

AD-A119 471

RAND CORP SANTA MONICA CA  
THE RELATION AMONG THE LIKELIHOOD RATIO-, WALD-, AND LAGRANGE MUL--ETC(U)  
APR 82 D F KOHLER  
RAND/P-6756

F/G 12/1

UNCLASSIFIED

NL

1 of 1

2 of 1

3 of 1

4 of 1

5 of 1

6 of 1

7 of 1

8 of 1

9 of 1

10 of 1

11 of 1

12 of 1

13 of 1

14 of 1

15 of 1

16 of 1

17 of 1

18 of 1

19 of 1

20 of 1

21 of 1

22 of 1

23 of 1

24 of 1

25 of 1

26 of 1

27 of 1

28 of 1

29 of 1

30 of 1

31 of 1

32 of 1

33 of 1

34 of 1

35 of 1

36 of 1

37 of 1

38 of 1

39 of 1

40 of 1

41 of 1

42 of 1

43 of 1

44 of 1

45 of 1

46 of 1

47 of 1

48 of 1

49 of 1

50 of 1

51 of 1

52 of 1

53 of 1

54 of 1

55 of 1

56 of 1

57 of 1

58 of 1

59 of 1

60 of 1

61 of 1

62 of 1

63 of 1

64 of 1

65 of 1

66 of 1

67 of 1

68 of 1

69 of 1

70 of 1

71 of 1

72 of 1

73 of 1

74 of 1

75 of 1

76 of 1

77 of 1

78 of 1

79 of 1

80 of 1

81 of 1

82 of 1

83 of 1

84 of 1

85 of 1

86 of 1

87 of 1

88 of 1

89 of 1

90 of 1

91 of 1

92 of 1

93 of 1

94 of 1

95 of 1

96 of 1

97 of 1

98 of 1

99 of 1

100 of 1

101 of 1

102 of 1

103 of 1

104 of 1

105 of 1

106 of 1

107 of 1

108 of 1

109 of 1

110 of 1

111 of 1

112 of 1

113 of 1

114 of 1

115 of 1

116 of 1

117 of 1

118 of 1

119 of 1

120 of 1

121 of 1

122 of 1

123 of 1

124 of 1

125 of 1

126 of 1

127 of 1

128 of 1

129 of 1

130 of 1

131 of 1

132 of 1

133 of 1

134 of 1

135 of 1

136 of 1

137 of 1

138 of 1

139 of 1

140 of 1

141 of 1

142 of 1

143 of 1

144 of 1

145 of 1

146 of 1

147 of 1

148 of 1

149 of 1

150 of 1

151 of 1

152 of 1

153 of 1

154 of 1

155 of 1

156 of 1

157 of 1

158 of 1

159 of 1

160 of 1

161 of 1

162 of 1

163 of 1

164 of 1

165 of 1

166 of 1

167 of 1

168 of 1

169 of 1

170 of 1

171 of 1

172 of 1

173 of 1

174 of 1

175 of 1

176 of 1

177 of 1

178 of 1

179 of 1

180 of 1

181 of 1

182 of 1

183 of 1

184 of 1

185 of 1

186 of 1

187 of 1

188 of 1

189 of 1

190 of 1

191 of 1

192 of 1

193 of 1

194 of 1

195 of 1

196 of 1

197 of 1

198 of 1

199 of 1

200 of 1

201 of 1

202 of 1

203 of 1

204 of 1

205 of 1

206 of 1

207 of 1

208 of 1

209 of 1

210 of 1

211 of 1

212 of 1

213 of 1

214 of 1

215 of 1

216 of 1

217 of 1

218 of 1

219 of 1

220 of 1

221 of 1

222 of 1

223 of 1

224 of 1

225 of 1

226 of 1

227 of 1

228 of 1

229 of 1

230 of 1

231 of 1

232 of 1

233 of 1

234 of 1

235 of 1

236 of 1

237 of 1

238 of 1

239 of 1

240 of 1

241 of 1

242 of 1

243 of 1

244 of 1

245 of 1

246 of 1

247 of 1

248 of 1

249 of 1

250 of 1

251 of 1

252 of 1

253 of 1

254 of 1

255 of 1

256 of 1

257 of 1

258 of 1

259 of 1

260 of 1

261 of 1

262 of 1

263 of 1

264 of 1

265 of 1

266 of 1

267 of 1

268 of 1

269 of 1

270 of 1

271 of 1

272 of 1

273 of 1

274 of 1

275 of 1

276 of 1

277 of 1

278 of 1

279 of 1

280 of 1

281 of 1

282 of 1

283 of 1

284 of 1

285 of 1

286 of 1

287 of 1

288 of 1

289 of 1

290 of 1

291 of 1

292 of 1

293 of 1

294 of 1

295 of 1

296 of 1

297 of 1

298 of 1

299 of 1

300 of 1

301 of 1

2

AD A119471

THE RELATION AMONG THE LIKELIHOOD RATIO-, WALD-, AND LAGRANGE  
MULTIPLIER TESTS AND THEIR APPLICABILITY TO SMALL SAMPLES

Daniel F. Kohler

April 1982

**DTIC**  
EXTRACTED  
SEP 22 1982  
i-1

DTIC FILE COPY

**DISTRIBUTION STATEMENT A**  
Approved for public release;  
Distribution Unlimited

P-6756

82 09 22 067

### The Rand Paper Series

Papers are issued by The Rand Corporation as a service to its professional staff. Their purpose is to facilitate the exchange of ideas among those who share the author's research interests; Papers are not reports prepared in fulfillment of Rand's contracts or grants. Views expressed in a Paper are the author's own, and are not necessarily shared by Rand or its research sponsors.

The Rand Corporation  
Santa Monica, California 90406

## 1. INTRODUCTION

The Lagrange multiplier test (LM), the likelihood ratio test (LR), and the Wald test (W), are frequently proposed as alternative means for testing parametric restrictions in a linear regression model. Since they all converge to the same limiting Chi-square distribution they have asymptotically the same power characteristics. Consequently the three tests are considered equivalent alternatives for large samples, and the choice among them is usually made on convenience grounds (e.g., ease of estimation).

However, Savin (1976) and Berndt and Savin (1977), among others, have shown that the numerical value of the LM test statistic is always less than that of LR, which in turn is less than that of W. This implies that if we use the same critical value, as seems indicated by the fact that all three tests converge to the same limiting Chi-square distribution, the Wald test will reject the null hypothesis most often, while the LM test will reject it least often. This difference in rejection probabilities raises the possibility of conflicting conclusions from the three tests.

In this paper we show that the three tests are monotonic functions of each other. This implies that they have the same power characteristics. If the critical values are specified such as to equate the probabilities of type I error, the probabilities of type II error will be equal as well, and conflicting results are impossible.

The critical values of the three tests are related to each other by the same monotonic functions as the three test statistics. The "conflicts" pointed out in the literature arise only when the exact

critical values for each test are replaced by the asymptotically justified critical value of a Chi-square distribution. The extent to which this critical value differs from the exact one varies from test to test.

Determining the exact critical values for finite samples is somewhat difficult, because the exact finite sample distributions of the LM and LR test are generally not known. However, under the assumption of normality, the W test can be transformed into an F-test by applying the standard degrees of freedom correction. Furthermore, since the LR and LM tests are functionally related to the W test, they can be transformed into the same test statistic. We can thus conduct an exact finite sample test, based either on the LM, LR or W statistic without running the risk of obtaining conflicting results.



Approved for DTIC DSIC Unannounced Justification		✓
By _____		
Distribution of Availability Code		
Dist	Special	
A		

2. THE FORM OF THE TESTS

Consider the model:

$$(1) \quad Y = X\beta + \varepsilon$$

$$(2) \quad \varepsilon \sim N(0, \sigma^2 I)$$

where  $X$  is of rank  $k$ .

Let  $\hat{\beta}$  and  $\hat{\sigma}^2$  be the maximum-likelihood estimators obtained by unconstrained maximization of the likelihood function, and  $\tilde{\beta}$  and  $\tilde{\sigma}^2$  the estimators obtained with the constraint  $r = R\tilde{\beta}$  imposed. Furthermore, we shall need an estimator for the Lagrange multiplier ( $\hat{\mu}$ ) and the ratio of the unconstrained to constrained maxima of the likelihood function ( $\lambda$ ). The three test statistics can be written as:

$$(3) \quad LR = -2 \log \lambda$$

$$(4) \quad W = [(r - R\hat{\beta})'(R(X'X)^{-1}R')^{-1}(r - R\hat{\beta})]/\hat{\sigma}^2$$

$$(5) \quad LM = [\hat{\mu}(R(X'X)^{-1}R')\hat{\mu}]/\hat{\sigma}^2$$

where  $R$  is the matrix of constraints and is of rank  $q$ .

To simplify the notation let  $A = R(1/n X'X)^{-1}R'$ . From the first-order conditions for maximizing the likelihood function subject to the constraint we can obtain an expression for  $\hat{\mu}$ :

$$(6) \quad \hat{\mu} = (R(X'X)^{-1}R')^{-1}(r - R\hat{\beta})$$

We can now rewrite Eqs. (4) and (5) as

$$(7) \quad W = n[(r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})]/\hat{\sigma}^2$$

$$(8) \quad LM = n[(r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})]/\tilde{\sigma}^2 .$$

This implies that we can express the LM test as

$$(9) \quad LM = \frac{\hat{\sigma}^2}{\tilde{\sigma}^2} W .$$

If we make use of the fact that the likelihood ratio  $\lambda = (\hat{\sigma}^2/\tilde{\sigma}^2)^{n/2}$  we can express Eq. (9) as

$$(10) \quad LM = \lambda^{2/n} \cdot W .$$

Equation (10) makes it obvious that the LM and W test statistics are asymptotically equivalent.

To express LR as a function of W consider  $\tilde{\sigma}^2 \equiv (Y - X\tilde{\beta})'(Y - X\tilde{\beta})/n$  and  $\tilde{\beta} = \hat{\beta} + n(X'X)^{-1}R'A^{-1}(r - R\hat{\beta})$ . Then, letting  $e = Y - X\hat{\beta}$ , we get by substitution:

$$(11) \quad \begin{aligned} \tilde{\sigma}^2 &= 1/n\{e'e - 2e'X(1/nX'X)^{-1}R'A^{-1}(r - R\hat{\beta}) \\ &\quad + (r - R\hat{\beta})'A^{-1}n[R(X'X)^{-1}X'X(X'X)^{-1}R']A^{-1}(r - R\hat{\beta})\} \\ &= \hat{\sigma}^2 + (r - R\hat{\beta})'A^{-1}(r - R\hat{\beta}) . \end{aligned}$$

We can, therefore, express Eq. (7) as

$$(12) \quad W = n(\tilde{\sigma}^2 - \hat{\sigma}^2)/\hat{\sigma}^2 = n(\tilde{\sigma}^2/\hat{\sigma}^2 - 1)$$

which is equivalent to

$$(13) \quad W = n(\lambda^{-2/n} - 1) .$$

Combining Eqs. (13) and (10) leads to

$$(14) \quad LM = n(1 - \lambda^{2/n}) .$$

Let  $\theta \equiv \lambda^{-2/n}$  which is larger or equal to unity with probability one. Then, from Eqs. (3), (13), and (14) the three tests can be rewritten as

$$(15) \quad LR = n \log \theta$$

$$(16) \quad W = n(\theta - 1)$$

$$(17) \quad LM = n(1 - 1/\theta) .$$

Thus, since  $(1 - 1/\theta) < \log \theta < \theta - 1$  we have the well-known inequality relationship

$$(18) \quad LM \leq LR \leq W .$$

3. THE THREE TESTS AS TRANSFORMATIONS OF THE F-TEST<sup>\*</sup>

If we divide Eq. (4) by  $q$  we get

$$(19) \quad W/q = [(r - R\hat{\beta})' (R(X'X)^{-1}R')^{-1} (r - R\hat{\beta})] / \hat{\sigma}^2 \cdot q$$

which is distributed as  $F(q, n - k)$ . From Eqs. (15) and (16) we can express  $W/q$  as a function of  $LR$ , i.e.,

$$(20) \quad W/q = n \cdot [\exp(LR/n) - 1] / q$$

which also follows an  $F(q, n - k)$  distribution. With Eqs. (19) and (20) we can thus transform either  $W$  or  $LR$  into an  $F$  statistic of which we know the exact finite sample distribution. This allows us to conduct exact finite sample tests.

Much of the justification of the LM test is based on the fact that it relies exclusively on parameter estimates from the restricted model. That makes it particularly attractive in situations where the unrestricted model is difficult or impossible to estimate. In the previous section we have given a formula (Eq. (9)) by which we could compute the  $W$  test statistic from the LM statistic. But it relies on the unconstrained estimate  $\hat{\sigma}^2$  which defeats the major advantage of the LM test. We have to find an alternative way of calculating an  $F$  statistic from the LM statistic.

Recall that the LM test is based on maximizing the likelihood function, subject to the constraint, by the Lagrange method and then

---

<sup>\*</sup> Throughout this section we assume that the F-test is defined. In particular we require that  $n - k > 0$ .

testing how close the estimated Lagrange multipliers ( $\hat{u}$ ) are to zero. We have shown that under the assumptions of the linear model this is equivalent to testing

$$(21) \quad LM = n[(r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})]/\tilde{\sigma}^2 .$$

But note that we can express  $\tilde{\sigma}^2$  as a function of the unconstrained parameters. In particular we have (see Eq. (11))

$$\tilde{\sigma}^2 = \hat{\sigma}^2 + (r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})$$

so that we can express the LM statistic as

$$(22) \quad LM = n \cdot [(r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})]/[\hat{\sigma}^2 + (r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})] .$$

Dividing LM by n and inverting results in

$$(23) \quad \frac{n}{LM} = 1 + \frac{\hat{\sigma}^2}{(r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})}$$

which can be transformed into

$$(24) \quad \left[ \frac{n}{LM} - 1 \right]^{-1} = \frac{(r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})}{\hat{\sigma}^2} = \frac{W}{n} .$$

For finite samples we can use the statistic

$$\frac{n}{q} \left[ \frac{n}{LM} - 1 \right]^{-1} = \frac{[(r - R\hat{\beta})'(R(X'X)^{-1}R')^{-1}(r - R\hat{\beta})]}{\hat{\sigma}^2 \cdot q} = \frac{W}{q}$$

which follows an F distribution with (q, n - k) degrees of freedom.

Like the W and LR test statistics, the LM statistic can thus be

adjusted for sample size and degrees of freedom to result in a test statistic applicable for finite samples. As it does not require estimating the unconstrained model, it is particularly useful in situations where only the constrained version can readily be estimated.

#### 4. CONCLUSIONS

An intuitively appealing analogy for the relationship between the three tests can be drawn from trigonometry. Testing linear restrictions can be compared to evaluating the angle  $\alpha$  in Fig. 1.

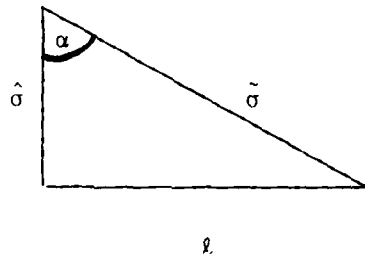


Fig. 1

The length  $l$  is equal to the square root of  $(r - R\hat{\beta})'(R(X'X)^{-1}R')^{-1}(r - R\hat{\beta})$ . The W test statistic corresponds to  $n \cdot [\text{tangent } (\alpha)]^2$ , and the LM statistic to  $n \cdot [\text{sine } (\alpha)]^2$ . Since in a triangle the sine of  $\alpha$  is always less than the tangent of  $\alpha$ , this is yet another way of demonstrating  $LM < W$ .

The likelihood ratio  $\lambda$  corresponds to  $\{[\text{cosine } (\alpha)]^{n/2}\}^2$ . Accordingly, the LR test statistic is  $-n \cdot \log[\text{cosine } (\alpha)]^2$ . We can now use the basic trigonometric identities to express any one of the tests as a function of the other two. In particular, we know that

---

\* Note that this quantity can also be calculated from the restricted regression estimate of  $\hat{u}$  (see Eq. (6)).

tangent ( $\alpha$ ) = sine ( $\alpha$ )/cosine ( $\alpha$ ) which, if raised to the power two, corresponds to Eq. (10).

Continuing with the trigonometric analogy, we see that Eq. (11) corresponds to the theorem of Pythagoras. Thus, if we have an estimate of any two of the three sides of the triangle in Fig. 1, either from the unconstrained regression ( $\hat{\sigma}, \hat{\beta}$ ) or from the constrained one ( $\hat{\sigma}, \hat{\mu}$ ), we can calculate the third side and then conduct any one of the three tests.

Note that  $\hat{\sigma}$  and  $(r - R'\hat{\beta})$  are orthogonal to each other. In statistical terms this means that they are independently distributed. We can thus compute the exact small sample distribution of  $n \cdot [\text{tangent } (\alpha)]^2$ , which turns out to be a multiple of an F distribution. For the other two tests this is not directly possible, because the components going into the calculation of sine ( $\alpha$ ) and cosine ( $\alpha$ ) are not independent of each other. However, by expressing these tests as deterministic functions of tangent ( $\alpha$ ), e.g., through the trigonometric relations described above, we can obtain test statistics with known finite sample distributions as well.

### References

- Aitchison, J. and Silvey, S. C. (1958), "Maximum-Likelihood Estimation of Parameters Subject to Restraints," Annals of Mathematical Statistics, 29, 813-828
- Berndt, E. and Savin, N. E. (1977), "Conflict Among Criteria for Testing Hypothesis in the Multivariate Linear Regression Model," Econometrica, 45, 1263-1278
- Breusch, T. S. (1979), "Conflict Among Criteria for Testing Hypothesis: Extension and Comments," Econometrica, 47, 203-207
- Breusch, T. S. and Pagan, A. R. (1980), "The Lagrange Multiplier Test and Its Application to Model Specification in Economics," Review of Economic Studies, XLVII, 239-253
- Hardy, G. H., Littlewood, J. E. and Polya, G. (1967), Inequalities, Cambridge University Press
- Kendall, M. G. and Stuart, A. (1973), The Advanced Theory of Statistics, Vol. 2, 3rd edition, Hafner, New York
- Kohler, D. F., "Conflict Among Testing Procedures in a Linear Regression Model," Wayne Economic Papers No. 81, Wayne State University, Detroit, 1979
- Pearson, C. E. (ed) (1974), Handbook of Applied Mathematics, Van Nostrand Reinhold Co., New York
- Savin, N. E. (1976), "Conflict Among Testing Procedures in a Linear Regression Model With Autoregressive Disturbances," Econometrica, 44, 1303-1315
- Silvey, S. D. (1959), "The Lagrange Multiplier Test," Annals of Mathematical Statistics, 30, 389-407
- Wald, A. (1943), "Tests of Hypotheses Concerning Several Parameters When the Number of Observations is Large," Transactions of the American Mathematical Society, 54, 476-482

DATA  
FILM  
0-8