Theor Appl Genet (1986) 72:587–591

Confidence intervals for heritability for two-factor mating design single environment linear models[*]

S. J. Knapp

Department of Crop Science, Oregon State University, Corvallis, OR 97331, USA

Received February 10, 1986; Accepted March 6, 1986
Communicated by A. R. Hallauer

Summary. Precision measurement is an essential part of heritability estimate interpretation. Approximate standard errors are commonly used as measures of precision for heritability on a progeny mean basis (H). Their derivation, however, is not inferred from the distribution theory for H. F-distribution based exact confidence intervals have been derived for some one-factor mating design H estimators. Extension of the confidence interval results from one-factor to two-factor mating designs is reported in this paper. Functions of heritability on a full-sib or half-sib progeny mean basis from nested or factorial mating design parameters were distributed according to the F-distribution. Exact confidence intervals were derived for heritability on a full-sib progeny mean basis. Exact confidence intervals for heritability on a half-sib progeny basis were adapted from previous results. Maize (*Zea mays* L.) data were used to estimate confidence intervals. Complete equations were given for interpolation in F-tables.

Key words: Random models – Genetic statistics – Maize – *Zea mays* L.

Introduction

The need for an analytical mechanism for partitioning genetic variation into additive and dominance components provided the impetus for the development of the two-factor mating designs (Comstock and Robinson 1948). There was also interest in estimating heritability and predicted response to selection from mating design statistics (Robinson et al. 1949).

Genetic theory for the one- and two-factor mating designs has produced several genetic statistics, including degree of dominance, heritability, and predicted response to selection. Distribution theory and confidence interval estimation procedures are lacking for some genetic statistics.

Exact confidence intervals have been developed for heritability on a progeny mean basis (H) for some one-factor mating design linear models (Knapp et al. 1985). Tai (1983) gives approximate confidence intervals for the same H estimator. The two confidence intervals are analytically different and the approximate intervals (Tai 1983) are considerably more narrow than the exact intervals (Knapp et al. 1985).

The research presented in this paper is an extension of the results of Knapp et al. (1985) to heritability estimators based on nested (Design I) and factorial (Design II) mating design parameters. The methods presented for constructing exact confidence intervals for H apply to several experimental designs with balanced data. The form of a linear model is obviously dependent upon experimental design. The mean squares form of H, however, is often not affected by experimental design differences. Experimental design invariable mean square functions of H provided the basis for the confidence intervals developed for H.

It was assumed that model effects were independent normal random variables with zero means and appropriate variances. Epistatic variances were not included in causal variance component expressions for simplicity. The results are limited to balanced data.

Confidence intervals

Nested mating design

The analysis of variance for the nested mating design for one environment in a replications-in-incomplete

---

[*] Oregon Agricultural Experiment Station Technical Paper No. 7659

Table 1. Replications-in-incomplete blocks experimental design analysis of variance for the nested mating design in one environment. Maize yield (lbs/plot) data from Comstock and Robinson (1948)

Source of variation	Degrees of freedom [a]	Mean square	Expected mean square [b]
Incomplete blocks (B)			
Replications/B			
Males/B	$df_m = b\,(m-1) = 36$	$M_m = 0.176$	$\theta_m = \sigma_e^2 + r\,\sigma_{f/m/b}^2 + r\,f\,\sigma_{m/b}^2$
Females/males/B	$df_{f/m} = b\,m\,(f-1) = 144$	$M_{f/m} = 0.069$	$\theta_{f/m} = \sigma_e^2 + r\,\sigma_{f/m/b}^2$
Residual			

[a] b is the number of incomplete blocks = 12, m is the number of males = 4, and f is the number of females = 4

[b] r is the number of replications = 2, σ_e^2 is the residual variance component, $\sigma_{f/m/b}^2$ is the females nested in males nested in incomplete blocks variance component, and $\sigma_{m/b}^2$ is the males nested in incomplete blocks variance component

blocks experimental design is given in Table 1. Heritability on a full-sib progeny mean basis for Design I (H_1) (Robinson et al. 1949) is

$$H_1 = \frac{\frac{1}{2}\sigma_A^2}{(\sigma_e^{2*} + \frac{1}{2}\sigma_A^2 + \frac{3}{4}\sigma_D^2)/r + \frac{1}{2}\sigma_A^2 + \frac{1}{4}\sigma_D^2}$$

$$= \frac{2\,\sigma_m^2}{\sigma_e^2/r + \sigma_{f/m}^2 + \sigma_m^2}$$

$$= 2\,[(\theta_m - \theta_{f/m})/r\,f]\,[\theta_{f/m}/r + (\theta_m - \theta_{f/m})/r\,f]^{-1}$$

$$= 2\,[f\,(\theta_m/\theta_{f/m} - 1)^{-1} + 1]^{-1} \qquad [1]$$

and is estimated by

$$\hat{H}_1 = 2\,[f\,(M_m/M_{f/m} - 1)^{-1} + 1]^{-1}$$

where σ_A^2 is additive genetic variance, σ_D^2 is dominance variance,

$$\sigma_e^2 = \sigma_e^{2*} + \tfrac{1}{2}\sigma_A^2 + \tfrac{3}{4}\sigma_D^2, \quad \sigma_m^2 = \tfrac{1}{4}\sigma_A^2,$$

and

$$\sigma_{f/m}^2 = \tfrac{1}{4}\sigma_A^2 + \tfrac{1}{4}\sigma_D^2;$$

other terms were defined in Table 1.

Heritability on a half-sib progeny mean basis for Design I (H_2) (El-Rouby and Penny 1967) is

$$H_2 = \frac{\frac{1}{4}\sigma_A^2}{(\sigma_e^{2*} + \frac{1}{2}\sigma_A^2 + \frac{3}{4}\sigma_D^2)/r\,f + (\frac{1}{4}\sigma_D^2 + \frac{1}{4}\sigma_A^2)/f + \frac{1}{4}\sigma_A^2}$$

$$= \frac{\sigma_m^2}{\sigma_e^2/r\,f + \sigma_{f/m}^2/f + \sigma_m^2} = [(\theta_m - \theta_{f/m})/r\,f]\,[\theta_m/r\,f]^{-1}$$

$$= 1 - \theta_{f/m}/\theta_m \qquad [2]$$

and is estimated by

$$\hat{H}_2 = 1 - M_{f/m}/M_m\,.$$

$\hat{H}_1$ and $\hat{H}_2$ simplify to functions of constants and singular F-statistics; $M_m/M_{f/m}$ for $\hat{H}_1$ and $M_{f/m}/M_m$ for $\hat{H}_2$. Given that

$$(M_m/\theta_m)/(M_{f/m}/\theta_{f/m}) \sim F\,(df_m, df_{f/m})$$

and

$$(M_{f/m}/\theta_{f/m})/(M_m/\theta_m) \sim F\,(df_{f/m}, df_m)$$

(Graybill 1976), then functions of H_1 and H_2 are distributed according to the F-distribution.

H_2, given in [2], has the same form as heritability on a progeny mean basis for the one-factor mating design; therefore, exact confidence intervals for H_2 are identical to those previously described (Knapp et al. 1985). In terms of Table 1 notation exact confidence intervals for H_2 are

$$P\,\{1 - [(M_m/M_{f/m})\,F_{1-\alpha/2:\,df_{f/m},\,df_m}]^{-1} \le H_2 \le 1 - [(M_m/M_{f/m})\,F_{\alpha/2:\,df_{f/m},\,df_m}]^{-1}\} = 1 - \alpha \qquad [3]$$

where $F_{1-\alpha/2:\,df_{f/m},\,df_m}$ and $F_{\alpha/2:\,df_{f/m},\,df_m}$ are values from the F-distribution such that the probabilities of exceeding these values are $1-\alpha/2$ and $\alpha/2$, respectively.

The derivation of exact confidence intervals for H_1 involved an extension of the results for H_2. Given that

$$P\,[F_{1-\alpha/2:\,df_{f/m},\,df_m} \le (M_{f/m}/\theta_{f/m})/(M_m/\theta_m) \le F_{\alpha/2:\,df_{f/m},\,df_m}] = 1 - \alpha$$

then exact confidence intervals for $\theta_m/\theta_{f/m}$ are

$$P\,[(M_m/M_{f/m})\,F_{1-\alpha/2:\,df_{f/m},\,df_m} \le \theta_m/\theta_{f/m} \le (M_m/M_{f/m})\,F_{\alpha/2:\,df_{f/m},\,df_m}] = 1 - \alpha\,. \qquad [4]$$

Exact confidence intervals for H_1 from [4] are

$$P\,\{2\,[f\,(M_m/M_{f/m}F_{1-\alpha/2:\,df_{f/m},\,df_m} - 1)^{-1} + 1]^{-1} \le H_1 \le 2\,[f\,(M_m/M_{f/m}\,F_{\alpha/2:\,df_{f/m},\,df_m} - 1)^{-1} + 1]^{-1}\} = 1 - \alpha. \qquad [5]$$

The exact confidence intervals, [3] and [5], are functions of the same F-statistic and tabled F-values.

Confidence intervals were estimated from the data in Table 1. Tabled F-values were obtained by interpolation because the F-values needed for the calculations were not included in available F-tables. Inter-

Table 2. Replications-in-incomplete blocks experimental design analysis of variance for the factorial mating design in one environment. Maize ear length (cm) data from Silva (1974)

Source of variation	Degrees of freedom[a]	Mean square	Expected mean square[b]
Incomplete blocks (B)			
Replications/B			
Males and females/B	$df_{m+f} = b\,(m+f-2) = 120$	$M_{m+f} = 6.944$	$\theta_{m+f} = \sigma_e^2 + r\,\sigma_{(mf)/b}^2 + r\,[(m+f)/2]\,\sigma_{(m+f)/b}^2$
Males/B	$df_m = b\,(m-1) = 60$	$M_m = 6.320$	$\theta_m = \sigma_e^2 + r\,\sigma_{(mf)/b}^2 + r\,f\,\sigma_{m/b}^2$
Females/B	$df_f = b\,(f-1) = 60$	$M_f = 7.568$	$\theta_f = \sigma_e^2 + r\,\sigma_{(mf)/b}^2 + r\,m\,\sigma_{f/b}^2$
Males × females/B	$df_{mf} = b\,(m-1)\,(f-1) = 180$	$M_{mf} = 2.311$	$\theta_{mf} = \sigma_e^2 + r\,\sigma_{(mf)/b}^2$
Residual			

[a] b is the number of incomplete blocks = 20, m is the number of males = 4, and f is the number of females = 4

[b] r is the number of replications = 2, σ_e^2 is the residual variance component, $\sigma_{(mf)/b}^2$ is the males × females nested in incomplete blocks variance component, $\sigma_{f/b}^2$ is the females nested in incomplete blocks variance component, $\sigma_{m/b}^2$ is the males nested in incomplete blocks variance component, and $\sigma_{(m+f)/b}^2$ is the males and females pooled nested in incomplete blocks variance component

polation equations are given in the appendix. F-values for 95% $(1-\alpha=0.95)$ exact confidence intervals were $F_{0.975:144,36} = 0.618$ and $F_{0.025:144,36} = 1.757$.

$$H_1 = 2\,[4\,(0.167/0.069 - 1)^{-1} + 1]^{-1} = 0.524$$

and $M_m/M_{f/m} = 2.42$. The lower 95% confidence limit for H_1 was $2\,\{4\,[(2.420)\,0.618 - 1]^{-1} + 1\}^{-1} = 0.221$ and the upper 95% confidence limit was

$$2\,\{4\,[(2.420)\,1.757 - 1]^{-1} + 1\}^{-1} = 0.897\,.$$

Similarly with $\hat{H}_2 = 1 - 0.069/0.167 = 0.587$ the lower 95% confidence limit for H_2 was $1 - [(2.420)\,0.618]^{-1} = 0.332$ and the upper 95% confidence limit was $1 - [(2.420)\,1.757]^{-1} = 0.765$.

Factorial mating design

The derivation of confidence intervals for H based on factorial mating design or Design II parameters for single environment linear models was similar to that for the nested mating design. The analysis of variance for the factorial mating design for one environment in a replications-in-incomplete blocks experimental design is given in Table 2. Heritability on a full-sib progeny mean basis for Design II (H_3), based on the pooled male plus female variance component, is

$$\begin{aligned} H_3 &= \frac{\frac{1}{2}\sigma_A^2}{(\sigma_e^{2}{*} + \frac{1}{2}\sigma_A^2 + \frac{3}{4}\sigma_D^2)/r + \frac{1}{2}\sigma_A^2 + \frac{1}{4}\sigma_D^2} \\ &= \frac{2\,\sigma_{m+f}^2}{\sigma_e^2/r + \sigma_{mf}^2 + 2\,\sigma_{m+f}^2} \\ &= \{2\,(\theta_{m+f} - \theta_{mf})/[r\,(m+f)/2]\} \\ &\quad \cdot \{2\,(\theta_{m+f} - \theta_{mf})/[r\,(m+f)/2] + \theta_{mf}/r\}^{-1} \\ &= [1 + [(m+f)/4]\,(\theta_{m+f}/\theta_{mf} - 1)^{-1}]^{-1}\,; \end{aligned} \qquad [6]$$

where $\theta_{m+f} = (\theta_m + \theta_f)/2$ and $\sigma_{m+f}^2 = (\sigma_m^2 + \sigma_f^2)/2$; terms not previously defined were given in Table 2. H_3 is estimated by

$$\hat{H}_3 = \{1 + [(m+f)/4]\,(M_{m+f}/M_{mf} - 1)^{-1}\}^{-1}\,.$$

M_m or M_f can substitute for M_{m+f}, however, M_{m+f} is the logical choice because of the increased degrees of freedom associated with the pooled estimate.

Heritability on a half-sib progeny mean basis for Design II (H_4) is

$$\begin{aligned} H_4 &= \frac{\frac{1}{4}\sigma_A^2}{(\sigma_e^{2}{*} + \frac{1}{2}\sigma_A^2 + \frac{3}{4}\sigma_D^2)/[r\,(m+f)/2] + \frac{1}{4}\sigma_D^2/[(m+f)/2] + \frac{1}{4}\sigma_A^2} \\ &= \frac{\sigma_{m+f}^2}{\sigma_e^2/[r\,(m+f)/2] + \sigma_{mf}^2/[(m+f)/2] + \sigma_{m+f}^2} \\ &= \{(\theta_{m+f} - \theta_{mf})/[r\,(m+f)/2]\}\,\{\theta_{m+f}/[r\,(m+f)/2]\}^{-1} = 1 - \theta_{mf}/\theta_{m+f} \end{aligned} \qquad [7]$$

and is estimated by

$$\hat{H}_4 = 1 - M_{mf}/M_{m+f}\,.$$

The derivation of confidence intervals for H_3 and H_4 is exactly analogous to that for H_1 and H_2. The form of H_4 in [7] is identical to that for H_2 in [2], therefore, exact confidence intervals for H_4 are

$$\begin{aligned} &P\,\{1 - [(M_{m+f}/M_{mf})\,F_{1-\alpha/2:\,df_{mf},\,df_{m+f}}]^{-1} \le H_4 \\ &\quad \le 1 - [(M_{m+f}/M_{mf})\,F_{\alpha/2:\,df_{mf},\,df_{m+f}}]^{-1}\} = 1 - \alpha\,. \end{aligned} \qquad [8]$$

The starting point for deriving exact confidence intervals for H_3 is, except for notation differences, identical to [4]. The equivalent equation for H_3 is

$$\begin{aligned} &P\,[(M_{m+f}/M_{mf})\,F_{1-\alpha/2:\,df_{mf},\,df_{m+f}} \le \theta_{m+f}/\theta_{mf} \\ &\quad \le (M_{m+f}/M_{mf})\,F_{\alpha/2:\,df_{mf},\,df_{m+f}}] = 1 - \alpha\,. \end{aligned} \qquad [9]$$

From [9] exact confidence intervals for H_3 are

$$\begin{aligned} P(\{1 + [(m+f)/4] \\ \cdot [(M_{m+f}/M_{mf})\, F_{1-\alpha/2:\, df_{mf},\, df_{m+f}} - 1]^{-1}\}^{-1} \\ \leq H_3 \leq \{1 + [(m+f)/4] \\ \cdot [(M_{m+f}/M_{mf})\, F_{\alpha/2:\, df_{mf},\, df_{m+f}} - 1]^{-1}\}^{-1}) = 1 - \alpha\,. \quad [10] \end{aligned}$$

The exact confidence intervals [8] and [10] for Design II H estimators are functions of the same F-statistic and tabled F-values.

Confidence intervals were estimated from the data in Table 2. Males plus females pooled variance component estimates were used in the calculations. F-values required for the calculations were obtained by interpolation for this example also. They were $F_{0.975:180,120} = 0.726$ and $F_{0.025:180,120} = 1.392$. The F-statistic was $M_{m+f}/M_{mf} = 6.944/2.311 = 3.005$.

$$H_3 = \{1 + [(4+4)/4]\,(6.944/2.311 - 1)^{-1}\}^{-1} = 0.501.$$

The lower 95% confidence limit for H_3 was $\{1 + 2\,[3.005\,(0.7264) - 1]^{-1}\}^{-1} = 0.372$. The upper 95% confidence limit was $\{1 + 2\,[3.005\,(1.392) - 1]^{-1}\}^{-1} = 0.614$. The estimate for H_4 was $H_4 = 1 - 2.311/6.944 = 0.667$. The lower 95% confidence limit for H_4 was $1 - [3.005\,(0.726)]^{-1} = 0.542$ and the upper 95% confidence limit was $1 - [3.005\,(1.392)]^{-1} = 0.761$.

Discussion

The exact confidence intervals H_2 and H_4 are a direct adaptation of exact confidence intervals for H for the one-factor mating design (Knapp et al. 1985). These confidence intervals are useful because they apply to several mating and experimental designs. The constancy of the simplified expected mean squares form of the estimators is the reason for the adaptability.

The exact confidence intervals for H_1 and H_3 involve extensions of previous results (Knapp et al. 1985). These confidence intervals are also adaptable to different experimental designs. The results necessary for the derivation of the confidence intervals are given in [1] and [6]. Equations [1] and [6] are simplified expected mean square expressions. They are functions of constants and singular variables distributed F.

The distributions of the H estimators, [1], [2], [6], and [7], are not known. It was shown that this does not prevent the development of exact confidence intervals for H. There are similar results for other random model statistics (Graybill 1976).

The confidence intervals are easily estimated when appropriate F-values are tabulated. Interpolation formulas must be used when they are not tabulated. Laubscher (1965) gives equations for interpolation in F-tables but does not give limits of some of the coefficients, which are necessary when specified degrees of freedom occur between given values and infinity in the F-table. Complete interpolation equations were given in the appendix to provide the equations necessary for confidence interval estimation.

Single environment heritability estimates, with rare exception, are biased by genotype × environment interaction. Single environment estimators were examined in this paper. Use of single environment point estimates is limited in practice (Dudley and Moll 1969). Single environment interval estimates, however, are useful because they measure the effect on precision of experimental design and half- and full-sib family number. Approximate confidence intervals are available for half-sib family H estimators for two-factor mating design linear models that include environmental effects (Knapp and Bridges unpublished manuscript).

Acknowledgements. I thank Dr. W. C. Bridges, Jr. for reviewing the manuscript and Dr. A. R. Hallauer for providing factorial mating design data.

References

Comstock RE, Robinson HF (1948) The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. Biometrics 4:254–266

Dudley JW, Moll RH (1969) Interpretation and use of heritability and genetic variances in plant breeding. Crop Sci 9:257–261

El-Rouby MM, Penny LH (1967) Variation and covariation in a high oil population of corn (*Zea mays* L.) and their implications in selection. Crop Sci 7:216–219

Graybill FA (1976) Theory and application of the linear model. Duxbury Press, Boston Mass

Knapp SJ, Stroup WW, Ross WM (1985) Exact confidence intervals for heritability on a progeny mean basis. Crop Sci 25:192–194

Laubscher NF (1965) Interpolation in F-tables. Am Stat 19: 39–40

Robinson HF, Comstock RE, Harvey PH (1949) Estimates of heritability and the degree of dominance in corn. Agron J 41:353–359

Scheffè H (1959) The analysis of variance. Wiley, New York

Silva JC (1974) Genetic and environmental variances and covariances estimated in the maize (*Zea mays* L.) variety, 'Iowa stiff stalk synthetic'. PhD Dissertation, Iowa State University Ames. Diss Abstr No 35/8-B

Tai GCC (1983) Confidence intervals of genetic parameters estimated from progeny test experiments in plant breeding. Can J Genet Cytol 25:502–508

Appendix

Two-sided confidence intervals require upper and lower percentage points from the F-distribution. The equation used for obtaining upper percentage points from lower percentage points or the reverse is

$$F_{\alpha:\, df_1,\, df_2} = (F_{1-\alpha:\, df_2,\, df_1})^{-1}$$

(Scheffè 1959).

Suppose that a tabled F-value, $F_{\alpha:\, df_1,\, df_2}$, is needed and df_1 and df_2 do not appear in the F-table. The interpolation formula for this case is

$$\begin{aligned} F_{\alpha:\, df_1,\, df_2} = {} & (1-J)\,(1-K)\, F_{\alpha:\, df_1',\, df_2'} + J\,(1-K)\, F_{\alpha:\, df_1',\, df_2''} \\ & + (1-J)\, K\, F_{\alpha:\, df_1'',\, df_2'} + J\, K\, F_{\alpha:\, df_1'',\, df_2''} \end{aligned}$$

where

$$J = [df_2''\,(df_2 - df_2')]\,[df_2\,(df_2'' - df_2')]^{-1}\,,$$

$$K = [df_1''\,(df_1 - df_1')]\,[df_1\,(df_2'' - df_1')]^{-1}\,,$$

df_1' is the degrees of freedom preceding df_1, df_1'' is the degrees of freedom following df_1; df_2' is the degrees of freedom preceding df_2, and df_2'' is the degrees of freedom following df_2 in the F-table (Laubscher 1965).

Suppose that df_1 is given in the F-table but not df_2. The interpolation equation for this case is

$$F_{\alpha:\, df_1,\, df_2} = (1-J)\, F_{\alpha:\, df_1,\, df_2'} + J\, F_{\alpha:\, df_1,\, df_2''}\,.$$

The equation for the reverse case, when df_2 is given but not df_1, is

$$F_{\alpha:\, df_1,\, df_2} = (1-K)\, F_{\alpha:\, df_1',\, df_2} + K\, F_{\alpha:\, df_1'',\, df_2}$$

(Laubscher 1965).

Equations are developed below for cases where df_1'' and/or df_2'' are infinite. The limit for K as df_1'' approaches ∞ is used when df_1'' is infinite. The reciprocal of K is

$$\begin{aligned} K^{-1} &= [(df_1'' - df_1)/df_1'']\,[df_1/(df_1 - df_1')] \\ &= (1 - df_1/df'')\,[df_1/(df_1 - df_1')]\,. \end{aligned}$$

Given that

$$\lim_{df_1'' \to \infty} 1 - df_1/df_1'' = 1$$

then

$$\lim_{df_1'' \to \infty} K^{-1} = df_1/(df_1 - df_1')$$

and

$$\lim_{df_1'' \to \infty} K = \left[\lim_{df_1'' \to \infty} K^{-1}\right]^{-1} = (df_1 - df_1')/df_1 = 1 - df_1'/df_1\,. \qquad [11]$$

The result for J is

$$\lim_{df_2'' \to \infty} J = (df_2 - df_2')/df_2 = 1 - df_2'/df_2\,. \qquad [12]$$

The limiting values, [11] and [12], are used accordingly in the interpolation equations.