

Kansas State University Libraries
New Prairie Press

Conference on Applied Statistics in Agriculture 2008 - 20th Annual Conference Proceedings

TESTING VARIANCE COMPONENTS BY TWO JACKKNIFE METHODS

Jixiang Wu

Johnie N. Jenkins

Jack C. McCarty

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), and the [Applied Statistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

Recommended Citation

Wu, Jixiang; Jenkins, Johnie N.; and McCarty, Jack C. (2008). "TESTING VARIANCE COMPONENTS BY TWO JACKKNIFE METHODS," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1093>

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.

Testing Variance Components by Two Jackknife Methods

Jixiang Wu, Johnie N. Jenkins, and Jack C. McCarty
J. Wu, Department of Plant and Soil Sciences, Mississippi State University, Mississippi State, MS 39762; , J.N. Jenkins and J. C. McCarty, USDA-ARS, Crop Science Research Lab, Mississippi State, MS 39762.

Corresponding authors:

Jixiang Wu
Department of Plant and Soil Sciences
Mississippi State University
P.O. Box 5367
Mississippi State, MS 39762
Email: Jixiang.wu@ars.usda.gov
Phone: 662-320-7425
Fax: 662-320-7528
Or Johnie N. Jenkins
USDA, ARS, 810 Highway 12 East
P.O. Box 5367
Mississippi State, MS 39762, USA
Email: johnie.jenkins@ars.usda.gov
Phone: 662-320-7386
Fax: 662-320-7528

ABSTRACT

The jackknife method, a resampling technique, has been widely used for statistical tests for years. The pseudo value based jackknife method (defined as pseudo jackknife method) is commonly used to reduce the bias for an estimate; however, sometimes it could result in large variation for an estimate and thus reduce the power for parameters of interest. In this study, a non-pseudo value based jackknife method (defined as non-pseudo jackknife method) was used for testing variance components under mixed linear models. We compared this non-pseudo value based jackknife method and the pseudo value based method by simulation regarding their biases, Type I errors, and powers. Our simulated results showed that biases obtained by the two jackknife methods are very similar; however, the non-pseudo value based method had higher testing powers than the pseudo value based method while the non-pseudo value based method had lower Type I error rates than the preset nominal probability values. Thus, we concluded that the non-pseudo value based jackknife method is superior to the pseudo value based method for testing variance components under a general mixed linear model.

1. INTRODUCTION

In statistical applications involving the point estimation of an unknown parameter θ , one needs to estimate the accuracy of $\hat{\theta}$ as an estimator of θ by an appropriate statistical test. The variance, the bias, and the mean squared error are commonly used to describe the accuracy of an estimate. However, many statistical tests are based on the assumption that the data follow particular distributions (i.e. normal distribution, binomial distribution, etc.). On the other hand, statistical tests for some parameters can not be easily derived. For example, covariance, proportion, and the ranked correlation coefficient are difficult to be statistically tested by parametric methods. The jackknife (Quenouille, 1949, 1956; Tukey, 1958; Miller 1964, 1968; Gray et al., 1972; Gray and Schucany, 1972; Shao, 1988; Shao and Tu, 1995) and the bootstrap (Efron, 1979; Davison and Hinkley, 1997) are two extensively used resampling methods not only for bias reduction but also for the development of nonparametric statistical tests for various types of parameters. The jackknife technique was initially introduced by Quenouille (1949) as a method for bias reduction and was later proposed as a method for robust interval estimation by Tukey (1958). Miller (1974a) gave a detailed review for this technique and also proved that the jackknife estimates of regression parameters in a linear regression model were asymptotically normally distributed under conditions that were not normally distributed (Miller, 1974b).

Little information has been reported for testing variance components under a general linear model using the jackknife method. For simple mixed linear models with balanced data under the normal distribution, the F-test can be used for a ratio of functions of variance components based on expected mean squares. The X^2 value can also be used to test the significance of a variance component for a complicated mixed linear model or an irregular data structure for large sample size (Searle et al., 1992). However, the power by the X^2 test is relatively low (Zhu, 1998). Thus, the jackknife was first applied to calculate the standard error for each point estimator through removal of one genotype or one replication each time (Zhu, 1998). The resampling method used by Zhu was the pseudo jackknife method which utilizes the pseudo values to determine a point estimator (i.e. mean) and its corresponding standard error (Tukey, 1958; Quenouille, 1956; Miller, 1997a, b). Then, an approximate t-test can be used to test the significance of a parameter (Miller, 1974a).

For a given sample size n with g subgroups, g pseudo values can be calculated accordingly (Miller, 1974b) and a pseudo jackknife estimate is obtained by averaging the g pseudo values. The number of pseudo value is called the number of jackknife times and is related to the number of subgroups g . For an experiment with replications or repeated plots, then g is the total number of replications or blocks, otherwise, it should be equal to $\binom{n}{r}$, where n is the total observations and r ($1 \sim n-1$) is the number of individual observations to be deleted at each time. It is easy to see that the number of jackknife times varies and can be very large when n is large. As we will show in Section 2, the pseudo values can be very sensitive and be out of reasonable ranges with increases in the number of jackknife times. For example, negative variances and out-of-range correlation

coefficients could be resulted. In addition, such sensitivity will definitely cause large variation for a pseudo value based estimate and low testing power accordingly. Thus, it will be helpful to develop a new resampling method which is reliable regardless of resampling time.

The objective of this study was to propose and evaluate a non-pseudo jackknife method by Monte Carlo simulation. The paper is organized as follows. In section 2 we briefly give an overview of the pseudo jackknife method and describe the non-pseudo jackknife method. In section 3, Monte Carlo simulation procedure for these two jackknife methods applied to two commonly used genetic models is described and the simulated results are reported. An actual example that included two cotton agronomic traits is analyzed by these two jackknife methods in Section 4. Some discussion and concluding points are presented in Section 5.

2. A BRIEF OVERVIEW OF JACKKNIFE METHODS

Let $\hat{\theta}$ be an estimator of the parameter θ based on the complete sample of size n with g subgroups. Let $\hat{\theta}_{-i}$ be the corresponding estimator based on the sample at the i^{th} deletion. Define

$$\tilde{\theta}_i = g\hat{\theta} - (g-1)\hat{\theta}_{-i} \quad (i=1, \dots, g) \quad (1)$$

The i^{th} deletion of the total could be one individual observation or several observations (Quenouelle, 1956; Miller, 1974a,b; Wu et al., 2006). The latter case is called group- or block-based jackknife if one replication or one block observations are deleted (Miller, 1974b; Wu et al., 2006). In equation (1) estimation $\tilde{\theta}_i$ is called the i^{th} pseudo value and the estimator in equation (2) is the jackknife estimator for the parameter θ , where θ can be a variance component, covariance component, correlation coefficient, or any other parameter of interest.

$$\tilde{\theta} = \frac{1}{g} \sum_{i=1}^g \tilde{\theta}_i = g\hat{\theta} - (g-1) \frac{1}{g} \sum_{i=1}^g \hat{\theta}_{-i} \quad (2)$$

We call $\tilde{\theta}$ in equation (2) a pseudo jackknife estimate. Then a t-test can be used to test significant deviation from a given parameter value, θ_0 with the degrees of freedom $g-1$ (Miller, 1974b). The equation (1) can be rewritten as follows,

$$\tilde{\theta}_i = g\hat{\theta} - (g-1)\hat{\theta}_{-i} = \hat{\theta} + (g-1)(\hat{\theta} - \hat{\theta}_{-i}) \quad (i=1, \dots, g) \quad (3)$$

Thus, it is obvious that pseudo value $\tilde{\theta}_i$ in equation (3) is related to choices for g . When g is large, slight differences between $\hat{\theta}$ and $\hat{\theta}_{-i}$ will cause unfavorable values. More importantly, it will potentially cause a large standard error for an estimate and thus decrease the power for the parameter being tested.

If we assume that the estimate $\hat{\theta}_{-i}$ in equation (1) for the i^{th} deletion is unbiased, then it is easy to prove that $\tilde{\theta}_-$ in equation (4) is unbiased too. It is often true if $\hat{\theta}$ is an

unbiased estimate of θ , then $\hat{\theta}_{-i}$ will be unbiased after a few individuals in the original data are deleted.

$$\bar{\hat{\theta}}_{-} = \frac{\sum_{i=1}^g \hat{\theta}_{-i}}{g} \quad (4)$$

We call $\bar{\hat{\theta}}_{-}$ in equation (4) a non-pseudo jackknife estimate of the parameter θ . For each non-normally distributed variable, based on the Central Limit Theorem, $\bar{\hat{\theta}}_{-}$ is approximately normally distributed when g is large (Wackerly et al., 1996). Thus, an approximate z-test can be used when g is large or t-test can be used to test significant deviation from a given parameter value, θ_0 , with the degrees of freedom $g-1$ (Daniel, 1995).

Considering the fact that the estimates $\hat{\theta}_{-i}$ for the i^{th} deletion are related to each other, a large Type I error can result. So it is necessary to adjust the Type I error as follows. Denote α to be a nominal value to be used for testing the significance of the parameter θ . Under a null hypothesis $H_0: \theta = \theta_0$, if the null hypothesis is rejected, then we will make a Type I error. Under an alternative hypothesis $H_1: \theta \neq \theta_0$, if we accept the null hypothesis, then we will make a Type II error. A high type I error indicates a null hypothesis is likely to be falsely rejected, while a high Type II error indicates a low power when an alternative hypothesis holds.

The probability for θ being falsely rejected under a null hypothesis (Type I error) is approximately as indicated in equation (5).

$$\alpha^* = 1 - (1 - \alpha)^{\rho^g} \quad (5)$$

where, ρ in equation (5) is data similarity between any two deletion sets, g is the resampling times, α is the Type I error for each test, and α^* is the Type I error caused by the non-pseudo jackknife method. To achieve $\alpha^* = \alpha_0$ (α_0 can be 0.05, 0.01, or 0.001), then equation (6) can be used to control the Type I error.

$$\alpha = 1 - (1 - \alpha_0)^{1/(\rho^g)} \quad (6)$$

For example, an original data set includes R replications, each time we remove one replication from the original data set, it is easy to derive that g is equal to R and ρ is equal to $1 - \frac{2}{R}$ ($R \geq 3$).

So far, we have described two types of jackknife methods: pseudo and non-pseudo methods. These two methods will be evaluated by the simulation data and real data using a mixed linear model approach. However, our emphasis is to use Monte Carlo simulation to compare these two jackknife methods regarding Type I errors and powers for variance components estimation through Monte Carlo simulation for two mixed linear model situations.

3. SIMULATION

3.1. Genetic models and experimental designs

For simplicity, in this study, two simple but widely used genetic models (genotype model and additive-dominance genetic model), where both include G×E effects, were chosen for simulation. To simplify, we call these two models genotype model and AD model, respectively. According to mixed linear model theory and genetic structure, these genetic models can be extended to complicated genetic models. The mixed linear model for the genotype model including G×E interaction effect is expressed as:

$$y_{hij} = \mu + E_h + G_i + GE_{hi} + B_{j(h)} + e_{hij} \quad (7)$$

Environmental effects (E_h) and the population mean (μ) in this model are considered as fixed effects, while the remaining effects are considered random.

The other genetic model, AD model, is a little more complicated. The data structure can involve parents and F_1 hybrids, parents and F_2 hybrids, or parents and F_3 hybrids. In this study, we only considered parents and F_2 hybrids to generate simulated data sets under multiple environments with a RCB design. So, the linear model for parents and F_2 hybrids are expressed as follows:

Parent:

$$y_{hiik(P)} = \mu + E_h + 2A_i + D_{ii} + 2AE_{hi} + DE_{hii} + B_{k(h)} + e_{hiik} \quad (8)$$

F_2 :

$$y_{hijk(F_2)} = \mu + E_h + (A_i + A_j) + \left(\frac{1}{4}D_{ii} + \frac{1}{4}D_{jj} + \frac{1}{2}D_{ij}\right) + (AE_{hi} + AE_{hj}) + \left(\frac{1}{4}DE_{hii} + \frac{1}{4}DE_{hjj} + \frac{1}{2}DE_{hij}\right) + B_{k(h)} + e_{hijk} \quad (9)$$

Here μ , the population mean, is a fixed effect; E_h , the environment effect, may be either random or fixed (but considered fixed in this study); A_i (or A_j) is the additive effect from parent i (or j); D_{ii} , D_{jj} , or D_{ij} is the dominance effect; AE_{hi} (or AE_{hj}) is additive × environment interaction effect; DE_{hii} , DE_{hjj} , or DE_{hij} is the dominance × environment interaction effect; $B_{k(h)}$ is the k -th block effect within environment h ; and e_{hijk} is the random error. All genetic effects, including the G×E effects, are assumed to be random.

Practically, a data set could involve unequal block numbers across environments or various types of data structures. On the other hand, when k individual observations (or one replication) are removed at each time, each new data set is unbalanced. The analysis of variance (ANOVA) or general linear model (GLM) approaches are fast but only applicable to balanced data or simple linear models. Thus, the mixed linear model approaches, which provide more flexibility for unbalanced data situations than the ANOVA or GLM approaches, are suggested. However, in this study, only one of several mixed linear model approaches, the minimum norm quadratic unbiased estimation (MINQUE) approach (Rao, 1971; Zhu, 1989), is employed for data analyses, including variance component estimation and random effect prediction. The two jackknife methods

previously discussed were used, where we removed one replication (block) observations from each original simulated data set at each time.

3.2. Simulation schemes

Two environments with four replications each with a RCB design structure were used for both models. Though simulations for various different sample sizes and preset parameters were conducted, since conclusions were similar, we only reported the results from the following configurations. For the genotype model, 20 genotypes were utilized. For the AD model, we used two females crossed with six male parents, containing a total of 20 entries (eight parents and 12 F₂ hybrids) were used. Two types of distributions, the normal distribution and uniform distribution were also used to generate the simulated data for analysis. The power was calculated through presetting all variance components at 20 for both models and the Type I error by presetting all variance components at 0 except the random error (set at 20) for these two genetic models. The two jackknife methods, pseudo-value and non-pseudo value based methods, as mentioned in Section 2, were then applied. Five hundred simulations were conducted for each scenario. The biases, powers, and Type I errors were obtained. All data analyses were conducted by computer programs developed by the authors of this manuscript using C++.

3.3. Simulation Results

The biases and Type I errors for the two different distributions by the two jackknife methods for each genetic model are summarized in Tables 1 and 2, respectively. There are no differences in bias between the two jackknife methods for either model (Tables 1 and 2), indicating that both jackknife methods were equally unbiased for the estimation of variance components. The pseudo jackknife method had lower Type I errors than non-pseudo jackknife method for both models. However, with few exceptions, the Type I errors obtained by the non-pseudo method were generally smaller than the three preset nominal probability values.

On the other hand, the biases and powers for the two models determined from the two different distributions are summarized in Tables 3 and 4. Again, the biases between the two jackknife methods were almost identical for the two genetic models by each distribution. However, the power for the non-pseudo method was much higher than that for the pseudo method.

In summary, our simulation results showed that the non-pseudo jackknife method had consistently higher power than the pseudo jackknife method with acceptable Type I errors for the different genetic models with different distributions.

4. APPLICATION

4.1. *Materials and experiments*

Eleven parental lines were selected from different breeding programs for use in this study. These lines were Acala Ultima (defined as parent 1, P1), Tamcot Prymaid (P2), Coker 315 (P3), Stoneville 825 (P4), FiberMax 966 (P5), M240RNR (P6), Paymaster HS26 (P7), Deltapine 90 (P8), Suregrow 747 (P9), PSC 355 (P10) and Stoneville 474 (ST474, P11) (Table 10). In 2004, these 11 parental lines were grown in two fields (Environments 1 and 2, respectively) at Mississippi State (33° 4' N, 88° 8' W). In 2005, the same lines were planted in one field (Environment 3) at Mississippi State. In each environment, a RCB design with four blocks was applied. Plot size was a single row 12 m in length with row spacing of 0.97 m. The planting was a solid row pattern. The stand density consisted of single plants spaced approximately 10 cm apart. Planting dates for environments 1, 2, and 3 were May 11, 2004; May 25, 2004; and May 13, 2005, respectively. Harvest dates for environments 1 and 2 were November 1, 2004 and for environment 3 October 17, 2005. Normal field practices were followed during the season. Prior to machine harvest, a twenty-five boll sample for each plot was hand harvested to determine lint percentage (LP, %). The amount of seed cotton per plot was obtained by machine picking and converted into lint yield per hectare (LY, kg ha^{-1}) accordingly.

Variance components were estimated for two agronomic traits by the MINQE approach (Rao, 1971; Zhu, 1989). The genotypic effects for two agronomic traits were predicted by the adjusted unbiased prediction method (Zhu, 1993). Two jackknife methods were employed to calculate the standard error for each parameter by removal of one replication (block) within each environment from the original data. An approximate t-test was used to test significance of each parameter (Miller, 1974b).

4.2. *Results*

Results in Table 5 showed that variance components for both traits calculated by both jackknife methods were similar. This agrees with our simulation results. Significance levels were higher for the non-pseudo value based jackknife method than for the pseudo value based jackknife method.

Predicted genotypic values for both agronomic traits by the two jackknife methods are summarized in Table 6. Predicted genotypic effects for the 11 parents by the two jackknife methods were very close, while non-pseudo value based method had higher significance levels than pseudo value based method. For example, genotypic effects of P3 and P8 for lint yield predicted by the pseudo value based method were not significant at probability 0.05 while the genotypic effects by the non-pseudo value based method for the same two parents were significant at 0.001.

5. CONCLUSION AND DISCUSSION

The jackknife approach is a powerful technique for bias reduction and interval tests for parameters of interest and its value has been recognized for decades. Usually the pseudo-values are used to calculate the jackknife estimates of a specific parameter (Quenouille, 1956; Miller, 1974a,b). However, with an increase in the number of jackknife times, the pseudo jackknife estimates of the parameters often are beyond the theoretical ranges, particularly for a variance component or a correlation coefficient. On the other hand, when some constraints for some pseudo values are applied, either over- or under-estimations could result. For example, if negative pseudo-values for a variance component are justified to zero, then this variance component could be overestimated. In addition, it could increase the Type I errors.

In this study, we compared two jackknife methods regarding the biases, Type I errors, and powers for variance components estimation. Our simulation results demonstrated that (1) the biases resulted of both jackknife methods are similar; (2) Type I errors for the non-pseudo jackknife method are less than the preset nominal probability levels with a few exceptions; and (3) powers for the non-pseudo jackknife method is higher than for the pseudo jackknife method. The results from an actual data set of two agronomic traits for cotton cultivars also supported the above conclusions. Additional simulations were also conducted for several other mixed linear models and the results were consistent (the results not shown). Therefore, we conclude that the non-pseudo jackknife method is more preferable than the pseudo method regarding variance components estimation and test.

ACKNOWLEDGMENTS

The authors want to take this opportunity to give our special appreciation to Drs. G. Milliken and J. Willers and one reviewer for their statistical suggestion and comments for improving the quality of this manuscript.

REFERENCES

- Daniel, W. W., 1995. Biostatistics: A foundation for analysis in the health sciences. 6th edition. John Wiley & Sons, Inc. New York, NY.
- Davison, A. C., and D. V. Hinkley, 1997. Bootstrap methods and their applications. Cambridge University Press, New York, NY.
- Efron, B., 1979. Bootstrap methods: another look at the jackknife. *Ann. Statist.* 7: 1-26.
- Gray, H. L., and W. R. Schucany, 1972. The Generalized Jackknife Statistics. Marcel Dekker, New York, NY.
- Gray, H. L., T. A. Watkins, and J. E. Admas, 1972. On the jackknife statistic, its extension, and its relation to en-transformation. *Ann. Math. Statist.* 43: 1-30.
- Miller, R. G., 1964. A trustworthy jackknife. *Ann. Math. Statist.* 35: 1594-1605.
- Miller, R. G., 1968. Jackknifing variances. *Ann. Math. Statist.* 39: 567-582.
- Miller, R.G., 1974a. An unbalanced jackknife. *Ann. of Statist.* 2: 880-891.
- Miller, R. G., 1974b. The jackknife-a review. *Biometrika* 61:1-15.
- Quenouille, M. H., 1949. Approximate tests of correlation in time-series. *J. R. Statist. Soc. B* 11: 68-84.
- Quenouille, M. H., 1956. Notes on bias in estimation. *Biometrika* 43 :353-360.
- Rao, C.R., 1971. Estimation of variance and covariance components MINQUE theory. *J. Multivar. Anal.* 1: 257-275.
- Searle, S. R., G. Casella, and C. E. McCulloch, 1992. Variance components. John Wiley and Sons, Inc., New York, NY.
- Shao, J., 1988. On resampling methods for variance and bias estimation in linear models. *Ann Statist.* 16: 986-1008.
- Shao, J. and D. Tu, 1995. The jackknife and bootstrap. Springer-Verlag, New York, NY.
- Tukey, J. W., 1958. Bias and confidence in not-quite large samples (Abstract). *Ann. Math. Statist.* 29: 614.
- Wackerly, D. D., W. S. Mendenhall, and R. L. Scheaffer, 1996. Mathematical statistics with applications, Wadsworth Publishing Company, Belmont, CA.
- Wu, J., J. N. Jenkins, J. C. McCarty, and D. Wu., 2006. Variance component estimation using the additive, dominance, and additive \times additive model when genotypes vary across environments. *Crop Sci.* 46: 174-179.

Zhu, J., 1989. Estimation of genetic variance components in the general mixed model. Ph.D. Dissertation, North Carolina State University, Raleigh, U.S.A. (Diss. Abstr. DA8924291).

Zhu, J., 1993. Methods of predicting genotype value and heterosis for offspring of hybrids. (Chinese). Journal of Biomathematics 8(1): 32-44.

Zhu, J., 1998. Analytical methods for genetic models. Press of China Agriculture, Beijing, China.

Table 1. Biases and Type I errors at three preset probability levels for three variance components for the genotype model by two jackknife methods under two distributions

Normal distribution					
Pseudo					
	True value	Bias	Type I error		
			0.001 [†]	0.010 [†]	0.050 [†]
V_G	0.000	0.410	0.000	0.000	0.004
V_{GE}	0.000	0.674	0.000	0.000	0.002
V_B	0.000	0.247	0.000	0.000	0.000
Non-pseudo					
	True value	Bias	Type I error		
			0.001 [†]	0.010 [†]	0.050 [†]
V_G	0.000	0.410	0.004	0.010	0.022
V_{GE}	0.000	0.674	0.000	0.002	0.012
V_B	0.000	0.247	0.000	0.000	0.006
Uniform distribution					
Pseudo					
	True value	Bias	Type I error		
			0.001 [†]	0.010 [†]	0.050 [†]
V_G	0.000	0.494	0.000	0.002	0.008
V_{GE}	0.000	0.693	0.000	0.000	0.004
V_B	0.000	0.242	0.000	0.000	0.000
Non-pseudo					
	True value	Bias	Type I error		
			0.001 [†]	0.010 [†]	0.050 [†]
V_G	0.000	0.494	0.004	0.010	0.040
V_{GE}	0.000	0.693	0.000	0.006	0.024
V_B	0.000	0.242	0.000	0.000	0.004

[†]: Preset nominal probability level

Table 2. Biases and Type I error rates for five variance components at three different probability levels for the AD model by two jackknife methods under two distributions.

Normal distribution					
Pseudo					
	True value	Bias	Type I error		
			0.001 [†]	0.010 [†]	0.050 [†]
V_A	0.000	0.180	0.000	0.000	0.000
V_D	0.000	1.438	0.000	0.000	0.002
V_{AE}	0.000	0.275	0.000	0.002	0.014
V_{DE}	0.000	2.299	0.000	0.000	0.004
V_B	0.000	0.136	0.000	0.000	0.000
Non-pseudo					
	True value	Bias	Type I error		
			0.001 [†]	0.010 [†]	0.050 [†]
V_A	0.000	0.170	0.000	0.006	0.020
V_D	0.000	1.420	0.000	0.006	0.022
V_{AE}	0.000	0.269	0.002	0.012	0.030
V_{DE}	0.000	2.183	0.000	0.004	0.008
V_B	0.000	0.136	0.000	0.002	0.010
Uniform distribution					
Pseudo					
	True value	Bias	Type I error		
			0.001 [†]	0.010 [†]	0.050 [†]
V_A	0.000	0.186	0.000	0.002	0.010
V_D	0.000	1.589	0.000	0.000	0.008
V_{AE}	0.000	0.289	0.000	0.002	0.008
V_{DE}	0.000	2.211	0.000	0.002	0.004
V_B	0.000	0.118	0.000	0.000	0.000
Non-pseudo					
	True value	Bias	Type I error		
			0.001 [†]	0.010 [†]	0.050 [†]
V_A	0.000	0.176	0.002	0.010	0.028
V_D	0.000	1.563	0.002	0.012	0.024
V_{AE}	0.000	0.291	0.002	0.010	0.038
V_{DE}	0.000	2.083	0.002	0.008	0.020
V_B	0.000	0.118	0.000	0.000	0.006

[†]: Preset nominal probability level

Table 3. Biases and powers at three different probability levels for four variance components for the genotype model by two jackknife methods under two distributions.

Normal distribution					
Pseudo					
	True value	Bias	Power		
			0.001 [†]	0.010 [†]	0.050 [†]
V_G	20	0.277	0.300	0.558	0.736
V_{GE}	20	0.271	0.140	0.556	0.866
V_B	20	-0.044	0.000	0.000	0.096
V_e	20	0.142	0.810	0.994	1.000
Non-pseudo					
	True value	Bias	Power		
			0.001 [†]	0.010 [†]	0.050 [†]
V_G	20	0.277	0.636	0.794	0.864
V_{GE}	20	0.271	0.704	0.920	0.978
V_B	20	-0.044	0.012	0.204	0.610
V_e	20	0.142	1.000	1.000	1.000
Uniform distribution					
Pseudo					
	True value	Bias	Power		
			0.001 [†]	0.010 [†]	0.050 [†]
V_G	20	1.177	0.312	0.632	0.826
V_{GE}	20	-0.404	0.126	0.504	0.882
V_B	20	-0.232	0.000	0.020	0.172
V_e	20	-0.119	0.952	1.000	1.000
Non-pseudo					
	True value	Bias	Power		
			0.001 [†]	0.010 [†]	0.050 [†]
V_G	20	1.177	0.736	0.872	0.930
V_{GE}	20	-0.404	0.692	0.950	0.990
V_B	20	-0.232	0.054	0.342	0.786
V_e	20	-0.119	1.00	1.000	1.000

Table 4. Biases and powers at three different probability levels for six variance components for the AD model by two jackknife methods under two distributions.

Normal distribution					
Pseudo					
	True value	Bias	Power		
			0.001 [†]	0.010 [†]	0.050 [†]
V_A	20.000	1.903	0.512	0.656	0.738
V_D	20.000	0.975	0.052	0.174	0.420
V_{AE}	20.000	-0.018	0.524	0.796	0.896
V_{DE}	20.000	0.603	0.002	0.074	0.256
V_B	20.000	-0.700	0.000	0.004	0.126
V_e	20.000	0.051	0.994	1.000	1.000
Non-pseudo					
	True value	Bias	Power		
			0.001 [†]	0.010 [†]	0.050 [†]
V_A	20.000	1.772	0.680	0.760	0.808
V_D	20.000	0.925	0.250	0.502	0.680
V_{AE}	20.000	-0.024	0.842	0.936	0.970
V_{DE}	20.000	0.739	0.116	0.346	0.596
V_B	20.000	-0.700	0.022	0.236	0.708
V_e	20.000	0.057	1.000	1.000	1.000
Uniform distribution					
Pseudo					
	True value	Bias	Power		
			0.001 [†]	0.010 [†]	0.050 [†]
V_A	20.000	0.675	0.500	0.652	0.752
V_D	20.000	0.154	0.042	0.186	0.394
V_{AE}	20.000	0.099	0.542	0.840	0.936
V_{DE}	20.000	0.172	0.008	0.056	0.250
V_B	20.000	0.398	0.004	0.016	0.220
V_e	20.000	0.015	1.000	1.000	1.000
Non-pseudo					
	True value	Bias	Power		
			0.001 [†]	0.010 [†]	0.050 [†]
V_A	20.000	0.831	0.702	0.792	0.824
V_D	20.000	0.243	0.246	0.496	0.676
V_{AE}	20.000	0.053	0.874	0.956	0.980
V_{DE}	20.000	0.083	0.100	0.366	0.594
V_B	20.000	0.398	0.050	0.366	0.796
V_e	20.000	0.012	1.000	1.000	1.000

Table 5. Estimated variance components, standard errors, and p values for two agronomic traits by two jackknife values under the genotype model.

Parameter	Lint percentage					
	Pseudo based method			Non-pseudo based method		
	Estimate	SE.	P-value	Estimate	SE.	P-value
V_G	3.051	0.300	0.000	3.197	0.095	0.000
V_{GE}	1.017	0.201	0.000	0.578	0.063	0.000
V_e	0.678	0.105	0.000	0.692	0.033	0.000
	Lint yield					
	Pseudo based method			Non-pseudo based method		
	Estimate	SE.	P-value	Estimate	SE.	P-value
V_G	34045	7151	0.001	34132	2252	0.000
V_{GE}	8080	9476	0.412	7818	2984	0.233
V_e	64389	14998	0.001	64396	4723	0.000

Table 6. Predicted genotypic effects (deviated from the population means), standard error, and p values for two agronomic traits by two jackknife values under the genotypic model.

	Lint percentage					
	Pseudo based method			Non-pseudo based method		
	Prediction	SE.	P value	Prediction	SE.	P value
P1	1.322	0.329	0.002	1.337	0.104	0.000
P2	-0.008	0.199	0.969	-0.008	0.063	1.000
P3	0.410	0.160	0.027	0.415	0.051	0.000
P4	-0.895	0.239	0.003	-0.906	0.075	0.000
P5	1.395	0.228	0.000	1.413	0.072	0.000
P6	-2.958	0.153	0.000	-2.994	0.048	0.000
P7	-2.670	0.310	0.000	-2.703	0.098	0.000
P8	-0.991	0.271	0.004	-1.003	0.085	0.000
P9	1.854	0.184	0.000	1.876	0.058	0.000
P10	0.053	0.230	0.822	0.055	0.073	0.999
P11	2.486	0.248	0.000	2.517	0.078	0.000

	Lint yield					
	Pseudo based method			Non-pseudo based method		
	Prediction	SE.	P value	Prediction	SE.	P value
P1	-190.4	51.9	0.000	-176.6	16.3	0.000
P2	-238.4	31.0	0.000	-221.1	9.8	0.000
P3	151.1	73.6	0.060	138.6	23.2	0.000
P4	-24.9	39.5	0.540	-22.8	12.4	0.660
P5	109.0	70.0	0.150	101.6	22.0	0.010
P6	-93.6	33.8	0.020	-86.6	10.7	0.000
P7	-346.7	47.4	0.000	-322.1	14.9	0.000
P8	173.2	91.2	0.080	161.8	28.7	0.000
P9	8.4	89.4	0.930	9.4	28.2	1.000
P10	193.3	83.7	0.040	177.7	26.4	0.000
P11	259.0	85.6	0.010	240.0	27.0	0.000