# STA511 Homework #4

Ezgi Karaesmen

11/11/15

1. The counts of hospital insurance policies reporting exactly $y_i$ claims during a particular year were given, where the observations are i.i.d. $Poisson(\lambda)$ and there are 9471 total observations.

   (a) Log-likelihood function was derived and plotted against the possible values of $\lambda$.

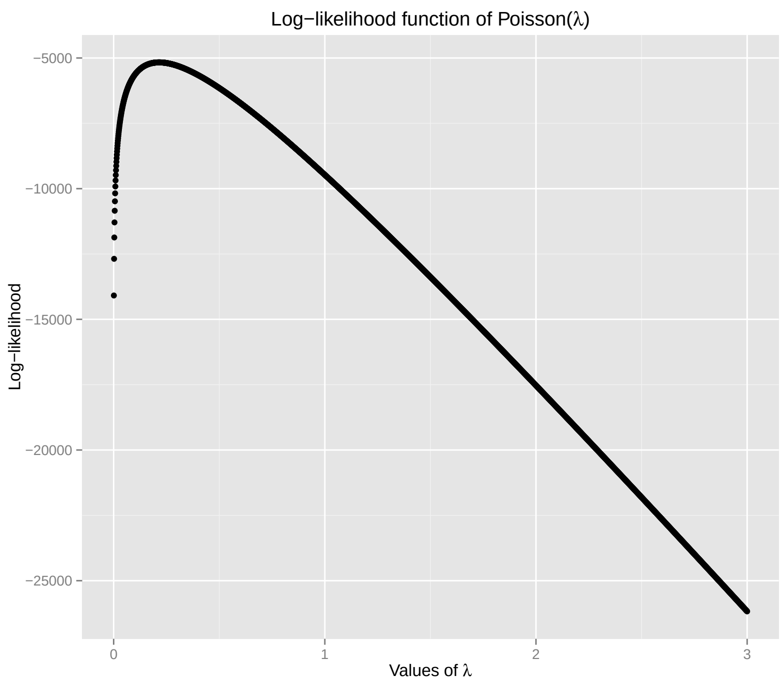


Figure 1: Log-likelihood function of $Poisson(\theta)$ for the possible values of $\lambda$.

The `R` code for the log-likelihood function and the plot is given below:

```
obs <- c(rep.int(0, 7840), rep.int(1, 1327), rep.int(2, 239), rep.int(3, 42),
         rep.int(4, 14), rep.int(5,4), rep.int(6,4), rep.int(7,1))

poisson.lik<-function(lambda){
  n<-length(obs)
  logl<-sum(obs)*log(lambda)-n*lambda
  return(logl)
}

pos.lam <- seq(0, 3, by=0.001)[-1] # Possible values for lambda
poisson.lik2 <- function(lambda){1*sapply(lambda, poisson.lik)}
l.pois <- poisson.lik2(lambda = pos.lam)
library(ggplot2)
qplot(pos.lam, l.pois, xlab = expression(paste("Values of ", lambda)), ylab = "Log-likelihood",
     type="l", main=expression(paste("Log-likelihood function of Poisson(", lambda, ")")))
```

(b) The MLE of $\lambda$ was determined via `nlminb` function in `R`, where the starting value for optimization was set to `3` (`start = 3`) and the negative log-likelihood function of the $Poisson(\theta)$ was set as the `objective` parameter. Consequently, $\hat{\lambda}_{MLE}$ was determined as `0.2151832`.

(c) The probability that a randomly selected policy has 2 claims, was computed as `0.01866955`. `R` code for the computation is given below:

```
g_lamb <- function(x, lambda){lambda^x*exp(-lambda)/factorial(x)}
g_lamb(2, mle_lamb) # mle_lamb = 0.2151832, as computed in the previous step
```

2. $X$ is defined as $X_1, ..., X_n \sim N(\theta, 1)$ and $Y_i$ is defined as:

$$Y_i = \begin{cases} 1 & \text{if } X_i > 0, \\ 0 & \text{if } X_i \leq 0. \end{cases}$$

(a) The MLE of $\theta$ was computed as follows:

$$f_{N(\theta,1)}(x) = \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}$$

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}$$

$$\ell(\theta) = \frac{-n}{2} log(2\pi) - \frac{\sum_{i=1}^{n}(X_i-\theta)^2}{2}$$

$$\ell'(\theta) = \sum_{i=1}^{n} X_i - n\theta = 0$$

$$\hat{\theta}_{MLE} = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X}$$

(b) Assuming $\Phi = Pr(Y_1 = 1)$, MLE of $\Phi$ was computed as follows:

According to the definition of Y, it is known that $Pr(Y_i = 0) = Pr(X_i > 0)$

And $Pr(X_i > 0) = 1 - Pr(X_i \leq 0)$, which can also be determined as,

$Pr(X_i \leq 0) = F(0)$, where F(x) is the cdf of $\sim N(\theta, 1)$.

Therefore, we can define $\Phi = 1 - Pr(X_i \leq 0)$ or $\Phi = 1 - F_{N(\theta,1)}(0 \mid \theta)$.

$$F_{N(\theta,1)}(0 \mid \theta) = P(X \leq 0) = P\left(\frac{X-\theta}{1} \leq \frac{0-\theta}{1}\right) = P(Z \leq -\theta), \quad \text{where} \quad Z \sim N(0,1)$$

$$P(Z \leq -\theta) = G(-\theta), \quad \text{where G is the cdf of standard normal}, N(0,1)$$

$$\text{Hence we can define} \quad \hat{\Phi}_{MLE} = 1 - G(-\hat{\theta}_{MLE}) = 1 - G(-\bar{X})$$

(c) The asymptotic standard error for $\hat{se}(\hat{\theta}_{MLE})$ can be computed using the Fisher Information as follows:

$$\hat{se}(\hat{\theta}_{MLE}) = \sqrt{\frac{1}{I_n(\hat{\theta}_{MLE})}}$$

$$\begin{aligned} I_n(\theta) &= -nE\left[\frac{\delta^2}{\delta\theta^2} logf(x|\theta)\right] \\ &= -nE\left[\frac{\delta^2}{\delta\theta^2} - \frac{1}{2}log(2\pi) - \frac{(x-\theta)^2}{2}\right] \\ &= -nE\big[-1\big] = n \end{aligned}$$

$$\hat{se}(\hat{\theta}_{MLE}) = \sqrt{\frac{1}{n}}$$

Consequently, the asymptotic standard error for $\hat{se}(\hat{\Phi}_{MLE})$ can be computed as follows:

$$\hat{se}(\hat{\Phi}_{MLE}) = \left|g'(\hat{\theta}_{MLE})\right| \cdot \hat{se}(\hat{\theta}_{MLE})$$

It was determined that $\Phi(\theta) = 1 - G(-\theta)$, where G is the CDF of standard normal, $\sim N(0,1)$. Therefore, $g'$ would be equal to pdf of $\sim N(0,1)$ which can be written as:

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

And finally $\hat{se}(\hat{\Phi}_{MLE})$ can be computed as,

$$\hat{se}(\hat{\Phi}_{MLE}) = \left| -\frac{1}{\sqrt{2\pi}} e^{-\bar{X}^2/2} \right| \cdot \sqrt{\frac{1}{n}}$$

3. Data was given as $X_1, X_2, ...X_n, Y_1, Y_2, ...Y_m$ where the $X_i$ come from model $f(x)$ and the $Y_i$ come from model $g(y)$. All $X_i$ are independent and all $Y_i$ are independent and any $X$ is independent from any $Y$. $f(x)$ and $g(x)$ are defined as:

$$f(x) = \frac{1}{\theta} e^{(-x/\theta)}\ , \quad x > 0$$

$$g(y) = e^{-5y/\theta} \cdot (1 - e^{5/\theta})^{1-y}\ , \quad y = \{0, 1\}$$

(a) As X and Y are independent their joint distribution can be computed as $h(x, y) = f(x) \cdot g(x)$. Hence the likelihood function for the data can be derived as follows:

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-X_i/\theta} \cdot \prod_{i=1}^{m} e^{-5Y_i/\theta} \cdot (1 - e^{5/\theta})^{1-Y_i}$$

$$L(\theta) = \frac{1}{\theta^n} e^{-\sum\limits_{i=1}^{n} X_i/\theta} \cdot e^{\sum\limits_{i=1}^{m} -5Y_i/\theta} \cdot \left(1 - e^{-5/\theta}\right)^{\sum\limits_{i=1}^{m}(1-Y_i)}$$

(b) Assuming that we have 10 observations from $f$ given by $2.8, 5.6, 24.7, 6.5, 1.6, 10.6, 1.0, 7.8, 7.2, 13.9$ and the following 15 observations from $g$ : $0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0$ , the MLE for $\theta$ was computed as `5.971734`. For the computation, log-likelihood function was derived and the maximum was computed via `optimize()` function.

The `R` code is presented below:

```
Xs <- c(2.8,5.6,24.7,6.5,1.6,10.6,1.0,7.8,7.2,13.9)
Ys <- c(0,0,0,1,1,1,0,0,1,0,0,0,0,0,0)

joint.lik<-function(theta){
  n<-length(Xs)
  m <- length(Ys)

  logl <- -n*log(theta, base = exp(1)) - sum(Xs)/theta - 5*sum(Ys)/theta
            + (m - sum(Ys))*log((1-exp(-5/theta)), base = exp(1))

  return(logl)
}

optimize(joint.lik, c(10,0), maximum=TRUE)
```