

List of errors and clarifications for *Analysis of Categorical Data with R*

August 17, 2020

First printing

Chapter 1

1. Page 9:

$$\begin{aligned}\widehat{Var}(\hat{\pi}) &= -E \left\{ \frac{\partial^2 \log [L(\pi|W)]}{\partial \pi^2} \right\}^{-1} \Bigg|_{\pi=\hat{\pi}} \\ &= -E \left\{ -\frac{W}{\pi^2} - \frac{n-W}{(1-\pi)^2} \right\}^{-1} \Bigg|_{\pi=\hat{\pi}} \\ &= \left[\frac{n}{\pi} + \frac{n}{1-\pi} \right]^{-1} \Bigg|_{\pi=\hat{\pi}} \\ &= \frac{\hat{\pi}(1-\hat{\pi})}{n}.\end{aligned}$$

2. Page 10, part 2.: The $w = 1$ should be $w = n$ in two locations. Therefore, replace part 2. with

When $w = 0$ or n , $\sqrt{\hat{\pi}(1-\hat{\pi})/n} = 0$ for $n > 0$. This leads to the lower and upper limits to be exactly the same (0 for $w = 0$ or 1 for $w = n$).

3. Page 17, midway down: Replace $\chi_{1-\alpha/2}^2$ with $\chi_{1-\alpha}^2$ twice as the critical value for the LRT test. _

4. Table 1.8: The overall total is 593, not 539.

5. Page 57, Exercise 23(e): should be $Z_{1-\alpha}$ not $Z_{1-\alpha/2}$

Chapter 2

1. Exercise 5

(a) Part (b): A hat symbol should be above π ; replace “Expected number of failures” with “Estimated expected number of O-ring failures”

(b) Part (e): Use the same values for `Temp` as in the observed data set for the simulation

2. Exercise 9, part (b): Replace the Figure number of 2.6 with 2.4

3. Exercise 10, part (c): Transforming the sensitivity values to the logit scale can be helpful too. Also, if one wanted to check for a two-way interaction between gender and specimen type *only*, the data could be aggregated over the symptomatic status prior to plotting.

4. Exercise 13, part (c): `mod.fit` is the model fit object

5. Exercise 14, part (b): Replace `polio$treatment` with `polio$Treatment`

6. Exercise 22

(a) Footnote: The referenced exercise corresponds to Section 6.5

(b) Part (d): Note that `set1` is a data frame containing the data and `LD.x` contains $\hat{x}_{0.9}$

(c) The `alpha = 0.95` argument in the code should be `alpha = 0.05`.

7. Exercise 27: Replace “presented” with “present”

Chapter 3

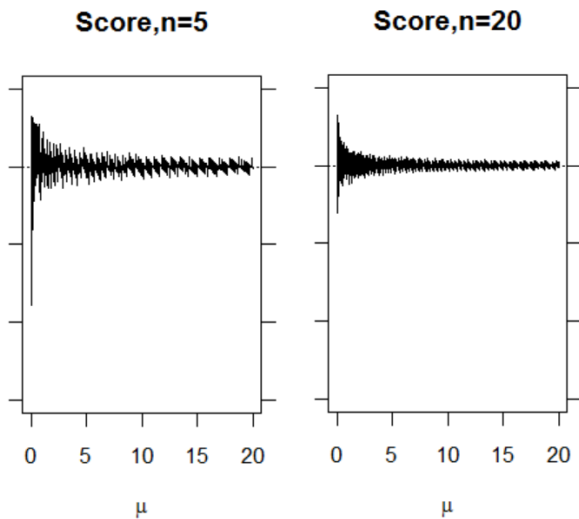
1. Page 147: “We need to estimate $I + J - 2$ parameters...” (currently $I - J - 2$)
2. Page 163, odds ratio interpretation: Replace “3.57 and 14.82” with “3.57 to 14.82”
3. Exercise 3, parts (b) and (c): Replace Equation 3.1 with Equation 3.2 to make the exercise a little easier
4. Exercise 5, part (b): The odds ratio to be derived should compare the cumulative probabilities
5. Exercise 11, part (b): Compute the odds ratios only for the multinomial regression model
6. Exercise 17, part (b): Replace $\log(\pi_j/\pi_{j+1})$ with $\log(\pi_{j+1}/\pi_j)$

Chapter 4

1. Page 200: The score interval should be $3.3 < \mu < 4.5$. This error was due to a missing $\wedge 2$ with the first and last `qnorm()` in the code. Below is the corrected code and output:

```
> (mu.hat + qnorm(p = c(alpha/2, 1 - alpha/2))^2/(2*n)) + qnorm(p = c(alpha/2, 1 -  
  alpha/2)) * sqrt((mu.hat + qnorm(p = 1 - alpha/2)^2/(4*n))/n)  
[1] 3.311097 4.534939
```

2. Page 202: The same score interval calculation error as described above occurred for this example as well. Below is the score portion of Figure 2.



The new version of Figure 4.3 is essentially indistinguishable from what was given in the book.

3. Page 213: The following code/output should be inserted after “Next, the model is fit and the summary is obtained.”

```
> M1 <- glm(formula = Birds ~ Loc, family = poisson(link = "log"), data = alldata)  
> summary(M1)  
Call: glm(formula = Birds ~ Loc, family = poisson(link = "log"), data = alldata)  
Deviance Residuals:  
    Min       1Q   Median       3Q      Max  
-3.4322  -0.7594  -0.1302   0.8874   3.1038  
  
Coefficients:  
      Estimate Std. Error z value Pr(>|z|)
```

(Intercept)	3.87640	0.07198	53.853	<2e-16	***
LocForA	0.90692	0.09678	9.371	<2e-16	***
LocForB	0.13094	0.09062	1.445	0.1485	
LocFrag	0.11874	0.09082	1.307	0.1911	
LocPasA	-0.20010	0.13356	-1.498	0.1341	
LocPasB	-0.23881	0.10844	-2.202	0.0277	*

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 216.944 on 23 degrees of freedom
 Residual deviance: 67.224 on 18 degrees of freedom
 AIC: 217.87

Number of Fisher Scoring iterations: 4

- Exercise 12, part (b) i.: There should be three separate models fit, one for each set of scores in the *GI* interaction. Each model should use 1-2-3-4-5 scores for *PI*.
- Exercise 32, part (b) ii.: Instead of writing out the estimated model, write out the model being fit in the form, $\mu_{ijkl} = \dots$
- Exercise 36: There are structural problems with the data that cause errors to appear upon fitting certain models

Chapter 5

- Page 317: Change “Overdispersion can affect multivariate regression models” to “Overdispersion can affect multinomial regression models”
- Exercise 16, part (c): Remove “, where `fit` is the model object”
- Exercise 19, part (b): Change “Estimate a Poisson regression model for y with either dose, log-dose, or a categorical version of dose...” to “Estimate three separate Poisson regression models for y with either dose, log-dose, or a categorical version of dose...”
- Exercise 23
 - Part (a): “linear form” means “as nominal main effects” and not as numerical scores
 - Part (d): The last sentence before part (i) should read, “Use the default number of groups (the integer portion of $M/5$ when $M \leq 100$).”
- Exercise 27
 - Part (a): The second sentence should read, “Use the default number of groups (the integer portion of $M/5$ when $M \leq 100$).”
 - The last paragraph should be indented one additional level so that it is included in part (b)

Chapter 6

- Page 424: “2 = no” should be “0 = no”
- Page 428: Replace the last three bullets with
 - (**x|b**): random effects are added to the intercept and to the regression coefficient for **x** and these random effects are correlated;
 - (**1|b**)+(0+**x|b**): both the intercept and the regression coefficient for **x** have *independent* random effects (e.g., Equation 6.18);
 - (**x1+x2|b**): The intercept and the regression coefficients for **x1** and **x2** have *correlated* random effects.
- Page 430: In the statement of the model, the intercept term should be -0.65 rather than 0.65
- Page 445: Equation 6.22 should be

$$\frac{\Gamma(n+a+b)}{\Gamma(w+a)\Gamma(n-w+b)}\pi^{w+a-1}(1-\pi)^{n+b-w-1}$$

5. Exercise 3 corresponding to Section 6.5:

(a) Part (a): Replace \hat{x}_π^2 with $\hat{\beta}_1^2$ in the denominator of $Var(\hat{x}_\pi)$

(b) Part (d): Replace \hat{x}_{90} with $\hat{x}_{0.9}$

6. Exercise 1 corresponding to Section 6.6: Replace “hepatitis B” with “hepatitis C”

Appendix A

None

Appendix B

None