

COMPREHENSIVE WRITTEN EXAMINATION, PAPER III

FRIDAY AUGUST 17, 2012, 9:00 A.M.–1:00 P.M.

STOR 665 QUESTIONS

1. [30 points] Box and Draper (1987) reported an experiment conducted to investigate the effects of three factors Length, Amplitude, Load on the Cycles to the failure of worsted yarn. Each of the regressors is coded in 3 levels. The data and the fitted GLM results are shown below.

```
> yarn
  length amplitude load cycles
1      0         0    0   674
2      0         0    1   370
3      0         0    2   292
4      0         1    0   338
5      0         1    1   266
6      0         1    2   210
7      0         2    0   170
8      0         2    1   118
9      0         2    2    90
10     1         0    0  1414
11     1         0    1  1198
12     1         0    2   634
13     1         1    0  1022
14     1         1    1   620
15     1         1    2   438
16     1         2    0   442
17     1         2    1   332
18     1         2    2   220
19     2         0    0  3636
20     2         0    1  3184
21     2         0    2  2000
22     2         1    0  1568
23     2         1    1  1070
24     2         1    2   566
25     2         2    0  1140
26     2         2    1   884
27     2         2    2   360
```

```
> yarn.mod=glm(cycles~length+amplitude+load, family=Gamma(link="log"), yarn)
```

```
> summary(yarn.mod)
```

Call:

```
glm(formula = cycles ~ length + amplitude + load, family = Gamma(link = "log"),  
     data = yarn)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-0.43391	-0.11553	-0.00922	0.10260	0.25342

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	6.52286	0.08026	81.270	< 2e-16 ***
length	0.84251	0.04192	20.100	4.34e-16 ***
amplitude	-0.63132	0.04192	-15.062	2.10e-13 ***
load	-0.38513	0.04192	-9.188	3.68e-09 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 0.0316243)

Null deviance: 22.88613 on 26 degrees of freedom
Residual deviance: 0.76939 on 23 degrees of freedom
AIC: 332.76

Number of Fisher Scoring iterations: 5

(a) [10 points]

Write down the statistical model mathematically, and derive the corresponding log-likelihood for estimating the parameters.

(b) [10 points]

Estimate the effect on the number of cycles to failure when increasing the length by one level, along with the corresponding 95% confidence intervals.

(c) [10 points]

Describe how one would decide whether the fitted model is adequate based on the fitted results. Do you think the fitted model is adequate and why?

2. [45 points] Consider observations y_{ij} , $i = 1, \dots, n$, $j = 1, \dots, n_i$, which follow the model below

$$y_{ij} = \beta + b_i + e_{ij},$$

where e_{ij} 's are iid $N(0, \sigma^2)$, β and b_1 are fixed, and the additional b_i 's satisfy

$$b_i = b_{i-1} + e_i$$

where e_i 's are iid $N(0, \sigma_b^2)$.

- (a) [5 points] Let y denote the response vector and $b = (b_1, \dots, b_n)'$. Denote

$$E(y) = X\beta + Zb.$$

Find out the expressions of X and Z .

- (b) [15 points] Write down the conditional log-likelihood function of b given b_1 , denoted as $\log p(b|b_1)$.
- (c) [10 points] Write down the log-likelihood function based on y and b .
- (d) [15 points] Given σ^2 and σ_b^2 , derive the system of linear equations that the maximum likelihood estimates of β and b satisfy.

3. [25 points] Consider the following model:

$$y_{ijk} = \mu + \alpha_i + \beta_{ij} + \epsilon_{ijk},$$

where $i = 1, \dots, a$, $j = 1, \dots, b$, $k = 1, \dots, n$, μ and α_i are fixed, ϵ_{ijk} 's are iid $N(0, \sigma^2)$, and β_{ij} 's are iid $N(0, \sigma_\beta^2)$.

- (a) [5 points]

Express the above model in matrix form.

- (b) [5 points]

Obtain the expression for the variance-covariance matrix of the response vector.

- (c) [15 points]

Derive ANOVA estimates for the variance components. *Need to show the derivation steps.*