CatData HW5

Matthew Vanaman 03-27-2019

3,20, 3,21, 3.22

3.19

(a)

Confirm the sizes of the deviances:

TrainCollisions \sim 1

[1] 35.11981

TrainCollisions ~ Year

[1] 23.51796

Get the difference in deviances and residual df between the models:

```
mod intercept only$deviance - mod with time$deviance
```

[1] 11.60185

```
mod intercept only$df.residual - mod with time$df.residual
```

[1] 1

The deviance approximates a chi-squared distribution when samples are large. Get the p-value for $\Delta \chi^2$:

```
pchisq(11.60185, df=1, lower.tail = FALSE)
```

[1] 0.0006588625

(b)

Get the likelihood χ^2 , or z^2 , and p-value with 1 degree of freedom:

```
(-0.0337/ 0.0130)^2
```

[1] 6.720059

```
pchisq(6.720059, df=1, lower.tail = FALSE)
```

[1] 0.009533447

(c)

The confidence intervals from the negative binomial model are around an additive, linear effect (because negative binomials are a special mixture of Poisson distributions). To get confidence intervals around a multiplicative effect, you exponentiate the intervals:

```
c(exp(-0.060), exp(-0.008))
```

[1] 0.9417645 0.9920319

3.20

(a)

Age	DeathRate_nonSmokers	DeathRate_Smokers	Ratio
35-44	0.106422604161124	0.610605453469956	0.174290294258501
45-54	1.1243324276211	2.4047354790973	0.467549315670744
55-64	4.90367775831874	7.19977631762897	0.681087514665125
65 - 74	10.8317214700193	14.6884624496565	0.737430585886317
75-84	21.203830369357	19.183750235095	1.10530162817521

Deathrates for smokers and non-smokers alike increase with age, as one would expect. Of interest is the fact that the ratio of death rats also increases, such that as people age, the death rates between those who smoke and those who don't approaches a 1:1 ratio (which would mean no effect of smoking on death rate beyond age) before exceeding a 1:1 ratio after age 75. This implies curvilinear relationship between smoking and age. You stand greater risk of death up until you are 75, but if you make it to 75 you may benefit from being smoker.