Longitudinal Homework 4

Tim Vigers

18 October 2019

1. Cereal data

a. Poisson regression

Model

$$Y_i \sim \text{Poisson}(e^{\beta_0+\beta_1*condition+\beta_2*sex+\beta_3*weight})$$

$$\text{Assuming that } Var(Y) = \mu$$

Where Y_i is the number of servings of breakfast cereal eaten in week one. The outcome is the same for all of the following models.

Results

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.9979487	0.1987601	5.020869	0.0000005
Cond	0.8645418	0.1349809	6.404921	0.0000000
Sex	-0.1988221	0.0822303	-2.417869	0.0156117
Wt1	0.1973105	0.1323242	1.491114	0.1359317

The results of Poisson models are interpreted as changes in rate ratio and are given above on the log scale. The interpretation of these results is that on average the experimental program group ate cereal at a rate 2.73 times higher (95% CI: 1.84, 3.12, $p < 0.0001$) than the control group when adjusting for sex and weight, and women ate cereal at a rate of 0.82 times (95% CI: 0.70 0.96, $p = 0.016$) the men's rate after adjusting for experimental condition and weight.

b. Overdispersion (quasilikelihood)

Model

$$Y_i \sim \text{Poisson}(e^{\beta_0+\beta_1*condition+\beta_2*sex+\beta_3*weight})$$

$$\text{Assuming that } Var(Y) = \phi\mu$$

Results

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.9979487	0.3410728	2.9259112	0.0042961
Cond	0.8645418	0.2316274	3.7324672	0.0003229
Sex	-0.1988221	0.1411073	-1.4090131	0.1620961
Wt1	0.1973105	0.2270686	0.8689464	0.3870665

On average, the experimental program group ate cereal at a rate 2.73 times higher (95% CI: 1.55, 3.85, p = 0.0003) than the control group when adjusting for sex and weight.

The estimates are the same as the Poisson model above, but the confidence intervals and p values have changed, because of the introduction of the scale parameter. The standard errors in this model are simply the standard errors from the Poisson model with $\phi = 1$ multiplied by the scale factor in the quasilikelihood model:

```r
scale <- 1.716
pois_ses <- summary(pois_mod)$coefficients[,2]
pois_ses
```

```
## (Intercept)        Cond         Sex         Wt1
##   0.1987601   0.1349809   0.0822303   0.1323242
```

```r
pois_ses * scale
```

```
## (Intercept)        Cond         Sex         Wt1
##   0.3410724   0.2316272   0.1411072   0.2270684
```

```r
summary(quasi_pois_mod)$coefficients[,2]
```

```
## (Intercept)        Cond         Sex         Wt1
##   0.3410728   0.2316274   0.1411073   0.2270686
```

The scale parameter above ($\sqrt{\phi}$) comes from SAS, but to determine it in R you would just divide the SEs from the quasilikelihood model by the SEs from the simple Poisson model. The point here is that the SEs have a linear relationship based on the scale parameter ϕ.

c. Normal error in linear predictor

Model

$$Y_i|\epsilon_i \sim \text{Poisson}(e^{\beta_0+\beta_1*condition+\beta_2*sex+\beta_3*weight+\epsilon_i}),\ \epsilon_i \sim \text{N}(0, \sigma^2)$$

Results

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.8817447	0.3242098	2.7196732	0.0065346
Cond	0.8551232	0.1970427	4.3397869	0.0000143
Sex	-0.1534800	0.1433302	-1.0708139	0.2842531
Wt1	0.1637159	0.2361203	0.6933583	0.4880847

On average the experimental program group ate cereal at a rate 2.35 times higher (95% CI: 1.60, 3.50, p < 0.0001) than the control group when adjusting for sex and weight.

d. Negative binomial

Model

$$Y_i \sim \text{Negative Binomial}(e^{\beta_0+\beta_1*condition+\beta_2*sex+\beta_3*weight}, k)$$

Results

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.1055072	0.3228816	3.4238780	0.0006173

	Estimate	Std. Error	z value	Pr(>|z|)
Cond	0.8573437	0.1954837	4.3857545	0.0000116
Sex	-0.1710841	0.1436796	-1.1907338	0.2337581
Wt1	0.0994404	0.2339898	0.4249773	0.6708532

On average, the experimental program group ate cereal at a rate 2.35 times higher (95% CI: 1.60, 3.47, $p < 0.0001$) than the control group when adjusting for sex and weight.

2. Model comparison

a. Chart

	Poisson Regression	Poisson QL	Poisson + NE	Neg Binom
Intercept (Beta (SE))	0.998 (0.199)	0.998 (0.341)	0.882 (0.324)	1.106 (0.323)
Cond (RR (95% CI))	2.374 (1.838, 3.123)	2.374 (1.546, 3.852)	2.352 (1.598, 3.46)	2.357 (1.603, 3.469)
Sex (RR (95% CI))	0.82 (0.697, 0.962)	0.82 (0.62, 1.079)	0.858 (0.648, 1.136)	0.843 (0.638, 1.114)
Wt (RR (95% CI))	1.218 (0.937, 1.575)	1.218 (0.774, 1.886)	1.178 (0.742, 1.871)	1.105 (0.689, 1.781)
Other Parameter	NA	sqrt(phi) = 1.716	0.543	0.316

b. Study results

All four models agree that condition is significantly associated with cereal consumption ($p < 0.001$ for all models). On average, the experimental program group ate cereal at a rate approximately 2.4 times higher than the control group when adjusting for sex and weight.

c. Model fit

Generally speaking the parameter estimates for these models are very close.

The Poisson regression with $\phi = 1$ has the same estimates as the quasilikelihood regression with $\sqrt{\phi} = 1.716$, but the standard errors are larger for the QL model as discussed in 1b. Because these data are overdispersed (mean of C1 = 6.55, var = 23.54), I would pick the QL model over the regular Poisson regression, because the scale parameter helps deal with the large variance. Also, sex is a significant covariate in the Poisson model, which is not the case in the other three models and suggests that the model isn't quite right.

The other three models are so close that I'm not sure how to pick between them, and all of the approaches make sense given the question of interest. The negative binomial model looks slightly better than the Poisson + normal error model by AIC, but I think these are all good models for these overdispersed data.