

Tilburg University

Jackknifing estimated weighted least squares

Kleijnen, J.P.C.; Karremans, P.C.A.; Oortwijn, W.K.; van Groenendaal, W.J.H.

Published in:
Communications in statistics: Part A: Theory and methods

Publication date:
1987

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Kleijnen, J. P. C., Karremans, P. C. A., Oortwijn, W. K., & van Groenendaal, W. J. H. (1987). Jackknifing estimated weighted least squares: JEWLS. *Communications in statistics: Part A: Theory and methods*, 16(3), 747-764.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

JACKKNIFING ESTIMATED WEIGHTED LEAST SQUARES: JEWLS

Jack P.C. Kleijnen
Peter C.A. Karremans
Wim K. Oortwijn
Willem J.H. Van Groenendaal

Department of Information Systems and Auditing
School of Business and Economics
Catholic University Brabant
5000 LE Tilburg
Netherlands

*Key Words and Phrases: regression analysis;
experimental design; variance heterogeneity;
nonnormality; non-linear estimators; confidence
intervals; t statistic.*

ABSTRACT

This paper investigates regression analysis of experimental designs with replications, assuming variance heterogeneity, possibly combined with nonnormality. These replications yield variance estimators which result in Estimated Weighted Least Squares (EWLS). Jackknifing yields confidence intervals for these nonlinear EWLS estimators. The validity

of these confidence intervals is examined in a Monte Carlo experiment. Jackknifed EWS estimators result in better confidence intervals than simple EWS.

1. INTRODUCTION

Although jackknifing is an "old" idea, introduced by Quenouille in 1949, the technique could not become popular until computers became widely available. And even nowadays jackknifing is not much applied. In this paper we apply jackknifing to the linear regression model with unequal error variances. If these variances were known, then Weighted Least Squares (WLS) would yield the Best (minimum variance) Linear Unbiased Estimator (BLUE). In practice these variances are unknown. However, we can easily estimate the error variances in experimental designs with replication (as is the case in simulation experiments in which we are particularly interested). These estimated variances result in the Estimated Weighted Least Squares (EWS) estimator, say $\tilde{\beta}$. This nonlinear estimator is unbiased under mild conditions; see Schmidt (1976, p. 71). And although the EWS estimator has smaller variance than the Ordinary Least Squares (OLS) estimator $\hat{\beta}$, the EWS confidence intervals hold only for large samples, i.e., for more than 25 replications per combination; see Kleijnen et al. (1985). Therefore we shall investigate whether Jackknifed Estimated Weighted Least Squares (JEWS) is a "jewel" indeed, i.e., yields valid confidence intervals.

2. DEFINITION OF JEWS

Consider the linear regression model

$$\tilde{y} = \tilde{X} \tilde{\beta} + \tilde{e} \quad (2.1)$$

where the underscore \sim denotes matrices (including vectors), and $\underline{y} = (y_1, \dots, y_N)'$, $\underline{X} = (x_{i',q})$ with $i' = 1, \dots, N$ and $q = 1, \dots, Q$, $\underline{\beta} = (\beta_1, \dots, \beta_Q)'$ and $\underline{e} = (e_1, \dots, e_N)'$. In experimental designs with replication, each combination i ($i = 1, \dots, n$) is replicated m_i times so that

$$N = \sum_{i=1}^n m_i. \tag{2.2}$$

We assume that $m_i > 2$ so that we have unbiased estimators $\hat{\sigma}_i^2$ of the error variances σ_i^2 :

$$\hat{\sigma}_i^2 = \frac{\sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2}{m_i - 1} \quad (i=1, \dots, n) \tag{2.3}$$

where we rearrange the N elements of the vector \underline{y} into a table with n rows and m_i elements in row i ; if $m_i = m$ (see Table 1 above eq. 2.9) we rearrange \underline{y} into an $n \times m$ matrix with elements y_{ij} ; obviously $\bar{y}_i = \sum_j y_{ij}/m_i$. Consequently, the EWLS estimator is

$$\tilde{\underline{\beta}} = (\underline{X}' \hat{\underline{\Omega}}^{-1} \underline{X})^{-1} \underline{X}' \hat{\underline{\Omega}}^{-1} \underline{y} \tag{2.4}$$

where $\hat{\underline{\Omega}}$ is a diagonal matrix with main-diagonal elements $\hat{\sigma}_1^2, \dots, \hat{\sigma}_1^2, \dots, \hat{\sigma}_n^2, \dots, \hat{\sigma}_n^2$ where $\hat{\sigma}_1^2$ occurs m_1 times, \dots , $\hat{\sigma}_n^2$ occurs m_n times. Using simple (but tedious) linear algebra we can prove that eq. (2.4) reduces to

$$\tilde{\underline{\beta}} = (\bar{\underline{X}}' \hat{\underline{D}}^{-1} \bar{\underline{X}})^{-1} \bar{\underline{X}}' \hat{\underline{D}}^{-1} \bar{\underline{y}} \tag{2.5}$$

where $\bar{\underline{y}} = (\bar{y}_1, \dots, \bar{y}_1, \dots, \bar{y}_n)'$, $\hat{\underline{D}}$ is an $n \times n$ diagonal matrix

with main diagonal elements $\hat{\sigma}_1^2/m_1, \dots, \hat{\sigma}_n^2/m_n$ and $\bar{\tilde{X}}$ is obtained from \tilde{X} by eliminating identical rows. The n different rows of $\bar{\tilde{X}}$ are specified by the experimental design (for example, n equals 2^k in a full factorial design with k factors; the 2^k design yields an $n \times Q$ matrix of independent variables $\bar{\tilde{X}}$ with $Q > k + 1 > n$). The asymptotic covariance matrix of $\tilde{\beta}$ (see Schmidt, 1976) is

$$\Omega_{\tilde{\beta}} = (\bar{\tilde{X}}' \underline{D}^{-1} \bar{\tilde{X}})^{-1} \quad (2.6)$$

where \underline{D} is diagonal with elements σ_i^2/m_i . To obtain confidence intervals for β_q we might replace \underline{D} in eq. (2.6) by $\hat{\underline{D}}$ (also see eq. 4.5). However, for a small number of replications this heuristic does not yield valid confidence intervals; see Kleijnen et al. (1985). Therefore we investigate jackknifing.

In general, jackknifing means that an estimator of some parameter is recomputed after deleting one or more observations; next those observations are again added and a different group of observations (with group size > 1) is deleted, which results in a new value for the estimator, and so on; see Miller (1974), Weber (1986).

We restrict our study to experimental designs with an equal number of replications: $m_i = m$. (If we permitted varying m_i , then it would be wise to replicate combinations with high variances more often; such an approach is investigated in Kleijnen and Van Groendendaal, 1986.) To apply jackknifing we delete replication j (where $j=1, \dots, m$) of each combination i ($i=1, \dots, n$); see column j in Table 1 or element

Table 1: Experimental data

Combination	Independent variables			Dependent variable
	1	...	Q	1 ... j ... m
1	x_{11}	...	x_{1Q}	$y_{11} \dots y_{1j} \dots y_{1m}$
⋮				
i	x_{i1}	...	x_{iQ}	$y_{i1} \dots y_{ij} \dots y_{im}$
⋮				
n	x_{n1}	...	x_{nQ}	$y_{n1} \dots y_{nj} \dots y_{nm}$

$(i-1)m + j$ of \bar{y} . Next we compute the variance estimator analogously to eq. (2.3):

$$\hat{\sigma}_{i(-j)}^2 = \frac{m}{\sum_{\substack{j'=1 \\ j' \neq j}}^m} (y_{ij'} - \bar{y}_{i(-j)})^2 / (m-2) \tag{2.7}$$

where

$$\bar{y}_{i(-j)} = \sum_{j'} y_{ij'} / (m-1) \tag{2.8}$$

These variance estimators yield m different $n \times n$ diagonal matrices $\hat{D}_{\sim j}$ with main - diagonal elements $\hat{\sigma}_{1(-j)}^2 / (m-1), \dots, \hat{\sigma}_{n(-j)}^2 / (m-1)$ where $j=1, \dots, m$. We also have m vectors with averaged responses $\bar{\underline{y}}_{\sim j} = (\bar{y}_{i(-j)})$. The $n \times Q$ matrix $\bar{\underline{X}}$ is not affected by this jackknifing. Hence eq. (2.5) becomes

$$\tilde{\underline{\beta}}_{\sim j} = (\bar{\underline{X}}' \hat{D}_{\sim j} \bar{\underline{X}})^{-1} \bar{\underline{X}}' \hat{D}_{\sim j} \bar{\underline{y}}_{\sim j} \quad (j=1, \dots, m) \tag{2.9}$$

Obviously these m estimators are dependent. Jackknifing proceeds as follows; see Miller (1974), Weber (1986). The original estimator and the m jackknifed estimators are linearly combined in the so-called pseudovalues

$$J_j = m \tilde{\beta} - (m-1) \tilde{\beta}_{-j} \quad (j=1, \dots, m) \quad (2.10)$$

where we suppress the index q of the Q parameters β . Obviously, if the estimators $\hat{\beta}$ and $\tilde{\beta}_{-j}$ are unbiased (as EWLS estimators are, under mild conditions), then the J_j remain unbiased. To derive a confidence interval we compute the traditional variance estimator of the pseudovalue:

$$\hat{\text{var}}(J) = \frac{\sum_{j=1}^m (J_j - \bar{J})^2}{m-1} \quad (2.11)$$

with $\bar{J} = \sum_j J_j / m$, and use the Student approximation

$$t_{m-1} \approx \frac{\bar{J} - \beta}{\{\hat{\text{var}}(J)/m\}^{1/2}}. \quad (2.12)$$

Whether it is correct to use this t approximation, we investigate in the following Monte Carlo experiment. (We shall also briefly discuss a JEWLS variant with only two, instead of m pseudovalue;s; see the end of Section 4.)

3. MONTE CARLO INPUTS

We use the following \bar{X} . Case 1 is a 2^3 full factorial design with main effects only besides the grand mean, i.e., \bar{X} is an orthogonal 8×4 matrix with elements $+1$ and -1 . The

values for the effects β are taken from a simulation study of the Rotterdam harbor (see Kleijnen et al., 1979): $\underline{\beta}' = (-1.42, -0.769, 13,4, -11.508)$. We quantify the degree of variance heterogeneity through

$$H = \frac{\max \sigma_i^2 - \min \sigma_i^2}{\min \sigma_i^2} \tag{3.1}$$

and fix H at 0, 10.83 and 1455 taken from Kleijnen et al. (1985). (If $H = 0$ then we take $\sigma_i^2 = 1$; if $H = 10.83$ then σ_i^2 equals 1, 2, 4, 5, 6, 7, 9, 11.83 respectively; if $H = 1455$ then σ_i^2 equals 93, 228.38, 821.78, 2809.64, 2567.11, 177.78, 15129, 576 respectively.) An increasing H means decreasing relative effects $\beta/(\Sigma \sigma_i^2/n)$. The number of replications m equals 4, 9 and 25 respectively. We study not only normally distributed errors terms but also asymmetric distributions. Erlang distributions have standardized skewness $\eta_3 = \mu_3/\sigma^3$ equal to 2 (exponential distribution), 0.8944 (sum of 5 exponentials) and 0.6325 (sum of 10 exponentials); see Hastings and Peacock (1975). The lognormal distribution has a standardized skewness which varies with the variance; so if $H \neq 0$ then η_3 varies with i where $i=1, \dots, n$; in Table 2 we shall display the standardized skewness averaged over the n combinations of independent variables $\bar{\underline{X}}$. We make all asymmetric distributions have means and variances equal to the means and variances of the corresponding normal distributions.

Case 2 concerns a 2^2 factorial design with $\underline{\beta}' = (1,1,1)$. If $H = 0$ then $\sigma_i^2 = 1$; if $H = 10.38$ then σ^2 equals 1, 4, 8, 11.38; if $H = 1289$ then σ^2 is 1, 200, 600, 1290.15.

		<u>$\alpha = 10\%$</u>		<u>$\alpha = 5\%$</u>		<u>$\alpha = 1\%$</u>	
		<u>EWLS</u>	<u>JEWLS</u>	<u>EWLS</u>	<u>JEWLS</u>	<u>EWLS</u>	<u>JEWLS</u>
		H_0	H'_0	H_0	H'_0	H_0	H'_0
<u>Case 1</u>		<u>Erlang; $\mu_3/\sigma_3 = 0.6325$</u>					
m=4	H=0						
	10		*	*		*	
	1455	*	*				
m=9	H=0						
	10						
	1455						
m=25	H=0						
	10						
	1455	*	*				
<u>Case 2</u>							
m=4	H=0						
	10				*		
	1289		*	*	*	*	*
m=9	H=0						
	10						
	1289						
m=25	H=0						
	10						
	1289						

(continued)

Table 2 (continued)

		<u>$\alpha = 10\%$</u>				<u>$\alpha = 5\%$</u>				<u>$\alpha = 1\%$</u>			
		<u>EWLS</u>		<u>JEWLS</u>		<u>EWLS</u>		<u>JEWLS</u>		<u>EWLS</u>		<u>JEWLS</u>	
		H_0	H'_0	H_0	H'_0	H_0	H'_0	H_0	H'_0	H_0	H'_0	H_0	H'_0
<u>Case 1</u>		<u>Erlang; $\mu_3/\sigma_3 = 0.8944$</u>											
m=4	H=0	*	*			*	*					*	*
	10	*	*										
	1455	*	*			*	*						
m=9	H=0		*										
	10	*	*									*	*
	1455		*			*	*						
m=25	H=0												
	10												
	1455												
<u>Case 2</u>													
m=4	H=0												
	10											*	*
	1289	*	*	*	*			*	*	*	*	*	*
m=9	H=0												
	10							*	*				
	1289	*	*		*	*	*	*	*	*	*	*	*
m=25	H=0							*	*				
	10												
	1289									*	*		

		<u>α = 10%</u>				<u>α = 5%</u>				<u>α = 1%</u>			
		<u>EWLS</u>		<u>JEWLS</u>		<u>EWLS</u>		<u>JEWLS</u>		<u>EWLS</u>		<u>JEWLS</u>	
		H ₀	H ₀ '	H ₀	H ₀ '	H ₀	H ₀ '	H ₀	H ₀ '	H ₀	H ₀ '	H ₀	H ₀ '
<u>Case 1</u>		<u>Erlang μ₃/σ³ = 2</u>											
m=4	H=0	*	*			*	*	*	*	*	*	*	*
	10	*	*	*	*	*	*	*	*	*	*	*	*
	1455	*	*	*	*	*	*	*	*	*	*	*	*
m=9	H=0	*	*	*	*	*	*			*	*		
	10	*	*	*	*	*	*	*	*	*	*	*	*
	1455	*	*	*	*	*	*	*	*	*	*	*	*
m=25	H=0	*	*			*	*			*	*		
	10	*	*			*	*			*	*		
	1455	*	*			*				*	*		
<u>Case 2</u>													
m=4	H=0	*	*	*	*	*	*	*	*	*	*	*	*
	10	*	*	*	*	*	*	*	*	*	*	*	*
	1289	*	*	*	*	*	*	*	*	*	*	*	*
m=9	H=0	*	*		*	*	*						
	10											*	*
	1289	*	*			*	*	*	*	*	*	*	*
m=25	H=0												
	10	*	*	*	*	*	*			*	*	*	*
	1289		*			*	*	*	*	*	*	*	*

(continued)

Table 2 (continued)

		<u>$\alpha = 10\%$</u>		<u>$\alpha = 5\%$</u>		<u>$\alpha = 1\%$</u>	
		<u>EWLS</u>	<u>JEWLS</u>	<u>EWLS</u>	<u>JEWLS</u>	<u>EWLS</u>	<u>JEWLS</u>
		H_0	H'_0	H_0	H'_0	H_0	H'_0
<u>Case 1</u>		<u>Lognormal (average skewness: $\bar{\eta}$)</u>					
m=4	H=0 ($\bar{\eta}=0.6080$)						
	10 (1.4878)	*	*	*	*	*	*
	1455 (30691.2)	*	*	*	*	*	*
m=9	H=0 ($\bar{\eta}=0.6080$)						
	10 (1.4878)	*	*			*	*
	1455 (30691.2)	*	*	*	*	*	*
m=25	H=0 ($\bar{\eta}=0.6080$)						
	10 (1.4878)	*	*			*	*
	1455 (30691.2)	*	*	*	*	*	*
<u>Case 2</u>							
m=4	H=0 ($\bar{\eta}=0.608$)					*	
	10 (1.5203)		*	*		*	*
	1289 (139.067)	*	*	*	*	*	*
m=9	H=0 ($\bar{\eta}=0.608$)					*	
	10 (1.5203)	*	*	*	*	*	*
	1289 (139.067)	*	*	*	*	*	*
m=25	H=0 ($\bar{\eta}= 0.608$)						
	10 (1.5203)	*	*			*	*
	1289 (139.067)	*	*	*	*	*	*

We use a multiplicative random number generator with multiplier 13^{13} and modulus 2^{59} , developed by NAG (Numerical Algorithms Group) in the United Kingdom. We never reset the random number seed. Consequently all results are independent, except for results on the same line in Tables 2 and 3; Tables 2 and 3 use the same responses y (hence these two tables have identical EWLS estimates).

4. MONTE CARLO OUTPUTS

Each Monte Carlo observation requires $n \times m$ independent samples from the error distribution (again see Table 1). These nm observations yield one EWLS estimate $\tilde{\beta}$ (see eq. 2.5) and m estimates $\tilde{\beta}_j$ (with $j=1, \dots, m$; see eq. 2.9) resulting in one JEWLS estimate \bar{J} (see eq. 2.10). The nm responses y_{ij} finally yield one set of Q confidence intervals for β_q (where $q = 1, \dots, Q$), using eq. (2.12).

Now we test if it is correct to base two-sided confidence intervals for the individual parameters β_q on the t statistic. Since we use only the tails of the t distribution, we estimate

$$P\left\{\frac{|\bar{J} - \beta|}{\{\text{var}(J)/m\}^{\frac{1}{2}}} > t_{m-1, \alpha/2}\right\} = \alpha^* \tag{4.1}$$

where we still suppress the index q and we estimate α^* through (say) $\hat{\alpha}$, using Monte Carlo experimentation (see below). We formulate two related null-hypotheses:

$$H_0: E(\hat{\alpha}) = \alpha \text{ versus } H_1: E(\hat{\alpha}) \neq \alpha \tag{4.2}$$

and

$$H_0': E(\hat{\alpha}) < \alpha \text{ versus } H_1': E(\hat{\alpha}) > \alpha. \quad (4.3)$$

where $\hat{\alpha}$ is an unbiased estimator of α^* in eq. (4.1), and α is defined by

$$P\{|t_{m-1}| > t_{m-1, \alpha/2}\} = \alpha \quad (4.4)$$

Obviously H_0 and H_0' require a two-sided and a one-sided test respectively.

The test statistic for H_0 and H_0' is the binomial variable $\hat{\alpha}$ based on 150 Monte Carlo observations "per situation", i.e., per combination of Case 1 or 2 (2^3 or 2^2 design) with a specific variance heterogeneity H , number of replications m , and distribution type; see Table 2.

We could approximate the binomial variable $\hat{\alpha}$ through the normal distribution $N(\hat{\alpha}, \hat{\alpha}(1-\hat{\alpha})/150)$. A problem arises if $\hat{\alpha} = 0$ (which may occur especially if α in eq. 4.1 is small, say, 1%); if $\hat{\alpha} = 0$ then $\text{var}(\hat{\alpha}) = 0$ and H_0 of eq. (4.2) is automatically rejected (not H_0' of eq. 4.3). Therefore we use the normal approximation $N(\alpha, \alpha(1-\alpha)/150)$ where α is specified by H_0 (or H_0').

For α (the error rate used to derive a two-sided confidence interval per parameters β_q) we select the traditional values 1%, 5% and 10%. Because there are Q parameters β we apply the Bonferroni inequality, i.e., we test H_0 and H_0' with a type I error rate of $0.05/Q$ so that the experiment-wise error rate is 0.05 at most; see Miller (1981). So a "situation" yields significantly bad results if at least one

of the Q parameters β_q results in tail behavior significantly deviating from the t distribution.

To compute the JEWLS estimate we also have to compute the EWLS estimate (see eqs. 2.10 and 2.5). So without much extra effort we can test the tail behavior of EWLS; eq. (4.1) becomes

$$P\left\{\frac{|\tilde{\beta} - \beta|}{\{\hat{\text{var}}(\tilde{\beta})\}^{\frac{1}{2}}} > t_{m-1, \alpha/2}\right\} = \alpha \tag{4.5}$$

where we suppress the index q ; $\tilde{\beta}$ is the q^{th} element of $\tilde{\beta}$ in eq. (2.5); $\hat{\text{var}}(\tilde{\beta})$ follows from the asymptotic covariance matrix in eq. (2.6) where we replace \underline{D} by $\hat{\underline{D}}$.

The above reasoning yields Table 2 where an asterisk (*) means that we reject H_0 or H_0' (using an experimentwise error rate of 5%). We interpret Table 2 as follows. In case of normality, JEWLS gives excellent results if there are more than 4 replications ($m = 9$ or 25). The fact that in case 1 (2^3 design) with $m = 4$ H_0' is rejected more often than H_0 , suggests that if the α error is not realized, then the actual error rate tends to be higher than the nominal α value.

Table 2 clearly shows that as the asymmetry increases, JEWLS yields poorer confidence intervals. JEWLS remains better than EWLS.

We also investigate a less computer-intensive JEWLS variant. Instead of deleting a single replication resulting in m pseudovalues (see eq. 2.10) we now delete half the replications (if m is odd we round $m/2$ downwards) which results

in only two pseudovalues ($m = 2$ in eqs. 2.9 through 2.12). Consequently the confidence intervals for β_q are based on a single degree of freedom. So $t_{m-1; \alpha/2}$ is high. It is possible that $\hat{\text{var}}(J)$ compensates; also see eq. (4.1). Actually our results (not displayed) show longer confidence intervals for β_q (when compared to JEWLS based on m pseudovalues). And these longer confidence intervals do not improve the validity of the t statistic; see Table 3.

5. CONCLUSIONS

JEWLS requires more computing than EWLS, but JEWLS yields better confidence intervals. More specifically, in case of normality EWLS yields valid confidence intervals only if the number of replications is "high" (also see Kleijnen et al., 1985); JEWLS requires fewer replications. In case of severe asymmetry, JEWLS performs better than EWLS, but not well enough.

REFERENCES

- Hastings, N.A.J. & Peacock, J.B., (1975). Statistical Distributions; A Handbook for Students and Practitioners. Butterworths & Co., London.
- Kleijnen, J.P.C., Van den Burg, A.J. & Van der Ham, R.T., (1979). Generalization of simulation results: practicality of statistical methods. European Journal of Operational Research, 3, pp. 50-64.
- Kleijnen, J.P.C., Cremers, P. & van Belle, F., (1985). The power of weighted and ordinary least squares with esti-

mated unequal variances in experimental design. Communications in Statistics, Simulation and Computation, B14, no. 1, pp. 85-102.

Kleijnen, J.P.C. & Van Groenendaal, W., (1986). Regression analysis of factorial designs with sequential replication. Catholic University Brabant.

Miller, R.G., (1974). The jackknife - a review. Biometrika, 61, pp. 1-15.

Miller, R.G., (1981). Simultaneous Statistical Inference. Revised second edition, Springer - Verlag, New York.

Schmidt, P., (1976). Econometrics. Marcel Dekker, Inc., New York.

Weber, N.C. (1986). On the jackknife and bootstrap techniques for regression models. Pacific Statistical Congress, edited by I.S. Francis, B.F.J. Manly and F.C. Lam, Elsevier (North-Holland).

Received by Editorial Board member August, 1986; Revised November, 1986.

Recommended by A. M. Kshirsagar; University of Michigan, Ann Arbor, MI.

Refereed Anonymously.