EFFECT OF NON-NORMALITY ON PLANS FOR SAMPLING INSPECTION BY VARIABLES

By NANI GOPAL DAS

St. Xavier's College, Calcutta

and

SUJIT KUMAR MITRA

Indian Statistical Institute

SUMMARY. This paper deals with a quantitative study on the effects of non-normality on the operating characteristcs of published variables sampling inspection plans.

1. Introduction

All published plans for sampling inspection by variables assume a normal distribution for the variable item quality. Although much have been said in literature in defense of this assumption in the present context, no systematic attempt seems ever to have been made to assess the effect of non-normality in quantitative terms.

The first few terms of a Gram-Charlier series have provided a convenient alternative to the normal distribution in several studies on 'non-normality'. In the present paper, probabilities of rejection of lots of AQL quality and of acceptance of lots of LTPD quality (known as the producer's and the consumer's risks respectively) are obtained, for some typical variables sampling inspection plans, assuming for the variable item quality a non-normal distribution of this type. A comparison of these values with the values assured by the table of plans (see for example Eisenhart, Hastay and Wallis, 1947,) shows the extent to which the operating characteristic of plan may be sensitive to the assumption of normality.

2. The 'modus operandi' of a variables plan

We consider the situation where the item quality, specified in terms of an upper specification limit (U), is a variable while the quality of the lot is expressed as a proportion defective. The AQL and LTPD values are denoted by p_1 and p_2 ($p_1 < p_2$), and the producer's and the consumer's risks by α and β respectively.

To determine the acceptability of a lot, we measure the item quality x in a sample of n items, and compute its mean $\bar{x}$ and standard deviation s. The lot is accepted if $\bar{x}+k\,s \leqslant U$, and rejected, if otherwise.

A table of variables plan provides values of n and k that are required to be taken to meet given specifications on p_1, p_2, α and β.

The following plans, selected from Eisenhart, Hastay and Wallis (1947), are included in the present investigation. It is felt that they provide a fair cross-section of the range of values of p_1 and p_2 that are likely to be met in practice.

	$\alpha = 0.05,$		$\beta = 0.10$	
plan	p_1	p_2	table values n	table values k
I	0.001	0.0025	507	2.9311
II		0.005	161	2.8018
III		0.01	68	2.6725
IV	0.01	0.02	389	2.1733
V		0.03	137	2.0701
VI		0.05	54	1.9527
VII	0.05	0.07	601	1.5409
VIII		0.10	133	1.4408
IX		0.15	44	1.3120

3. Fractiles of a Gram-Charlier series

Consider the following density function for the distribution of a variable v $(-\infty < v < \infty)$

$$\psi(v) = \phi(v)\Big[1+\frac{\sqrt{\beta_1}}{6}\; H_3(v)+\frac{\beta_2-3}{24}\; H_4(v)\;\Big] \qquad \ldots \quad (3.1)$$

where $\phi(v)$ is the standard normal density function, $H_3(v)$ and $H_4(v)$ are the Hermite polynomials of degrees 3 and 4 respectively and β_1, β_2 have their usual interpretation. The probability integral is given by

$$\Psi(v) = \int_{-\infty}^{v} \psi(t)dt = \Phi(v)-\phi(v)\Big[\frac{\sqrt{\beta_1}}{6}\; H_2(v)+\frac{\beta_2-3}{24}\, H_3(v)\Big] \qquad (3.2)$$

where $\Phi(v)$ is the probability integral of the standard normal. The p-fractile v_p of v is defined by the equation

$$\Psi(v_p) = p. \qquad \ldots \quad (3.3)$$

4. The OC curve

Let us now assume that for the item quality x, the distribution of the standardised variable $v = (x-\mu) \div \sigma$ is given by (3.1). μ and σ are respectively the process average and the standard deviation of x. For a specified value p of the proportion defective, the values of U, μ and σ are connected by means of the equation

$$U = \mu+v_{1-p}\, \sigma. \qquad \ldots \quad (4.1)$$

Since $\bar{x}+ks$ has asymptotically a normal distribution with mean $\mu+k\sigma$ and standard error

$$\sigma\sqrt{\frac{1}{n}\Big\{1+k^2\Big(\frac{\beta_2-3}{4}\Big)+k\sqrt{\beta_1}\Big\}+\frac{k^2}{2(n-1)}} \qquad \ldots \quad (4.2)$$

the probability of acceptance of a lot containing a proportion p of defectives can be approximately evaluated as

$$L(p) = \Phi\left[(v_{1-p}-k)\Big/\sqrt{\frac{1}{n}\left\{1+k^2\left(\frac{\beta_2-3}{4}\right)+k\sqrt{\beta_1}\right\}+\frac{k^2}{2(n-1)}}\,\right]. \qquad \ldots \quad (4.3)$$

The inaccuracy introduced in this formula, for finite values of n, is under investigation. However, for the range of values of p_1, p_2 and n considered in this paper, the error would perhaps be of minor importance, as may be seen in the special case of $\beta_1 = 0$, $\beta_2 = 3$.

5. Calculation of producer's and consumer's risks

Computation of $L(p)$ in (4.3) presupposes a knowledge of v_{1-p}. The following formulae, connecting v_p with λ_p, the p-fractile of the standard normal distribution, derived on the lines of Cornish and Fisher (1937), are useful in this connection. The subscripts are dropped throughout for simplicity.

Case I : $\beta_1 = 0$, $d = (\beta_2-3) \div 24$.

$$v = \lambda + \sum_{i=1}^{4} d^i P_{1i}(\lambda) \qquad \ldots \quad (5.1)$$

where $P_{11}(\lambda) = \lambda^3 - 3\lambda$

$$P_{12}(\lambda) = -\frac{1}{2}(\lambda^7 - 12\lambda^5 + 33\lambda^3 - 18\lambda)$$

$$P_{13}(\lambda) = \frac{1}{6}(2\lambda^{11} - 47\lambda^9 + 333\lambda^7 - 891\lambda^5 + 837\lambda^3 - 162\lambda)$$

$$P_{14}(\lambda) = -\frac{1}{8}(2\lambda^{15} - 73\lambda^{13} + 904\lambda^{11} - 4990\lambda^9 + 13158\lambda^7 - 15849\lambda^5 + 7128\lambda^3 - 648\lambda)$$

Case II : $\beta_2 = 3$, $c = \sqrt{\beta_1} \div 6$.

$$v = \lambda + \sum_{i=1}^{4} c^i P_{2i}(\lambda) \qquad \ldots \quad (5.2)$$

where

$$P_{21}(\lambda) = \lambda^2 - 1$$

$$P_{22}(\lambda) = -\frac{1}{2}(\lambda^5 - 6\lambda^3 + 5\lambda)$$

$$P_{23}(\lambda) = \frac{1}{3}(\lambda^8 - 13\lambda^6 + 30\lambda^4 - 31\lambda^2 + 4)$$

$$P_{24}(\lambda) = -\frac{1}{24}(6\lambda^{11} - 127\lambda^9 + 760\lambda^7 - 1626\lambda^5 + 1258\lambda^3 - 271\lambda)$$

$$P_{23}(\lambda) = \frac{1}{30}(6\lambda^{14}-181\lambda^{12}+1723\lambda^{10}-6760\lambda^{8}+11780\lambda^{6}-8941\lambda^{4}$$
$$+2401\lambda^{2}-118)$$

$$P_{26}(\lambda) = -\frac{1}{720}(120\lambda^{17}-4778\lambda^{15}+64343\lambda^{13}-388404\lambda^{11}+1158505\lambda^{9}$$
$$-1755370\lambda^{7}+1313673\lambda^{5}-420848\lambda^{3}+41759\lambda).$$

The coefficients $P_{1i}(\lambda)$ in (5.1) and $P_{2i}(\lambda)$ in (5.2) are tabulated in Table 1, for select values of p. For the range of values of β_1, β_2 and p considered in this paper, the formulae (5.1) and (5.2) are seen to provide values of v_p to reasonable degree of accuracy (with at most an error of 1 in the third place of decimal) except for the following few cases :

		values of p			
		0.001	0.0025	0.005	0.01
Case I		$\beta_2>3.2$	$\beta_2>3.2$	$\beta_2>3.4$	$\beta_2>3.6$
Case II	$\mu_3>0$	$\beta_1>0.05$	$\beta_1>0.05$	$\beta_1>0.1$	$\beta_1>0.2$
	$\mu_3<0$	$\beta_1>0.01$	$\beta_1>0.05$	$\beta_1>0.1$	$\beta_1>0.1$

In these cases the values of v_p are obtained by inverse interpolation in a table of $\Psi(v)$ which is easily constructed using a table of the normal probability integral and its derivatives (Harvard University, 1952).

Values of producer's risk $1-L(p_1)$ and of consumer's risk $L(p_2)$, obtained by using formula (4.3), are shown in Tables 2 and 3 for Case I and Case II respectively. The nominal values of α and β are $\alpha = 0.05$ and $\beta = 0.10$ in each case. In each table a 'check' column is introduced which corresponds to a normal distribution. Deviations noticed in the values of α and β in this column, from the nominal values, are ascribable to the inherent error in the asymptotic formula (4.3) for finite values of n.

6. Concluding remarks

The following observations are based on Tables 2 and 3. It is seen that

(1) The effect of asymmetry on both α and β, has been in general more severe than that of kurtosis. This confirms similar findings of Pearson (1929) and Gayen (1949), in a different connection.

(2) Distortions have been generally more marked, when the plan requires a large sample size and also when lot specifications are strict (as indicated by low values for p_1 and p_2).

(3) A positively skew distribution for the item quality favours the producer at the expense of the consumer, while reverse is the case with negative skewness.

TABLE I. VALUES OF $P_{1i}(\lambda)$ AND $P_{2i}(\lambda)$

$p =$	0.001	0.0025	0.005	0.01	0.02	0.03	0.05	0.07	0.10	0.15
$\lambda =$	3.09023	2.80703	2.57583	2.32635	2.05375	1.88079	1.64485	1.47579	1.28155	1.03643
Case I : $\beta_1 = 0$										
$P_{11}(\lambda)$	20.23058	13.69075	9.30287	5.61091	2.50123	1.01071	−0.48434	−1.21316	−1.73987	−1.09597
$P_{12}(\lambda)$	−113.82529	19.37675	45.37350	37.64483	17.72173	6.73305	−2.67110	−5.37316	−5.29260	−2.60878
$P_{13}(\lambda)$	−7,962.0439	−2,906.1806	−831.0626	8.3849	102.3741	43.2649	−14.5483	−20.8488	−7.1652	10.7130
$P_{14}(\lambda)$	194,224.9027	−11,100.6895	−16,040.3319	−3,014.3309	326.1967	202.5489	−77.7959	−60.8098	43.7341	55.6718
Case II : $\beta_2 = 3$										
$P_{21}(\lambda)$	8.54953	6.87044	5.63400	4.41190	3.21768	2.53739	1.70554	1.17796	0.64238	0.07419
$P_{22}(\lambda)$	−60.00944	−27.80216	−11.80486	−2.11376	2.58430	3.49002	3.21840	2.46205	1.38206	0.15094
$P_{23}(\lambda)$	86.55145	−107.06701	−114.65055	−74.75904	−30.63988	−12.16602	0.67079	3.22525	2.65633	0.30657
$P_{24}(\lambda)$	6,908.7422	3,137.0089	1,120.0062	165.7213	−81.3629	−71.3923	−23.7563	−2.7083	4.0743	0.6212
$P_{25}(\lambda)$	−124,379.80	−16,095.080	2,408.130	2,902.031	640.170	35.289	−80.240	−33.191	2.033	1.255
$P_{26}(\lambda)$	564,020.24	−233,570.87	−102,595.49	−12,329.87	3,180.36	1,611.80	−22.09	−108.21	−10.25	2.53

TABLE 2. VALUES OF PRODUCER'S RISK (α), CONSUMER'S RISK (β)

(Case I : $\beta_1 = 0$)

nominal values: $\alpha = 0.05$, $\beta = 0.10$

plan	p_1	p_2	n	k		$\beta_2 = 2.8$	2.9	2.95	3.0	3.05	3.1	3.2	3.4	3.6	3.8	4.0
I	.001	.0025	507	2.0311	α	.553	.210	.112	.050	.020	.007	.001	.000	.000	.000	.000
					β	.006	.028	.050	.100	.104	.240	.450	.824	.062	.093	.990
II	.001	.005	101	2.8018	α	.243	.110	.078	.050	.031	.019	.007	.001	.000	.000	.000
					β	.037	.062	.079	.099	.122	.149	.212	.371	.640	.054	.784
III	.001	.01	68	2.6725	α	.163	.097	.073	.055	.040	.030	.017	.008	.003	.001	.001
					β	.060	.075	.083	.092	.102	.113	.137	.104	.262	.340	.420
IV	.01	.02	389	2.1733	α	.112	.077	.062	.050	.030	.031	.018	.005	.001	.000	.000
					β	.060	.078	.088	.099	.112	.125	.156	.233	.333	.453	.685
V	.01	.03	137	2.0761	α	.080	.064	.057	.050	.044	.038	.028	.015	.007	.003	.001
					β	.083	.091	.095	.100	.104	.108	.118	.138	.161	.188	.218
VI	.01	.05	54	1.0527	α	.072	.063	.050	.055	.051	.047	.039	.027	.018	.011	.007
					β	.089	.091	.093	.094	.095	.096	.097	.101	.103	.105	.107
VII	.05	.07	601	1.5499	α	.039	.044	.047	.050	.053	.056	.063	.078	.097	.119	.145
					β	.120	.112	.106	.100	.094	.088	.077	.057	.041	.028	.019
VIII	.05	.10	133	1.4408	α	.043	.046	.048	.050	.052	.054	.058	.067	.077	.088	.100
					β	.115	.107	.103	.100	.096	.092	.085	.072	.061	.050	.041
IX	.05	.15	44	1.3120	α	.048	.051	.053	.054	.056	.057	.061	.067	.075	.083	.091
					β	.099	.095	.093	.091	.089	.087	.083	.075	.069	.062	.057

TABLE 3. VALUES OF PRODUCER'S RISK (α), CONSUMER'S RISK (β)

(Case II : $\beta_2 = 3$)

nominal values : $\alpha = 0.05$, $\beta = 0.10$

plan	p_1	p_2	n	k		negative skewness (μ_3)						positive skewness (μ_3)				
					$\beta_1 =$	.3	.2	.1	.05	.01	0	.01	.05	.1	.2	.3
I	.001	.0025	567	2.0311	α	1.000	1.000	1.000	.994	.409	.050	.002	.000	.000	.000	.000
					β	.000	.000	.000	.000	.004	.100	.431	.817	.938	.988	.996
II	.001	.005	101	2.8018	α	1.000	.990	.951	.724	.223	.050	.011	.002	.001	.000	.000
					β	.000	.000	.000	.003	.029	.099	.225	.423	.561	.713	.793
III	.001	.01	68	2.6725	α	.982	.929	.715	.448	.153	.055	.021	.008	.005	.003	.002
					β	.000	.002	.007	.018	.049	.092	.154	.249	.320	.431	.503
IV	.01	.02	380	2.1733	α	.998	.978	.815	.548	.189	.050	.009	.001	.000	.000	.000
					β	.000	.000	.000	.003	.028	.099	.248	.509	.698	.874	.943
V	.01	.03	137	2.0761	α	.836	.683	.433	.268	.115	.050	.020	.006	.003	.001	.000
					β	.001	.002	.009	.021	.054	.100	.166	.276	.373	.614	.810
VI	.01	.05	54	1.9527	α	.401	.375	.240	.165	.092	.055	.032	.017	.011	.006	.004
					β	.009	.016	.030	.044	.069	.094	.124	.169	.208	.269	.321
VII	.05	.07	601	1.5499	α	.783	.624	.393	.250	.113	.050	.019	.004	.001	.000	.000
					β	.000	.001	.005	.015	.047	.100	.184	.334	.470	.667	.800
VIII	.05	.10	133	1.4408	α	.225	.180	.131	.102	.070	.050	.034	.021	.014	.007	.004
					β	.016	.025	.041	.056	.078	.100	.124	.159	.189	.235	.270
IX	.05	.15	44	1.3120	α	.108	.098	.085	.076	.064	.054	.046	.038	.030	.022	.017
					β	.041	.051	.063	.071	.082	.091	.099	.109	.117	.127	.136

(4) Similarly when lot specifications are strict, a symmetrical leptokurtic distribution for the item quality means for the producer a risk lower than what he was willing to accept and for the consumer a risk higher than what he would expect. The situation is however exactly the reverse when the item quality has a symmetrical platykurtic distribution.

Opposite, though milder, effects are noticed in plans with more liberal lot specifications.

Finally, the authors would like to point out that even though the present paper has been prepared in the context of an acceptance sampling plan, the problem could also be viewed as one of testing of hypothesis or of confidence interval estimation [(Jennett and Welch, 1939); (Johnson and Welch, 1940)]. One would, for example, be considering essentially the same problem while testing the hypothesis that the unknown proportion defective p contained in a lot, is equal to p_1 as against the specified alternative $p = p_2$. The producer's and the consumer's risks would in such a case correspond to the probabilities of Type I and Type II errors respectively.

REFERENCES

CORNISH, E. A. and FISHER, R. A. (1937) : Moments and cumulants in the specification of distributions. *Rev. de l' Inst. int. de stat.* 5, 307.

EISENHART, C., HASTAY, M. W. and WALLIS, W. A., editors (1947) : *Techniques of Statistical Analysis*, (Chapter I), McGraw-Hill Book Co. Inc., New York.

GAYEN, A. K. (1949) : The distribution of Student's *t* in random samples of any size drawn from non-normal universes. *Biometrika*, 36, 353.

JENNETT, W. J. and WELCH, B. L. (1939) : The control of proportion defective as judged by a single quality characteristic varying on a continuous scale. *J. Roy. Stat. Soc.*, Supplement, 6, 80.

JOHNSON, N. L. and WELCH, B. L. (1940) : Applications of the non-central *t*-distribution. *Biometrika*, 31, 362.

PEARSON, E. S. (1929) : The distribution of frequency constants in small samples from non-normal symmetrical and skew populations. *Biometrika*, 21, 259.

HARVARD UNIVERSITY, COMPUTATION LABORATORY (1952) : *Tables of the Error Function and of its First Twenty Derivatives*, Annals of the Computation Laboratory of Harvard University, 23, Harvard University Press, Cambridge (Massachusetts).

Paper received : July, 1964.