Poisson Regression: The Training Data

Office workers at a large insurance company are randomly assigned to one of 3 computer use training programmes, and their number of calls to IT support during the following month is recorded. Additional information on each worker includes years of experience and score on a computer literacy test (out of 100). It is reasonable to model calls to IT support as a Poisson process, and the question is whether training programme affects the rate of the process.

Could test $H_0: \lambda_1=\lambda_2=\lambda_3$ with a likelihood ratio test, but ...
```r
> model1 = glm(Support ~ Program, family=poisson)
> summary(model1)

Call:
glm(formula = Support ~ Program, family = poisson)

Deviance Residuals:
   Min       1Q   Median       3Q      Max
-2.8531  -0.6319  -0.0348   0.4552   3.1765

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)    1.403643   0.049567  28.318   <2e-16 ***
ProgramB      -0.159488   0.073066  -2.183   0.0291 *
ProgramC      -0.004926   0.070185  -0.070   0.9440

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 330.39  on 299  degrees of freedom
  Residual deviance: 324.26  on 297  degrees of freedom
  AIC: 1250.2

Number of Fisher Scoring iterations: 4

> anova(model1,test="Chisq") # Overall likelihood ratio test
Analysis of Deviance Table

Model: poisson, link: log
Response: Support

Terms added sequentially (first to last)

             Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
NULL          299 330.39
Program 2    297 324.26 0.04684 *
             ---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```
> # Include covariates
> model2 = glm(Support ~ Score+Experience+Program, family=poisson)
> summary(model2)

Call:
  glm(formula = Support ~ Score + Experience + Program, family = poisson)

Deviance Residuals:
     Min       1Q   Median       3Q      Max
-2.9625  -0.6957  -0.1018   0.5362   2.9386

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept) 1.992744   0.159223 12.515  < 2e-16 ***
Score       -0.009205   0.003019  -3.049  0.00230 **
Experience  -0.028014   0.010317  -2.715  0.00662 **
ProgramB    -0.170519   0.073163  -2.331  0.01977 *
ProgramC    -0.007833   0.070218  -0.112  0.91118

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 330.39  on 299  degrees of freedom
Residual deviance: 305.90  on 295  degrees of freedom
AIC: 1235.8

Number of Fisher Scoring iterations: 4

> anova(model2,test="Chisq") # Sequential
Analysis of Deviance Table

Model: poisson, link: log

Response: Support

Terms added sequentially (first to last)

        Df Deviance Resid. Df Resid. Dev Pr(>Chi)
NULL     299 330.39
Score    1  9.9766       298     320.41 0.001585 **
Experience 1  7.6333       297     312.78 0.005730 **
Program  2  6.8767       295     305.90 0.032118 *

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1