A NOTE ON THE UNBIASED ESTIMATION OF POWERS OF THE VARIANCE AND ITS APPLICATIONS

NAUNIHAL SINGH

Defence Science Laboratory, Delhi

ABSTRACT

This paper is concerned with the unbiased estimation of integral powers of the variance from a sample taken from a large population. The method characterises the distribution free estimation. It also gives the variance for the estimated powers of the variance together with a few examples wherein the results obtained find their applications.

INTRODUCTION

In certain situations it is of interest, sometimes necessary, to study the problem like— that of estimation of higher integral powers of the variance. In such cases one should consider the estimation of the *r th* power of the variance σ^2 rather than estimating the variance itself and then raising it to the *r th* power. It is essential also in view of the fact that sample estimate of the variance, as well known, is always biased, consequently the higher powers of the estimated variance will involve greater amount of the bias, making it too faulty for all practical purposes. Therefore, a function as discussed in this paper certainly gives an unbiased estimate of the variance as well as for its higher powers. The underlying assumption is that the population mean is either zero or known, for instance suppose we are interested in estimating the radius of a sphere which includes say α % of the probability mass following orthogonal tivariate normal law. The function as will be shown in the text involves higher powers of the variance therefore requires unbiased estimators for them. Similarly the correct estimation of the error function and its complementary would require essentially the unbiased estimators for the higher powers of the variance, as illustrated under *Estimation and Variance of Functions Involving* $\hat{\mu}_2^{\,r}$. Classer[1] has considered an unbiased estimator for powers of the arithematic mean and the present work it based more or less on parallel lines.

UNBIASED ESTIMATOR $\hat{\mu}_2^{\,r}$

This has been obtained with the help of binary symmetric functions which are a special case of symmetric functions typified (Fisher[2]) by

$$\begin{pmatrix} \pi_1 & \pi_2 & & \pi_s \\ p_1 & p_2 & \ldots\ldots\ldots & p \end{pmatrix} = \sum x_i^{p_1} \; x_j^{p_1} \ldots\ldots\ldots x_q^{p_2} \; x_r^{p_2} \ldots\ldots\ldots x_u^{p_s} \; x_v^{p_s} \qquad (1)$$

where summation takes place over all suffixes $i, j, \ldots\ldots\ldots u, v$ which are different

The number $\sum_{j=1}^{s} \pi_j = \overline{v}$, say is called the order of the symmetric function and

$$\sum_{j=1}^{s} \pi_j \, p_j = v,$$ say, is called the weight. There are n/v terms in the summation on the right of (1), where n is the number of possible different suffixes, by putting $p_1 = 2,\ p_2 = p_3 = \ldots\ldots p_s = 0$ and $\pi_1 = n = \gamma$ we get from (1) the binary symmetric function

$$\xi_r = \Sigma\, X_{i1}^{2}\ X_{i2}^{2}\ \ldots\ldots\ldots\ X_{ir}^{2} \tag{2}$$

where summation extends over all permutations of all sets of r observations in a sample of size N ($r<N$). When $x's$ are measured from the mean μ, (2) can be rewritten as

$$\xi_r = \Sigma\, X_{i1}^{2}\ \ X_{i2}^{2}\ \ldots\ldots\ X_{ir}^{2} \tag{3}$$

where $X_i = x_i - \mu$.

The expectation of (2) gives

$$E\left(\xi_r\right)\frac{N\,!}{N-r\,!}\ \mu_2^{\,r}. \tag{4}$$

therefore the unbiased estimator for μ_2^r is

$$\mu_2^{\,r} = \frac{\overline{N-r}\,!}{N\,!}\sum X_{i1}^{2}\ X_{i2}^{2}\ \ldots\ldots\ldots\ X_{ir}^{2} \tag{5}$$

The one part symmetric function or power sums are defined as (David and Kendall[3])

$$S_t = \sum_{i=1}^{N} X_i^{t} \tag{6}$$

for given values of r , the estimator (5) can be expressed in terms of power sums. For $r = 1, 2, 3, 4$ and 6, we have

$$\left.\begin{aligned}
\hat{\mu}_2 &= \frac{S_2}{N} \qquad , \quad \hat{\mu}_2^{\,2} = \frac{1}{N^{[2]}}\left\{S_2^{\,2} - S_4\right\}\\
\hat{\mu}_2^{\,3} &= \frac{1}{N^{[3]}}\left\{S_2^{\,3} - 3S_2\,S_4 + 2S_6\right\}\\
\hat{\mu}_2^{\,4} &= \frac{1}{N^{[4]}}\left\{S_2^{\,4} - 6S_2^{\,2}\,S_4 + 3S_4^{\,2} + 8S_2S_6 - 6S_8\right\}\\
\hat{\mu}_2^{\,6} &= \frac{1}{N^{[6]}}\Big\{S_2^{\,6} - 15\,S_{10}\,S_2 - 180S_8\,S_4 - 45\,S_4^{\,2}\,S_2^{\,2}\\
&\quad - 15\,S_4\,S_2^{\,4} + 360\,S_6^{\,2} + 180S_6\,S_4\,S_2 + 30S_4^{\,3}\\
&\quad + 40S_6S_2^{\,3} - 215\,S_{12}\Big\}
\end{aligned}\right\} \tag{7}$$

VARIANCE OF THE UNBIASED ESTIMATOR $\hat{\mu}_2^r$

$$\text{Var}\ (\hat{\mu}_2^r) = E\ (\hat{\mu}_2^r - \mu_2^r)^2$$

$$= \left\{ \frac{\overline{N-r}\ !}{N\ !} \right\}^2 E \left\{ \sum X_{i1}^2\ X_{i2}^2 \ldots\ldots\ldots X_{ir}^2 \right\}^2 - \mu_2^{2r} \qquad (8)$$

It is evident that

$$\left\{ \sum X_{i1}^2\ X_{i2}^2 \ldots\ldots\ldots X_{ir}^2 \right\}^2 = \sum X_{i1}^2\ X_{i2}^2 \ldots\ldots\ldots X_{i2r}^2,$$

therefore

$$\text{Var}\ (\mu_2^r) = \left\{ \frac{\overline{N-r}\ !}{N\ !} \right\}^2 E \sum_{\beta=0}^{r} \frac{r^2 !}{(\overline{r-\beta}\ !)^2\ \beta\ !}$$

$$\times \left\{ \sum X_{i1}^4\ X_{i2}^4 \ldots\ldots\ldots X_{i\beta}^4\ X_{i\beta+1} \ldots\ldots\ldots X_{i2\gamma-2\beta} \right\} - \mu_2^{2r} \qquad (9)$$

The coefficient $\frac{r^2\ !}{(\gamma-\beta)\ !\ ^2\ \beta\ !}$ arises because there are $\binom{r}{\beta}^2$ combinations of β pairs of the equal subscripts, $\binom{r}{\beta}$ from each of the set of V and then must be multiplied by $\beta\,!$, the number of ways in which the β pairs may be formed. Then

$$\text{Var} \left(\hat{\mu}_2^r \right) = \left\{ \frac{\overline{N-r}\ !}{N\ !} \right\}^2 \sum_{\beta=0}^{\gamma} \frac{r\ !^2\ N\ !\ \mu_4^{\beta}\ \mu_2^{2r-2\beta}}{\{ r-\beta\ ! \}^2\ \beta\ !\ (N-2r+\beta)\ !} - \mu_2^{2r}$$

$$= \sum_{\beta=0}^{r} \frac{\binom{r}{\beta} \binom{N-r}{r-\beta} \mu_4^{\beta}\ \mu_2^{2r-2\beta}}{\binom{N}{r}} - \mu_2^{2r} \qquad (10)$$

$(N \geq 2r)$

for $r = 1, 2, 3$ and 4 we have

$$\text{Var}\ (\hat{\mu}_2) = \frac{1}{N} \left\{ \mu_4 - \mu_2^2 \right\} \qquad (11)$$

$$\text{Var}\ (\mu_2^2) = \frac{1}{N^{[2]}} \left[\left\{ (N-2)^{[2]} - N^{[2]} \right\} \mu_2^4 + 2\ (N-2)\ \mu_4 \mu_2^2 + 2\mu_4^2 \right] \qquad (12)$$

$$\text{Var} \left(\hat{\mu}_2^3 \right) = \frac{1}{N^{[3]}} \left[\left\{ (N-3)^{[3]} - N^{[3]} \right\} \mu_2^6 \right.$$

$$\left. + 9\ (N-3)^{[2]} \mu_4 \mu_2^4 + 18\ (N-3)\ \mu_4^2 \mu_2^2 + 6\mu_4^3 \right] \qquad (13)$$

$$\text{Var}\left(\hat{\mu_2^4}\right) = \frac{1}{N^{[4]}}\left[\left\{(N-4)^{[4]} - N^{[4]}\right\}\mu_2^8 + 16\,(N-4)^{[3]}\,\mu_4\mu_2^6 + 72\,(N-4)^{[2]}\,\mu_4^2\mu_2^4 + 96\,(N-4)\,\mu_4^3\mu_2^3 + 24\mu_4^4\right] \tag{14}$$

for $r = 1$ the result (11) tallies with that given by Kendall [4].

The above expressions involve second moments of degree eight and fourth moment of degree four. The unbiased estimators for μ_4^S ($S = 1,2,3,4$) are given below while for μ_2^r the estimate μ_2^r is to be substituted from (5).

$$\left.\begin{aligned}
\hat{\mu}_4 &= \frac{S_\mu}{N}, \quad \hat{\mu_4^2} = \frac{1}{N^{[2]}}\left[S_4^2 - S_8\right] \\
\hat{\mu_4^3} &= \frac{1}{N^{[3]}}\left\{S_4^3 - 3S_8S_4 + 2\,S_{12}\right\} \\
\hat{\mu_4^4} &= \frac{1}{N^{[4]}}\left\{S_4^4 + 20\,S_{12}S_4 + 18\,S_8^2 - 12\,S_4^2S_8 - 27\,S_{16}\right\}
\end{aligned}\right\} \tag{15}$$

while for $\hat{\mu_2^S}$ (S=1,2,3,4,6) are given in the previous section. As it looks from the last expression of (7) the evaluation of $\hat{\mu_2^8}$ will still be complicated. The alternative is to neglect the term involving $\hat{\mu_2^8}$ when N is sufficiently large since the coefficient of the term will be very near to zero.

ESTIMATION AND VARIANCE OF FUNCTIONS INVOLVING $\hat{u_2^r}$

Consider the function

$$\phi = \sum_{n=0}^{r} K_n \left(\mu_2\right)^{-n} \tag{16}$$

The approximate unbiased estimator of (16) will be given by

$$\hat{\phi} = \sum_{n=0}^{r} K_n \left(\hat{\mu_2^{-n}}\right) + 0\left(\frac{1}{n}\right) \tag{17}$$

and the estimate of its variance

$$\text{Var}\;(\hat{\phi}) = \left\{\sum_{n=0}^{r} \frac{n^2\,K_n^2}{\{\hat{\mu_2^{n+1}}\}^2} + \sum_{p=0}^{r}\sum_{q=0}^{r} \frac{p\;q\;K_p\,K_q}{\hat{\mu_2^{p+1}}\,\hat{\mu^{q+1}}}\right\}\left(S_4 - \frac{S_2^2}{N}\right)\frac{1}{N^{[2]}} \qquad p \neq q \tag{18}$$

for example, let us consider the function

$$\frac{\sqrt{2\pi}}{100}\,\alpha = \frac{r^3}{3\sigma3} - \frac{r^5}{10\sigma^5} + \frac{r^7}{56\sigma^7} \tag{19}$$

where $r = \sqrt{x^2 + y^2 + z^2}$, x, y and z are normal variates with means equal to zero and variances equal to σ^2. Its admissible root gives the spherical probable error (Singh[5]) when $\alpha = 50$.

In order to determine the value of left hand side for given r we will have to make use of estimators for higher powers of variance and it will be given by

$$\hat{\alpha} = \frac{100}{\sqrt{2\pi}}\left\{\frac{r^3}{3\hat{\mu}_2} - \frac{1}{10}\,\frac{r^5}{\hat{\mu_2^2}} + \frac{1}{56}\,\frac{r^7}{\hat{\mu_2^3}}\;\cdots\cdots\right\}\frac{1}{\hat{\sigma}} \tag{20}$$

Although the above series and the following (21), (22), (23), and (28) are infinite series but they are rapidly convergent that is if S_n and S_{n+1} denote the partial sums of a series $F(x/\sigma)$ including n and $n+1$ terms, respectively, the true value of $F(z/\sigma)$ lies between S_n and S_{n+1}, therefore the difference between $F(x/\sigma)$ or S_n is less than the first term in the series that is neglected.

And its variance on omitting higher terms.

$$\text{Var}\ (\hat{\alpha}) = \frac{10^4}{8\pi}\left\{\frac{r^4}{\hat{\mu}_2^{4}} + \frac{1}{4}\frac{r^{10}}{\hat{\mu}_2^{6}} - \frac{r^8}{2\hat{\mu}_2^{5}}\ \cdots\cdots\right\}\frac{S_4 - S_2^2/N}{N^{[2]}\ \hat{\mu}_2} \tag{21}$$

The correct evaluation of α would be helpful in testing the normality of a trivariate normal sample which is being discussed by the author in a separate paper. Similarly unbiased estimators and variance of the expression of error function can be had from

$$\wedge\left(\frac{x}{\sigma}\right) \simeq \frac{2}{\sqrt{\pi}}\left[x - \frac{x^3}{\hat{\mu}_2} + \frac{x^5}{10\hat{\mu}_2^{2}} - \frac{1}{4^2}\frac{x^7}{\hat{\mu}_2^{3}}\ \cdots\cdots\cdots\right]\frac{1}{\hat{\sigma}} \tag{22}$$

and the variance

$$\text{Var}\left(\hat{H}\left(\frac{\alpha}{\sigma}\right)\right) = \frac{1}{N^{[2]}\ \pi\hat{\mu}_2}\left[\frac{x^2}{\hat{\mu}_2^{2}} + \frac{x^6}{\hat{\mu}_2^{4}} + \frac{x^{10}}{4\hat{\mu}_2^{6}} - \frac{x^4}{\hat{\mu}_2^{4}} + \frac{x^6}{2\hat{\mu}_1^{5}} - \frac{1}{2}\frac{x^8}{\hat{\mu}_1^{6}}\ \cdots\cdots\cdots\right]\left(S_4 - \frac{S_2^2}{N}\right) \tag{23}$$

If the function is of the form,

$$\phi = \sum_{n=0}^{m} K_n\ \mu_2^{n} \tag{24}$$

The unbiased estimator for ϕ is given by

$$\hat{\phi} = \sum_{m=0}^{m} K_n\ \hat{\mu}_2^{n} \qquad \text{if} \qquad N \geqslant m \tag{25}$$

Where $\hat{\mu}_2^{n}$ is given by equations (7). The variance of $\hat{\phi}$ is

$$\text{Var}\ (\hat{\phi}) = \sum_{n=0}^{m} K_n^{\ 2}\ \text{Var}\ (\hat{\mu}_2^{n}) + \sum_{n=0}^{m}\ \sum_{p=0}^{m} K_n\ K_p\ \text{cov}\ (\hat{\mu}_2^{n}, \hat{\mu}_2^{p})_{n\neq p} \tag{26}$$

$\text{Var}\ (\hat{\mu}_2^{n})$ is given by equation (10) and

$\text{Cov}\ (\hat{\mu}_2^{n}\ \hat{\mu}_2^{p})$ by

$$\sum_{\beta=0}^{p}\frac{\binom{n}{\beta}\binom{N-n}{p-\beta}}{\binom{N}{p}}\ \mu_4^{\beta}\ \mu_2^{n+p-2\beta} - \mu_2^{n+p}\ ,\quad N \geqslant n+p,\ n > p \tag{27}$$

For illustration let us take the complementary error function

$$1 - \hat{H}\left(\frac{x}{\sigma}\right) \simeq \frac{1}{\sqrt{\pi}} e^{-\frac{x^2}{\hat{\sigma}^2}} \left[\frac{1}{x} - \frac{\hat{\sigma^2}}{2x^3} + \frac{1\cdot3\cdot\hat{\sigma^4}}{4x^5} - \frac{1\cdot3\cdot5\hat{\sigma^6}}{8x^7} + \frac{1\cdot3\cdot5\cdot7\hat{\sigma^8}}{16x^9} \cdots\cdots \right] \hat{\sigma} \quad (28)$$

where x is the normal variate with mean zero and variance σ^2. The expression (28) is a polynonial in σ^2. The unbiased estimator and variance of (28) may be obtained by making use of equations (5, 10, 25 and 28) and unbiased estimate of σ is obtained from

$$\hat{\sigma} = C\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$$

where $C = \sqrt{\frac{n}{2}}\,\Gamma\,(n/2)\,/\,\Gamma\,\frac{1}{2}(n+1)$

The term e^{-x^2} may separately be evaluated from any suitable table. An unbiased estimators for the normal theory error function has been given by Barton[6]. The present paper provides approximate unbiased estimators by distribution-free approach. The bias may not be significant for the evident reason that normal error function series are rapidly convergent.

Similarly unbiased estimates of higher powers of variance have their applications while estimating dispersion by the use of higher order mean square successive differences in cases when strong trend is present in the mean values and the method of first differences does not adequately eliminate the trend as discussed by Moore[7]. The mean square successive second difference is defined by Kamat[8] and Rao[9] as

$$\delta^2 = (n-2)^{-1} \sum_{i=1}^{n-2} (x_{i+2} - 2x_{i+1} + x_i)^2 \quad (29)$$

The absolute moments $\mu_r\,(S^2)$, $\{ r = 2, 3, \ldots\ldots\ldots \}$, of δ^2 involve $\mu_2{}^S$, $(S = 1, 2, \ldots\ldots\ldots)$ the moments about the mean $\mu_1{}^1$ of the sampled population which should positively be replaced by their unbiased estimates while evaluating the moments of δ^2—, in order to get better estimate of the moments of δ^2.

Acknowledgements—The author is thankful to Dr. P. V. Krishna Iyer, Assistant Director, DSL for his encouragement and kind supervision. Thanks are also due to Dr. V. Ranganathan, Director, Defence Science Laboratory, for giving permission to publish the paper.

REFERENCES

1. GLASSER, G. J., *J. Roy. Statist. Soc.*, **23**, 154 (1961).
2. FISHER, R. A., *Proc. Camb. Phil. Soc.* (1929).
3. DAVID, F. B. and KENDALL, M. G., *Biometrika*, 36, 431 (1949).
4. KENDALL, M. G., "*Adv. Theory of Statist*", **1**, 207, (1949).
5. SINGH NAUNIHAL, *Nature*, Feb, 10, (1962).
6. BARTON, D. E., *Biometrika*, **48**, 227 (1961).
7. MOORE, P.G. *J. Amer. Statist. Soc.*, **50** 434 (1955).
8. KAMAT, A.R. *Biometrika*, **41**, 1 (1954).
9. RAO, J.N. K., *J. Amer. Statist. Soc.*, **54**, 801 (1959).