Lecture10 Handout

Elizabeth Colantuoni

4/27/2021

I. Objectives and acknowledgements

Upon completion of this session, you will be able to do the following:

- Understand and explain the Poisson distribution
- Specify a log-linear regression model and interpret its coefficients as log relative rates (risks)
- Specify and maximize the log-likelihood function for a log-linear regression
- Understand and explain the concept of quasi-likelihood estimation for cout data, with emphasis on more flexible mean-variance relationships and robust variance estimates

These lecture notes were adapted from materials written by Dr. Scott Zeger for prior iterations of 140.654.

The goal of the materials is to provide a detailed introduction to log-linear models for count data.

II. Introduction

A. What is a count variable?

A count variable is one that takes on any integer value including 0.

Answer: 0, 1, 2, ..., 237,. . . 32,098,. . . .

B. Example of count variables

Count variables drive questions in many aspects of public health and medical research.

- Number of non-accidental deaths each day in Chicago
- Number of days of work missed due to illness in the Health Intervew Survey (HIS)
- Number of myocardial infarctions (MIs) in (t,t+dt) out of N(t) patients at risk
- Number of subjects in each gender x history of hypertension category in a 2x2 contingency table

C. Characteristics of count variables

Count variables are:

- Non-negative integers
- Variability tends to increase as mean increases
- Effects of predictors tend to be multiplicative (relative changes)

EXAMPLE: Numbers of Non-accidental Death per Day in Chicago, 1987-1994

| Season	Mean	Variance	Variance/Mean
Winter (Dec-Feb)	122	177.6	1.45
Summer (June-Aug)	107	128.4	1.20

III. Poisson process

A Poisson process is a process that defines how observations of a particular event of interest occur within time (or space)

- $Pr(\text{Event in (t,t+dt)}) = \lambda dt$
- $Pr(2 \text{ or more events in (t,t+dt)})$ is neglible
- The chance of an event in one interval is independent of past history of events – "memoryless process"

In a Poisson process, the event times in an interval [0,T] are uniformly distributed, that is, have equal chance of occurring anywhere in the part of the interval.

The figure below displays 10 realizations of a Poisson process with mean 25 events per unit time.

10 Realizations of Poisson Process
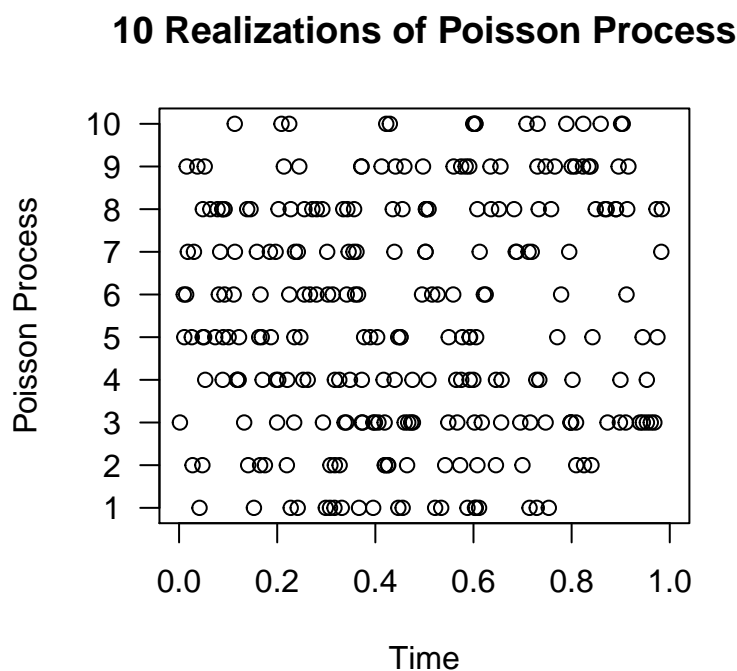
The number of events X in any interval [0,T] has a Poisson distribution with mean λT. The probability distribution is given by:

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

The mean and variance of X are λT.

For the 10 sequences above, with mean 25, the numbers of events are: 21, 22, 37, 30, 28, 23, 23, 37, 27, 16.

Like the Gaussian model for continuous responses, the Poisson distribution is the canonical model for counts but seldom realized by real data. More often, the variance in a sample of counts is greater than the mean. We say the distribution is "over-dispersed" relative to the Poisson. See a simple model for over-dispersed counts below.

A. Derivation of the Poisson distribution

Take the range $[0, T]$ and break it into n bins, t to $t + dt$. Let $Y_j = 1$ if event in j^{th} bin, 0 otherwise for $j = 1, ..., n$. Then $E(Y_j) = \lambda dt = \frac{\lambda T}{n}$

Assumptions:

- $E(Y_j) = Pr(Y_j) = \lambda dt = \frac{\lambda T}{n}$
- $Pr(Y_j > 1) \approx 0$
- $Y_j \perp Y_{j'}$, for all j and j'.

Then $Y = \sum_{j=1}^{n} Y_j \sim P(\mu), \mu = \lambda T$, where λ is the rate per unit "time" or "exposure" and T is the "total time" or "total exposure".

We know that Y will follow a Binomial distribution:

$$Pr(Y = y|\mu) = \binom{n}{y}\left(\frac{\lambda T}{n}\right)^y\left(1 - \frac{\lambda T}{n}\right)^{n-y}$$

To derive the probability density function for the Poisson distribution, let n the number of intervals within T go to infinity.

$$\lim_{n\to\infty}\binom{n}{y}\left(\frac{\lambda T}{n}\right)^y\left(1 - \frac{\lambda T}{n}\right)^{n-y} = \lim_{n\to\infty}\frac{n!}{y!(n-y)!}\frac{(\lambda T)^y}{n^y}\left(1 - \frac{\lambda T}{n}\right)^{n-y}$$

$$= \lim_{n\to\infty}\left(\frac{n}{n}\frac{(n-1)}{n}...\frac{(n-y+1)}{n}\right)\frac{(\lambda T)^y}{y!}\left(1 - \frac{\lambda T}{n}\right)^{n}\left(1 - \frac{\lambda T}{n}\right)^{-y}$$

$$\text{NOTE: } e^x = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n$$

$$= \frac{(\lambda T)^y}{y!}e^{-\lambda T}$$

$$= \frac{\mu^y}{y!}e^{-\mu}$$

IV. Log-linear models for contingency tables.

Historically, the most important application of the Poisson model is for the analysis of contingency tables.

Consider the 2x3 table below of smoking (0-never, 1-former, 2-current) against lung cancer/copd (0-no, 1-either lc, copd or laryngeal cancer) taken from NMES.

```
load('./nmes.rdata')
d1 = nmes
d1[d1=='.'] = NA
d1$smokestatus = ifelse(d1$eversmk==0,0,ifelse(d1$former==1,1,2))
d1 = d1[!is.na(d1$lc5) & !is.na(d1$smokestatus),]

conttable = table(d1$lc5,d1$smokestatus)
conttable
```

```
##
##       0    1    2
##   0 5036 3328 2915
##   1   23  122   72
```

To ask whether smoking is related to lc/copd, we could use logistic regression and express the log odds of lc/copd as a function of two indicator variables for smoking and then test whether smoking is related to lc/copd via a likelihood ratio test.

```
fit.ext = glm(lc5~as.factor(smokestatus),data=d1,family="binomial")
fit.null = glm(lc5~1,data=d1,family="binomial")
library(lmtest)
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
test = lrtest(fit.ext,fit.null)
```

The likelihood ratio test to determine whether lc/copd status is independent of smoking status (2 df test) yielded a chi-square statistic of 124.99 and p-value <0.001.

An alternate, equivalent approach is to assume the contingency table counts are Poisson with mean that we model on a log scale. We take as a null model that the expected counts are determined by the rates of lc/copd and smoking and that there is no interaction, that is, the rate of disease is the same for smokers and non-smokers. We then extend the model by adding two interaction terms: former x lc/copd and current x lc/copd to allow the rates of disease to be different for smokers.

```
c.data = as.data.frame(cbind(c(rep(0,3),rep(1,3)),
              rep(c(1,2,3),2),
              c(conttable[1,],conttable[2,])))
names(c.data) = c("lc5","smokestatus","y")
row.names(c.data) = NULL
c.data
```

```
##   lc5 smokestatus    y
## 1   0           1 5036
## 2   0           2 3328
## 3   0           3 2915
## 4   1           1   23
## 5   1           2  122
## 6   1           3   72
```

```
p.fit.null = glm(y~lc5 + as.factor(smokestatus),data=c.data,family="poisson")
p.fit.ext = glm(y~lc5*as.factor(smokestatus),data=c.data,family="poisson")
p.test = lrtest(p.fit.ext,p.fit.null)

summary(p.fit.null)
```

```
##
## Call:
## glm(formula = y ~ lc5 + as.factor(smokestatus), family = "poisson",
##     data = c.data)
##
## Deviance Residuals:
##      1        2        3        4        5        6
##  1.0265  -0.9804  -0.2887  -8.9165   6.2782   1.9934
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)             8.50987    0.01412  602.73   <2e-16 ***
## lc5                    -3.95080    0.06853  -57.65   <2e-16 ***
## as.factor(smokestatus)2 -0.38279    0.02208  -17.34   <2e-16 ***
## as.factor(smokestatus)3 -0.52690    0.02307  -22.83   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 14506.82  on 5  degrees of freedom
## Residual deviance:   124.99  on 2  degrees of freedom
## AIC: 180.86
##
## Number of Fisher Scoring iterations: 5
```

```
summary(p.fit.ext)
```

```
##
```

```
## Call:
## glm(formula = y ~ lc5 * as.factor(smokestatus), family = "poisson",
##     data = c.data)
##
## Deviance Residuals:
## [1]  0  0  0  0  0  0
##
## Coefficients:
##                            Estimate Std. Error z value Pr(>|z|)
## (Intercept)                 8.52437    0.01409 604.930  < 2e-16 ***
## lc5                        -5.38887    0.20899 -25.785  < 2e-16 ***
## as.factor(smokestatus)2    -0.41424    0.02234 -18.543  < 2e-16 ***
## as.factor(smokestatus)3    -0.54674    0.02327 -23.493  < 2e-16 ***
## lc5:as.factor(smokestatus)2  2.08277    0.22842   9.118  < 2e-16 ***
## lc5:as.factor(smokestatus)3  1.68791    0.24064   7.014 2.31e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1.4507e+04  on 5  degrees of freedom
## Residual deviance: 7.7938e-14  on 0  degrees of freedom
## AIC: 59.867
##
## Number of Fisher Scoring iterations: 3
```

p.test

```
## Likelihood ratio test
##
## Model 1: y ~ lc5 * as.factor(smokestatus)
## Model 2: y ~ lc5 + as.factor(smokestatus)
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1   6 -23.933
## 2   4 -86.429 -2 124.99  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

A. Poisson and Binomial Distribution Link

In fact, there is a strong connection between Poisson and logistic regression. If Y_1 and Y_2 are independent Poisson with means λ_1 and λ_2, respectively, then given the total $Y_1 + Y_2$, Y_1 has a binomial distribution with probability $\lambda_1/(\lambda_1 + \lambda_2)$. The conditional likelihood for a Poisson model given the total number of events is the same as a logistic model likelihood function.

$$
\begin{aligned}
Pr(Y_1|Y_1 + Y_2) \quad &= \quad \frac{Pr(Y_1 \text{ and } Y_1 + Y_2)}{Pr(Y_1 + Y_2)} \\[2mm]
&= \quad \frac{Pr(Y_1, Y_2)}{Pr(Y_1 + Y_2)} \\[2mm]
&= \quad \frac{\lambda_1^{Y_1} e^{-\lambda_1}}{Y_1!} \frac{\lambda_2^{Y_2} e^{-\lambda_2}}{Y_2!} \bigg/ \frac{(\lambda_1+\lambda_2)^{Y_1+Y_2} e^{-(\lambda_1+\lambda_2)}}{(Y_1+Y_2)!} \\[2mm]
&= \quad \frac{(Y_1+Y_2)!}{Y_1! Y_2!} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{Y_1} \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{Y_2} \\[2mm]
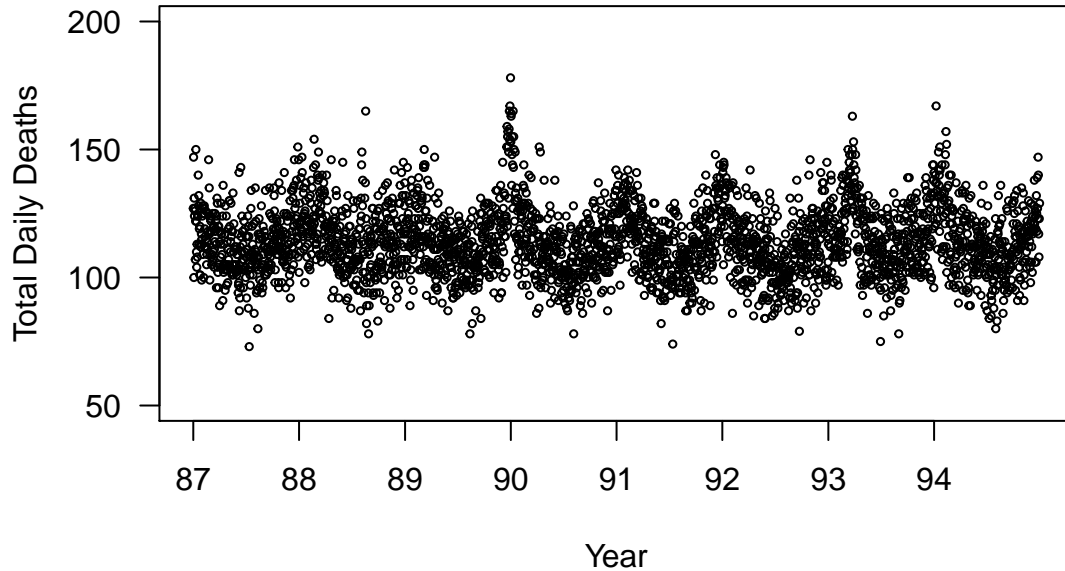&\sim \quad Binomial(Y_1 + Y_2 | \tfrac{\lambda_1}{\lambda_1+\lambda_2})
\end{aligned}
$$

In a 2x2 table, $y_j \sim P(\lambda_j)$ and $y_j|y_j + y_{j'} \sim Binomial(y_j + y_{j'}, \frac{\lambda_j}{\lambda_j+\lambda_{j'}})$.

y_1	y_2	$y_1 + y_2$
y_3	y_4	$y_3 + y_4$

V. Overdispersed counts

It is unusual to encounter a data set of counts that follow the Poisson distribution as we illustrated above for the daily deaths in Chicago. Why are the number of deaths per day not Poisson?

Total Non–accidental Daily Deaths – Chicago, 1987 – 1994

Suppose the rate of death varies from one day to the next because of measured and unmeasured factors. We will show how to account for measured covariates below, but now assume there are no predictors of mortality. The expected number of deaths might still vary from day to day. Let λ_t be the expected number on day t and suppose $\lambda_t = \lambda \epsilon_t$ where ϵ_t is a Gamma random variable with mean 1 and variance ϕ. Then the number of events on day t follows a negative binomial distribution with mean λ and variance $\lambda + \lambda^2\phi$. Note that the variance is now greater than the mean. Hence, the effect of variation in the expected number of events is to produce over-dispersion relative the Poisson.

VI. Log-linear regression for counts

We can express the log-linear regression model for a count variable as:

- Y is a count variable, e.g. number of events
- $\mu = E(Y|X) = N\lambda$ is the expected number of events given exposure N with λ the rate of events or expected number of events per unit exposure time

The model is:

$$\log(\mu) \quad = \quad \log(N\lambda)$$

$$= \quad \log(N) + \log(\lambda)$$

$$\text{Let } \log(\lambda) \quad = \quad X_i^!\beta$$

$$\log(\mu) \quad = \quad \log(N) + X_i^!\beta$$

In the model above, the term $\log(N)$ is referred to as the **offset**.

Consider a simple log-linear model where $X_i = 0$ vs. 1 a binary indicator for a particular risk factor vs. not.

$$\log(\mu_i) = \log(N_i) + \beta_0 + \beta_1 X_i$$

Then,

- β_0 is the log expected number of events per unit exposure time among persons with $X_i = 0$
- $exp(\beta_0)$ is the expected number of events per unit exposure time among persons with $X_i = 0$
- $\beta_0 + \beta_1$ is the log expected number of events per unit exposure time among persons with $X_i = 1$
- $exp(\beta_0 + \beta_1)$ is the expected number of events per unit exposure time among persons with $X_i = 1$
- The expected number of events for subject i is $N_i exp(\beta_0 + \beta_1 X_i)$
- $exp(\beta_1) = \frac{exp(\beta_0+\beta_1)}{exp(\beta_0)}$ is the relative expectied number of events per unit time comparing persons with $X_i = 1$ to those with $X_i = 0$

A. Example

For insulin-dependent diabetic patients, in addition to monitoring blood sugar for hyperglycemia (i.e. high blood sugar), patients record periods of hypoglycemia (i.e. low blood sugar).

Consider a hypothetical example of a study of insulin-dependent diabetic patients followed for 4 weeks after acquiring an insulin pump. The patients record and report the total number of hypoglycemic episodes during the 4 week follow-up. The goal of the analysis is to compare the total number of hypoglycemic episodes for male and female diabetic patients.

1. Example: Constant exposure time

We can describe the data as Y_i is the number of hypoglycemic episodes reporting during the 4 week follow-up period. Then $Y_i \sim P(\mu_i)$ with model:

$$Log(E(Y_i)) = Log(\mu_i) = \beta_0 + \beta_1 male_i$$

Interpretations:

- $\hat{\beta}_0$ is the logarithm of the mean number of hypoglycemic episodes during the 4-week follow-up among females. The mean number of hypoglycemic episodes among females during the follow-up is $exp(\hat{\beta}_0) = exp(2.52) = 12.4$.

- $\hat{\beta}_0 + \hat{\beta}_1$ is the logarithm of the mean number of hypoglycemic episodes during the 4-week follow-up among males. The mean number of hypoglycemic episodes among males during the follow-up is $exp(\hat{\beta}_0 + \hat{\beta}_1) = exp(2.52 + 0.20) = 15.2$.

- $\hat{\beta}_1$ is the difference in the log mean number of hypoglycemic episodes during the 4 week follow-up comparing males to females OR the log relative mean number of hypoglycemic episodes during the 4 week follow-up comparing males to females.

- $exp(\hat{\beta}_1) = exp(0.20) = 1.22$ represents the relative mean number of hypoglycemic episodes comparing males to females. The mean number of hypoglycemic episodes during the 4-week follow-up is 22% greater for males compared to females.

2. Example: Non-constant exposure time

Now suppose that not all patients were able to be followed for the entire 4-week period; patients were followed from 10 to 28 days. Patients report the number of hypoglycemic episodes within the duration of the patient's specific follow-up.

Now, we will describe the data as Y_i is the number of hypoglycemic episodes, N_i is the total number of days during the maximum of 4-week follow-up. Then $Y_i \sim P(\mu_i = N_i \lambda_i)$ with model:

$$
\begin{aligned}
Log(E(Y_i)) &= Log(\mu_i) \\
&= Log(N_i \lambda_i) \\
&= Log(N_i) + Log(\lambda_i) \\
&= Log(N_i) + \beta_0 + \beta_1 male_i
\end{aligned}
$$

Then,

- for patient i, the expected number of hypoglycemic episodes is $N_i \lambda_i$ where N_i is the total follow-up time in days for patient i and λ_i is the risk of a hypoglycemic episode per unit time / per day.

- β_0 is the logarithm of the risk of a hypoglycemic episode in a day for females.

- $\beta_0 + \beta_1$ is the logarithm of the risk of a hypoglycemic episode in a day for males.

- $exp(\beta_1)$ is the relative risk of a hypoglycemic episode in a day comparing males to females OR the relative expected number of hypoglycemic episodes comparing males and females who have the same duration of follow-up.

```
set.seed(1346)
n = 100
N = sample(seq(10,28),size=n,replace=TRUE)
male = rbinom(N,1,0.5)
```

```
Y= rpois(n,exp(log(N) + log(0.7)+0.2*male))
fit = glm(Y~male,family="poisson",offset=log(N))
summary(fit)$coefficients
```

```
##               Estimate Std. Error   z value     Pr(>|z|)
## (Intercept) -0.2752677 0.03603750 -7.638368 2.199923e-14
## male         0.1142061 0.05012278  2.278527 2.269520e-02
```

```
expected.Y = fit$fitted
predicted.lambda = exp(fit$coefficients[1] + male*fit$coefficients[2])
head(cbind(N,Y,male,expected.Y,predicted.lambda))
```

```
##     N  Y male expected.Y predicted.lambda
## 1 17 19    1   14.47107        0.8512397
## 2 22 18    0   16.70611        0.7593688
## 3 19 16    1   16.17355        0.8512397
## 4 19 15    1   16.17355        0.8512397
## 5 22 13    0   16.70611        0.7593688
## 6 25 18    1   21.28099        0.8512397
```

Interpretations:

- $\hat{\beta}_0$ is the logarithm of the daily risk of a hypoglycemic episode among females. The estimated daily risk of a hypoglycemic episode among females is $exp(\hat{\beta}_0) = exp(-0.31) = 0.73$.

- $\hat{\beta}_0 + \hat{\beta}_1$ is the logarithm of the daily risk of a hypoglycemic episode among males. The estimated daily risk of a hypoglycemic episode among males is $exp(\hat{\beta}_0 + \hat{\beta}_1) = exp(-0.31 + 0.23) = 0.93$.

- $\hat{\beta}_1$ is the difference in the log daily risk of a hypoglycemic episode comparing males to females OR the log relative risk comparing males to females.

- $exp(\hat{\beta}_1) = exp(0.23) = 1.26$ represents the relative daily risk of a hypoglycemic episode comparing males to females. OR: For patients with the same duration of follow-up, the expected number of hypoglycemic episodes is 26% greater for males compared to females.

B. Estimation

The likelihood function is:

$$L(\beta|Y) = \prod_{i=1}^{n} \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}$$

The log-likelihood is:

$$logL(\beta|Y) = \sum_{i=1}^{n}(-\mu_i) + y_i log(\mu_i) - log(y_i!)$$

Before deriving the score equations, note that:

- $\mu_i = N_i e^{X_i'\beta}$

- $\frac{\partial \mu_i}{\partial \beta} = N_i e^{X_i'\beta} X_i' = \mu_i X_i'$

- $\frac{\partial log(\mu_i)}{\partial \beta} = X_i'$

The score equation is:

$$
\begin{aligned}
\frac{\partial logL(\beta|Y)}{\partial \beta} &= \sum_{i=1}^{n}\left(-\frac{\partial \mu_i}{\partial \beta}\right) + y_i \frac{\partial log(\mu_i)}{\partial \beta} \\[2mm]
&= \sum_{i=1}^{n}(-\mu_i X_i^!) + y_i X_i^! \\[2mm]
&= \sum_{i=1}^{n} X_i^!(y_i - \mu_i)
\end{aligned}
$$

Similar to all generalized linear models, $\hat{\beta} \sim N(\beta, (X^! diag(\hat{\mu})X)^{-1})$. Inferences for β are the same as before.

C. Robust Variance Estimates for Over-dispersion

Count data is almost always over-dispersed, i.e. $Var(Y_i) > E(Y_i)$.

Solution: Assume $E(Y_i|X_i) = \mu_i = N_i e^{X_i^! \beta}$ and $Var(Y_i|X_i) = \mu_i \phi$.

We can estimate ϕ by:

$$
\hat{\phi} = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \bigg/ (n-p)
$$

which is the Pearson residual estimate of ϕ.

Alternatively, you can use the deviance estimator as:

$$
\hat{\phi} = 2\sum_{i=1}^{n} [Y_i log(Y_i/\mu_i) - (Y_i - \mu_i)] \bigg/ (n-p)
$$

Either is fine for computing the robust variance estimate.

1. Example: Chicago mortality

Here we will consider a log-linear model to describe the expected daily non-accidental mortality in Chicago as a function of particulate pollution level (PM10), temperature (current temp and average temp of prior 3 days, each modeled by a natural spline with 3 degrees of freedom), and a step function of calendar time with indicator variables for either: year, season or month. We will extract the PM10 coefficients and standard error estimates based on the Poisson model assumption and allowing for overdispersion in the data.

The PM10 coefficient is scaled to represent the relative increase in daily non-accidental deaths per 10 unit change in PM10.

```
data=read.table("./chicago.txt",sep=",",header=T)
#The data contains 7 variables:
#date -> form is YY-MM-DD
#total -> total deaths on that day
#temp -> average daily temp on that day (F)
#day -> counter of days in dataset (1 - 2292)
#year -> year of study (87 to 94)
#month -> month (1-12)
#pm10 -> measure of particulate on the day in micrograms per cubic meter
```

```r
#Compute average daily temp over last three days
data$avgtemp = NULL
for(i in 1:nrow(data)) {
  if(i==1) data$avgtemp[i] = data$temp[1]
  if(i==2) data$avgtemp[i] = mean(data$temp[1:2])
  if(i>2) data$avgtemp[i] = mean(data$temp[(i-2):i])
  }
data$season = ifelse(data$month>=3 & data$month<=5,1,
                    ifelse(data$month>=6 & data$month<=8,2,
                          ifelse(data$month>=9 & data$month<=11,3,4)))
data$season = (data$year-87)*4+data$season
data$month = (data$year-87)*12+data$month

library(splines)
fit.poisson.year = glm(total~ pm10+ns(temp,3)+ns(avgtemp,3)+as.factor(year),
                  data=data,family="poisson")
fit.poisson.season = glm(total~ pm10+ns(temp,3)+ns(avgtemp,3)+as.factor(season),
                  data=data,family="poisson")
fit.poisson.month = glm(total~ pm10+ns(temp,3)+ns(avgtemp,3)+as.factor(month),
                  data=data,family="poisson")

fit.robust.month = glm(total~ pm10+ns(temp,3)+ns(avgtemp,3)+as.factor(month),
                  data=data,family="quasipoisson")
fit.robust.season = glm(total~ pm10+ns(temp,3)+ns(avgtemp,3)+as.factor(season),
                  data=data,family="quasipoisson")
fit.robust.year = glm(total~ pm10+ns(temp,3)+ns(avgtemp,3)+as.factor(year),
                  data=data,family="quasipoisson")

out = cbind(rbind(summary(fit.poisson.year)$coeff[2,1:2]*10,
                  summary(fit.poisson.season)$coeff[2,1:2]*10,
                  summary(fit.poisson.month)$coeff[2,1:2]*10),
            rbind(summary(fit.robust.year)$coeff[2,1:2]*10,
                  summary(fit.robust.season)$coeff[2,1:2]*10,
                  summary(fit.robust.month)$coeff[2,1:2]*10))
out = as.data.frame(round(out,5))
names(out) = c("Poisson beta","Poisson SE","Robust beta","Robust SE")
out
```

```
##   Poisson beta Poisson SE Robust beta Robust SE
## 1      0.00349    0.00104     0.00349   0.00116
## 2      0.00229    0.00107     0.00229   0.00117
## 3      0.00178    0.00111     0.00178   0.00118
```