APPLIED STATISTICS
TUTORIAL 11 SOLUTIONS

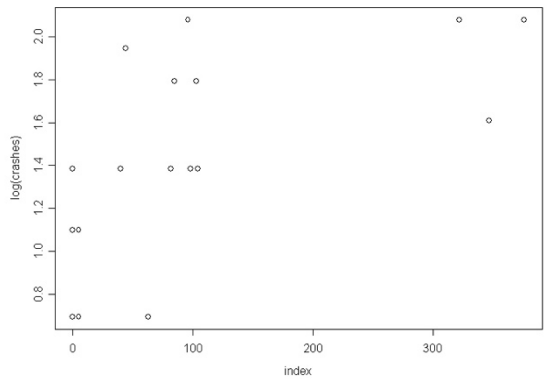
Question 1 (ex from Chapter 22 of the class text)

Some sociologists suspect that highly publicised suicides may trigger additional suicides. In one investigation of this hypothesis, D.P. Phillips collected information about 17 airplane crashes that were known (because of notes left behind) to be murder-suicides. For each of these crashes, Phillips reported an index of the news coverage (circulation in nine newspapers devoting space to the crash multiplied by length of coverage) and the number of multiple- fatality plane crashes during the week following the publicized crash. The data are contained in "ex2216.csv". Is there evidence that the mean number of crashes increases with increasing levels of publicity of a murder-suicide? Provide evidence that your model is appropriate.

Display 22.12 Multiple-fatality plane crashes in the week following a murder–suicide by plane crash, and the amount of newspaper coverage given the murder–suicide

Index of coverage	Number of crashes	Index of coverage	Number of crashes
376	8	63	2
347	5	44	7
322	8	40	4
104	4	5	3
103	6	5	2
98	4	0	4
96	8	0	3
85	6	0	2
82	4		

```
plane=read.table("ex2216.csv",header=T,sep=",")
index=plane$.INDEX
crashes=plane$.CRASHES
plot(index,log(crashes),xlab="index",ylab="log(crashes)")
```



This figure suggests that perhaps a linear term is probably all that is needed. However, to be safe, we will also try including a quadratic term.

```
plane.pois=glm(crashes~index+I(index^2),family=poisson(link=log))
summary(plane.pois)
```

```
Call: glm(formula = crashes ~ index + I(index^2), family = poisson(link =
   log))
Deviance Residuals:
       Min         1Q      Median        3Q      Max
 -1.304483 -0.6099821 0.005270715 0.4396175 1.358694

Coefficients:
                     Value    Std. Error   t value
(Intercept)  1.09556769579 0.24078786631  4.549929
      index  0.00715469510 0.00414915356  1.724375
 I(index^2) -0.00001342052 0.00001051015 -1.276911

(Dispersion Parameter for Poisson family taken to be 1 )

    Null Deviance: 15.2953 on 16 degrees of freedom

Residual Deviance: 8.138918 on 14 degrees of freedom
```

This output shows that the quadratic term in index is not needed. We will refit the model with this term removed.

```
Call: glm(formula = crashes ~ index, family = poisson(link = log))
Deviance Residuals:
       Min         1Q     Median        3Q      Max
 -1.197427 -0.3978249 -0.1766068 0.3537311 1.491924

Coefficients:
                  Value   Std. Error  t value
(Intercept) 1.309863910 0.1580044623 8.290044
      index 0.001993332 0.0008161018 2.442503

(Dispersion Parameter for Poisson family taken to be 1 )

    Null Deviance: 15.2953 on 16 degrees of freedom

Residual Deviance: 9.794001 on 15 degrees of freedom
```

The fitted model is:

$$\log(\hat{\mu}(\text{crashes}))=1.31+0.002\text{index}$$

The deviance goodness-of-fit statistic is 9.79 with a p-value of 0.83. This does not provide any evidence that the model is inappropriate. The test statistic for index of 2.44 shows that increases in media coverage are associated with increases in the number of crashes.