STA 303 H1S / 1002 HS – Winter 2012 – Assignment 2

SOLUTIONS

1. (a) For ESWL, 82.6% were a success while for open surgeries, 78.0% were a success, supporting the paper's claim that ESWL works better. However, there is no statistical evidence that there is indeed a difference in the probability of having a successful surgery between the two methods ($p = 0.1285$ for the chi-square test and $p = 0.1281$ for the likelihood ratio test) so the paper's claim is overstated.

   (b) For large stones, ESWL had a 68.8% success rate while open surgeries had a 73.0% success rate. For small stones, ESWL had a 86.7% success rate while open surgeries had a 93.1% success rate. However, there is no statistical evidence for either stone size that there is indeed a difference in the probability of having a successful surgery between the two methods (for large stones: $p = 0.4580$ for the chi-square test and $p = 0.4615$ for the likelihood ratio test; for small stones: $p = 0.1051$ for the chi-square test and $p = 0.0881$ for the likelihood ratio test).

   So when considering outcomes by stone size, the findings contradict the paper's claim. Instead, it appears that open surgery seems to be better than ESWL when looking only at the success rates. But there is no statistical evidence in favour of either method.

   (c) Looking more closely at the data, it appears that stone size is what matters and either treatment is more successful on small stones. Regardless of method, 88.2% $((81 + 234)/(81 + 234 + 6 + 36))$ of treatments on small stones are successful while 72.0% $((192 + 55)/(192 + 55 + 71 + 25))$ of treatments on large stones are successful.

   Moreover, the treatments were disproportionately assigned to the patients. 75.6% $((234 + 36)/(234 + 36 + 81 + 6))$ of patients with small stones had ESWL while only 23.3% $((55 + 25)/(55 + 25 + 192 + 71))$ of patients with large stones had ESWL.

   So the apparent success of ESWL in part (a) is really due to the fact that success is more likely with small stones, and patients with small stones were more likely to have ESWL.

2. (a) The coefficient of the interaction term is not statistically significantly different from 0 ($p = 0.3288$). So there is no evidence that the way the odds of a successful surgery differ between stone sizes differs between treatment groups. This is consistent with the answer to question 1 part (b) where the analysis showed that, for both large and small stones, open surgery had larger success rates but there was no statistically significant difference between the success rates for either stone size.

   (b) Since we are now interested in the main effects of type of surgery and stone size, we should consider the model without the interaction term. The coefficient of the term for surgical method is not statistically significantly different from 0 ($p = 0.1189$), so for each stone size, there is no evidence of a difference in odds of success between the surgical methods. This is consistent with the lack of a significant finding for surgical method in question 1 part (b).

   The coefficient of the term for stone size is statistically significantly different from 0 ($p < 0.0001$) so, for each stone size, there is strong evidence that the odds of success differ with stone size. Although no statistical test for this was carried out, in question 1 part (c) the difference in success rates between stone sizes was noted.

(c) Since there is no evidence of a significant interaction, we will use the model without the interaction term.

(1) For stones of the same size category, the odds ratio for the odds of having a successful surgery for ESWL versus open surgery is 0.700; so the odds of success for patients having ESWL are 70% of the odds of success for patients having open surgery. The 95% confidence interval for the odds ratio is $(0.447, 1.096)$. An odds ratio of 1 indicates that there is no difference in the odds of success for the surgical methods. Since 1 is in the confidence interval, this is consistent with the finding in part (b) of no difference between the surgical methods for patients with the same stone size classification.

(2) For patients undergoing the same surgical method, the odds ratio for the odds of having a successful surgery for large versus small kidney stones is 0.283.; so the odds of success for patients with large stones are 28% of the odds of success for patients with small stones. The 95% confidence interval for the odds ratio is $(0.177, 0.453)$. Since 1 is not in the confidence interval, this is consistent with the finding in part (b) of a significant difference in the odds of success between the stone size classifications for patients having the same surgical method.

3. (a) The Poisson regression model with the three-way interaction and the logistic regression model with the two-way interaction both have deviances of 0. This is because both are the saturated model, fitting the data perfectly. For the Poisson regression model there are 8 observations (2 outcomes $\times$ 2 surgical methods $\times$ 2 stone sizes) and 8 parameters in the model (1 intercept + 3 dummy variables, one for each of the main effects (each having 2 categories) + 3 two-way interactions (3 choose 2) + 1 three-way interaction). For the logistic regression model there are 4 observations (2 surgical methods $\times$ 2 stone sizes) and 4 parameters in the model (1 intercept + 2 dummy variables, one for each of the main effects (each having 2 categories) + 1 two-way interactions).

For the models with the highest order interaction removed, the deviance is the difference in $-2\times$ the log-likelihood between this model and the model containing the interaction term.

(b) Note that, in addition to the deviance, the p-values correspond between the Poisson and logistic regression models but the interpretations differ slightly.

In the Poisson regression, the three-way interaction term's coefficient is not statistically significantly different from 0 ($p = 0.3288$). Thus the interaction between surgical method and outcome (that is, how the mean count of successful versus not successful outcomes differs with surgical method) does not differ between the two stone size categories. This is consistent with the answer to question 2 part (a).

For the Poisson regression model without the three-way interaction, we consider the results of the three two-way interactions:

• The coefficient of the interaction term between surgical method and outcome is not statistically significantly different from 0 ($p = 0.1189$). Thus there is no evidence that the mean count of successful versus non-successful outcomes differs with surgical method. This is consistent with the lack of a difference in the odds of success between surgical methods found in question 2 part (b).

• The coefficient of the interaction term between outcome and stone size is statistically significantly different from 0 ($p < 0.0001$). Thus there is strong evidence that the

mean count of successful versus non-successful outcomes differs with stone size. This is consistent with the difference in the odds of success between stone sizes found in question 2 part (b).

• The coefficient of the interaction term between surgical method and stone size is statistically significantly different from 0 ($p < 0.0001$). Thus there is strong evidence that the mean count of patients having EWSL versus open surgery differs with stone size. This was not considered in the logistic regression analysis (but is consistent with the differences in the percentages of small versus large stone patients who had EWSL discussed in question 1 part (c)).