TR/17/85 May 1985

Properties of statistical inference procedures for a gamma regression model.

by

A. M. Al-Abood and D. H. Young

Properties of Statistical Inference Procedures For A Gamma Regression Model

by

D.H. Young and A.. Al-Abood

SUMMARY

A gamma regression model with an exponential link function for the means is considered. Approximations to the percentiles of the distributions of the maximum likelihood and weighted least squares estimators of the regression coefficients are presented and evaluated for the case of a single explanatory variable. These are used to develop approximate confidence interval and hypothesis testing pr-cedures for the regression coefficients which are assessed by simulation. Finally, the null distribution properties of goodness of fit tests for the exponential link function are investigated.

CONTENTS

1. Introduction

Let Y_1, $Y_2,\ldots,$ Y_g represent g independent gamma random variables where Y_i has density

$$f_i(y) = \frac{r_i}{\mu_i}\left(\frac{r_i y}{\mu_i}\right)^{ri-1} \frac{\exp(-r_i y_i/\mu_i)}{(r_i - 1)!}, \quad y > 0 \tag{1.1}$$

where the shape parameter r_i is a known positive integer. The mean $\mu_i.$ is assumed to depend on the values $x_{i1}.$, $x_{i2}.$, ..., x_{ik} for k explanatory variables through the model

$$\mu i = \exp\ (x_i\beta)\ , \quad i = 1,\ldots\ldots g \tag{1.2}$$

where $x_i = (1, x_{i1},\ldots, x_{ik})$ and $\beta' = (\beta_0, \beta_1,\ldots, \beta_k$ is a vector of regression coefficients with unknown values.

This model has a number of important applications. For example, consider accelerated life testing (Lawless (1976)) in which there are g groups of items, the ith group containing n_i items and having values $x_{11}.$,...,x_{ik} for k test variables. Suppose that there is type II censoring within groups and let $Y_{i(1)}. < Yi_{(2)}, < \ldots < Yi_{(ri)}$ represent the observed order statistics in the ith group, the remaining $n_i - r_i$ observations being right censored at the value $Yi_{(ri)}.$, ... If the underlying distributions are exponential and the means are given by (1.2), then the statistics

$$Yi = r_i^{-1}\left\{\sum_{j=1}^{ri} Y_{i(j)} + (n_i - r_i)Y_i(r_i)\right\}, \qquad i = 1,\ldots, g \tag{1.3}$$

are distributed independently with p.d.f's given by (1.1). The $\{Y_i\}$ are sufficient unbiased and minimum variance estimators of $\{\mu_i.\}$ and provide the basic observations for a likelihood analysis.

The model also arises in connection with the analysis of intervals between events in g Poisson processes, where the rates λ_i - depend on explanatory variables. If the observation $Y_i.$ for the ith process denotes the time from the origin to the r_ith following event, then Y_i has p.d.f. given by (1.1) with $\mu_i = r_i\lambda_i^{-1}$. The model with a single explanatory variable is considered by Cox and Lewis (1966) in the analysis of trend in a single Poisson process.

Two well-known methods of estimation of the regression coefficients are maximum likelihood and weighted least squares. The ML estimates are given by the solution of the $k + 1$ equations

$$\sum_i r_i x_{ir} y_i \exp(-x_i' \hat{\beta}) = \sum_i r_i x_{ir}, \quad r = 0,1,\ldots, k \tag{1.4}$$

and an iterative solution is required. The WLS method (Kahn (1979)) yields the non-iterative solution

$$\hat{\beta}_W = (x' D_W x)^{-1} X' D^W z \tag{1.5}$$

where $\underset{\sim}{D}_W = \text{diag}\{\psi'(r_1), \ldots, \psi'(r_2)\}, z_i = \log y_i + \log r_i - \psi(r_i)$, $\underset{\sim}{X}$ is the design matrix and where $\psi(\cdot)$, $\psi'(\cdot)$ are the digamma and trigamma functions, respectively.

Asymptotically, the ML and WLS have the same efficiency as $r_i. \to \infty$, $i = 1, \ldots g$, but for small values of the $\{r_i\}$, some loss of efficiency does occur in using WLS. Abood and Young (1985) contrast the moment properties of the estimators for small to moderate values of the $\{r_i\}$ and propose modifications to the ML estimators leading to bias reduction and improved mean square error efficiency.

In this report, we consider statistical inference procedures based on the ML and WLS estimates of. $\underset{\sim}{\beta}$ The main objective is to assess the performance of the procedures by means of a large scale simulation investigation for the case of a single explanatory variable. In Sections 2 and 3, approximations to the percentiles of the distributions of the ML and WLS estimators of $\underset{\sim}{\beta}$ for the case when $\underset{\sim}{\beta} = \underset{\sim}{0}$ are presented. Monte Carlo results assessing the percentile approximations are given in section 4. Confidence interval estimation procedures are discussed in section 5 and test procedures for hypotheses concerning the regression coefficients are evaluated in section 6. Finally, null distribution properties of goodness of fit statistics for testing the assumption of the link function given by (1.2) are presented in section 7.

2. Approximations To The Distribution Of The MLEstimator

Under the gamma regression model, we have

$$E\left(- \frac{\partial^2 L}{\partial \beta_r \partial \beta_s} \right) = \sum_i r_i x_{ir} x_{is} = I_{rs} \quad \text{say,} \tag{2.1}$$

where L denotes the log-likelihood. This leads to the well-known first order expression for the covariance matrix of $\underset{\sim}{B}$

$$\text{cova} \quad \hat{\underset{\sim}{\beta}} = (\underset{\sim}{X}' \underset{\sim}{D} \underset{\sim}{X})^{-1} = \underset{\sim}{I}^{-1} \tag{2.2}$$

where $\underset{\sim}{D} = \text{diag}(r_1, r_2, \ldots, r_g)$ and where in the usual notation, $\underset{\sim}{I} = ((Irs))$ denotes the information matrix. For 'large' values of the $\{r_i\}$, we have the ordinary normal approximation

$$\hat{\beta}_r \text{ approx } N(\beta_r, I^{rr}), \quad r = 0,1,\ldots, k \tag{2.3}$$

where I^{rs} denotes the element in the(r+1)st row and (s+1)st column of $\underset{\sim}{I}^{-1}$. The ML estimators are asymptotically unbiased but will show some bias for small values of $\{r_i\}$. Abood and Young (1985) show that the biases $b_r = E(\hat{\beta}_r) - \beta_r$ are to order R^{-1}, where $R = \sum_i r_i$, given by

$$b_r = -\frac{1}{2} \sum_{s=0}^{k} \sum_{t=0}^{k} \sum_{u=0}^{k} I^{rs} i^{tu} k_{stu}, \quad r = 0,1,\ldots, k \tag{2.4}$$

where $K_{rst} = \sum_i r_i x_{ir} x_{is} x_{it}$. Using a normal approximation with bias correction we take

$$\hat{\beta}_r \underset{\sim}{\text{approx}} \quad N(\beta_r + b_r, I^{rr}), \quad r = 0,1,\ldots, k \tag{2.5}$$

Statistical inferences for the β_r based on the ML estimators $\hat{\beta}_r$ are made using the distribution properties of the random variables $\hat{\beta}_r - \beta_r$ r = 0,1,..,k. These random variables are distributed independently of $\underset{\sim}{\beta}$" and we have the exact distribution result

$$\hat{\beta} - \beta_r \overset{d}{\sim} \hat{\beta}_r^{(0)}, \quad r = 0,1,\ldots, k \tag{2.6}$$

where $\hat{\beta}_r^{(0)}$ denotes the ML estimator of βr when $\underset{\sim}{\beta} = \underset{\sim}{0}$

Approximations to the percentiles of the distribution of $\hat{\beta}_r^{(0)}$ are needed for both hypotheses testing and confidence interval estimation. If we let $b_r(\alpha)$ denote the 100α percentile of the distribution of $\hat{\beta}_r^{(0)}$ the ordinary normal approximation is

$$b_r(\alpha) \approx u_\alpha (I^{rr})^{1/2}, \qquad r = 0,1,\ldots,k \tag{2.7}$$

where u_α denotes the 100α percentile of the N(0,1) distribution. The normal approximation with bias correction is

$$b_r(\alpha) \approx b_r + u_\alpha (I_{rr})^{1/2}, \qquad r = 0,1,\ldots,k \tag{2.8}$$

3. Approximations To The Distribution Of The WLS Estimator

The WLS estimators are linear functions of log-gamma random variables and hence have the advantage over the ML estimators that their moments are known exactly for all values of the $\{r_i\}$. In particular, the covariance matrix is given by

$$\mathrm{cov}\left(\hat{\underset{\sim}{\beta}}_w\right) = \left(\underset{\sim}{x}' \underset{\sim}{D} \; L \; \underset{\sim}{X}\right)^{-1} = \underset{\sim}{V}^{-1} \text{ say} \tag{3.1}$$

where

$$Vrs = \sum_i x_{ir} x_{is} / \psi'(r_i) \tag{3.2}$$

Using an ordinary normal approximation, we take

$$\hat{\beta}_{wr} \underset{\sim}{d} N(\beta_r, V^{rr}). \tag{3.3}$$

Writing $\hat{\beta}_{wr} = \sum_i a_{ir} Z_i$, then from akhn (1979) the skewness and kurtosis coefficients of $\hat{\beta}_{wr}$ are

$$\gamma_{ir} = \frac{\sum_i (\pm)_{air}{}^3 \psi^{(2)}(r_i)}{\left\{\sum_i a_{ir}{}^2 \psi^{(1)}(r_i)\right\}^{3/2}}, \qquad \gamma_{2r} = \frac{\sum_i a^4{}_{ir} \psi^{(3)}(r_i)}{\left\{\sum_i a^2{}_{ir} \psi(1)(r_i)\right\}^2} \tag{3.4}$$

where $\Psi(s)(\cdot)$ denotes the sth derivative of the digamma function and $(\pm)$ is the sign of a_{ir}. These coefficients may be used in an Edgeworth series representation to provide skewness and kurtosis corrections to the normal approximation to the c.d.f. of $\hat{\beta}_{wr}$.

For the random variables $\hat{\beta}_{wr} - \beta r'$ r = 0,1,...,k we have the exact distribution result

$$\hat{\beta}_{wr} - \beta_r \underset{\sim}{d} \hat{\beta}_{wr}^{(0)} \tag{3.5}$$

Where $\hat{\beta}_{wr}^{(0)}$ denotes the WLS estimator of β_r when $\underset{\sim}{\beta} = \underset{\sim}{0}$. Denoting the 100α Percentile of the distribution of $\hat{\beta}_{wr}^{(0)}$ by $b_{wr}(\alpha)$, the ordinary normal approximation is

$$b_{wr}(\alpha) \approx u_{\alpha} (V_{rr})^{1/2}, \qquad r = 0,1,....., \quad k. \tag{3.6}$$

The normal approximation with skewness correction is

$$b_{wr}(\alpha) \approx (V^{rr})^{1/2} \left\{ u_{\alpha} + \frac{1}{6} (u_{\alpha}^{2} - 1)\gamma_{1,r} \right\}, \quad r = 0,1,............k \tag{3.7}$$

and the normal approximation with skewness and kurtosis correction is

$$b_{wr}(\alpha) \approx (V_{rr})^{1/2} \left\{ u_{\alpha} + \frac{1}{6} (u_{\alpha}^{2} - 1)\gamma_{1r} + \frac{1}{24} (u_{\alpha}^{3} - 3u_{\alpha})\gamma_{2r} \right\} \quad r = 0,1,.....k \tag{3.8}$$

4. <u>Monte Carlo Results</u>

In order to assess the various approximations to the percentiles of the distributions of, $\hat{\beta}_r^{(0)}$ and $\hat{\beta}_{wr}^{(0)}$ we consider the particular case when there is a single explanatory variable x, the means being given by

$$\mu_i = = \exp(\beta_0 + \beta_1 x_1 \quad , \qquad i = 1,...,g \quad . \tag{4.1}$$

Without loss of generality we suppose that the x values are centred such that $\sum_i r_i x_i = 0$ In this case the results for the estimators take on particularly simple forms. For the ML estimators we have

$$\mathrm{Var}_a\left(\hat{\beta}_0\right) R^{-1}, \qquad \mathrm{var}_a\left(\hat{\beta}_1\right) = \left(\sum_i r_i x_i^2\right)^{-1} \tag{4.2}$$

$$b_o = -R^{-1}, \qquad b_i = -\frac{1}{2}\left(\sum_i rix_i^3\right) \Big/ \left(\sum_i r_i x_i^2\right)^2 \tag{4.3}$$

The WLS estimators are

$$\hat{\beta}_{wo} = \frac{\sum_i \{z_i/\psi\psi(r_i)\}}{\sum_i \{1/\psi/(r_i)\}}, \qquad \hat{\beta}_{w1} = \frac{\sum_1 \{x_i z_i/\psi\psi(r_i)\}}{\sum_i \{x_1^2/\psi\psi(r_i)\}} \tag{4.4}$$

with

$$\mathrm{var}(\hat{\beta}_{wo}) = D^{-1}\sum_i x_1^2/\psi'(r_i), \qquad \mathrm{var}(\hat{\beta}_{w1}) = D^{-1}\sum_i \{1/\psi/(r_i)\} \tag{4.5}$$

where

$$D = \sum_i \{1/\psi/(r_i)\}\left\{\sum_i x_1^2/\psi\psi(r_i)\right\} - \left\{\sum_i xi/\psi i\,(r_i)\right\}^2 \tag{4.6}$$

The skewness and kurtosis coefficients of the WLS estimators are then

$$\gamma_{10} = \frac{\sum_i \psi^{(2)}(r_i)/\left\{\psi^{(1)}(ri\}\right\}^3}{\sum_i 1/\psi^{(1)}(r_i)^{3/2}}, \qquad \gamma_{20} = \frac{\sum_i \psi^{(3)}(r_i)/\{\psi(1)(r_i)\}^4}{\left\{\sum_i 1/\psi/\psi(1)_i)\right\}^2} \tag{4.7}$$

$$\gamma_{11} = \frac{\sum_i x\frac{3}{1}\psi^{(2)}(r_i)/\left\{\psi^{(1)}(r_i)\right\}^3}{\left\{\sum_i x\frac{2}{1}/\psi^{(1)}(r_i)\right\}^{3/2}}, \quad \gamma_{21} = \frac{\sum_i x\frac{4}{1}\psi^{(3)}(r_i)/\left\{\psi^{(1)}(r_i)\right\}^4}{\left\{\sum x\frac{2}{1}/\psi^{(1)}(r_i)\right\}^2} \tag{4.8}$$

In order to assess the accuracy of the percentile approximations for $\hat{\beta}_r$ and $\hat{\beta}_{wr}$' a large scale simulation investigation was made for the case when the explanatory variable x has equally spaced values with $x_i = i - \frac{1}{2}(g+1)$, i = 1,...,g. Equal alues for the shape parameter were taken with r_i = r = 1(1)10(2)20(5)50 for i = 1,...,g with g = 5,10 and r_i = r = 6(1)10 for g = 6(1)9. A simulation run-size of 4000 was used.

For the given values of x, we have $\sum_i x\frac{3}{1} = 0$ and hence the approximating bias b_1 and the skewness coefficient γ_{11} are both zero. Hence the bias correction and skewness correction approximations will only apply to $b_0(\alpha)$ and $b_{wo}(\alpha)$, respectively.

The broad conclusions from the investigation are

i) For the ML estimator $\hat{\beta}_0^{(0)}$, the use of the bias correction gives a marked improvement in the approximation to the percentiles.

ii) The ordinary normal approximation tends to underestimate the upper percentiles and to overestimate the lower percentiles of the distribution of $\hat{\beta}_1^{(0)}$ for r = 1,2 but gives satisfactory results for larger values of r.

iii) For the WLS estimator $\hat{\beta}_{wo}^{(0)}$ the use of the skewness correction to the normal approximation appears to be worthwhile for small values of r and small values of the tail probabilities.

iv) The ordinary normal approximation to the percentiles of the distribution of $\hat{\beta}_{w1}^{(0)}$ work satisfactorily for all values of r.

These findings are illustrated in tables 1 and 2 which gives the upper and lower percentiles respectively for $\hat{\beta}_0^{(0)}$ and $\hat{\beta}_1^{(0)}$ and in tables 3 and 4 which give the upper and lower percentiles respectively for $\hat{\beta}_{wo}^{(0)}$ and $\hat{\beta}_{w1}^{(0)}$, for the cases r = 1(1)6(2)10, g = 5 and α = 0.10,0.05,0.01.

Table 1

Upper percentiles of th ML estimators $\hat{\beta}_0^{(0)}$ and $\hat{\beta}_1^{(0)}$ when g = 5 and x. $=1-\frac{1}{2}$ (g+1), i - 1,...,g. Actual tail probabilities associated with the percentiles are shown in parentheses.

r	1	2	3	4	5	6	8	10
b_0(0.90)	0.374	0.304	0.267	0.234	0.208	0.195	0.179	0.160
App(2.7)	0.573	0.405	0.331	0.287	0.256	0.234	0.203	0.181
	(0.040)	(0.049)	(0.060)	(0.065)	(0.064)	(0.068)	(0.076)	(0.076)
App(2.8)	0.373	0.305	0.264	0.237	0.216	0.201	0.178	0.161
	(0.101)	(0.100)	(0.103)	(0.098)	(0.093)	(0.095)	(0.102)	(0.098)
b_0(0.95)	0.524	0.402	0.344	0.315	0.277	0.263	0.241	0.212
App(2.7)	0.736	0.520	0.425	0.368	0.329	0.300	0.260	0.233
	(0.014)	(0.020)	(0.023)	(0.027)	(0.026)	(0.029)	(0.037)	(0.036)
App(2.8)	0.536	0.420	0.358	0.318	0.289	0.267	0.235	0.213
	(0.047)	(0.043)	(0.046)	(0.048)	(0.044)	(0.047)	(0.054)	(0.049)
b_{o0}(0.99)	0.792	0.619	0.505	0.449	0.402	0.373	0.343	0.311
App(2.7)	1.040	0.736	0.601	0.520	0.465	0,425	0.368	0.329
	(0.0013)	(0.0040)	(0.0025)	(0.0043)	(0.0043)	(0.0048)	(0.0060)	(0.0070)
App(2.8)	0.840	0.636	0.534	0.470	0.425	0.391	0.343	0.309
	(0.0075)	(0.0085)	(0.0063)	(0.0075)	(0.0083)	(0.0083)	(0.0125)	(0.0108)
b_1 (0.90)	0.457	0.308	0.234	0.209	0.178	0.161	0.148	0.131
App(2.7)	0.405	0.287	0.234	0.203	0.181	0.166	0.143	0.128
	(0.124)	(0.114)	(0.100)	(0.107)	(0.097)	(0.094)	(0.109)	(0.105)
b1(0.95)	0.591	0.393	0.303	0.265	0.235	0.216	0.191	0.162
App(2.7)	0.520	0.368	0.300	0.260	0.233	0.212	0.184	0.165
	(0.074)	(0.060)	(0.051)	(0.053)	(0.052)	(0.052)	(0.056)	(0.048)
b_1(0.99)	0.919	0.572	0.454	0.379	0.326	0.297	0.272	0.235
App(2.7)	0.736	0.520	0.425	0.368	0.329	0.300	0.260	0.233
	(0.0253)	(0.0178)	(0.0148)	(0.0125)	(0.0095)	(0.0095)	0.0128	(0.0113)

Table 2

Lower percentiles for the ML estimators $\hat{\beta}_0^{(0)}$ and $\hat{\beta}_1^{(0)}$ when $g = 5$ and $x_i = i - \frac{1}{2}(g+1)$, $i = 1,...,g$.

r	1	2	3	4	5	6	8	10
$b_0(0.10)$	-0.859	-0.536	-0.413	-0.345	-0.305	-0.271	-0.234	-0.207
App(2.7)	-0.573	-0,405	-0.331	-0.287	-0.256	-0.234	-0.203	-0.181
	(0.229)	(0.178)	(0.159)	(0.152)	(0.148)	(0.142)	(0.139)	(0.134)
App(2.8)	-0.773	-0.505	-0.398	-0.337	-0.296	-0.267	-0.288	-0.201
	(0.131)	(0.166)	(0.112)	(0.106)	(0.107)	(0.106)	(0.107)	(0.106)
$b_0(0.05)$	-1.061	-0.676	-0.510	-0.439	-0.384	-0.347	-0.292	-0.261
App(2.7)	-0.736	-0.520	-0.425	-0.368	-0.329	-0.300	-0.260	-0.233
	(0.145)	(0.106)	(0.093)	(0.087)	(0.085)	(0.079)	(0.078)	(0.075)
App(2.8)	-0.936	-0.620	-0.491	-0.418	-0.369	-0.334	-0.285	-0.253
	(0.078)	(0.066)	(0.059)	(0.060)	(0.058)	(0.058)	(0.053)	(0.057)
$b_0(0.01)$	-1.466	-0.956	-0.727	-0.633	-0.545	-0.491	-0.422	-0.367
App(2.7)	-1.040	-0.736	-0.601	-0.520	-0.465	-0.425	-0.368	-0.329
	(0.0540)	(0.0340)	(0.0258)	(0.0260)	(0.0230)	(0.0218)	(0.0205)	(0.0185)
App(2.8)	-1.240	-0.836	-0.667	-0.570	-0.505	-0.458	-0.393	-0.349
	(0.025)	(0.019)	(0.016)	(0.017)	(0.016)	(0.016)	(0.014)	(0.014)
$b_1(0.10)$	-0.456	-0.308	-0.245	-0.208	-0.188	-0.171	-0.141	-0.129
App(2.7)	-0.405	-0.287	-0.234	-0.203	-0.181	-0.166	-0.143	-0.128
	(0.125)	(0.117)	(0.116)	(0.107)	(0.107)	(0.109)	(0.098)	(0.100)
$b_1(0.05)$	-0.601	-0.399	-0.320	-0.260	-0.242	-0.222	-0.187	-0.169
App(2.7)	-0.520	-0.368	-0.300	-0.260	-0.233	-0.212	-0.184	-0.165
	(0.075)	(0.065)	(0.063)	(0.050)	(0.059)	(0.057)	(0.053)	(0.053)
$b_1(0.01)$	-0.890	-0.610	-0.471	-0.397	-0.366	-0.317	-0.270	-0.239
App(2.7)	-0.736	-0.520	-0.425	-0.368	-0.329	-0.300	-0.260	-0.233
	(0.023)	(0.019)	(0.016)	(0.014)	(0.015)	(0.015)	(0.012)	(0.013)

Table 3

Upper percentiles of the WLS estimators $\hat{\beta}_{w0}^{(0)}$ and $\hat{\beta}_{w1}^{(0)}$ when $g = 5$ and $x_i = i - \frac{1}{2}(g + 1)$, $i = 1,\ldots,g$.

r	1	2	3	4	5	6	8	10
$b_{w0}(0.90)$	0.676	0.441	0.355	0.294	0.256	0.237	0.207	0.184
App(3.6)	0.735	0.460	0.360	0.305	0.270	0.244	0.209	0,186
	(0.082)	(0.089)	(0.094)	(0.093)	(0.087)	(0.094)	(0.095)	(0.098)
App(3.7)	0.704	0.447	0.352	0.299	0.265	0.240	0.206	0.184
	(0.092)	(0.097)	(0.104)	(0.096)	(0.091)	(0.096)	(0.104)	(0.100)
$b_{w0}(0.95)$	0.854	0.559	0.439	0.380	0.331	0.303	0.264	0.235
	0.944	0.591	0.462	0.392	0.346	0.313	0.268	0.239
App(3.6)	(0.033)	(0.039)	(0.042)	(0.046)	(0.043)	(0.042)	(0.045)	(0.047)
App(3.7)	0.860	0.555	0.440	0.376	0.334	0.303	0.261	0.233
	(0.049)	(0.051)	(0.049)	(0.054)	(0.048)	(0.050)	(0.053)	(0.051)
$b_{w0}(0.99)$	1.160	0.774	0.590	0.520	0,457	0.417	0.369	0.330
App(3.6)	1.334	0.835	0.654	0.554	0.489	0.443	0.380	0.337
	(0.0018)	(0.0055)	(0.0043)	(0.0065)	(0.0065)	(0.0060)	(0.0080)	(0.0083)
App(3.7)	1.291	0.743	0.596	0.513	0.457	0.416	0.360	0.322
	(0.0075)	(0.0120)	(0.0095)	(0.0115)	(0.0100)	(0.0108)	(0.0115)	(0.0123)
$b_{w1}(0.90)$	0.505	0.319	0.239	0.213	0.183	0.166	0.151	0.131
App(3.6)	0.520	0.362	0.255	0.216	0.191	0.173	0.148	0.132
	(0.094)	(0.093)	(0.086)	(0.098)	(0.092)	(0.090)	(0.103)	(0.100)
$b_{w1}(0.95)$	10.684	0.410	0.314	0.271	0.238	0.214	0.193	0.165
App(3.6)	0.667	0.418	0.327	0.277	0,245	0.222	0.190	0.169
	(0.053)	(0.047)	(0.044)	(0.047)	(0.046)	(0.045)	(0.053)	(0.047)
b_{w1} (0.99)	1.014	0.606	0.455	0.388	0.338	0.302	0.277	0.239
App(3.6)	0.943	0.591	0.462	0.392	0,346	0.313	0.268	0.239
	(0.0160)	(0.0115)	(0.0098)	(0.0095)	(0.0093)	(0.0078)	(0.0125)	(0.0100)

Table 4

Lower percentiles of the WLS estimators $\hat{\beta}_{w0}^{(0)}$ and $\hat{\beta}_{w1}^{(0)}$ when g = 5 and $x_i = i - \frac{1}{2}(g+1)$, i = 1,... ,g,

r	1	2	3	4	5	6	8	10
$b_{wo}(0.10)$	-0.744	-0.478	-0.362	-0.307	-0.275	-0.246	-0.215	-0.190
App(3.6)	-0.735	-0.460	-0.360	-0.305	-0.270	-0.244	-0.209	-0.186
	(0.109)	(0.107)	(0.106)	(0.101)	(0.106)	(0.104)	(0.109)	(O.105)
App(3.7)	-0.767	-0.474	-0.369	-0.312	-0.275	-0.248	-0.212	-0.188
	(0.096)	(0.101)	(0.097)	(0.099)	(0.100)	(0.097)	(0.104)	(0.103)
b_{wo} (0.05)	-1.008	-0.609	-0,476	-0.401	-0.361	-0.326	-0.274	-0.257
App(3.6)	-0.944	-0.591	-0.462	-0.392	-0.346	-0.313	-0.268	-0.239
	(0.062)	(0.058)	(0.054)	(0.055)	(0.058)	(0.057)	(0.053)	(0.055)
App(3,7)	-1.027	-0.626	-0.485	-0.408	-0.359	-0.324	-0.276	-0.245
	(0.046)	(0.046)	(0.047)	(0.047)	(0.051)	(0.052)	(0.048)	(0.051)
$b_{wo}(0.01)$	-1.477	-0.929	-0.675	-0.599	-0.505	-0.477	-0.400	-0.368
App(3.6)	-1.334	-0.835	-0.654	-0.554	-0.489	-0.443	-0.380	-0.337
	(0.0168)	(0.0173)	(0.0138)	(0.0135)	(0.0128)	(0.0143)	(0.0130)	(0.0138)
App(3.7)	-1.549	-0.928	-0.711	-0.596	-0.522	-0.470	-0.399	-0.353
	(0.0073)	(0.0100)	(0.0088)	(0.0100)	(0.0088)	(0.0105)	(0.0100)	(0.0115)
$b_{w1}(0.10)$	-0.486	-0.318	-0.242	-0.214	-0.192	-0.180	-0.144	-0.131
App(3.6)	-0.520	-0.326	-0.255	-0.216	-0.191	-0.173	-0.148	-0.132
	(0.087)	(0.093)	(0.097)	(0.097)	(0.101)	(0.111)	(0.093)	(0,099)
b_{w1} (0.05)	-0.648	-0.419	-0.339	-0.272	-0.244	-0.232	-0.191	-0.171
App(3.6)	-0.667	-0.418	-0.327	-0.277	-0.245	-0,222	-0.190	-0.169
	(0.046)	(0.051)	(0.056)	(0.047)	(0.049)	(0.058)	(0.051)	(0.052)
b_{w1} (0.01)	-0.937	-0.630	-0.476	-0.412	-0.360	-0.327	-0.269	-0.240
App(3.6)	-0.943	-0.591	-0.462	-0. 39 2	-0.346	-0.313	-0.268	-0.239
	(0.0090)	(0.0128)	(0.0110)	(0.0110)	(0.0118)	(0.0125)	(0.0103)	(0.0103)

5. Confidence Interval Estimation

In this section, we consider confidence interval estimation for the regression coefficients $\{\beta_j\}$, based on the use of the ML and WLS estimates.

If the exact percentiles of the distributions of $\hat{\beta}_j$ andwere $\hat{\beta}_{wj}$ were known, confidence intervals with confidence coefficients equal to the nominal confidence coefficients $1-\alpha$ could be found. The two-sided central $100(1-\alpha)\%$ confidence intervals for β_j. based on the ML and WLS estimators would be given by

$$\left\{\hat{\beta}_j - b_j(1-\tfrac{1}{2}\alpha),\ \hat{\beta}_j - b_j(\tfrac{1}{2}\alpha)\right\},\quad \left\{\hat{\beta}_{wj} - b_{wj}(1-\tfrac{1}{2}\alpha),\ \hat{\beta}_{wj} - b_{wj}(\tfrac{1}{2}\alpha)\right\} \tag{5.1}$$

respectively. For one-sided $100(1-\alpha)\%$ intervals, the lower confidence bounds for β_j- would be

$$\hat{\beta}_j - b_j(\alpha)\ ,\ \hat{\beta}_{wj} - b_j(\alpha) \tag{5.2}$$

respectively. Upper confidence bounds could be found similarly.

Since the exact distributions of $\hat{\beta}_j$ and $\hat{\beta}_{wj}$ are unknown, approximate confidence intervals may be found using the percentile approximations developed in sections 2 and 3. Based on the ML estimate, the <u>ordinary normal</u> approximating confidence interval for B_J is given by

$$\left\{\hat{\beta}_j - u_{1-\frac{1}{2}\alpha}(I^{jj})^{1/2},\ \hat{\beta}_j + u_{1-\frac{1}{2}\alpha}(I^{jj})^{1/2}\right\} \tag{5.3}$$

With bias correction, the approximate interval is

$$\hat{\beta}_j - b_j - u_{1-\frac{1}{2}\alpha}(I_{jj})^{1/2},\ \hat{\beta}_j - b_j + u1_{-\frac{1}{2}\alpha}\ (I^{jj})^{1/2} \tag{5.4}$$

Similarly, using the WLS estimate the ordinary normal approximating confidence interval for β_j. is

$$\hat{\beta}_{wj} - u_{1-\frac{1}{2}\alpha}(V^{jj})^{1/2}\ ,\ \hat{\beta}_{wj} + u_{1-\frac{1}{2}\alpha}(V^{jj})^{1/2} \tag{5.5}$$

With skewness correction, the interval becomes

$$\hat{\beta}_{wj} - (y_{jj})^{1/2}\left\{(u_{1-\frac{1}{2}\alpha} + \frac{1}{6}\gamma_j(u^2_{1-\frac{1}{2\alpha}} - 1)\right\}, \hat{\beta}_{wj} + (V_{jj})^{1/2}\left\{u_{1-\frac{1}{2}\alpha} + \frac{1}{6}\gamma_{1j}(u^2{}_{1-\frac{1}{2}\alpha} - 1)\right\}$$

(5.6)

where γ_{1j} . is given by (3.4). Kurtosis correction could also be applied but will not be considered here as the correction led to little improvement in the percentile approximations.

For one-sided 100(1-α)% confidence intervals, the ordinary normal approximations to the upper confidence bounds are

$$\hat{\beta}_j + u_{1-\alpha}(I^{jj})^{1/2}, \qquad \hat{\beta}_{wj} + u_{1-\alpha}(V^{jj})^{1/2} \tag{5.7}$$

for ML and WLS estimation, respectively. With bias correction for the ML estimators and skewness correction for the WLS estimators, the approximate upper confidence bounds are

$$\hat{\beta}_j - b_j + u_{1-\alpha}(I^{jj})^{1/2}, \quad \hat{\beta}_{wj} + (v^{jj})^{1/2}\left\{u_{1-\alpha} + \frac{1}{6}\gamma_{1j}(u^2_{1-\alpha} - 1)\right\} \tag{5.8}$$

respectively.

When assessing properties of approximating confidence intervals two properties are particularly important, namely their average width and the deviation of the actual confidence coefficients from their nominal values 1-α. In the present case, the widths of the confidence intervals are non-random and for estimation of $ß_j$. are

$$W_{ij} = 2u_{1-\frac{1}{2}\alpha}(I^{jj})^{1/2}, W_{zj} = 2u_{1-\frac{1}{2}\alpha}(v_{jj}), W_{3j} = 2(v^{jj})^{1/2} \ \{u_{1-\frac{1}{2}\alpha} + \frac{1}{6}\gamma_{ij}(u^2_{1-\frac{1}{2}\alpha} - 1)\}$$

for ML estimation, for WLS estimation without skewness correction and for WLS estimation with skewness correction, respectively. The ratios of the widths of the confidence intervals are therefore

$$W_{2j}/w_{1j} = (V^{jj}/I^{jj})^{1/2}, w_{3j}/w_{1j} = (v^{jj}/I^{jj})^{1/2} \ \{1 + \frac{1}{6}\gamma_{1j}(u_{1-\frac{1}{2}\alpha} - u^{-1}_{1-\frac{1}{2}\alpha})\} \tag{5.9}$$

for j = 0,1 respectively.

We now consider the special case of a single explanatory variable with r_i = r, i - 1,….,g and $\sum_{i=1}^{g} x_i = 0$. We have

$$\frac{W_{20}}{w_{10}} = \frac{W_{21}}{w_{11}} = \frac{w_{31}}{w_{11}} = \{r\psi r\ (r)\}^{1/2} \tag{5.10}$$

and

$$w_{30}/w_{10} = \{r\psi\ '(r)\}^{1/2}\ \{1 + \frac{1}{6}\psi^{(2)}\ (r)(u_{1-\frac{1}{2}\alpha} - u^{-1}_{1-\frac{1}{2}\alpha})\ /\ [g^{1/2}\ \{\psi^{(1)}\ (r)\}^{3/2}\ ]\ . \tag{5.11}$$

Table 5 shows values of $\{r\psi'(r)\}^{1/2}$ and W_{30}/W_{10} for r = 1(1)4(2)8(4)20, g = 5,10 and α = 0.10, 0.05, 0.01. The results show that for r ≤ 5, the widths of the confidence intervals based on WLS without skewness correction are considerably larger than those based on ML estimation. This property is much less marked when skewness correction is applied
and when a is very small WLS leads to a small reduction in the widths of the confidence intervals.

Table 5

Ratio of widths of approximate confidence intervals for $ß_0$ and $ß_1$ based on ML and WLS estimation

r	$\{r\psi'(r)\}^{\frac{1}{2}}$	W_{30}/W_{10} : g = 5 α = 0.10	0.05	0.01	W_{30}/W_{10} : g = 10 α = 0.10	0.05	0.01
1	1 .28	1.17	1.12	1.04	1.20	1.17	1.11
2	1.14	1.07	1.04	0.99	1.09	1.07	1.03
3	1.09	1.04	1.02	0.98	1.05	1.04	1.01
4	1.07	1.02	1.00	0.97	1.03	1.02	1.00
6	1.04	1.01	1.00	0.97	1.02	1.01	0.99
8	1.03	1.00	0.99	0.97	1.01	1.00	0.99
12	1.02	1.00	0.99	0.97	1.00	1.00	0.99
16	1.02	1.00	0.99	0.97	1.00	1.00	0.99
20	1.01	0.99	0.99	0.98	1.00	1.00	0.99

Values of the actual confidence coefficients as estimated by the simulation investigation described in section 4 are shown in tables 6 and 7 for estimation of $ß_0$ and $ß_1$, respectively, for the two-sided case with nominal confidence coefficients 1-α = 0.90, 0.95, 0.99. Tables 8 and 9 show the corresponding coefficients for the one-sided case. For $ß_0$, bias and skewness corrections are examined but for $ß_1$ these corrections are zero under the given configuration of values for the single explanatory variable.

The broad conclusions reached from the results in tables 6-9 are as follows.

(i) For two-sided confidence interval estimation of β_0 , the ordinary normal approximation based on ML leads to confidence coefficients systematically smaller than the nominal values, and the bias correction leads to a worthwhile improvement in the control of the confidence coefficient. Using WLS estimation skewness correction does not lead to any improvement, and the overall control of the confidence coefficient is better for WLS than for ML.

(ii) For two-sided confidence interval estimation of β_1 , the procedure based on WLS estimation gives a slightly better performance than that based on ML estimation, for very small values of r. In general, both methods provide, very satisfactory results.

(iii) For one-sided confidence interval estimation of β_0, the procedure based on the ordinary ML estimator leads to confidence coefficients which are much larger than the nominal values, particularly for small values of r. With bias correction, excellent control over the confidence coefficients is obtained. For the WLS procedures, skewness correction leads only to a marginal improvement, but the control of the confidence coefficient is very good.

(iv) For one-sided confidence interval estimation of β_1 the WLS estimation procedure has a slightly better performance than ML for $r = 1,2$. For larger values of r, both methods provide excellent control over the confidence coefficient.

Table 6

Estimated confidence coefficients for approximate 100(1-α)% two-sided confidence intervals for $ß_0$ based on (i) ordinary ML, (ii) ML with bias correction, (iii) ordinary WLS, (iv) WLS with skewness correction.

	g = 5							
1-α=0.90 r	1	2	3	4	6	8	12	20
(i)	0.841	0.875	0.884	0.886	0.892	0.885	0.892	0.896
(ii)	0.876	0.892	0.896	0.893	0.896	0.893	0.894	0.897
(iii)	0.906	0.908	0.906	0.900	0.901	0.898	0.904	0.895
(iv)	0.895	0.895	0.895	0.892	0.895	0.886	0.895	0.894
1-α=0.95								
(i)	0.901	0.924	0.937	0.938	0.943	0.943	0.943	0.944
(ii)	0.932	0.942	0.949	0.946	0.947	0.946	0.951	0.948
(iii)	0.953	0.954	0.957	0.951	0.950	0.946	0.953	0.947
(iv)	0.932	0.935	0.942	0.939	0.941	0.941	0.943	0.942
1-α=0.99								
(i)	0.966	0.977	0.982	0.982	0.985	0.986	0.987	0.988
(ii)	0.980	0.983	0.988	0.986	0.989	0.987	0,989	0.988
(iii)	0.990	0.988	0.991	0.989	0.990	0.988	0.990	0.989
(iv)	0.971	0.976	0.980	0.980	0.982	0.984	0.986	0.987

	g=10							
1-α=0.99 r	1	2	3	4	6	8	12	20
(i)	0.872	0.884	0.891	0.884	0.898	0.890	0.887	0.897
(ii)	0.885	0.890	0.895	0.892	0.898	0.895	0.888	0.897
(iii)	0.906	0.900	0.905	0.898	0.900	0.897	0.893	0.894
(iv)	0.926	0.910	0.907	0.898	0.902	0.894	0.888	0.894
1-α=0.95								
(i)	0.924	0.937	0.942	0.944	0.946	0.945	0.944	0.944
(ii)	0.938	0.946	0.947	0.945	0.951	0.948	0.945	0.946
(iii)	0.955	0.953	0.958	0.946	0.951	0.949	0.942	0.948
(iv)	0.961	0.954	0.953	0.948	0.947	0.946	0.943	0.944
1-α=0.99								
(i)	0.978	0.982	0.985	0.986	0.987	0.987	0.987	0.988
(ii)	0.988	0.985	0.989	0.987	0.989	0.989	0.990	0.989
(iii)	0.989	0.991	0.991	0.987	0.989	0.990	0.991	0.989
(iv)	0.990	0.984	0.986	0.986	0.987	0.986	0.985	0.987

Table 7

Estimated confidence coefficients for approximate 100(1-α)% two—sided confidence intervals for $ß_1$ based on (i) ML, (ii) WLS

g = 5

1-α	r	1	2	3	4	6	8	12	20
0.90	(i)	0.852	0.875	0.881	0.898	0.882	0.891	0.896	0.894
	(ii)	0.901	0.903	0.900	0.907	0.898	0.897	0.896	0.897
0.95	(i)	0.912	0.930	0.935	0.944	0.943	0.944	0.943	0.947
	(ii)	0.946	0.951	0.951	0.954	0.949	0.951	0.946	0.948
0.99	(i)	0.969	0.978	0.983	0.985	0.989	0.986	0.988	0.988
	(ii)	0.987	0.985	0.988	0.988	0.990	0.988	0.989	0.988

g = 10

1- α	r	1	2	3	4	6	8	12	20
0.90	(i)	0.872	0.891	0.892	0.892	0.892	0.898	0.897	0.904
	(ii)	0.902	0.904	0.901	0.901	0.895	0.904	0.899	0.903
0.95	(i)	0.929	0.946	0.941	0.944	0.940	0.944	0.951	0.948
	(ii)	0.946	0.949	0.950	0.951	0.946	0.947	0.951	0.951
0.99	(i)	0.978	0.985	0.987	0.988	0.988	0.988	0.989	0.989
	(ii)	0.989	0.989	0.988	0.989	0.990	0.988	0.990	0.989

Table 8

Estimated confidence coefficients for approximate 100(1-α)% one-sided confidence intervals for β_0 based on (i) ordinary ML, (ii) ML with bias correction, (iii) ordinary WLS, (iv) WLS with skewness correction

g = 5

1-α=0.90	r 1	2	3	4	6	8	12	20
(i)	0.960	0.951	0.941	0.935	0.933	0.925	0.922	0.916
(ii)	0.900	0.901	0.897	0.902	0.906	0.899	0.898	0.902
(iii)	0.918	0.911	0.904	0.907	0.907	0.899	0.898	0.903
(iv)	0.908	0.904	0.896	0.904	0.904	0.896	0.894	0.900
1-α=0.95								
(i)	0.986	0.981	0.977	0.973	0.971	0.963	0.969	0.959
(ii)	0.954	0.957	0.954	0.952	0.954	0.947	0.953	0.949
(iii)	0.968	0.961	0.958	0.954	0.958	0.950	0.957	0.948
(iv)	0.951	0.949	0.951	0.946	0.950	0.944	0.952	0.945
1-α=0.99								
(i)	0.999	0.996	0.998	0.996	0.995	0.994	0.993	0.992
(ii)	0.993	0.992	0.995	0.993	0.992	0.990	0.992	0.988
(iii)	0.998	0.995	0.996	0.994	0.994	0.991	0.992	0.989
(iv)	0.988	0.988	0.991	0.989	0.989	0.988	0.991	0.988

g = 10

1-α=0.90	r 1	2	3	4	6	8	12	20
(i)	0.949	0.934	0.936	0.924	0.921	0.917	0.910	0.908
(ii)	0.907	0.903	0.905	0.398	0.896	0.900	0.893	0.897
(iii)	0.912	0.909	0.908	0.903	0.901	0.898	0.896	0.898
(iv)	0.905	0.906	0.904	0.900	0.898	0.897	0.894	0.897
1-α=0.95								
(i)	0.977	0.973	0.970	0.962	0.966	0.962	0.955	0.952
(ii)	0.953	0.951	0.954	0.946	0.953	0.951	0.947	0.949
(iii)	0.961	0.958	0.958	0.951	0.959	0.954	0.947	0.951
(iv)	0.949	0.953	0.953	0.946	0.954	0.951	0.945	0.947
1-α=0.99								
(i)	0.998	0.996	0.995	0.994	0.994	0.993	0.993	0.993
(ii)	0.993	0.992	0.992	0.990	0.992	0.990	0.990	0.991
(iii)	0.994	0.998	0.994	0.992	0.992	0.991	0.991	0.990
(iv)	0.987	0.991	0.991	0.988	0.990	0.989	0.988	0.989

Table 9

Estimated confidence coefficients for approximate 100(1-α)% one-sided confidence intervals for ß$_1$ based on (i) ML, (ii) WLS

g = 5

1-α. \ r		1	2	3	4	6	8	12	20
0.90	(i)	0.876	0.886	0.900	0.893	0.906	0.888	0.893	0.902
	(ii)	0.904	0.907	0.914	0.902	0.910	0.895	0.900	0.905
0.95	(i)	0.926	0.940	0.949	0.947	0.949	0.944	0.949	0.948
	(ii)	0.947	0.953	0.957	0.954	0.955	0.947	0.952	0.948
0.99	(i)	0.975	0.982	0.988	0.988	0.991	0.987	0.991	0.989
	(ii)	0.984	0.989	0.990	0.991	0.992	0.988	0.992	0.990

g = 10

1-a \ r		1	2	3	4	6	8	12	20
0.90	(i)	0.881	0.894	0.904	0.896	0.897	0.894	0.897	0.897
	(ii)	0.904	0.903	0.904	0.894	0.903	0.899	0.899	0.900
0.95	(i)	0.932	0.944	0.953	0.946	0.947	0.948	0.944	0.951
	(ii)	0.950	0.949	0.955	0.949	0.950	0.951	0.947	0.948
0.99	(i)	0.982	0.987	0.991	0.989	0.990	0.987	0.990	0.987
	(ii)	0.988	0.989	0.991	0.990	0.991	0.990	0.990	0.989

6. Tests Of Hypotheses Concerning The Regression Coefficients

In regression problems we are often interested in testing the hypothesis that a particular subset of the explanatory variables have no effect. Without loss of generality, we shall take the subset to contain the last k-ℓ variables, so that we wish to test the hypothesis H_0: $ß_j = 0$ for $j = \ell+1, \ell+2, \ldots, k$. We shall write $\underset{\sim}{\beta}' = (\underset{\sim}{\beta}'_1, \underset{\sim}{\beta}'_2$ where

$$\underset{\sim}{\beta}'_i = (\beta_0, \beta_1, \ldots \beta_\ell), \qquad \underset{\sim}{\beta}'_2 = (\beta_{\ell+1}, \beta_{\ell+2}, \ldots, \beta_k) \,. \tag{6.1}$$

When H_0 : $\underset{\sim}{\beta}_2 = \underset{\sim}{0}$ is true ,we have

$$\mu_i = \exp(\underset{\sim}{x}_i \underset{\sim}{\beta}_i), \quad i = 1,\ldots\ldots, g \tag{6.2}$$

which, we shall refer to as the restricted model. The model $\mu_i = \exp(\underset{\sim}{x}_i \underset{\sim}{\beta})$, $i = 1,..,g$ will be called the full model. We firstly develop tests based on the maximum likelihood estimators.

The log-likelihood under the restricted model is

$$L(\underset{\sim}{\beta}) = c - \sum_{i=1}^{g} ri \ (\log \ \mu_i + \mu_i^{-1} y_i) \tag{6.3}$$

where

$$c = \sum_{i=1}^{g} r_i \log r_i + \sum_{i=1}^{g} (r_i - 1)\log y_i - \sum_{i=1}^{g} \log (r_i - 1)!$$

is a constant not depending on $\underset{\sim}{\beta}$ We let $\hat{\underset{\sim}{\beta}}$ and $\hat{\mu} = \exp(\underset{\sim}{x}_i \hat{\underset{\sim}{\beta}})$ denote the ML estimates of $\underset{\sim}{\beta}$ and μ_i under the full model. The estimates are given by the solution of the k+1 equations

$$\sum_{i=1}^{g} r_i x_{ij} y_i \ \exp(-\underset{\sim}{x}'_i \ \hat{\underset{\sim}{\beta}}) = \sum_{i=1}^{g} r_i x_{ij}, \quad j = 0,1,\ldots.., \ k. \tag{6.4}$$

Under the restricted model, the ML estimate $\hat{\underset{\sim}{\beta}}_1$ is given by the solution of the ℓ+1 equations

$$\sum_{i=1}^{g} r_1 x_{ij} y_i \exp(-\underset{\sim}{x}' i \ \hat{\underset{\sim}{\beta}}) = \sum_{i=1}^{g} r_i x_{ij}, \quad j = 0,1,\ldots\ldots, \ell \tag{6.5}$$

Setting $\hat{\mu}_{10} = \exp(\underset{\sim}{x}'_i \ \hat{\underset{\sim}{\beta}}_1)$, i = 1,...,g, the likelihood ratio statistic for comparing the full and restricted models is

$$S_1 = 2 \sum_{i=1}^{g} r_i \ \underset{\sim}{x}'_i \left\{ (\hat{\underset{\sim}{\beta}}_1 - \hat{\underset{\sim}{\beta}}) + y_i(\hat{\mu}_{10}^{-1} - \hat{\mu}_i^{-1}) \right\}. \tag{6.6}$$

The statistics S_1 is taken to be approximately distributed as x^2 with k-ℓ degrees of freedom if H_o is true.

The statistic S_1 takes on a particularly simple form when the x's are centred such that $\sum_{i=1}^{g} r_i x_{ij} = 0$ for j = 1,....,k. Under this condition use of the first likelihood equation in (6.4) gives

$$\sum_i r_i y_i \hat{\mu}_i^{-1} = R, \qquad L(\hat{\underset{\sim}{\beta}}) = c - R(\hat{\beta}_0 + 1) \tag{6.7}$$

Similarly, use of the first likelihood equaton in (5.5) gives

$$\sum_i r_i y_i \hat{\mu}_{i0}^{-1} = R, \qquad L(\hat{\underset{\sim}{\beta}}_1) = c - R(\hat{\beta}_0 + 1) \tag{6.8}$$

where $\hat{\beta}_n$ denotes the ML estimate of $ß_0$ under the restricted model.

Hence we may write

$$S_1 = 2R(\hat{\beta}_0 - \hat{\beta}_0) \tag{6.9}$$

Al-Abood and Young (1985) show that the bias of $\hat{\beta}_0$ to order R^{-1} is -(k+1)/(2R) when the centering conditions for the x's hold. This result also holds when H_o is true. Under H_0, the bias of $\hat{\beta}_0$ is $-(\ell+1)/(2R)$ to the same order of approximation and hence to order (1)

$$E(S_1) = k - \ell \tag{6.10}$$

Which agrees with the first moment of the approximating chi-square Distribution

We now consider the special case when $\ell = k - 1$ and we are testing $H_0^{(k)}$: $ß_k = 0$. In this case, S_1 has approximately a non-central x^2 distribution with 1 degree of freedom and non-centrality parameter

$$\lambda_K = \beta_k^2 / I^{kk} \tag{6.11}$$

where I^{kk} is the (k+1)st element in the diagonal of the inverse of the information matrix. The approximate test procedure using a double-tailed test with significance level α is

$$\text{reject} \quad H_0^{(k)} \text{ if } S_1 > X_1^2(1 - \alpha) \tag{6.12}$$

where $x_v^2(1-\alpha)$ denotes the upper $100\alpha\%$ point of the X_v^2 distribution.

If we let $Y_k(S_1)$ denote the power of the test based on S_1, then using the results that S_1 is aproximately distributed as U^2 where $U \sim N(\lambda_k^{\frac{1}{2}}, 1)$ and that $X_1^2(1-\alpha) = u^2_{1-\frac{1}{2}\alpha}$, we obtain the power approximation

$$\gamma_k(S_1) \approx 1 - \varphi\left\{u_{1-\frac{1}{2}\alpha} - \beta_k (I^{kk})^{-\frac{1}{2}}\right\} + \varphi\left\{u_{\frac{1}{2}\alpha} - \beta_k (I^{kk})^{-\frac{1}{2}}\right\}. \tag{6.13}$$

When $\beta_k = 0$, this power is α as required.

An alternative test procedure can be made by taking $\left(\hat{\beta}_k - \beta_k\right)/\left(I^{kk}\right)^{\frac{1}{2}}$ to be approximately distributed as N(0,1). To test $H_0^{(k)}: \beta_k, = 0$ against the two-sided alternative $\beta_k, \neq 0$, we use the test statistic.

$Z_1 = \hat{\beta}_k/(I^{kk})^{\frac{1}{2}}$ and

$$\text{reject} \quad H_0^{(k)} \text{ if } |\, z_i \,| > U_{1-\frac{1}{2}\alpha} \tag{6.14}$$

The power of the test is

$$\gamma_k(z_1) = p\left\{Z_1 - \beta_k (I^{kk})^{-\frac{1}{2}} > u_{1-\frac{1}{2}\alpha} - \beta_k (I^{kk})^{-\frac{1}{2}}\right\}$$

$$+ P\left\{Z_1 - \beta_k (I^{kk})^{-\frac{1}{2}} < u_{\frac{1}{2}\alpha} - \beta_k (I^{kk})^{-\frac{1}{2}}\right\}$$

$$\approx 1 - \varphi\left\{u_{1-\frac{1}{2}\alpha} - \beta_k (I^{kk})^{-\frac{1}{2}}\right\} + \varphi\left\{u_{\frac{1}{2}\alpha} - \beta_k (I^{kk})^{-\frac{1}{2}}\right\}. \tag{6.15}$$

Asymptotically ($r_i \cdot \to \infty$, i = 1,...,g) the powers of the S_1 and Z_1 tests are therefore equivalent.

We now develop tests which utilise the weighted least squares estimators which are derived from the linear model representation

$$Z_i \; \underset{\sim}{x}'_i \; \underset{\sim}{\beta} + \in_i, \qquad i = 1,\ldots,g \tag{6.16}$$

where $Z_i = \log Y_i - \psi(r_i.) + \log r_i$ and

$$E(\varepsilon_i) = 0, \quad var(\varepsilon_i) - \psi'(r_i), \quad cov(\varepsilon_i, \varepsilon_j.) = 0 \tag{6.17}$$

for $i \neq j = 0,1,\ldots,k$. In matrix notation we have

$$\underset{\sim}{Z} = \underset{\sim}{X}\,\underset{\sim}{\beta} + \underset{\sim}{E}, \quad E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}, \quad cov(\underset{\sim}{\varepsilon}) = \underset{\sim}{D}_w^{-1} \tag{6.18}$$

Where $\underset{\sim}{D}_w = \text{diag}\{\Psi'(r_1),\ldots\ldots, \psi'(r_g)\}$

Under the full model, the WLS estimator $\underset{\sim}{\hat{\beta}}_W$ is obtained as the value of $\underset{\sim}{\beta}$ which minimises $R(\underset{\sim}{Z} - \underset{\sim}{z}\,\underset{\sim}{\beta})'\,\underset{\sim}{D}_w\,(\underset{\sim}{Z} - \underset{\sim}{x}\,\underset{\sim}{\beta})$ and the solution is given; by (1.5). The generalised residual sum of squares about the fitted full model is

$$R(\hat{\underset{\sim}{\beta}}_w) = (\underset{\sim}{Z} - \underset{\sim}{x}\,\hat{\underset{\sim}{\beta}})'\,\underset{\sim}{D}_w(\underset{\sim}{Z} - \underset{\sim}{x}\,\underset{\sim}{\beta}_w). \tag{6.19}$$

Writing $\underset{\sim}{Z} - \underset{\sim}{X}\,\hat{\underset{\sim}{\beta}}_w = (\underset{\sim}{I} - \underset{\sim}{X}\,\underset{\sim}{M}^{-1}\,\underset{\sim}{X}'\,\underset{\sim}{D}_w)\,\underset{\sim}{z}$ where $\underset{\sim}{M} = \underset{\sim}{X}\,\underset{\sim}{D}_w\,\underset{\sim}{X}$, we obtain

$$R(\hat{\underset{\sim}{\beta}}) = \underset{\sim}{z}\,'\underset{\sim}{A}\,\underset{\sim}{Z} \tag{6.20}$$

Where $\underset{\sim}{A} = \underset{\sim}{D}_w - \underset{\sim}{D}_w\,\underset{\sim}{X}\,\underset{\sim}{M}^{-1}\,\underset{\sim}{X}'\,\underset{\sim}{D}_w$. Hence

$$\left\{R(\hat{\underset{\sim}{\beta}}_w)\right\} = E\left\{(\underset{\sim}{\beta}'\,\underset{\sim}{X}' + \underset{\sim}{\varepsilon})\,\underset{\sim}{A}(\underset{\sim}{x}\,\underset{\sim}{\beta} + \underset{\sim}{\varepsilon})\right\} = E(\underset{\sim}{\varepsilon}'\,\underset{\sim}{A}\,\underset{\sim}{\varepsilon})$$

$$= tr(\underset{\sim}{A}\,\underset{\sim}{D}_w^{-1}) = tr(\underset{\sim}{I}_g - \underset{\sim}{D}_w\,\underset{\sim}{X}\,\underset{\sim}{M}^{-1}\,\underset{\sim}{X}'\ .$$

Since $tr(\underset{\sim}{I}_g) = g$ and $tr\,(\underset{\sim}{D}_w\,\underset{\sim}{X}\,\underset{\sim}{M}^{-1}\,\underset{\sim}{X}') = tr(\underset{\sim}{I}_{k+1}) = k + 1$, we have

$$E\left\{R(\hat{\underset{\sim}{\beta}}_w)\right\} = g - k - 1 \tag{6.21}$$

This result holds for all $\underset{\sim}{\beta}$ and hence in particular when $H_0 : \underset{\sim}{\beta}_2 = \underset{\sim}{0}$ is true.

Similarly, for the restricted model $\underset{\sim}{Z} = \underset{\sim}{X}_1\,\underset{\sim}{\beta}_1 + \underset{\sim}{\varepsilon}_1$ the generalised residual sum of squares about the LS fitted model is

$$R(\hat{\underset{\sim}{\beta}}_{w1}) = (\underset{\sim}{Z} - \underset{\sim}{x}_1\,\hat{\underset{\sim}{\beta}}_{wi})'\,\underset{\sim}{D}_w(\underset{\sim}{Z} - \underset{\sim}{X}_1\,\hat{\underset{\sim}{\beta}}_{wi}) \tag{6.22}$$

Where $\hat{\underset{\sim}{\beta}}_1 = (\underset{\sim}{x}'\,\underset{\sim}{D}_w\,\underset{\sim}{X}_1)^{-1}\,\underset{\sim}{X}'_1\,\underset{\sim}{D}_w\,\underset{\sim}{Z}$. Settings $\underset{\sim}{A}_1 = \underset{\sim}{D}_w - \underset{\sim}{D}_w\,\underset{\sim}{X}_1\,M_1^{-1}\,\underset{\sim}{X}'_1\,\underset{\sim}{D}_w$ where $\underset{\sim}{M}_1 = \underset{\sim}{X}_1{}'\,\underset{\sim}{D}_w\,\underset{\sim}{X}_1$, we may write

$$R(\hat{\underset{\sim}{\beta}}_{w1}) = \underset{\sim}{Z}'\,\underset{\sim}{A}_1\,\underset{\sim}{Z}\,. \tag{6.23}$$

We Have

$$\begin{aligned} E\{R(\hat{\underset{\sim}{\beta}}_{wl}) &= \underset{\sim}{\beta}'\,\underset{\sim}{A}'\,\underset{\sim}{X}\,\underset{\sim}{\beta} + E(\underset{\sim}{\varepsilon}'\,\underset{\sim}{A}_1\,\underset{\sim}{\varepsilon}) \\ &= \underset{\sim}{\beta}'\,\underset{\sim}{X}'\,\underset{\sim}{A}_1\,\underset{\sim}{X}\,\underset{\sim}{\beta} + g - \ell - 1 \end{aligned} \tag{6.24}$$

A straightforward calculation shows that

$$\underset{\sim}{\beta}'\,\underset{\sim}{X}'\,\underset{\sim}{A}_1\,\underset{\sim}{X}\,\underset{\sim}{\beta} = \underset{\sim}{\beta}'_2\,\underset{\sim}{X}'_2\,\underset{\sim}{A}_1\,\underset{\sim}{X}_2\,\underset{\sim}{\beta}_2\,. \tag{6.25}$$

Hence we have the exact expectation results

$$E\{R(\hat{\beta}_{w1}) \mid H_0\} = g - \ell - 1. \tag{6.26}$$

Using the extra sum of squares principle, a suitable test statistic for testing H_0 is

$$S_2 = R(\hat{\underset{\sim}{\beta}}_{w1}) - R(\hat{\underset{\sim}{\beta}}_{w}) \tag{6.27}$$

for which we have the exact expectation result

$$E(S_2) = k - \ell + \underset{\sim}{\beta}_2{}' \underset{\sim}{X}'_2 \underset{\sim}{A}_1 \underset{\sim}{X}_2 \underset{\sim}{\beta}_2 \ . \tag{6.28}$$

This gives $E(S_2) = k - \ell$ when H_0 is true.

The exact distribution of S_2 is unknown and an approximation is required. Taking the $\{\varepsilon_i\}$ which are independently distributed as log-gamma random variables to be approximately distributed as $N(0,\psi'(r_i))$, we obtain the approximation

$$S_2 \underset{\sim}{\text{approx}} x'^2_{k-2}(\underset{\sim}{\beta}'_2 \underset{\sim}{X}_2{}' \underset{\sim}{A}_1 \underset{\sim}{X}_2 \underset{\sim}{\beta}_2) \tag{6.29}$$

where $X'^2_v(\lambda)$ represents the non-central chi-square distribution with v degrees of freedom and non-centrality parameter λ.

We now consider the special case when ℓ = k. - 1 and we are testing $H_0^{(k)} : \beta_k = 0$. In this case S_2 is approximately distributed as non-central x^2 with 1 degree of freedom and non-centrality parameter

$$\lambda_{wk} = \beta_k^2/V^{kk} \tag{6.30}$$

where V^{kk} is the (k+1)st diagonal element in the inverse of $\underset{\sim}{X}' \underset{\sim}{D}_w \underset{\sim}{X}$.
he test procedure is

$$\text{reject } H_0^{(k)} \text{ if } S_2 > X_1^2(1-\alpha) \tag{6.31}$$

where $X_V^2(1 - \alpha)$ is the upper 100α% point of the distribution of X_V^2
The approximate power of the test is

$$\gamma_k(S_2) \approx 1 - \varphi\left\{u_{1-\frac{1}{2}} - \beta_k(V^{kk})^{-\frac{1}{2}}\right\} + \varphi u_{\frac{1}{2}\alpha} - \beta_k(V^{kk})^{-\frac{1}{2}}. \qquad (6.32)$$

An alternative test procedure is to use the test statistic $Z_2 = \hat{\beta}_{wk}/(V^{kk})^{\frac{1}{2}}$ and

$$\text{reject } H_0^{(k)} \text{ if } \mid Z_2 \mid > u_{1-\frac{1}{2}\alpha}. \qquad (6.33)$$

The exact powers of the tests based on S_2 and Z_2 are equal so a choice of test can be made on grounds of computational simplicity.

In order to examine the power properties of the tests based on the S_1 , S_2 , Z_1 and Z_2 statistics and to assess the adequacy of the approximating powers given by ((6.13),(6.15,(6.32) and (6.34), a simulation investigation has been made for the case of a single explanatory variable when the means $\{\mu_i.\}$ satisfy the model defined in (3.1). Equal values for the shape parameters were taken with $r_i = r = 1(1)10(2)20$ for $i = 1,...,g$ with $g = 5,10$. Equally spaced values $x_i = i - \frac{1}{2}(g+1)$ were used for the explanatory variable. Values $\beta_1 = \log\theta/(g-1)$ were used giving $\max_i\mu_i/\min_i\mu_i. = \theta$, for $\theta = 1(1)5$. The run-size was 2000 in each case.

The broad conclusions reached from the investigation are

(i) The use of the S_1 and S_2 tests lead to excellent control over the significance levels for all values of r. The actual significance levels of the Z_1 -test are much larger than the nominal values for $r = 1,2,3$ but are satisfactory for larger values of r.

(ii) For the very small values of r the power of the Z_1-test is greater than that of the S_1 -test but this seems to simply reflect the differences in the actual significance levels of the tests. The power differences between the two tests are very small for $r > 3$.

(iii) The power performance of the S_1-test is markedly better than that of the S_2 -test for $r = 1,2$ but the power advantage diminshes rapidly with increasing values of r,

(iv) The power approximation given by (6.13) gives a slight overestimation of the power, particularly for large values of θ and small values of r. However, the results are generally very encouraging.

These findings are illustrated in tables 10, 11, 12 which show the estimated powers as obtained by simulation together with approximating powers for the S_1 , S_2 and Z_1 tests respectively for g = 5 and nominal significance levels α = 0.10, 0.05, 0.01.

Table10

Powers of the S_1-test for g = 5 and nominal significance levels α as obtained by (i) simulation, (ii) approximation (6.13)

	θ	1		2		3		4		5	
	r	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
	1	0.109	0.100	0.146	0.151	0.232	0.225	0.279	0.295	0.311	0.357
	2	0.106	0. 100	0.209	0.200	0.326	0.341	0.447	0.463	0.544	0.562
α=0.10	3	0.102	0.100	0.231	0.248	0.433	0.445	0.576	0.600	0.683	0.712
	4	0.095	0.100	0.296	0.295	0.525	0.537	0.687	0.708	0.794	0.816
	5	0.104	0.100	0.355	0.339	0.595	0.617	0.773	0.790	0.861	0.885
	1	0.058	0.050	0.077	0.085	0.142	0.140	0.177	0.195	0.209	0.247
	2	0.053	0.050	0.118	0.121	0.225	0.233	0.336	0.341	0.413	0.436
α=0.05	3	0.048	0.050	0.149	0.158	0.317	0.325	0.445	0.475	0.572	0.596
	4	0.049	0.050	0.193	0.195	0.391	0.412	0.561	0.592	0.685	0.721
	5	0.047	0.050	0.234	0.232	0.480	0.493	0.682	0.688	0.790	0.812
	1	0.011	0.010	0.017	0.022	0.039	0.044	0.060	0.070	0.064	0.096
	2	0.009	0.010	0.033	0.036	0.075	0.089	0.137	0. 152	0.200	0.219
α=0.01	3	0.010	0.010	0.057	0.052	0.126	0.142	0.233	0.249	0.322	0.355
	4	0.009	0.010	0.068	0.070	0. 187	0.201	0.324	0.351	0.466	0.488
	5	0.009	0.010	0.091	0.089	0.242	0.263	0.431	0.450	0.571	0.606

Table 11

Powers of the S_2-test for g = 5 and nominal significance level α as obtained by (i) simulation, (ii) approximation (6.32)

	θ	1		2		3		4		5	
	r	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
	1	0.098	0.100	0.124	0.131	0.170	0.177	0.212	0.221	0.224	0.261
	2	0.100	0.100	0.178	0.178	0.290	0,290	0.394	0.391	0.473	0.476
α=0.10	3	0.104	0.100	0.213	0.226	0.395	0.398	0.524	0.540	0.650	0.648
	4	0.097	0.100	0.277	0.273	0.493	0.495	0.665	0.660	0.772	0.771
	5	0.100	0.100	0.339	0.318	0.577	0.580	0.756	0.753	0.847	0.855
	1	0.056	0.050	0.065	0.071	0.101	0.104	0.134	0.137	0.149	0.168
	2	0.051	0.050	0.102	0.105	0.190	0.191	0.277	0.276	0.352	0.354
α=0.05	3	0.049	0.050	0.136	0.141	0.274	0.282	0.396	0.415	0.522	0.526
	4	0.047	0.050	0.175	0.177	0.366	0.371	0.537	0.539	0.672	0.666
	5	0.048	0.050	0.219	0.214	0.447	0,455	0.650	0.644	0.772	0.772
	1	0.016	0.010	0.020	0.017	0.032	0.029	0.049	0.043	0.057	0.057
	2	0.013	0.010	0.034	0.030	0.063	0.068	0.107	0.113	0.157	0.161
α=0.01	3	0.010	0.010	0.048	0.045	0.110	0.116	0.205	0.203	0.278	0.291
	4	0.012	0.010	0.059	0.061	0.163	0.172	0.297	0.302	0.423	0.426
	5	0.010	0.010	0.083	0.079	0.221	0.233	0.413	0.403	0.545	0.551

Table 12

Powers of the Z_1 -test for g = 5 and nominal significance level α as obtained by (i) simulation, (ii) approximation (6.15)

	θ	1		2		3		4		5	
	r	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
	1	0.153	0.100	0.195	0.151	0.272	0.225	0.322	0.295	0.350	0.357
	2	0.132	0.100	0.231	0.200	0.358	0.341	0.478	0.463	0.568	0.562
α=0.10	3	0.122	0.100	0.247	0.248	0.447	0.445	0.582	0.600	0.711	0.712
	4	0.102	0.100	0.305	0.295	0.532	0.537	0.705	0.708	0.805	0.816
	5	0.106	0.100	0.352	0.339	0.607	0.617	0.783	0.790	0.865	0.885
	1	0.087	0.050	0.127	0.085	0.182	0.140	0.234	0.195	0.254	0.247
	2	0.072	0.050	0.137	0.121	0.249	0.233	0.358	0.341	0.437	0.436
α=0.05	3	0.064	0.050	0.159	0.158	0.332	0,325	0.468	0.475	0.588	0.596
	4	0.052	0.050	0.205	0.195	0.415	0.412	0.591	0.592	0.710	0.721
	5	0.052	0.050	0.246	0.232	0.499	0.493	0.685	0.688	0.799	0.812
	1	0.031	0.010	0.040	0.022	0.080	0.044	0.096	0.070	0.118	0.096
	2	0.022	0.010	0.049	0.036	0.102	0.089	0.159	0.152	0.232	0,219
α=0.01	3	0.014	0.010	0.061	0.052	0.151	0.142	0.255	0.249	0.354	0.355
	4	0.013	0.010	0.073	0.070	0.199	0.201	0.347	0.351	0.492	0,488
	5	0.013	0.010	0.093	0.089	0.252	0.263	0.457	0.450	0.603	0.606

7. Goodness Of Fit Tests For The Exponential Link Function

Finally, we consider tests of fit for the assumed exponential link function for the means as given by (1.2) against general alternatives. Two tests are examined, the first utilising the ML estimates and providing the likelihood ratio test, the second test being based on the WLS estimates.

When no model is imposed on the means $\{\mu_{i\cdot}\}$, the ML estimates are $\hat{\mu}_i = Y_i$, i = 1,.....,g, Using (6.3) and the first equation in (6.4), the likelihood ratio statistic for testing the exponential link function is $D = 2\sum_1 (\underset{\sim}{X}_i{}'\hat{\underset{\sim}{\beta}} - \log Y_i)$. With loss of generality we shall assume that the x's re centred such that $\sum_i r_i x_{ij} = 0$, $j = 1,....,\ k$. In this case, D akes the simple form

$$D = 2(R\hat{\beta}_0 - \sum_{i=1}^{g} r_i \log Y_i). \quad (7.1)$$

Using the well-known result for the expectation of a log gamma random variable, we have

$$E(\log Y_i) = \log\mu_i + \psi(r_i) - \log r_i$$

$$= \beta_0 - \frac{1}{2r_i} - \frac{1}{12r_i^2} + \frac{1}{120r_i^4} + 0(r_1^{-6}) \quad (7.2)$$

if the exponential link function is correct. Setting

$$E(\hat{\beta}_0) \approx \beta_0 - (k + 1)/2R$$

we obtain

$$E(D) \approx g - k - 1 + \frac{1}{6}\sum_{i=1}^{g} r_i^{-1} \quad (7.3)$$

ignoring terms of $0(r_i^{-3})$ and smaller terms

The usual procedure is to refer the statistic D to the chi-square Distribution with g-k-1 degees of freedom . The form of (7.3) Suggests the use of the modified statistic D* = D/(1+c), where

$$C = \frac{1}{6(g - k - 1)}\sum_{i=1}^{g} r_i^{-1} \quad (7.4)$$

and to take D* as approximately distributed as chi-square with g-k— 1 degrees of freedom if the exponential link function is correct.

To assess the effect of the modification, moments and critical values of the null distributions of the statistics D and D * have been estimated by simulation for the model $\mu_i. = \exp(\beta_0 + \beta_1 x_i.)$, with $x_i = i - \frac{1}{2}(g+1)$, $i = 1,\ldots,g$, for $r_i- = r = 1(1)10(2)20$ and $g = 5,10$. The results showed that the null distribution of D* approaches the X^2_{g-2} distribution much more rapidly than the distribution of D. Use of D* therefore leads to much better control over the significance level of the test for small values of the shape parameter. These findings are illustrated in table 13 which shows the means and variances of D and D* and in table 14 which shows the estimated upper 10%, 5% and 1% critical values of the null distributions of D and D* for

$r_i = r = 1(1)5$, $i = 1,...,g$ and $g = 5,10$. The estimated significance levels associated with the chi-square approximating critical values are shown in parentheses.

Table13

Means and variances of the D and D* statistics when $r_i = r$, $i = 1,...,g$

	g = 5				g= 10			
	mean		variance		mean		variance	
r	D	D*	D	D*	D	D*	D	D*
1	3.800	2.938	25.894	12.984	9.512	7.872	21.576	14.777
2	3.396	2.943	7.494	5.778	8.838	8.004.	18.969	15.558
3	3.217	2.944	6.963	5.823	8.597	8.039	18.401	15.676
4	3.205	2.996	6.643	5.833	8.466	8.047	18.255	16.392
5	3.170	3.003	6.456	5.895	8.431	8.094	17.436	16.069
χ^2_{g-2}	3.000	3.000	6.000	6.000	8.000	8.000	16.000	16.000

Table14

Upper 100 % points of the null distributions of the D and D* statistics when $r_i = r$, $i = 1,\dots,g$

g = 5

	α = 0.10		α = 0.05		α = 0.01	
r	D	D*	D	D*	D	D*
1	7.52(0,165)	5.88(0.082)	9.31(0.088)	7.28(0.041)	13.56(0.024)	10.61(0.0063)
2	7.00(0.132)	6.14(0.095)	8.68(0.068)	7.62(0.046)	12.51(0.017)	10.98(0.0090)
3	6.72(0.122)	6.15(0.096)	8.20(0.062)	7.51(0.044)	12.36(0.017)	11.31(0.0100)
4	6.70(0.118)	6.27(0.101)	8.40(0.062)	7.86(0.051)	12.01(0.014)	11.23(0.0093)
5	6.61(0.116)	6.26(0.101)	8.17(0.061)	7.74(0.049)	11.98(0.011)	11.18(0.0088)
$\chi^2_3(\alpha)$	6.25	6.25	7.82	7.82	11.34	11.34

g = 10

	α = 0.10		α = 0.05		α = 0.01	
r	D	D*	D	D*	D	D*
1	15.89(0.182)	13.15(0.094	18.04(0.107	14.93(0.042	23.18(0.031	19.19(0.0085)
2	14.78(0.143)	13.38(0.101	17.17(0.076	15.55(0.051	21.79(0.019	19.74(0.0090)
3	14.31(0.134)	13.38(0.100	16.51(0.071	15.44(0.049	21.43(0.017	20.04(0.0098)
4	14.14(0.119)	13.44(0.102	16.64(0.068	15.81(0.055	21.65(0.018	20.58(0.0125)
5	14.23(0.120)	13.66(0.104	16.40(0.067	15.74(0.054	20.80(0.014	19.97(0.0098)
$\chi^2_8(\alpha)$	13.36	13.36	15.51	15.51	20.09	20.09

If the gamma regression model is fitted by WLS using the transformed observations $Z_i = \log Y_{i.} - \psi(r_i) + \log r_i$, $i = 1,\ldots,g$, a goodness of fit statistic is provided by $R(\hat{\beta}_W)$ which was defined in (5.18). The statistic has exact expectation g-k-1 and its distribution approaches the x^{2}- distribution with g-k-1 degress of freedom as the $\{r_i\}$ increase, if the assumed model is correct.

Although the means of the exact distribution of $R(\hat{\beta}_W)$ and the approximation x^2 -distribution agree, the variances are not equal. To demonstrate this, consider the case when the shape parameters are equal, that is, $r_i = r$ for $i = 1,\ldots,g$. In this case the $\{Z_i\}$ are identically distributed as log-gamma random variables with skewness and kurtosis coefficients given by

$$\gamma_1(Z_i) = \psi^{(2)}(r) / \left\{ g^{\frac{1}{2}} \{\psi'(r)\} \right\}^{\frac{3}{2}}, \quad \gamma_2(Z_i) = \psi^{(3)}(r) / \left[\{\psi'(r)\}^2 g \right] \tag{7.5}$$

for I= 1,……..,g (kahn (1979)). The WLS estimator is the same as the OLS estimator of $\underset{\sim}{\beta}$ and using results from atiqullah (1962) we have

$$\begin{aligned} \operatorname{var}\{R(\beta_{\sim W})\} &= \{\psi'(r)\}^2 \operatorname{Var}\left\{ \sum_{i=1}^{g} zi - xi'\beta_{\sim W})^2 \right\} \\ &= 2(g-k-1) + \gamma_2(Z_i) \sum_{i=1}^{g} (1-h_{ii})^2 \end{aligned} \tag{7.6}$$

where h_{ii} is the ith diagonal element in the hat matrix $\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'$
We therefore have

$$\operatorname{Var}\{R(\hat{\beta}_{\sim W})\} = 2(g - k - 1)(1 + c) \tag{7.7}$$

where

$$c = \frac{1}{2} \frac{\psi^{(3)}(r) \sum_i (1 - h_{ii})^2}{\{\psi'(r)\}^2 (g - k - 1)} \tag{7.8}$$

The form of (7.7) leads us to consider the modified statistic $R^*(\hat{\beta}_{\sim W}) =$ $R^*(\hat{\beta}_{\sim W}) + b$, where the constants a and b are selected to give agreement between the exact mean and variance of $R^*(\hat{\beta}_{\sim W})$ and the corresponding moments of the approximating $X^2{}_{g-k-1}$ distribution. This gives

$$R^*(\hat{\beta}_{\sim W}) = (1 + c)^{\frac{1}{2}} [R(\hat{\beta}_{\sim W}) + (g + k - 1)\{(1 + C)^{\frac{1}{2}} - 1\}]. \tag{7.9}$$

To examine the adequacy of the chi-square approximation to the null distributions of $R(\hat{\beta}_W)$ and $R^*(\hat{\beta}_W)$, the moments and critical values of the distributions have been estimated by simulation for the model $\mu_{i.} = \exp(\beta_0 + \beta_1 x_{i.})$ with $x_{i.} = i - \frac{1}{2}(g+1)$, $i = 1,..,g$ and $r_i = r = 1(1)10(2)20$, $g = 5,10$,

The results show that the mean and variance of the distribution of $R^*(\hat{\beta}_W)$, as obtained by simulation are very close to the corresponding moments of the approximating x^2-distributions. The use of the modified statistic $R^*(\hat{\beta}_W)$, leads to better control of the significance levels for small values of α and small values of r, particularly for the larger value of g. These findings are illustrated in table 15 which shows the means and variances of $R(\hat{\beta}_W)$, and $R^*(\hat{\beta}_W)$, and in table 16 which shows the estimated upper 10%, 5 % and 1 % critical values of the null distributions of $R(\hat{\beta}_W)$, and $R^*(\hat{\beta}_W)$, for $r_{i.} = r = 1(1)5$, $i = 1,...,g$ and $g = 5,10$. The estimated significance levels associated with the chi-square approximating critical values are shown in parentheses.

Table 15

Means and variances of the $R(\hat{\beta}_{\sim W})$ and $R^*(\hat{\beta}_{\sim W})$ statistics when $r_i = r$, $i = 1,...,g$,

	g = 5				g = 10			
	Mean		Variance		Mean		Variance	
r	$R(\hat{\beta}_{\sim W})$	$R^*(\hat{\beta}_{\sim W})$	$R(\hat{\beta}_{\sim W})$	$R^*(\hat{\beta}_{\sim W})$	$R(\hat{\beta}_{\sim W})$	$R^*(\hat{\beta}_{\sim W})$	$R(\hat{\beta}_{\sim W})$	$R^*(\hat{\beta}_{\sim W})$
1	2.989	2.992	9.915	5.583	8.025	8.018	31.472	15.962
2	3.018	3.016	8.814	6.369	8.049	8.040	24.054	16.244
3	2.988	2.990	7.760	6.225	8.061	8.053	20.735	15.844
4	3.011	3.011	6.965	5.902	8.066	8.059	20.562	16.579
5	3.013	3.012	6.641	5.881	8.104	8.096	19.467	16.560
x^2_{g-2}	3.000	3.000	6.000	6.000	8.000	8.000	16.000	16.000

Table 16

Upper 100 % points of the null distribution of the $R(\hat{\beta}_{\sim w})$ and $R^*(\hat{\beta}_{\sim w})$ statistic when $r_i = r$, $i = 1,\ldots,g$

$g = 5$

r	$\alpha = 0.10$ $R(\hat{\beta}_{\sim w})$	$R^*(\hat{\beta}_W)$	$\alpha = 0.05$ $R(\hat{\beta}_{\sim w})$	$R^*(\hat{\beta}_W)$	$\alpha - 0.01$ $R(\hat{\beta}_{\sim w})$	$R^*(\hat{\beta}_W)$
1	6.52(0.110)	5.64(0.077)	8.88(0.068)	7.41(0.043)	16.29(0.028	12.97(0.015)
2	6.49(0.109)	5.97(0.086)	8.30(0.059)	7.51(0.044)	13.50(0.019)	11.93(0.013)
3	6.45(0.109)	6.09(0.094)	8.28(0.056)	7.65(0.045)	12.97(0.018)	11.93(0.013)
4	6.41(0.106)	6.14(0.095)	8.11(0.057)	7.71(0.047)	12.17(0.014)	11.44(0.011)
5	6.37(0.104)	6.16(0.095)	8.00(0.054)	7.78(0.048)	11.94(0.013)	11.37(0.011)
$X_3^2(\alpha)$	6.25	6.25	7.82	7.82	11.34	11.34

$g = 10$

r	$\alpha = 0.10$ $R(\hat{\beta}_{\sim w})$	$R^*(\hat{\beta}_W)$	$\alpha = 0.05$ $R(\hat{\beta}_{\sim w})$	$R^*(\hat{\beta}_W)$	$\alpha = 0.01$ $R(\hat{\beta}_{\sim w})$	$R^*(\hat{\beta}_W)$
1	14.96(0.131)	12.96(0.089)	18.84(0.090)	15.72(0.053)	28.73(0.042)	22.76(0.019)
2	14.04(0.119)	12.96(0.089)	17.07(0.071)	15.45(0.049)	25.38(0.027)	22.28(0.016)
3	13.92(0.117)	13.18(0.095)	16.72(0.066)	15.62(0.052)	23.87(0.023)	21.93(0.013)
4	13.84(0.113)	13.28(0.098)	16.67(0.065)	15.83(0.053)	23.17(0.023)	21.70(0.014)
5	13.99(0.115)	13.52(0.104)	16.36(0.064)	15.70(0.053)	21.71(0.018)	20.63(0.012)
$X_3^2(\alpha)$	13.36	13.36	15.51	15.51	20.09	20.09

References

Al-Abood, A.M. and Young, D.H. (1985). Moment properties of maximum likelihood and least squares estimators for an exponential regression model with type II censoring. Technical Report TR/14/85, Brunei University.

Atiquallah, M. (1962). The estimation of residual variance in quadratically balanced least squares problems and the robustness of the F-test. Biometrika, 49, 83-91.

Cox, D.R. and Lewis, P.A. (1966). The statistical analysis of a series of events. London, Methuen.

Kahn, H.D. (1979). Least squares estimation for the inverse power law for accelerated life tests. Appl. Statist. 28, No.1, 40-46.

Lawless, J.F, (1976). Confidence interval estimation in the inverse power law model. Appl. Statist., 25, No.2, 128-138.