Zdzisław Rychlik; Dominik Szynal

Inflated truncated negative binomial acceptance sampling plan

Aplikace matematiky, Vol. 22 (1977), No. 3, 157–165

Persistent URL: http://dml.cz/dmlcz/103688

INFLATED TRUNCATED NEGATIVE BINOMIAL ACCEPTANCE SAMPLING PLAN

ZDZISŁAW RYCHLIK and DOMINIK SZYNAL

(Received February 28, 1973)

1. Introduction and preliminaries. In a great number of problems of quality control, the rejection or the acceptance criteria depends on finding the Kth defective or Kth nondefective of items in a sample respectively. Some sampling plans, to solve these problems of quality control, assume that the number of inspected items is a constant (see e.g. [6], [7], [8]) while others assume it is a random variable ([1], [3], [4]). To reach decisions with respect to acceptance or rejection of a lot, single-stage as well as many-stage sampling plans are used. For instance in [8] are given maximum likelihood estimates of the fraction defective (p) under the following three sampling plans:

1°. Inspect a random sample of n units from the lot. Accept the lot if there are fewer than k defectives. Reject the lot if there are k or more defectives.

2°. Inspect randomly selected units of the lot one at a time until either k defectives have been observed or until n units have been inspected. Reject the lot if k defectives are observed. Accept the lot if n units are inspected, provided that the number of defectives observed is less than k.

3°. Inspect randomly selected units of the lot one at a time until either k defectives or $n - k + 1$ nondefectives have been observed. Reject the lot if there are k defectives.

In all these plans k and n are predetermined numbers. In general, k will be much less than n.

Considerations of this note are based on inspections from a sequence of m lots of inspected items and our curtailed sampling plans are similar to those in [8] (the same as in [1]). Thus, we deal with an attribute acceptance plan in which randomly selected individual units from a lot are inspected in sequence until either:

1. an accumulated total of k defectives is found, in which case the lot is rejected, or until

2. an accumulated total of K nondefectives is found, in which case the lot is accepted.

In such sampling plans, the number Y of inspected items is a random variable, taking the values k, $k + 1, \ldots, n$, where

$$n = k + K - 1 \,. \tag{1}$$

In [1] the probability p of selecting a defective in a single trial is assumed constant from trial to trial and trials are assumed stochastically independent. Under those conditions the joint probability $[Y = y]$, and that a lot will be rejected and the joint probability that $[Y = y]$, and that a lot will be accepted are probability functions of one-parameter negative binomial distributions [1], i.e.

$$f(y \cap R; p) = \binom{y-1}{k-1} p^k q^{y-k} \,, \quad y = k \,, k + 1, \ldots, n \,, \tag{2}$$

and

$$f(y \cap A; p) = \binom{y-1}{K-1} q^K p^{y-K} \,, \quad y = K \,, K + 1, \ldots, n \,, \tag{3}$$

where A and R are events denoting acceptation and rejecting sample respectively.

The above given distribution of the random variable Y can be applied in the case when the quality control was accomplished directly after the production cycle or when the article was not subjected to deterioration. However, in practical applications it needs maximum likelihood estimation of the fraction defective when the quality control is carried out after a certain period of time or after transport, during which the number of defectives can increase.

Thus, in such a case, the maximum likelihood estimate of the fraction defective is not well described by the formulas given in [1] and [8]. In this case the number of defectives in a sample is better described by the so-called inflated negative binomial distribution introduced in this note.

Let us note here that examples of phenomena well described by the inflated binomial or Poisson distributions may be found in [5], [9], [10].

Our extensions of attribute acceptance plans considered in [1] go in two directions:

(a) p is a value of the random variable P with distribution function $F(p)$,

(b) the random variable Y is being well described by functions of probability functions of negative binomial distributions except for k which is inflated, that is, there are more observations than can be expected on the basis of the distribution (2).

The aim of this paper is to give maximum likelihood estimations of the process (or lot) average proportion (p) of defectives and the proportion (α) of the population which follows the negative binomial distribution based on attribute samples that have been curtailed either with the rejection of a lot on finding the kth defective or with the acceptance of it on finding the Kth nondefective. The maximum likelihood estimations, $\hat{p}$ and $\hat{\alpha}$ are based on inspections from a sequnece of m lots of the inspected items. Using the method of linearization, we obtain linear estimates of p and α. Moreover, we find the asymptotic variance and covariance of considered estimators.

2. The Probability Function of the Random Variable Y. The probability function of Y, the number of inspected items, may be expressed as

$$f(y;p,\alpha)=f(y\cap R;p,\alpha)+f(y\cap A;p,\alpha)\,, \tag{4}$$

where A, R and p are as above, and α, $0<\alpha\leqq 1$, denotes the proportion of the population which follows the negative binomial distribution.

Under the conditions (a) and (b)

$$f(y\cap R;p,\alpha)=\begin{cases}1-\alpha+\alpha p^k\,, & y=k\,,\\ \alpha\binom{y-1}{k-1}p^kq^{y-k}\,, & y=k+1,\ldots,n\,,\end{cases} \tag{5}$$

$$f(y\cap A;p,\alpha)=\alpha\binom{y-1}{K-1}q^Kp^{y-K}\,,\quad y=K\,,\quad K+1,\ldots,n\,, \tag{6}$$

where here, and in what follows, $q=1-p$.

But, by (4)

$$f(y;p,\alpha)=\begin{cases}f(y\cap R;p,\alpha)\,, & y=k\,,\quad k+1,\ldots,K-1\,,\\ f(y\cap R;p,\alpha)+f(y\cap A;p,\alpha)\,, & y=K,\ldots,n\,,\end{cases}$$

so that

$$f(y;p,\alpha)=\begin{cases}1-\alpha+\alpha p^k\,, & y=k\,,\\ \alpha\binom{y-1}{k-1}p^kq^{y-k}\,, & y=k+1,\ldots,K-1\,,\\ \alpha\binom{y-1}{k-1}p^kq^{y-k}+\alpha\binom{y-1}{K-1}q^Kp^{y-K}\,, & y=K,\ldots,n\,,\\ 0 \text{ elsewhere.}\end{cases} \tag{7}$$

If $\alpha=1$, the above distribution reduces to the one considered by [1].

From (7) we have

$$f(y;\alpha)=\begin{cases}1-\alpha+\alpha\int_0^1 p^k\,\mathrm{d}F(p)\,, \quad y=k\,,\\ \alpha\binom{y-1}{k-1}\int_0^1 p^kq^{y-k}\,\mathrm{d}F(p)\,, \quad y=k+1,\ldots,K-1\,,\\ \alpha\binom{y-1}{k-1}\int_0^1 p^kq^{y-k}\,\mathrm{d}F(p)+\alpha\binom{y-1}{K-1}\int_0^1 q^Kp^{y-K}\,\mathrm{d}F(p)\,,\\ \quad y=K,\ldots,n\,.\end{cases} \tag{8}$$

On the basis of (5), the probability of rejecting a lot is given by

$$P(R) = 1 - \alpha + \alpha \int_0^1 p^k \, \mathrm{d}F(p) + \alpha \sum_{y=k+1}^{n} \binom{y-1}{k-1} \int_0^1 p^k q^{y-k} \, \mathrm{d}F(p) =$$

$$= 1 - \alpha + \alpha \sum_{y=k}^{n} \binom{y-1}{k-1} \int_0^1 p^k q^{y-k} \, \mathrm{d}F(p) .$$

Using the formula

$$\sum_{z=k}^{n} \binom{n}{z} p^z q^{n-z} = \sum_{y=k}^{n} \binom{y-1}{k-1} p^k q^{y-k} \quad ([6]-[7]) ,$$

we obtain

$$P(R) = 1 - \alpha + \alpha \sum_{z=k}^{n} \binom{n}{z} \int_0^1 p^z q^{n-z} \, \mathrm{d}F(p) .$$

Let us consider some particular cases:

(i) If P is uniformly distributed, i.e.

$$g(p) = \begin{cases} 1 & \text{if} \quad 0 < p < 1 , \\ 0 & \text{if} \quad p \leqq 0 , \quad p \geqq 1 , \end{cases}$$

then

$$(9) \qquad f(y, \alpha) = \begin{cases} 1 - \alpha + \alpha/(k+1) , & y = k , \\ \alpha k / y(y+1) , & y = k+1 , \quad k+2, \ldots, K-1 , \\ \alpha(k+K)/y(y+1) , & y = K , \quad K+1, \ldots, n , \end{cases}$$

and

$$P(R) = 1 - \alpha + \alpha(n-k)/(n+1) .$$

(ii) If P is beta distributed, i.e.

$$g(p) = \begin{cases} \dfrac{1}{B(a, b)} p^{a-1}(1-p)^{b-1} & \text{if} \quad 0 < p < 1 , \quad a > 0 , \quad b > 0 , \\ 0 \quad \text{if} \quad p \leqq 0 , \quad p \geqq 1 , \end{cases}$$

where

$$B(a, b) = \int_0^1 p^{a-1}(1-p)^{b-1} \, \mathrm{d}p ,$$

then

$$(10) \; f(y, \alpha) = \begin{cases} 1 - \alpha + \alpha \, B(a+k, b)/B(a, b) , \quad y = k , \\ \alpha \dbinom{y-1}{k-1} B(a+k, b+y-k)/B(a, b) , \quad y = k+1, \ldots, K-1 , \\ \alpha \dbinom{y-1}{k-1} B(a+k, b+y-k)/B(a, b) + \\ \quad + \alpha \dbinom{y-1}{K-1} B(a-K+y, b+K)/B(a, b) , \quad y = K, \ldots, n , \end{cases}$$

and

$$P(R) = 1 - \alpha + \alpha \sum_{z=k}^{n} \binom{n}{z} B(a + z, b + n - z)/B(a, b).$$

3. Average Sample Size. In the case, when p is fixed, we have

$$\mathrm{E}Y = (1 - \alpha + \alpha p^k) k + \alpha \sum_{y=k+1}^{n} y \binom{y-1}{k-1} p^k q^{y-k} + \alpha \sum_{y=K}^{n} y \binom{y-1}{K-1} q^K p^{y-K}.$$

Hence we obtain

$$\mathrm{E}Y = (1 - \alpha) k + \alpha \left\{ \sum_{y=k}^{n} y \binom{y-1}{k-1} p^k q^{y-k} + \sum_{y=K}^{n} y \binom{y-1}{K-1} q^K p^{y-K} \right\} =$$

$$= (1 - \alpha) k + \alpha \frac{k}{p} \left[1 - S(p, n+1, k) + \alpha \frac{K}{q} S(p, n+1, k-1) \right],$$

where $n = K + k - 1$, and

$$S(p, n, k) = \sum_{x=0}^{k} \binom{n}{x} p^x q^{n-x}.$$

Putting $\alpha = 1$, we get the value given in [1].

In the case, when (8) takes place, we obtain

$$\mathrm{E}Y = (1 - \alpha) k + \alpha k \int_0^1 p^{-1} [1 - S(p, n+1, k)] \, \mathrm{d}F(P) +$$

$$+ \alpha K \int_0^1 q^{-1} S(p, n+1, k-1) \, \mathrm{d}F(p).$$

When P has uniform or beta distribution, we obtain, accordingly to (9) and (10),

$$\mathrm{E}Y = (1 - \alpha) k + \alpha k \sum_{y=k}^{n} 1/(y+1) + \alpha K \sum_{y=K}^{n} 1/(y+1)$$

and

$$\mathrm{E}Y = (1 - \alpha) k + \alpha k \sum_{y=k}^{n} \binom{y}{k} B(a + k, b + y - k)/B(a, b) +$$

$$+ \alpha K \sum_{y=K}^{n} \binom{y}{K} B(a - K + y, b + K)/B(a, b) \quad \text{respectively.}$$

4. Estimation of the Parameters α and p. Let us suppose, similarly to [1], that m lots have been subjected to inspection in accordance with the curtailed plan described above. Let r_0 be the number of the lots that were accepted, r_1 be the number of the lots whose sample results contained defectives and nondefectives, and let r_2 be the number of lots whose sample results contained only defectives, so that

$$m = r_0 + r_1 + r_2 .$$

Let the number of defectives found and the number of items inspected be recorded for each lot. The sample data then consist of paired values (z_1, y_1), (z_2, y_2), ..., (z_{r_0}, y_{r_0}), (k, y_{r_0+1}), (k, y_{r_0+2}), ..., $(k, y_{r_0+r_1})$, $(k, y_{r_0+r_1+1} = k)$, ..., $(k, y_{r_0+r_1+r_2} = k)$, where z_i $(i = 1, 2, \ldots, r_0)$ is the number of defectives found in the ith accepted lot $(z_i < k)$ and k, of course, is the number of defectives found in each rejected lot and $y_l(l = 1, 2, \ldots, r_0 + r_1 + r_2)$ is the number of inspected items. Thus z is a constant equal to k for indices $r_0 + 1$ to $r_0 + r_1 + r_2$ and y is a constant also equal to k for indices $r_0 + r_1 + 1$ to $r_0 + r_1 + r_2$.

The likelihood function for such a sample is given by

$$L[(z_1, y_1), \ldots, (z_m, y_m)] = \alpha^{r_0+r_1}(1 - \alpha + \alpha p^k)^{r_2} \prod_{j=1}^{r_0} \binom{y_j - 1}{K - 1} q^K p^{y_j - K} \prod_{i=r_0+1}^{r_1} \binom{y_i - 1}{k - 1} p^k q^{y_i - k} . \tag{11}$$

Taking logarithms of (11), differentiating it with respect to α and p, equating to zero, and solving the resulting equations for α and p, we obtain

$$\hat{\alpha} = (m - r_2)/m(1 - \hat{p}^k) , \tag{12}$$

$$\hat{p} = \frac{\sum_{j=1}^{r_0} y_j - (K + k) r_0 + k(m - r_2)/(1 - \hat{p}^k)}{\sum_{i=1}^{m} y_i - km + k(m - r_2)/(1 - \hat{p}^k) .} \tag{13}$$

The maximum likelihood estimates obtained from (12) and (13) are not linear in α and p. Thus, there is some trouble with their calculation. In order to get the pilot estimates $\bar{\alpha}$ and $\bar{p}$ of α and p respectively, we can make a linearization similarly to [10]. Pilot estimations $\bar{\alpha}$ and $\bar{p}$ of α and p can be obtained from the following equations

$$f_k = 1 - \alpha + \alpha p^k , \tag{14}$$

$$f_{k+1} = \alpha k p^k q , \tag{15}$$

where f_k and f_{k+1} are observed relative frequencies for k defectives, and k defectives and one nondefective respectively in a sample of size n, while $1 - \alpha + \alpha p^k$ and $\alpha k p^k q$ are the respective probabilities of k defectives, and k defectives and one nondefective, obtained from (5).

Eliminating α from (14) and (15), we have

$$f_{k+1}(1 - p^k) = (1 - f_k)(1 - p) k p^k . \tag{16}$$

Hence

(17) $$1 - p^k = (1 - f_k)(1 - p)\,k/[(1 - f_k)(1 - p)\,k + f_{k+1}]\,.$$

Putting (17) into (12) and (13), we obtain linear equations for estimates $\bar{\alpha}$ and $\bar{p}$ of α and p respectively

(18) $$\bar{\alpha} = (m - r_2)\,[f_{k+1} + k(1 - f_k)(1 - p)]/mk(1 - f_k)(1 - \bar{p})\,,$$

and

(19)
$$\bar{p} = \frac{\sum_{j=1}^{r_0} y_j - (K + k)\,r_0 + (m - r_2)\,[f_{k+1} + k(1 - f_k)(1 - \bar{p})]/(1 - f_k)(1 - \bar{p})}{\sum_{i=1}^{m} y_i - km + (m - r_2)\,[f_{k+1} + k(1 - f_k)(1 - \bar{p})/(1 - f_k)(1 - \bar{p})}\,.$$

Hence

(20) $$\bar{\alpha} = \frac{(m - r_2)\,[(\sum_{i=1}^{m} y_i - km)\,f_{k+1} + k(1 - f_k)\,(\sum_{i=r_0+1}^{r_1} y_i + Kr_0 - kr_1)]}{mk[(1 - f_k)\,(\sum_{i=1}^{m} y_i + Kr_0 - kr_1) - (m - r_2)\,f_{k+1}]}\,,$$

(21) $$\bar{p} = \frac{(1 - f_k)\,(\sum_{j=1}^{r_0} y_j - Kr_0 + kr_1) + (r_0 + r_1)\,f_{k+1}}{(1 - f_k)\,(\sum_{i=1}^{m} y_i - kr_2)}\,,$$

The asymptotic variance and covariance of the estimates $\hat{\alpha}$ and $\hat{p}$ are given by the matrix $\mathbf{M}$ [2], where

$$\mathbf{M} = \begin{bmatrix} -\mathrm{E}\left(\dfrac{\partial^2 \log L}{\partial \alpha^2}\right), & -\mathrm{E}\left(\dfrac{\partial^2 \log L}{\partial \alpha\, \partial p}\right) \\ -\mathrm{E}\left(\dfrac{\partial^2 \log L}{\partial \alpha\, \partial p}\right), & -\mathrm{E}\left(\dfrac{\partial^2 \log L}{\partial p^2}\right) \end{bmatrix}^{-1} = \begin{bmatrix} \sigma^2(\hat{\alpha}), \operatorname{Cov}(\hat{\alpha}, \hat{p}) \\ \operatorname{Cov}(\hat{\alpha}, \hat{p}), \sigma^2(\hat{p}) \end{bmatrix}.$$

Taking into account the equalities

$$\partial^2 \log L/\partial \alpha^2 = -m(1 - \hat{p}^k)/\hat{\alpha}(1 - \hat{\alpha} + \hat{\alpha}\hat{p}^k)\,,$$

$$\partial^2 \log L/\partial \alpha\, \partial p = mk\hat{p}^{k-1}/(1 - \hat{\alpha} + \hat{\alpha}\hat{p}^k)\,,$$

$$\partial^2 \log L/\partial p^2 = m\hat{\alpha}k^2\hat{p}^{k-2}(1 - \hat{\alpha})/(1 - \hat{\alpha} + \hat{\alpha}\hat{p}^k) - \sum_{i=1}^{m} y_i - mk(1 - \hat{\alpha})/\hat{p}\hat{q}\,,$$

we have

$$\sigma^2(\hat{\alpha}) = \hat{\alpha}\{[\mathrm{E}Y - k(1 - \hat{\alpha})]\,(1 - \hat{\alpha} + \hat{\alpha}\hat{p}^k) - k^2\hat{\alpha}(1 - \hat{\alpha})(1 - \hat{p})\,\hat{p}^{k-1}\}/\Delta$$

$$\operatorname{Cov}(\hat{\alpha}, \hat{p}) = k\hat{\alpha}(1 - \hat{p})\,\hat{p}^k/\Delta\,,$$

and

$$\sigma^2(\hat{p}) = \hat{p}(1 - \hat{p})(1 - \hat{p}^k)/\Delta ,$$

where

$$\Delta = m\{[EY - k(1 - \hat{\alpha})](1 - \hat{p}^k) - k^2\hat{\alpha}(1 - \hat{p})\hat{p}^{k-1}\} .$$

The mathematical expectation EY was given in Section 3. However, for a sufficiently large number of lots, the mean of the observed values Y should provide a reasonable approximation to EY. Thus, in this case, we have

$$\sigma^2(\hat{\alpha}) \approx \hat{\alpha}\{[\sum_{i=1}^{m} y_i - mk(1 - \hat{\alpha})](1 - \hat{\alpha} + \hat{\alpha}\hat{p}^k) - k^2\hat{\alpha}(1 - \hat{\alpha})(1 - \hat{p})\hat{p}^{k-1}\}/m\Delta_1 ,$$

$$\mathrm{Cov}(\hat{\alpha}, \hat{p}) \approx k\hat{\alpha}(1 - \hat{p})\hat{p}^k/\Delta_1$$

and

$$\sigma^2(\hat{p}) \approx \hat{p}(1 - \hat{p})(1 - \hat{p}^k)/\Delta_1 ,$$

where

$$\Delta_1 = [\sum_{i=1}^{m} y_i - mk(1 - \hat{\alpha})](1 - \hat{p}^k) - mk^2\hat{\alpha}(1 - \hat{p})\hat{p}^{k-1} .$$

After using the linearization (14) and (15), we get

$$\sigma^2(\bar{\alpha}) = \bar{\alpha}^2\{[EY - k(1 - \bar{\alpha})]\bar{p}f_k - k(1 - \bar{\alpha})f_{k+1}\}/\Delta_2 ,$$

$$\mathrm{Cov}(\bar{\alpha}, \bar{p}) = \bar{\alpha}\bar{p}f_{k+1}/\Delta_2 , \quad \sigma^2(\bar{p}) = \bar{p}^2\bar{q}(1 - f_k)/\Delta_2 ,$$

where

$$\Delta_2 = m\{[EY - k(1 - \bar{\alpha})]\bar{p}(1 - f_k) - k\bar{\alpha}f_{k+1}\} .$$

References

[1] *C. A. Cohen:* Curtailed attribute sampling. Technometrics 12 (1970), 295—298.

[2] *H. Cramér:* Mathematical methods of statistics. Princeton Univ. Press, Princeton 1946.

[3] *C. C. Craig:* The ASN for truncated single and double attribute acceptance sampling plans. ASQC Convention Transactions (1968), 63—67.

[4] *C. C. Craig:* The average sample number for truncated single and double attribute sampling plans. Technometrics 10 (1968), 685—692.

[5] *K. N. Pandey:* On generalized inflated Poisson distribution. J. Scient. Res. Banaras Hindu Univ. 15 (1964—1965), 157—162.

[6] *G. P. Patil:* On the evaluation of the negative binomial distribution with example. Technometrics 2 (1960), 501—505.

[7] *G. P. Patil:* Note on the equivalence of the binomial and inverse binomial acceptance sampling plans and an acknowledgement. Technometrics 5 (1963), 119—121.

[8] *A. G. Phatak, M. M. Bhatt:* Estimation of fraction defective in curtailed sampling plans. Technometrics 9 (1967), 219—228.

[9] *M. P. Singh:* Inflated binomial distribution. J. Scient. Res. Banaras Hindu Univ. 16 (1965 to 1966), 87—90.

[10] *S. N. Singh:* Probability models for the variation in the number of births per couple. J. Am. Statist. Ass. 58 (1972), 721—727.

Souhrn

VÝBĚROVÝ PŘEJÍMACÍ PLÁN, ZALOŽENÝ NA „ROZŠÍŘENÉM" USEKNUTÉM NEGATIVNÍM BINOMICKÉM ROZDĚLENÍ

ZDISŁAW RYCHLIK a DOMINIK SZYNAL

Článek obsahuje maximálně věrohodné odhady průměrného podílu p zmetků ve výrobním procesu (nebo dodávce) a podílu α základního souboru, který má negativní binomické rozdělení.

Odhady jsou založeny na výběrové kontrole srovnáváním z posloupnosti m dodávek předložených ke kontrole při cenzorování výběru buď zamítnutím dodávky při výskytu k-tého zmetku nebo přijetím dodávky při výskytu K-tého vyhovujícího kusu ve výběru. Dále jsou uvedeny lineární odhady podílů p a α a příslušné asymptotické variance a kovariance.

Authors' address: Dr. *Zdzisław Rychlik*, Doc. dr. *Dominik Szynal*, Institute of Mathematics, Maria Curie-Skłodowska University, 20-031 Lublin, ul. Nowotki 10.