PubH 7407, Spring 2008: Exam II

There are three problems, each with multiple parts; each part is worth 5 points. Please start a new page for each problem. Put your name at the top of each page (5 points).

1. Consider the following $3 \times 4 \times 4$ table containing data on dumping severity S (1=none, 2=slight, and 3=moderate dumping severity), operation type O, across four hospitals H. The four operations, treatments for duodenal ulcer patients, increase in their invasiveness (1=drainage and vagotomy, 2=25% resection and vagotomy, 3=50% resection and vagotomy, and 4=75% resection); dumping severity is an undesirable side effect of the operation.

	Dumping Severity											
	Hospital 1			Hospital 2			Hospital 3			Hospital 4		
Operation	1	2	3	1	2	3	1	2	3	1	2	3
1	23	7	2	18	6	1	8	6	3	12	9	1
2	23	10	5	18	6	2	12	4	4	15	3	2
3	20	13	5	13	13	2	11	6	2	14	8	3
4	24	10	6	9	15	2	7	7	4	13	6	4

A log-linear model was fit to the cell counts of this contingency table using the following SAS code:

```
proc genmod; class operation hospital severity;
 model count=operation|severity hospital / dist=poi link=log type3;
```

(a) Write down the model being fit here.

Either:

$$\log(\mu_{ij}) = \lambda + \lambda_i^O + \lambda_j^S + \lambda_k^H + \lambda_{ij}^{OS},$$

or

$$\log(n\pi_{ij}) = \lambda + \lambda_i^O + \lambda_j^S + \lambda_k^H + \lambda_{ij}^{OS},$$

for Poisson or multinomial sampling, respectively.

Here's some of the output:

```
                Criteria For Assessing Goodness Of Fit

Criterion                    DF           Value        Value/DF
Deviance                     33         21.7329          0.6586
Pearson Chi-Square           33         22.4491          0.6803

                 LR Statistics For Type 3 Analysis

                                      Chi-
Source                        DF    Square    Pr > ChiSq
operation                      3      3.88        0.2750
severity                       2    142.40        <.0001
operation*severity             6     10.88        0.0922
hospital                       3     28.05        <.0001
```

(b) Is there any evidence of lack of fit? Note that $P(\chi^2_{33} \le 21.7329) = 0.067$.

No, p-value is $P(\chi^2_{33} > 21.7329) = 1 - 0.067 = 0.933$. I would say there's reasonable replication here, but we do not have the table of expected cell counts under the model, so we cannot say for sure if the rule of thumb is satisfied. However, $21.7329 < 33 < 2(33)$ so the model fits fine.

(c) What is the shorthand for this model? Use O, S, and H for operation, severity, and hospital.

$[H][OS]$.

(d) Draw an association graph for the final model above (I am opting to include the `operation*severity` interaction in the model). What is the independence or conditional independence interpretation?

$(O, S) \perp H$. The association graph connects O to S, but H is not connected to either S or S.

(e) Can you collapse over hospital and retain valid interpretation between operation and severity?

Yes; the theorem in the book says we can collapse. In fact hospital is independent of severity and operation, so none of the log-linear model parameters will change for S and O.

A proportional odds (cumulative logit) model was also fit to these data using the following code:

```
proc logistic; class operation hospital / param=ref; freq count;
 model severity=operation / link=clogit;
```

with output:

```
                          Class Level Information

                   Class         Value     Design Variables
                   operation     1           1      0      0
                                 2           0      1      0
                                 3           0      0      1
                                 4           0      0      0

                 Score Test for the Proportional Odds Assumption

                      Chi-Square       DF     Pr > ChiSq
                          4.0213        3         0.2592

                    Analysis of Maximum Likelihood Estimates

                                        Standard          Wald
     Parameter      DF    Estimate       Error     Chi-Square     Pr > ChiSq
     Intercept 1     1     -0.0185      0.1868         0.0098         0.9210
     Intercept 2     1      1.7377      0.2149        65.3775         <.0001
     operation 1     1      0.5998      0.2787         4.6313         0.0314
     operation 2     1      0.5916      0.2724         4.7168         0.0299
     operation 3     1      0.1705      0.2601         0.4298         0.5121

                           Odds Ratio Estimates

                                  Point          95% Wald
            Effect             Estimate      Confidence Limits
            operation 1 vs 4      1.822       1.055       3.146
            operation 2 vs 4      1.807       1.059       3.082
            operation 3 vs 4      1.186       0.712       1.975
```

(f) Write down the fitted model.

$$\log \frac{P(S \leq j)}{P(S > j)} = -0.0185I\{j=1\}+1.7377I\{j=2\}+0.5998I\{O=1\}++0.5916I\{O=2\}+0.1705I\{O=3\}.$$

(g) The model that also included hospital as a predictor gave the following Type III effects:

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
operation	3	7.3954	0.0603
hospital	3	2.6870	0.4424

Is this result ($p = 0.4424$) surprising in light of part (c)?

No, we would accept that hospital does not significantly affect severity with operation in the model. This would certainly happen if $S \perp H|O$, which we see is the case from part (c). In fact, $S \perp H$ holds, which is stronger.

(h) Do any of operations 1, 2, or 3 have significantly different outcome odds compared to the most invasive operation 4? Elaborate.

Yes, the table of odds ratio estimates indicate that operation types 1 and 2 significantly increase the odds of less dumping severity by 80%.

Finally, a baseline-category logit model was fit by replacing `link=clogit` with `link=glogit` above with the following output:

Analysis of Maximum Likelihood Estimates

Parameter	severity	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	1.1977	0.2853	17.6297	<.0001
Intercept	2	1	0.8650	0.2980	8.4244	0.0037
operation 1	1	1	0.9673	0.4905	3.8883	0.0486
operation 1	2	1	0.5213	0.5171	1.0163	0.3134
operation 2	1	1	0.4569	0.4159	1.2065	0.2720
operation 2	2	1	-0.2945	0.4574	0.4144	0.5197
operation 3	1	1	0.3778	0.4265	0.7846	0.3757
operation 3	2	1	0.3390	0.4440	0.5828	0.4452

Odds Ratio Estimates

Effect	severity	Point Estimate	95% Wald Confidence Limits	
operation 1 vs 4	1	2.631	1.006	6.880
operation 1 vs 4	2	1.684	0.611	4.640
operation 2 vs 4	1	1.579	0.699	3.568
operation 2 vs 4	2	0.745	0.304	1.826
operation 3 vs 4	1	1.459	0.632	3.366
operation 3 vs 4	2	1.404	0.588	3.351

(i) Write down the fitted model.

$$\log \frac{P(S=1)}{P(S=3)} = 1.1977 + 0.9673 I\{O=1\} + 0.4569 I\{O=2\} + 0.3778 I\{O=3\}.$$

$$\log \frac{P(S=2)}{P(S=3)} = 0.8650 + 0.5213 I\{O=1\} - 0.2945 I\{O=2\} + 0.3390 I\{O=3\}.$$

(j) Do any of operations 1, 2, or 3 have significantly different outcome odds compared to operation 4? Elaborate.

Yes. According to the odds ratio estimates table,

$$\frac{P(S=1|O=1)/P(S=3|O=1)}{P(S=1|O=4)/P(S=3|O=4)} = 2.631,$$

and this estimate is significantly not equal to one. The odds of none versus moderate dumping increase by a factor of 2.6 times for operation 1 versus 4.

(k) Find $P(S = 1|O = 1)$ from both the cumulative logit and baseline-category logit models. Are the probabilities similar?

From the cumulative logit model,

$$\begin{aligned} P(S = 1|O = 1) &= P(S \leq 1|O = 1) \\ &= \frac{e^{-0.0185+0.5998}}{1 + e^{-0.0185+0.5998}} \\ &= 0.6414. \end{aligned}$$

The baseline-category logit model gives

$$\begin{aligned} P(S = 1|O = 1) &= \frac{\frac{P(S=1|O=1)}{P(S=3|O=1)}}{\frac{P(S=1|O=1)}{P(S=3|O=1)} + \frac{P(S=2|O=1)}{P(S=3|O=1)} + \frac{P(S=3|O=1)}{P(S=3|O=1)}} \\ &= \frac{e^{1.1977+0.9673}}{e^{1.1977+0.9673} + e^{0.8650+0.5213} + 1} \\ &= 0.6354. \end{aligned}$$

The probabilities agree closely and are approximately 0.64.

---

2. Shaquille O'Neal played for Orlando Magic, Los Angeles Lakers, and Miami Heat. He also starred in the movie Kazaam (1996) as a rapping genie. The following is data on free throw attempts during the 2000 N.B.A. playoffs; i is the game, y_i is the number made out of attempts n_i.

i	y_i	n_i	i	y_i	n_i	i	y_i	n_i	i	y_i	n_i
1	4	5	2	5	11	3	5	14	4	5	12
5	2	7	6	7	10	7	6	14	8	9	15
9	4	12	10	1	4	11	13	27	12	5	17
13	6	12	14	9	9	15	7	12	16	3	10
17	8	12	18	1	6	19	18	39	20	3	13
21	10	17	22	1	6	23	3	12			

The following SAS code fits two models without random game effects and one model with random game effects:

```
data shaq;
 input game made attempts @@;
 datalines;
 1 4 5 2 5 11 3 5 14 4 5 12 5 2 7 6 7 10 7 6 14 8 9 15
 9 4 12 10 1 4 11 13 27 12 5 17 13 6 12 14 9 9 15 7 12 16 3 10
 17 8 12 18 1 6 19 18 39 20 3 13 21 10 17 22 1 6 23 3 12
;
proc logistic descending; model made/attempts = / aggregate=(game) scale=none; run;
proc nlmixed; parms alpha=-0.17;
 eta=alpha; p=exp(eta)/(1+exp(eta)); model made~binomial(attempts,p);
proc nlmixed qpoints=100; parms alpha=-0.17;
 eta=alpha+u; p=exp(eta)/(1+exp(eta)); model made~binomial(attempts,p);
 random u ~ normal(0,var) subject=game;
```

Here is some of the output (in order):

The LOGISTIC Procedure

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	40.0206	22	1.8191	0.0108
Pearson	35.5109	22	1.6141	0.0342

The NLMIXED Procedure

Fit Statistics

-2 Log Likelihood	101.3
AIC (smaller is better)	103.3

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
alpha	-0.1761	0.1167	23	-1.51	0.1448	0.05	-0.4175	0.06528	-1.5E-7

The NLMIXED Procedure

Fit Statistics

-2 Log Likelihood	100.3
AIC (smaller is better)	104.3

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
alpha	-0.1883	0.1495	22	-1.26	0.2209	0.05	-0.4983	0.1217	1.786E-7
var	0.1571	0.1959	22	0.80	0.4312	0.05	-0.2492	0.5634	-1.06E-6

(a) According to the deviance and Pearson goodness of fit tests, is the model without random effects adequate?

Based on the p-values 0.01 and 0.03, there appears to be some lack of fit. There also appears to be sufficient replication. Assuming a success rate of $\hat{\pi} = e^{-0.1761}/[1+e^{-0.1761}] = 0.456 \approx 0.5$ we obtain expected cell counts on the order of half of the number of trials going into each game for both successes and failures. This gives expected cell counts $2.5, 5.5, 7, 6, \ldots, 3, 6$ across the 23 games (each number is roughly the expected number for both successes and failures). $17/23 = 74\%$ of the attempts have $n_i \geq 10$ and hence expected cell counts $E(y_i|\pi = 0.5) \geq 5$. I'd say the p-values can be trusted here.

(b) Which model is preferred according to AIC?

The model without random effects; 103.3 is smaller than 104.3.

(c) State the GLMM fit in the second PROC NLMIXED statement. In this model, provide an approximate large sample test of $H_0 : \sigma = 0$. Note that $P(\chi^2_1 > 1.0) \approx 0.32$. What do you conclude about free throws across games from this hypothesis test?

The model is

$$y_i \sim \text{bin}(n_i, \pi_i), \quad \log\left(\frac{\pi_i}{1-\pi_i}\right) = \alpha + u_i, \quad u_1, \ldots, u_{23} \overset{iid}{\sim} N(0, \sigma^2).$$

The difference in -2 times the log-likelihood is $101.3 - 100.3 = 1.0$. The p-value is $0.5P(\chi^2_1 > 1.0) = 0.16$, so we do not reject $H_0 : \sigma = 0$. Accepting that $\sigma = 0$ implies that all $u_i = 0$, and so the game outcomes y_i are *independent*.

(d) Does the model including random effects appear to help?

Neither AIC or testing $H_0 : \sigma = 0$ supports the random effects model.

---

3. A contingency table has the factors A, B, C, D, and E. The log-linear model $[AB][AC][AD][DE]$ fits the data well.

(a) Draw an association graph for this model.

(b) True or false: $B \perp D|E$.

False

(c) True of false: $C \perp E|D$.

True

(d) We are only interested in factors A,C, and E. Can we collapse the table over B and D and still retain valid log-linear model inferences among A, C, and E? For example, do we have to worry about Simpson's paradox occurring?

We cannot collapse in this way.
