2930	10.2	10.3	2.96	2.97	1.47	1.47
75	0.4514	15.2	15.2	4.53	4.52	2.26	2.26
50	0.6826	22.2	21.8	6.77	6.74	3.40	3.39
25	0.9882	30.6	29.7	9.66	9.57	4.88	4.86
10	1.355	39.7	37.8	13.0	12.8	6.64	6.59
5	1.635	46.0	43.3	15.5	15.2	7.96	7.89
1	2.343	59.9	54.7	21.5	20.9	11.2	11.1

ON THE DETERMINATION OF ATTRIBUTE SAMPLING PLANS OF GIVEN STRENGTH

by
Anders Hald
Institute of Mathematical Statistics
University of Copenhagen

SUMMARY. Simple and accurate approximation formulas for the determination of single sampling attribute plans in the binomial and hypergeometric case are given based on the Poisson solution. The solution is generalized to multiple sampling.

KEY WORDS. Attribute sampling plans. Single sampling. Multiple sampling. Producer's and consumer's risk. Poisson. Binomial. Hypergeometric.

In a previous paper (Hald, 1967) we have given a survey of exact and approximative solutions to the problem of determining single sampling attribute plans with given producer's and consumer's risk. The purpose of the present note is to present a new approximation which is at the same time simpler and more accurate than the previous ones.

Let n denote the sample size and c the acceptance number for a single sampling plan so that the binomial operating characteristic (OC) equals $B(c,n,p)$. The corresponding OC functions derived from the Poisson and the hypergeometric distributions will be denoted by $G(c,np)$ and $H(c,n,Np,N)$, respectively.

Ideally we want to determine (n,c) so that the producer's risk equals α and the consumer's risk equals β, but since n and c have to be integers it is usually not possible to find a plan satisfying these requirements exactly. We therefore reformulate the problem as follows:

Let $P(p)$ denote the OC function. Determine (n,c) so that $1-P(p_1) \leqq \alpha$, $P(p_2) \leqq \beta$, and c is as small as possible, where p_1, p_2, α and β are given numbers, $0<p_1<p_2<1$ and $0<\beta<1-\alpha<1$.

The main tool for solving this problem is the auxiliary function $m_P(c)$, say, defined as solution to the equation $G(c,m_P) = P$ for $0 \leqq P \leqq 1$. The function $m_P(c) = \frac{1}{2}\chi^2_{1-P}(2c+2)$ has been extensively tabulated, also for non-integral values of c. Furthermore we need the function

$$R(c;\alpha,\beta) = m_\beta(c)/m_{1-\alpha}(c)$$

which is a decreasing function of c. Since (α,β) in the present context are given we shall write $R(c)$ only.

Consider first the problem of solving the two equations $G(c,n\lambda_1) = 1-\alpha$ and $G(c,n\lambda_2) = \beta$, where λ_1 and λ_2 are given positive numbers, $\lambda_1 < \lambda_2$. Allowing c and n to be non-integral we have $n\lambda_1 = m_{1-\alpha}(c)$ and $n\lambda_2 = m_\beta(c)$ so that

$$R(c) = \lambda_2/\lambda_1 \quad \text{and} \quad n = m_{1-\alpha}(c)/\lambda_1 = m_\beta(c)/\lambda_2.$$

These two equations determine c and n uniquely. They are simple to solve by means of a table of $m_P(c)$. It will be shown how they may be generalized to cover the three problems considered.

For integral values of c and n it is easy to see that the two equations above must be replaced by two inequalities. We then get the well-known Poisson solution:

The smallest acceptance number satisfying $1 - G(c,np_1) \leqq \alpha$ and $G(c,np_2) \leqq \beta$ is uniquely determined by the inequality $R(c-1) > p_2/p_1 \geqq R(c)$ and the corresponding sample size satisfies the inequality $m_\beta(c)/p_2 \leqq n \leqq m_{1-\alpha}(c)/p_1$.

Since the binomial and hypergeometric distributions may be approximated by the Poisson distribution it is to be expected that a simple modification of the Poisson solution above will lead to a satisfactory approximation for the two other cases.

Let p_P be defined by the equation $B(c,n,p_P) = P$ for $0 \leqq P \leqq 1$. The two requirements to the OC function may then be written as $p_{1-\alpha} \geqq p_1$ and $p_\beta \leqq p_2$. Using the well-known approximation formula

$$p_P \simeq m_P/\{n + \frac{1}{2}(m_P - c)\} \quad , \quad m_P = m_P(c),$$

we get

$$p_1\{n + \frac{1}{2}(m_{1-\alpha} - c)\} \leqq m_{1-\alpha} \quad \text{and} \quad p_2\{n + \frac{1}{2}(m_\beta - c)\} \geqq m_\beta \ .$$

Solving these inequalities we obtain the following <u>approximation to the binomial solution</u>:

The smallest acceptance number satisfying $1 - B(c,n,p_1) \leqq \alpha$ and $B(c,n,p_2) \leqq \beta$ may be found approximately from the inequality

$$R(c-1) > (p_2/p_1)\{1 + \frac{1}{2}(p_2 - p_1)\} \geqq R(c)$$

and the corresponding sample size satisfying

$$(m_\beta/p_2) - \frac{1}{2}(m_\beta - c) \leqq n \leqq (m_{1-\alpha}/p_1) + \frac{1}{2}(c - m_{1-\alpha}) \ .$$

Comparing with the Poisson solution it will be seen that the ratio p_2/p_1 of the two quality levels has been multiplied by the factor $1 + \frac{1}{2}(p_2 - p_1)$ which is of no importance for small values of p_1 and p_2 but for large values it may lead to a smaller value of c than the Poisson solution. The interval for n will normally be larger than in the Poisson case. The solution is just as simple to compute as the Poisson solution.

In the previous paper 27 examples of the determination of the binomial c were given (Table 4) for values of c between 1 and 51. In 17 out of these 27 cases the approximation formula gave the correct value of c and in the remaining 10 cases the error was one unit only. The new formula, which is much simpler than the previous one, leads to the correct result in 25 of the 27 cases and an error of one unit in the remaining 2 cases.

For the hypergeometric distribution we use the approximation $H(c,n,Np,N) \simeq G(c,m)$ with $m = (2Np - c)(n - \frac{1}{2}c)/(2N - Np - n + 1)$, see Molenaar (1970, p. 139). By the same procedure as for the binomial we find the following approximation to the hypergeometric solution:

The smallest acceptance number satisfying $1 - H(c,n,Np_1,N) \leqq \alpha$ and $H(c,n,Np_2,N) \leqq \beta$ may be found approximately from the inequality

$$R(c-1) > (p_2/p_1)\{1 + \tfrac{1}{2}(p_2 - p_1) + (c/2N)(p_1^{-1} - p_2^{-1})\} \geqq R(c)$$

and the limits for the corresponding sample size may be found from the expression

$$n = \{(2N+1)m - Np(m-c) - \tfrac{1}{2}c^2\} / \{2Np + m - c\} ,$$

where the lower limit is obtained from (m_β, p_2) and the upper limit from $(m_{1-\alpha}, p_1)$.

Note that c has to be determined iteratively starting from the binomial c. However, only one or two iterations is usually needed. The reader may check this on Example 5 in the previous paper and also convince himself that the new formula is much simpler to use than the previous one.

For $p_2 < 0.2$ and values of (α,β) customarily used the approximation formulas to the binomial and hypergeometric solution will in most cases give the correct value of c and occasionally a value one unit in error. It should be noted that the formulas are self-correcting in the sense that if the value found for c is too small then the lower limit for n becomes larger than the upper limit thus indicating that c should be increased by 1. Of course, for c = 0 or 1 and for small values of N it is advisable to check the solution by means of the exact formulas.

A k-stage sampling plan is characterized by $3k-1$ parameters, viz. the k sample sizes, the k acceptance numbers and the $k-1$ rejection numbers. We shall denote these parameters by (n_1,b), where n_1 is the size of the first sample and b is a vector of dimension $3k-2$ consisting of the ratios of the $k-1$ sample sizes to n_1 and the $2k-1$ acceptance and rejection numbers. With this notation the OC function for the multiple sampling plan computed from the Poisson distribution may be written as $P_k(\lambda) = G_k(b,n_1\lambda)$, say, corresponding to $P(\lambda) = G(c,n\lambda)$ for single sampling. The auxiliary function $v_P(b)$, say, is defined as solution to the equation $G_k(b,v_P) = P$, $0 \leq P \leq 1$, and $R_k(b;\alpha,\beta) = v_\beta(b)/v_{1-\alpha}(b)$. Hence, a multiple sampling plan of strength $(\lambda_1,\alpha,\lambda_2,\beta)$ must satisfy the two equations

$$R_k(b) = \lambda_2 / \lambda_1 \quad \text{and} \quad n_1 = v_{1-\alpha}(b) / \lambda_1 = v_\beta(b) / \lambda_2 \ .$$

Suppose now that we have a collection of multiple sampling plans characterized by $b_1, b_2, \ldots$ and ordered so that $R_k(b_1) > R_k(b_2) > \ldots$. Amomg these plans the plan having a producer's risk smaller than or equal to α and a consumer's risk smaller than or equal to β and satisfying the risk requirements as nearly as possible is uniquely determined by the inequality $R_k(b_{i-1}) > \lambda_2/\lambda_1 \geqq R_k(b_i)$ and the first sample size satisfies the inequality $v_\beta(b_i)/\lambda_2 \leqq n_1 \leqq v_{1-\alpha}(b_i)/\lambda_1$.

In another paper (Hald, 1975) it has been shown that the approximate relationship between the binomial and the Poisson OC fractiles used above for single sampling is valid also for multiple sampling. Hence, the binomial multiple sampling plan may be found from the Poisson table of multiple sampling plans by a procedure analogous to the one for single sampling. It is conjectured that the same is true also in the hypergeometric case.

REFERENCES.

Hald, A. (1967): The determination of single sampling attribute plans with given producer's and consumer's risk. Technometrics, 9, 401 - 415.

Hald, A. (1975): An approximation to the binomial OC by means of the Poisson OC for multiple sampling plans. Preprint No. , 1975, Institute of Mathematical Statistics, University of Copenhagen.

Molenaar, W. (1970): Approximations to the Poisson, binomial and hypergeometric distribution functions. Mathematical Centre Tracts, No. 31. Amsterdam.