Improvement by Multi-Stage Selection in Non-Normal Large Populations

M. Gopinath Rao[1*] and L.S.Singh[2]
[1]Department of Agricultural Statistics, University of Agricultural Sciences, Bangalore, India
[2]Department of Mathematics and Statistics, Ram Manohar Lohia Avadh University, Faizalabad, U.P., India
[*]Corresponding Author: E-mail: mgrao2000@rediffmail.com

Received: 05, March 2015/ Revised: 18, September 2015/ Accepted: 25, September 2015

Abstract

Multistage selection based on the covariates of the breeding value is studied in non-normal large populations. Expressions for expectation and variance of the target variable are given for retained sub population in the case of lognormal, skew normal and pareto distributions. Numerical illustration for optimum level of cullings for fixed overall intensity of selection is discussed.

Keywords: Selection intensity, response to selection, non-normal distributions

1. Introduction

In plant and animal breeding breeder saves for reproduction a fraction of the population through truncation selection. The aim of selection is to improve the breeding value (Falconer, 1989). Since breeding value is not observable traits correlated with the target variable. Further selection is practiced in two or more stages for the covariates of the target variable may not be available simultaneously (Jain and Amble, 1962, Norrell, Arnason and Hugason, 1991). The distribution of the target variable and the other traits are assumed to be multivariate normal (Bhat, 1990; Malhotra, 1973; and Young, 1974). In such cases the expression for genetic gain on the assumption of normality is not strictly applicable. Hence in this paper expressions for response to selection and variance of the target variable after selection when the traits follow non-normal distribution are given for large populations.

Quantitative traits like foreheads of crabs (Kapteyn, 1903), aflatoxin in peanut (Quesenberry, Whitaker and Dickens, 1976) wool yield have been represented by lognormal distribution. Moreover lognormal distribution is more realistic representations of characters like weight, height and density than is normal distribution since these traits take only non-negative values (Johnson, Kotz and Balakrishnan, 1994). Another distribution to deal with non-normal data with the problem of moderate skewness is skew-normal distribution (Azzalinni and Dalla Valle, 1996; Arnold and Beaver, 2000). Pareto distribution (Mardia, Kent and Bibby, 1979) is a distribution with positive skewness and mathematical tractability like normal. In all the three distributions correlation between variates has reasonably wide range unlike many other multivariate distributions. Hence these three distributions have been considered.

2. General theory

Let X_1 and X_2 be two observable variables which are available for selection to improve unobservable variable Y (also denoted by X_3). Assume X_1, X_2 and Y follow trivariate distribution. Let $f(x_1, x_2, y)$ be the joint density function of X_1, X_2 and Y. µ and denote the mean and variance of Y in the unselected population. First selection is made on the basis of X_1 while the second selection is made on X_2.

Consider a two stage selection programme in a population in which variables X_1 and X_2 are observed. The truncation points c_i (i=1,2) depend on the proportions p_i (i=1,2) retained in the successive stages and satisfy

$$p_1 = \int_{c_1}^{\infty} f_1(x_1)\, dx_1$$

$$p = p_1 p_2 = \int_{c_2}^{\infty}\int_{c_1}^{\infty} f_2(x_1, x_2)\, dx_1 dx_2 \tag{1}$$

where $f_1(x_1)$ is the marginal density of X_1 and $f_2(x_1, x_2)$ is the joint density of the variables X_1 and X_2 and p is the proportion for the whole selection.

The selection is aimed at improving the genetic variable Y which is not directly measurable (Norrell, Arnason, Hugason, 1991; Smith and Quaas, 1982). The improvement is measured through genetic gain which is the difference between the means of breeding values in the selection group and the population as a whole. The mean of Y after two stage selection is

$$E(Y / X_1 > c_1, X_2 > c_2) = \frac{1}{p}\int_{c_2}^{\infty}\int_{c_1}^{\infty}\int_{-\infty}^{\infty} y\, f(x_1, x_2, y)\, dy\, dx_1\, dx_2 \tag{2}$$

The genetic gain after two stage selection programme is

$$\Delta G = \frac{1}{\sigma}\left[E(Y / X_1 > c_1, X_2 > c_2) - \mu\right] \tag{3}$$

The variance of Y after selection is

$$Var(Y / X_1 > c_1, X_2 > c_2)$$
$$= E\left[Y^2 / X_1 > c_1, X_2 > c_2\right] - \left[E(Y / X_1 > c_1, X_2 > c_2)\right]^2 \qquad (4)$$

This theory on genetic gain follows from Cochran's (1951) results which showed that for multistage selection in an infinite population that genetic gain is optimized if the successive conditional means of the target variable with respect to available measurements on X_1 and X_2 are used in selection.

3. Results in Non-normal populations

The non-normal distributions considered are lognormal, skew-normal and Pareto.

3.1 Lognormal distribution

The density of trivariate lognormal distribution is

$$f(x_1, x_2, x_3) = \frac{1}{(2\pi)^{\frac{3}{2}} |R|^{\frac{1}{2}}} \exp\left[-\frac{\ln \underset{\to}{x}' R^{-1} \ln \underset{\to}{x}}{2 x_1 x_2 x_3} \right], \quad x_1, x_2, x_3 > 0 \qquad (5)$$

where R is the correlation matrix

$$R = \begin{bmatrix} 1 & r_{12} & r_{13} \\ r_{12} & 1 & r_{23} \\ r_{13} & r_{23} & 1 \end{bmatrix} \quad and \; \ln x = \begin{bmatrix} \ln x_1 \\ \ln x_2 \\ \ln x_3 \end{bmatrix}$$

The mean and variance of X_3 are

$E(X_3) = e^{1/2}$

$V(X_3) = e(e-1)$ (6)

The truncation points c_1 and c_2 are determined for given values of p_1 and p_2, the proportions retained in the two stages of selection. It can be shown that

$p_1 = \Psi(\ln c_1)$

$p = \Psi(\ln c_1, \ln c_2; r_{12})$ (7)

where $\Psi(\ln c_1)$ and $\Psi(\ln c_1, \ln c_2; r_{12})$ are the upper tail probabilities of univariate normal N(0,1) and bivariate normal $N(0, 0, 1, 1, r_{12})$ respectively. It can further be shown that

$$E[Y \mid X_1 > c_1, X_2 > c_2] = \frac{e}{p}\ \psi(\ln c_1 - 2r_{13},\ \ln c_2 - 2r_{23}) \tag{8}$$

and

$$E[Y^2 \mid X_1 > c_1, X_2 > c_2] = \frac{e^2}{p} \psi(\ln c_1 - 2r_{13},\ \ln c_2 - 2r_{23}) \tag{9}$$

the genetic gain and variation if genotypic variable after two-stage selection can be worked out using (3) and (4).

3.2 Skew-normal distribution

The expression for density of skew-normal random vector **X** (Azzalini and Dalla Valle, 1996) is

$$f_3(\boldsymbol{x}) = 2\phi_3(\boldsymbol{x};\Omega)\Phi(\boldsymbol{\alpha}'\boldsymbol{x}) \tag{10}$$

where ϕ_3 (. ; **Ω**) is the three dimensional normal density with zero mean vector and correlation matrix **Ω**. Φ is the standard normal N(0,1) distribution function and $\boldsymbol{\alpha}$ is the vector of shape parameter.

The mean and variance of X_3 (Azzalini and Dalla Valle, 1996) are

$$E(X_3) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \delta_3$$

and

$$V(X_3) = 1 - \left(\frac{2}{\pi}\right) \quad \delta_3^2$$

where

$$\delta_3(\lambda) = \frac{\lambda}{(1+\lambda_3^2)^{\frac{1}{2}}}$$

The cut-off points c_1 and c_2 are found for given values of $p_1,$ p the proportions retained in the two-stage selection programme as

$$p_1 = 2\int_{c_1}^{\infty} \phi(x_1)\,\Phi\,(\lambda_1 x_1)\,dx_1$$

$$p = \int_{c_2}^{\infty}\int_{c_1}^{\infty} \phi(x_1, x_2, \omega_{12})\,\Phi\,(\alpha_1 x_1 + \alpha_2 x_2)\,dx_1\,dx_2 \tag{11}$$

The method of m.g.f (Tallis, 1961; Gopinath Rao, Singh and Nagamani, 2014) is used to find the mean and second raw moment of genotypic variable Y and the m.g.f of trivariate skew normal density is

$$m(t) = 2p^{-1}\int_{c_2}^{\infty}\int_{c_1}^{\infty}\int_{-\infty}^{\infty} e^{t'x}\phi_3(x;\Omega)\ \Phi(\alpha' x)\,dx$$

$$= 2p^{-1}(2\pi)^{-3/2}|\Omega|^{-1/2}\int_{c_2}^{\infty}\int_{c_1}^{\infty}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}x'\Omega^{-1}x - 2t'x\right]\ \Phi(\alpha' x)\ dx$$

where **t** is the column vector of t_s (s = 1, 2, 3) and taking X_3 as Y and it can be shown that

$$E(Y \mid X_1 > c_1, X_2 > c_2) = \frac{1}{p}\left[\frac{2\left(\alpha_1\omega_{13} + \alpha_2\omega_{23} + \alpha_3\right)}{(2\pi)^{\frac{2}{3}}\left(1-\omega_{12}^2 + \alpha_3^2 \mid \Omega \mid\right)^{\frac{1}{2}}} \right.$$

$$\int_{c2}^{\infty}\int_{c1}^{\infty} \exp\left[\frac{-\left(Q_1u_1^2 + Q_2u_2^2 - 2Q_3u_1u_2\right)}{2\left(1-\omega_{12}^2+\alpha_3^2|\Omega|\right)}\right] du_1\, du_2$$

$$+ 2\omega_{13}\phi(c_1)\int_{c_2^*}^{\infty} \phi(v_1)\Phi(g_1 + h_1v_1)\, dv_1$$

$$\left. + 2\omega_{23}\phi(c_2)\int_{c_1^*}^{\infty} \phi(v_2)\Phi(g_2 + h_2v_2) dv_2 \right] \quad (12)$$

and

$$E(Y^2 \mid X_1 > c_1, X_2 > c_2) = \frac{1}{p}\left[p + \frac{2\left(\alpha_1\omega_{13} + \alpha_2\omega_{23} + \alpha_3\right)^2}{(2\pi)^{\frac{2}{3}}\left(1-\omega_{12}^2 + \alpha_3^2 \mid \Omega \mid\right)^{\frac{1}{2}}} \right.$$

$$\int_{c2}^{\infty}\int_{c1}^{\infty} \exp\left[\frac{-\left(Q_1u_1^2 + Q_2u_2^2 - 2Q_3u_1u_2\right)}{-2\left(1-\omega_{12}^2+\alpha_3^2|\Omega|\right)}\right] du_1\, du_2$$

$$+ 2\omega_{13}\left(\alpha_1\omega_{13} + \alpha_2\omega_{23} + \alpha_3\right)\phi\left(c_1\right)\int_{c_2^*}^{\infty} \phi\left(v_1\right)\Phi\left(g_1 + h_1v_1\right) dv_1$$

$$+ 2\omega_{23}\left(\alpha_1\omega_{13} + \alpha_2\omega_{23} + \alpha_3\right)\phi\left(c_2\right)\int_{c_1^*}^{\infty} \phi\left(v_2\right)\Phi\left(g_2 + h_2v_2\right) dv_2$$

$$\frac{+ 2\omega_{13}^2\left(\omega_{12}\omega_{23} - \omega_{13}\right)\alpha_3}{\left(1-\omega_{12}^2\right)}|\Omega|^{\frac{1}{2}}\int_{c_2}^{\infty} \phi(c_1, u_2; \omega_{12})\,\phi\left(\frac{g_*}{\left(1+h_*^2\right)^{\frac{1}{2}}}\right) du_2$$

$$+ 2c_1\,\phi\left(c_1\right)\omega_{13}^2\int_{c_2^*}^{\infty} \phi\left(v_1\right)\Phi\left(g_1 + h_1v_1\right) dv_1 - 2\omega_{13}^2\omega_{12}\frac{\phi(c_1)\phi\left(c_2^*\right)}{\left(1-\omega_{12}^2\right)^{\frac{1}{2}}}\Phi\left(g_1 + h_1c_2^*\right)$$

$$+ \frac{2\omega_{13}^2\omega_{12}\phi\left(c_1\right)h_1}{\left(1-\omega_{12}^2\right)^{\frac{1}{2}}}\int_{c_2^*}^{\infty} \phi\left(v_1\right)\phi\left(g_1 + h_1v_1\right)dv_1$$

$$\frac{+ 2\omega_{23}^2\left(\omega_{12}\omega_{13} - \omega_{23}\right)\alpha_3|\Omega|^{\frac{1}{2}}}{\left(1-\omega_{12}^2\right)}\int_{c_2}^{\infty} \phi\left(u_1, c_2; \omega_{12}\right)\phi\left(\frac{g_{**}}{\left(1+h_*^2\right)^{\frac{1}{2}}}\right)du_1$$

$$+ 2c_2\phi\left(c_2\right)\omega_{23}^2\int_{c_1^*}^{\infty} \phi\left(v_2\right)\Phi\left(g_2 + h_2v_2\right)dv_2$$

$$-2\omega_{23}^{2}\omega_{12}\frac{\phi\left(c_1^*\right)\phi\left(c_2\right)}{\left(1-\omega_{12}^{2}\right)^{\frac{1}{2}}}\ \Phi\left(g_2+h_2c_1^*\right)$$

$$+\frac{2\omega_{23}^{2}\,\omega_{12}\phi\left(c_2\right)h_2}{\left(1-\omega_{12}^{2}\right)^{\frac{1}{2}}}\int_{c_1^*}^{\infty}\phi(v_2)\,\Phi(g_2+h_2v_2)\,dv_2+4\omega_{13}\omega_{23}\,\phi(c_1,c_2;\omega_{12})$$

$$\Phi\left(\frac{\alpha_1c_1+\alpha_2c_2+\alpha_3\left(c_1\omega^*+c_2\omega_1^*\right)}{\left(1+h_*^2\right)^{\frac{1}{2}}}\right)$$

$$+\frac{2\omega_{12}\left(\alpha_2\omega_{23}+\alpha_3\right)}{\sqrt{2\pi\,Q_2}}\phi\left(c_1\left\{\frac{Q_1Q_2-Q_3^2}{Q_2\left(1-\omega_{12}^2+\alpha_3^2\,|\Omega|\right)}\right\}^{\frac{1}{2}}\right)$$

$$\psi\left[\frac{(Q_2c_2-Q_3c_1)}{Q_2^{\frac{1}{2}}\,(1-\omega_{12}^2+\alpha_3^2\,|\Omega|)^{\frac{1}{2}}}\right]$$

$$+\frac{2\omega_{12}(\alpha_1\omega_{13}+\alpha_3)}{\sqrt{2\pi\,Q_1}}\phi\left[c_2\left\{\frac{(Q_1Q_2-Q_3^2}{Q_1\,(1-\omega_{12}^2+\alpha_3^2\,|\Omega|)}\right\}^{\frac{1}{2}}\right]$$

$$\psi\left[\frac{(Q_1c_1-Q_3c_2)}{Q_1^{\frac{1}{2}}\,(1-\omega_{12}^2+\alpha_3^2\,|\Omega|)^{\frac{1}{2}}}\right]\Bigg] \qquad (13)$$

Where $|\Omega| = 1-\omega_{12}^2-\omega_{13}^2-\omega_{23}^2+2\omega_{12}\omega_{13}\,\omega_{23}$ $\quad \psi\,(.)=1-\Phi\,(.)$

$Q_1 = 1+\alpha_1^2+\alpha_2^2-\alpha_1^2\omega_{12}^2-\alpha_3^2\omega_{23}^2-2\alpha_1\alpha_3(\omega_{12}\omega_{23}-\omega_{13})$

$Q_2 = 1+\alpha_2^2+\alpha_3^2-\alpha_2^2\omega_{12}^2-\alpha_3^2\omega_{13}^2-2\alpha_2\alpha_3(\omega_{12}\omega_{13}-\omega_{23})$

$Q_3 = \omega_{12}+\alpha_1\alpha_2-\alpha_1\alpha_2\omega_{12}^2+\alpha_1\alpha_3(\omega_{23}-\omega_{12}\omega_{13})$

$+\alpha_2\alpha_3\,(\omega_{13}-\omega_{12}\omega_{23})-\alpha_3^2(\omega_{12}-\omega_{13}\omega_{23})$

$g_* = c_1(\alpha_1+\alpha_3\omega_1^*)+u_2\left(\alpha_2+\alpha_3\omega_1^*\right)$

$g_{**} = c_2(\alpha_2+\alpha_3\omega_1^*)+u_1\left(\alpha_1+\alpha_3\omega^*\right)$

$$h_*=\frac{\alpha_3\,|\,\Omega\,|^{1/2}}{\left(1-\omega_{12}^2\right)^{1/2}},\quad g_1=\frac{c_1\left(\alpha_1+\alpha_3\omega^*\right)+\left(\alpha_2+\alpha_3\omega_1^*\right)c_1\omega_{12}}{\left(1+h_*^2\right)^{1/2}}$$

$$h_1=\frac{\left(\alpha_2+\alpha_3\omega_1^*\right)\left(1-\omega_{12}^2\right)^{1/2}}{\left(1+h_*^2\right)^{1/2}},\ v=\frac{u_2-c_1\omega_{12}}{\left(1-\omega_{12}^2\right)^{1/2}},\ c_2^*=\frac{c_2-c_1\omega_{12}}{\left(1-\omega_{12}^2\right)^{1/2}},$$

$$\omega^*=\frac{\left(\omega_{13}-\omega_{12}\omega_{23}\right)}{\left(1-\omega_{12}^2\right)},\quad \omega_1^*=\frac{\left(\omega_{23}-\omega_{12}\omega_{23}\right)}{\left(1-\omega_{12}^2\right)},$$

$$g_2 = \frac{c_2\left(\alpha_2+\alpha_3\omega_1^*\right)+\left(\alpha_1+\alpha_3\omega^*\right)c_2\omega_{12}}{\left(1+h_*^2\right)^{1/2}}, \quad h_2 = \frac{\left(\alpha_1+\alpha_3\omega^*\right)\left(1-\omega_{12}^2\right)^{1/2}}{\left(1+h_*^2\right)^{1/2}}$$

$$u_2 = \frac{u_1-c_1\omega_{12}}{\left(1-\omega_{12}^2\right)^{1/2}}, \quad c_1^* = \frac{c_1-c_2\omega_{12}}{\left(1-\omega_{12}^2\right)^{1/2}}$$

The single and double integrals in the expression can be numerically evaluated (Yakowitz and Szidarovszky, 1990) for known α_i's and ω_{ij}'s.(12) and (13) can be used to find the genetic gain and variance of Y for skew-normal distribution using (3) and (4).

3.3 Pareto distribution

The density function of trivariate Pareto distribution is

$$f(x_1,x_2,x_3) = \frac{\theta(\theta+1)(\theta+2)}{a_1a_2a_3\left\{\sum_{i=1}^{3}\frac{x_i}{a_i}-2\right\}^{\theta+3}}, \quad x_i > a_i > 0,\ i=1,\ 2,3,\ \ \theta>0 \tag{14}$$

The mean and variance of X_3 are

$$E(X_3) = \frac{a_3\theta}{(\theta-1)}$$

$$V(X_3) = \frac{a_3^2\theta}{(\theta-1)(\theta-2)},\ \theta>2 \tag{15}$$

The truncation points c_1 and c_2 are related to proportions retained p_1 and p as

$$p_1 = \frac{a_1^\theta}{c_1^\theta}$$

$$p = \frac{a_1^\theta a_2^\theta}{\left(a_2c_1+a_1c_2-a_1a_2\right)^\theta} \tag{16}$$

The mean and second raw moment after selection can be shown to be

$$E(Y \mid X_1>c_1,\ X_2>c_2) = \frac{(a_1a_2)^{\theta-1}a_3\left(a_2c_1+\ a_1c_2+\ \theta a_1a_2-\ 2a_1a_2\right)}{p(\theta-1)\left(a_2c_1\ +\ a_1c_2\ -\ a_1a_2\right)^\theta} \tag{17}$$

and

$$E\left(Y^2 \mid X_1>c_1,\ X_2>c_2\right) = \frac{1}{p}\left[\frac{a_3^2}{\left\{\frac{c_1}{a_1}+\frac{c_2}{a_2}-1\right\}^\theta} + \frac{2a_3}{(\theta-1)\left\{\frac{c_1}{a_1}+\frac{c_2}{a_2}-1\right\}^{\theta-1}}\right.$$

$$+ \frac{2}{(\theta-1)(\theta-2)\left\{\frac{c_1}{a_1}+\frac{c_2}{a_2}-1\right\}^{\theta-2}}\Bigg] \tag{18}$$

The response to selection and variance of the genotypic variable Y can be found from the above using (3) and (4).

4. Numerical illustration

The most important question in the context of sequential selection is the optimum levels of cullings in each of the two stages. One of the strategies is to achieve maximum genetic gain for a fixed overall proportion selected. This strategy is relevant in plant breeding work (Finney, 1984) one method of arriving at an approximate solution is an empirical approach of considering a range of sets of values p_1 and p_2 and choosing the set which maximises the genetic gain.

Let r_{12}, r_{13} and r_{23} be the correlation coefficients between X_1 and X_2, X_1 and X_3 and X_2 and X_3 respectively, with $r_{12} = 0.5$, $r_{13}= 0.4$ and $r_{23}= 0.6$. Suppose an overall proportion of 60% is be retained in the first and second stage respectively have been considered for sets of values covering the entire range. Genetic gain for lognormal distribution is presented in Table 1.

Table 1: Expected Genetic Gain for Lognormal Distribution

Proportion selected		Point of truncation		Genetic gain
P_1	P_2	$\ln c_1$	$\ln c_2$	
1.0	0.60	- 4.00	-0.25	0.3935
0.90	0.66	-1.28	-0.32	0.5535
0.80	0.75	-0.84	-0.47	0.5317
0.75	0.80	-0.67	-0.59	0.5119
0.60	1.00	-0.25	-4.00	0.3935
$r_{12} = 0.5$,	$r_{13} = 0.4$,	$r_{23} = 0.6$;	Mean(Y)= $\sqrt{e}$	Var(Y)= $e\,(e-1)$

It is seen that the combination of proportion selected p_1=0.9, p_2=0.66 is optimum as the genetic gain is maximum in this case.

In the case of skew-normal distribution the values of the parameter considered are as follows: $\delta_i = 0.5$ (i=1, 2, 3), the correlation matrix of Z (see Azzalini and Dalle Valle (1996) is

$$R = \begin{bmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix}$$

The value of skewness parameter λ in the univariate case is 0.5774 and the value of α_i (i = 1, 2) in the bivariate skew normal distribution is 0.3698 with ω = 0.625 and for trivariate skew-normal α_i = 0.27216 (i= 1, 2, 3) and the correlation matrix Ω is

$$\Omega = \begin{bmatrix} 1 & 0.625 & 0.625 \\ 0.625 & 1 & 0.625 \\ 0.625 & 0.625 & 1 \end{bmatrix}$$

The conditional expectation of the genetic variable Y is found from (12). The expression has three terms which involve incomplete intergrals. The evaluation of these integrals is done by numerical methods (Yakowitz and Szidarovszky, 1990). The double integral of the first term is evaluated numerically by Simpson's 1/3 rule.

Genetic gains have been determined for the same set of proportions which were considered in the case of lognormal distribution and are presented in table 2.

Table 2: Expected genetic gain for skew-normal distribution

p_1	p_2	c_1	c_2	c_1^*	c_2^*	$T_1^\dagger$	$T_2^\dagger$	$T_3^\dagger$	$T=T_1+T_2+T_3$	$E(Y \mid X_1 > c_1,\ X_2 > c_2)$	*Genetic gain*
1.0	0.60	$-\infty$	0.1615	$-\infty$	∞	0.11879	0.00000	0.26437	0.38316	0.638595	0.261429
0.90	0.66	-0.7750	0.0805	-1.0573	0.7236	0.10117	0.04028	0.23283	0.37428	0.623808	0.245304
0.80	0.75	-0.3750	-0.0500	-0.4404	0.2362	0.10053	0.09814	0.18310	0.38177	0.636279	0.258903
0.75	0.80	-0.2220	-0.1590	-0.1571	-0.0259	0.10054	0.13301	0.14990	0.38345	0.639070	0.261947
0.60	1.00	0.1615	$-\infty$	∞	$-\infty$	0.11879	0.26437	0.00000	0.38316	0.638595	0.261429

Mean (Y)=0.398862 *Var* (Y)=0.9170109

† The entries in T_1, T_2 and T_3 are the values for three terms in (12)

The combination of proportion p_1 = 0.75 and p_2 = 0.80 is optimum for skew-normal distribution.
In the case of Pareto distribution the values of the parameter $a_1 = a_2 = a_3 = 1$ and $\theta = 3$ are considered. It is found that the genetic gain for this strategy is the same for different sets of proportions considered and hence there is no particular combination for which gain is optimum.

References

1. Arnold, B.C. and R.J. Beaver., 2000: Hidden truncation models. *Sankhya A 62, 23-35.*

2. Azzalini, A., 1985: A class of distributions which includes the normal ones. *Scand. J. Statist. 12, 171-178.*

3. *Azzalini, A., and A. Dalla Valle, 1996: The multivariate skew-normal distribution.* Biometrika 83(4), 715-726. DOI: 10.1093/biomet/83.4.715

4. Bhat, Sudha G., 1990: Characterisation of distributions of some related traits in poultry. M.Sc. dissertation submitted to UAS, Bangalore.

5. Cochran, W.G., 1951: Improvement by means of selection. *Proc. Second Berkeley Symp. Math. Stat. Prob.,* 449-470.

6. Falconer, D.S., 1989: *Introduction to Quantitative Genetics. 3rd* edn., Longman, Burnt Mill.

7. Finney, D.J., 1984: Improvement by Planned multistage selection. *Jou. Am. Stat. Assoc. 79(3), 501-509.* DOI: 10.1080/01621459.1984.10478076

8. Gopinath Rao, M, Singh L. S and Nagamani,, M. K., 2014: The moments of non-truncated variable of a truncated bivariate skew-normal distribution. *Int. J. Agri. Stat. Sci.,* 10(2), 367-369.

9. Jain, J.P., and V.N. Amble., 1962: Improvement through selection at successive stages. *J. Ind. Soc. Ag. Stat 14, 88-109.*

10. Johnson, N.L., S. Kotz and N. Balakrishnan., 1994: Continuous Univariate Distributions Vol l, John Wiley, New York.

11. Kapteyn, J.C., 1903: Skew Frequency Curves in Biology and Statistics. Groningen, Nordhooft.

12. Malhotra, P.K., 1974: Selection indices in poultry. Diploma thesis in Agriculture and Animal husbandry Statistics, Institute of Agricultural Research Statistics, New Delhi.

13. Mardia, K.V., J.T. Kent and J.M. Bibby, 1979: *Multivariate Analysis. Academic Press,* London.

14. Norrell, L.T., Arnason, T.K. Hugason, 1991: *Multistage index selection in finite populations,* Biometrics 47, 205-221. DOI: 10.2307/2532507

15. Quesenberry, C.P., T.B. Whitaker and J.W. Dickens., 1976: On testing normality using several samples. An analysis of peanut aflatoxin data. *Biometrics 32,* 753-759. DOI: 10.2307/2529261

16. Smith, S.P., and R.L.Quaas., 1982: Optimal truncation points for independent culling level selection involving two traits. *Biometrics 38,*975-980. DOI: 10.2307/2529877

17. Tallis, G.M. (1961), The moment generating function of the truncated multinormal distribution. *Journal of the Royal Statistical Society Series B. 23,* 223-229.

18. Yakowitz. S., and F. Szidarovzsky, 1990: *An Introduction to Numerical Computations.* Macmillan Publishing Company, New York.

19. Young, J.C., 1974: Multiple-stage screening from non-normal populations. *Biometrika 61(1),* 155-164. DOI: 10.1093/biomet/61.1.155