STA 511 Homework #4

Suruchi Jaikumar Ahuja

October 25 2015

1. The counts of a hospial insurance policies reporting y_i claims are

y_i	count
0	7840
1	1327
2	239
3	42
4	14
5	4
6	4
7	1

   (a) The Log - Likelihood Function (Refer to Figure 1)

   R Code

```
install.packages("asbio")
library(asbio)
library(ElemStatLearn)
library(MASS)

X <- c(rep(0,7840), rep(1,1327), rep(2,239),
        rep(3,42), rep(4, 14), rep(5,4), rep(6, 4), rep(7,1))
n <- length(X)
negloglike<-function(lam)
{
 sum(X) *log(lam) -n* lam + sum(log(factorial(X)))
}
```

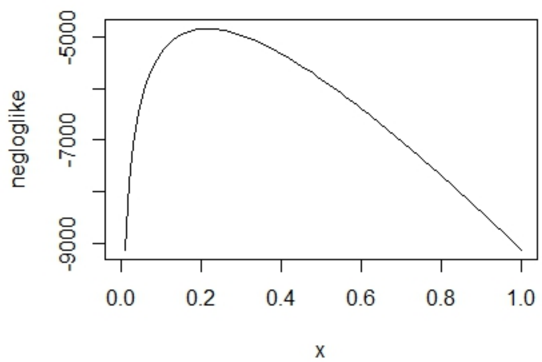

Figure 1: Plot of Log Likelihood Function

```
# Now we perform the optimization on the negative log like function.
out<-nlm(negloglike,p=c(0.5), hessian = TRUE)
#nlm is a nonlinear  minimization function
mean(X)
plot(negloglike)
```

.

Output :
The mean of x is 0.2151832

```
 > out
$minimum
[1] -47342038

$estimate
[1] 5000.5

$gradient
[1] -9470.592

$hessian
              [,1]
[1,] -8.143346e-05
```

```
$code
[1] 5

$iterations
[1] 5

To access the elements of out :
> out$estimate
[1] 5000.5
>  out$hessian
              [,1]
[1,] -8.143346e-05
```

(b) Finding the MLE of λ using computational methods

A likelihood for a statistical model is defined by the same formula as the density, but the roles of the data x and the parameter θ are interchanged
$L_x(\theta) = f_\theta(x)$.
The so-called method of maximum likelihood uses as an estimator of the unknown true parameter value,
the point $\hat{\theta}_x$ that maximizes the likelihood Lx. This estimator is called the maximum likelihood estimator (MLE). The R function nlm minimizes arbitrary functions written in R.
So to maximize the likelihood, we use the negative of the log likelihood in nlm.

```
poisson.LL<-function(lam) sum(log(dpois(X,lam)))
poisson.negloglik<-function(lam) -poisson.LL(lam)
nlm(poisson.negloglik,4,hessian=T)->out1
```

Output :

```
> out1
$minimum
[1] 5508.31

$estimate
[1] 0.2151827
```

```
$gradient
[1] 0.0006020855

$hessian
         [,1]
[1,] 43972.98

$code
[1] 1

$iterations
[1] 11
```

(c) Estimating the probability that a randomly selected policy has 2 claims,
$g(\lambda) = Pr(\lambda_i = 2)$

$\lambda = 0.215$

$$Pr(\lambda_i = 2) = \frac{\lambda^y e^{-\lambda}}{y!}$$

$$\rightarrow \frac{(0.215)^2 e^{(-0.215)}}{2!}$$

$\rightarrow 0.0186411$

2. Let $X_1, ...X_n \sim N(\theta, 1)$

$$f(x) = \begin{cases} 1 & : x_i > 0 \\ 0 & : x_i \leq 0 \end{cases}$$

(a) Maximum Likelihood Estimator of θ

$$f(x_i; \mu, \sigma_2) = \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{x_i - \mu^2}{2\sigma^2})$$

$L[\theta] = \prod_{i=1}^{n} f(x_i)$

$= \prod_{i=1}^{n} \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$

Since the σ = 1

$= \prod_{i=1}^{n} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\theta)^2}{2\sigma^2}}$

$= 0 - \dfrac{2(\bar{X}_i - \theta)(-1)}{2}$

$(x_i - \hat{\theta}) \to 0$ (Setting the value to 0)

$\hat{\theta}_{MLE} = \dfrac{\Sigma_{i=1}^{n} X_i}{n} = \bar{X}_i$

(b) Maximum likelihood Estimator of ϕ

Here we have Pr(Y_1 = 1) = Pr($X_1 > 0$) = 1 - Pr($X_1 \leq 0$)

So we have g(θ) = 1− f(0)

$\hat{\phi}_{MLE} = g(\hat{\phi}_{MLE})$

$Pr(\theta, 1)(X \geq 0) = Pr(\dfrac{X-\theta}{1} \leq \dfrac{0-\theta}{1})$

$= 1 - Pr(Z \leq -\theta)$

Here Z is a standard normal distribution N(0,1)

$\to \hat{\phi}_{MLE} = 1 - \phi(-\theta)$

Note : The cdf of certain normal distribution cannot be calculated as you cant take the anti-deprivative of the normal distribution pdf.
So the function has to be made into a standard normal, which can be done using the pnorm function in R.
A certain probablity value for a variable is obtained, but the actual cdf cannot be obtained.
It has to be experimentally calculated by converting the N(θ,1) distribution into a standard.

(c) Computing the asymptotic error for θ and ϕ

The Fischer information is a way of measuring the amount of information that an observable random variable X carries about an unknown parameter θ of a distribution that models X.

So we have, $I_n(\theta) = V(\theta)[\Sigma_{i=1}^{n} S(X_i;\theta)]$

$$\hat{e}rror = \sqrt{\frac{1}{-nE[\frac{d^2}{d\theta^2} log f(x|\theta)]}}$$

$$I_n(\theta) = -nE[\tfrac{d^2}{d\theta^2} log(x|\theta)]$$

On differentiating the above equation,

$$-nE[\frac{d}{d\theta}(x-\theta)]$$

$$-nE(-1) = nE$$

So now the asymptotic error for θ is,

$$\rightarrow \sqrt{\frac{1}{n}}$$

And now for the asymptotic error for ϕ

$$\hat{e}rror = \left| g\prime(\hat{\theta}_{MLE}) \right| [\hat{e}rror(\hat{\theta}_{MLE})]$$

$$g\prime(\hat{\theta}) = -\frac{1}{\sqrt{2\pi}}e^{\frac{\theta-\mu^2}{2}}$$

$$= \left|\frac{1}{\sqrt{2\pi}}e^{\frac{(-\theta)^2}{2}}\right|$$

$$\rightarrow \left|\frac{1}{\sqrt{2\pi}}e^{\frac{(-\theta)^2}{2}}\right|\sqrt{\frac{1}{n}}$$

So now substitute θ = x;

$$\rightarrow \left|\frac{1}{\sqrt{2\pi}}e^{\frac{(-x)^2}{2}}\right|\sqrt{\frac{x}{n}}$$

3. Consider the given data
   X_1,X_2, X_3, . . . X_n,
   Y_1, Y_2, . . .Y_m
   where the Xs come from model f andthe Ys come from model g.

   All Xs are independent and all Ys are independent and any X is independent from any Y .

   Now f(x) = $\frac{1}{\theta}e^{(-\frac{x}{\theta})}$ for x > 0 and

   g(y) = $e^{(-\frac{5y}{\theta})}(1 - e^{(\frac{5}{\theta})})^{(1-y)}$ where y is either 0 or 1.

   (a) The Joint Distribution → f(x).g(x)

$$(\frac{1}{\theta}e^{(-\frac{x}{\theta})}) * (e^{(-\frac{5y}{\theta})}(1 - e^{(\frac{5}{\theta})})^{(1-y)})$$

$$f_{x,y}(X, Y) = \frac{1}{\theta}e^{\dfrac{-x-5y}{\theta}}(1 - e^{-\frac{5}{\theta}})^{(1-y)}$$

$L(\theta) = \prod_{i=1}^{n} (\frac{1}{\theta} e^{\frac{-x-5y}{\theta}} (1 - e^{-\frac{5}{\theta}})^{(1-y)})$

$(\theta) = \frac{1}{\theta^{n_1}} \left(e^{-\sum_{i=1}^{n_1} \frac{x_i}{\theta}} \right) \left(e^{-5\sum_{i=1}^{n_2} \frac{y_i}{\theta}} \left(1 - e^{\frac{-5}{\theta}} \right)^{\left(n_2 - \sum_{i=1}^{n_2} y_2 \right)} \right)$

(b) Now we have 10 observations from f - 2.8, 5.6, 24.7, 6.5, 1.6, 10.6, 1.0, 7.8, 7.2,13.9 and 15 observations from g: 0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0. Now to compute the MLE for .

```
 x <- c(2.8,5.6,24.7,6.5,1.6,10.6,1.0,7.8,7.2,13.9)
 nx <- length(x)
 y <- c(0,0,0,1,1,1,0,0,1,0,0,0,0,0,0)
 ny <- length(y)
 mlefunc <- function(theta)
 {
  -nx * log(theta) - sum(x)/theta - 5 * sum(y)/theta +
                         (ny-sum(y)) * log(1-exp(-5/theta))
  }
opt <- optimize(func1, interval=c(0,10),maximum=TRUE)
```

Output :

The maximum Likelihood estimator for θ is estimated to be 5.971734