A NOTE ON TESTS OF SIGNIFICANCE AND CONFIDENCE INTERVALS FOR ESTIMATES OF GENETIC CHANGE INVOLVING OFFSPRING OF RAMS FROM CONTROL AND SELECTION FLOCKS

C.Z. Roux

Animal and Dairy Science Research Institue, P.O. Box X2, Irene, Republic of South Africa, 1675.

OPSOMMING: BETEKENISTOETSE EN BETROUBAARHEIDSINTERVALLE VIR BERAMINGS VAN GENETIESE VERANDERING DEUR GEBRUIK TE MAAK VAN DIE NAGESLAG VAN RAMME UIT KONTROLE- EN SELEKSIEKUDDES

Betekenistoetse word gegee om te toets of die beramings van 'n moontlike verskil in teelwaardes van kontrole en seleksiekuddes, wat voorgestel is vir Prestasiegetoetsde Merinos, betekenisvol is. Die gewone t-toets tussen gemiddeldes is geskik om tussen die ongeweegde gemiddeldes van nageslag van ramme uit die kontrole en seleksiekuddes te toets, mits die variansies en vryheidsgrade vanaf die ramgemiddeldes verkry word, waar 'n ramgemiddelde 'n ongeweegde gemiddelde tussen lammers van die twee geslagte is. Hierdie metode is ook toepaslik om tussen toetse van verskillende seleksiekuddes te onderskei, mits twee verskillende stelle ramme uit die kontrolekudde gebruik is. In die geval waar dieselfde stel ramme gebruik is, moet 'n toets tussen gemiddeldes gebruik word wat gebaseer is op die aanname dat die variansies van die betrokke verdelings verskil. Die variansies van die kontroleramme word dan bereken op die verskil tussen die nageslag-gemiddeldes van elke kontroleram in die twee seleksiekuddes. Die teoretiese vorm van die variansies van die betrokke gemiddeldes in die toets van verskille tussen teelwaardes toon dat daar 'n groter voordeel in doeltreffende eksperimentele ontwerp behaal kan word deur die aantal ramme te vermeerder in plaas van die aantal nageslag per ram.

SUMMARY:

Tests of significance are provided for estimates of possible differences in genetic merit between control and selection flocks of Performance Tested Merinos. The t-test for differences between means provides a test for a significant difference between unweighted offspring means of rams from the control and selection flocks respectively, provided that the variances and d.f. are obtained from the ram means, where a ram mean is an unweighted average of the offspring means of both sexes. This procedure is also applicable to the results of comparisons with controls on different selection flocks, if two different sets of control flock rams were used. In the situation where the same set was used, a test for differences between means for distributions with different variances is applicable. Here the variance of the control rams is calculated from the differences, for each ram, between their offspring means in the two selection flocks. The theoretical form of the variances of the offspring means relevant to a test of differences in breeding value, shows a greater advantage to efficient experimental design by an increase in the number of rams rather than an increase in the number of offspring per ram.

A novel scheme by which breeders can estimate genetic progress due to artificial selection has been evolved for Performance Tested Merinos (see van der Merwe & Poggenpoel, 1977). Two groups of rams, one from a control flock and the other from a selection flock, are mated to 2 groups of ewes from the selection flock, chosen at random but of comparable age. Let x_s be the unweighted mean of the sire/sex progeny subgroups of rams from the selection flock and x_c the unweighted mean of the sire/sex progeny subgroups of rams from the control flock. The difference in breeding value between the 2 flocks is then estimated by $2 (x_s - x_c)$. The expected production of the control flock in the same environment as the selection flock is estimated by $2 x_c - x_s$. The percentage improvement is then estimated by $200 (x_s - x_c) / (2 x_c - x_s)$.

Tests of Significance for a Single Selection Flock

It is clear that a test of significance for genetic improvement is equivalent to a test of significance for a difference between x_c and x_s. An ordinary t-test will do. As the variability between offspring due to both rams and ewes should be taken into account the variances of x_c and x_s should be calculated from unweighted averages between male and female offspring for each ram. Thus the number of degrees of freedom for the t-test is equal to $n_c + n_s - 2$, where n_c and n_s are the number of rams from the control and selection flocks.

To obtain robustness aginst deviations from the underlying assumptions of normality and homogeneous variances designs with $n_c = n_s$ are strongly recom-

mended; in this situation the 2 variances can generally be combined in the usual manner.

Tests of Significance for Two Selection Flocks

Let y_c and y_s be the means in a second selection flock. Here one would like to test $(x_s - x_c) - (y_s - y_c)$ or $(x_s - y_s) - (x_c - y_c)$ for significance. If x_c and y_c contain different rams, no problem exists, and the procedure is entirely analogous to the single flock situation. If x and y contain the same rams, a problem exists since var $(x_s - y_s) \neq$ var $(x_c - y_c)$. Var $(x_c - y_c)$ can be calculated from the difference, for each ram, between offspring means in the two flocks. Any of the accepted methods in the literature for testing between means from populations with different variances can then be employed, since no test exists which is clearly superior to all others. A convenient test is described in the Biometrika Tables, Vol. 1, Table 11 (Pearson & Hartley, 1970). If these tables are unavailable, a test due to Banerji (1960) appeals. The proposed test at the α % level is:

$$[(x_s - y_s) - (x_c - y_c)] > [(t_s^2 \text{ var}(x_s - y_s) + t_c^2 \text{ var}(x_c - y_c)]^{\frac{1}{2}}$$

where t_s and t_c are the upper $(\alpha/2)$- points of the t distribution for $(n_s (x) + n_s (y) - 2)$ and $(n_c - 1)$ d.f. respectively, with $n_s (x)$ and $n_s (y)$ being the numbers of rams from the 2 selection flocks and n_c the number from the control flock.

Optimal Experimental Design

Assume that $n_c = n_s = n$ and that each ram has m offspring. Then var $(x_i) = (\sigma_p^2 /n) [h^2/4 + 3h^2/4m]$, $i = c, s$; where h^2 is the heritability of the trait under consideration and σ_p^2 is its phenotypic variance. This form of the variance shows that generally there is greater advantage to increasing n rather than m, since

$$\text{var}(x_i) \to 0 \text{ if } n \to \infty$$

but

$$\text{var}(x_i) \to \sigma_p^2 h^2/4n \text{ if } m \to \infty .$$

If estimates of h^2 and σ_p^2 are available values of m and n can be calculated for confidence intervals of specific length for 2 $(x_s - x_c)$. Assume $h^2 = 0,36$ and $\sigma_p^2 = 0,15$, $n = m = 10$. Hence, as follows from the relationship var $[2 (x_s - x_c)] = 8$ var (x_i), s.d. $[2 (x_s - x_c)] = 0,12$ with a 95% confidence interval equal to $\pm 0,12$ x 2,10 $= \pm 0,25$, where $t_{18} = 2,10$; $t_{18} = $ t with 18 d.f. It follows that $n = m = 10$ would be enough to declare a difference of $\pm 0,25$ units between $2x_c$ and

$2x_s$ significant at the 5% level. Similar calculations can be done for all traits of interest if likely values of the relevant parameters are known. The t-values should have degrees of freedom $2n - 2$, if the procedure of the first part of the paper is followed.

Combining Estimates of Variance

In a situation of similar environments and genotypes it is likely that variances will be homogeneous. Here increased accuracy in testing due to combined estimates can be exploited. In the situation where the same rams from the control flock are used in q selection flocks the analysis of variance on ram means would be described by:

Source	d.f.	m.s.
Flocks	q - 1	F
Rams	n - 1	R
Rams x Flocks	(q - 1) (n - 1)	F x R.

An estimate of var $(x_c - y_c)$ would then be equal to 2 (F x R)/n, with (q - 1) (n - 1) degrees of freedom for the relevant test. In the situation where different rams from the control flock are used in all selection flocks var $(x_s - x_c - y_s + y_c)$ would be estimated by R $[1/n_c (x) + 1/n_c (y) + 1/n_s (x) + 1/n_s (y)]$ with R the mean square between ram means, within selection and control groups, within selection flocks, with $\Sigma [(n_s (i) - 1) + (n_c (i) - 1)]$ degrees of freedom $(i = 1,2, . ., q)$ as is usual in nested analysis of variance. If always $n_c (i) = n_s (i) = n$, the relevant variance would be estimated by $4R/n$ with $2q (n - 1)$ d.f.

Note on the Use of Ratios

Preferential use of a ratio like 2 $(x_s - x_c) / (2x_c - x_s)$ really implies that a multiplicative model is appropriate. Let $u = \ln x$, and let u_s and u_c be the unweighted means of the logarithms, similar to the usage on the arithmetic scale. Then tests of significance can be done on $u_s - u_c$ in the manner indicated previously and exp 2 $(u_s - u_c)$ would be an estimate of (the mean of the selected flock) / (the mean of the control flock in the same environment). Confidence intervals can similarly be detransformed from the logarithmic scale. The advantage to this procedure is that statistical tests are possible on a ratio which is directly useful in the comparison of different selected flocks. In contrast, no theory exists for tests or intervals on 2 $(x_s - x_c) / (2x_s - x_s)$ under the natural assumption, for multiplicativety, of log-normal distributions. The possibility exists that rankings of different flocks may not coincide between 2 $(x_s - x_c)/(2x_c - x_s)$ and exp 2 $(u_s - u_c)$. Hence the estimate should be preferred for which the requisite statistical theory exists.

References

BANERJI, S.K., 1960. Approximate confidence interval for linear functions of means of k populations when the population variances are not equal. *Sankhyā* 22, 357.

PEARSON, E.S. & HARTLEY, H.O., 1970. Biometrika tables for Statisticians. Volume 1. Third edition, with additions. Cambridge University Press, London.

VAN DER MERWE, C.A. & POGGENPOEL, D.G., 1977. The practical application of scientific principles in Merino sheep breeding. *S. Afr. J. Anim. Sci.* 7, 71.