FSRM/MSDS 581 Homework 10
Due on Wednesday 11:59pm EST, 12/9

Note: Questions 2(c,d,e) and Question 3 are NOT due.

1. Let (x_i, Y_i) $(i = 1, \ldots, n)$ be independent following the **Poisson regression** model (without intercept)

$$Y_i \sim Poission(\lambda x_i)$$

Treat $x_1, \ldots, x_n$ as given constants $(x_i > 0)$.

(a) Find the mean, variance and risk under squared error loss of the least square (LS) estimator

$$\hat{\lambda}_1 = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$$

[Hint: Since Y_i are independent, we have $Var(\sum_{i=1}^n x_i Y_i) = \sum_{i=1}^n x_i^2 Var(Y_i)$]

(b) Find the mean, variance and risk under squared error loss of the MLE

$$\hat{\lambda}_2 = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i}$$

(c) Use Cauchy-Schwarz inequality $(a_1^2 + \ldots + a_n^2)(b_1^2 + \ldots + b_n^2) \geq (a_n b_n + \ldots + a_n b_n)^2$ to show that under squared error loss, $R(\hat{\lambda}_2) \leq R(\hat{\lambda}_1)$ for all $\lambda > 0$ and positive $x_1, \ldots, x_n$.

(d) Find the score function $S(\lambda)$

(e) Find the Fisher information $I(\lambda)$.

(f) The likelihood function satisfies the regularity conditions for the MLE to have asymptotic normality. Find the asymptotic distribution of the MLE.

(g) (Computer experiment) Set $\lambda = 1$. Fix $x_i = i$, for $i = 1, \ldots, 10$. Generate

$$Y_i \sim Poission(\lambda x_i), \quad i = 1, \ldots, 10,$$

and obtain the corresponding $\hat{\theta}_1$ and $\hat{\theta}_2$. Do this 1000 times to obtain 1000 pairs of $(\hat{\theta}_1, \hat{\theta}_2)$. Calculate the following items using the 1000 $\hat{\theta}_1$'s and $\hat{\theta}_2$'s and put the results in a table.

1. the average bias (as an estimate of the mean bias)
2. the sample standard deviation (as an estimate of the standard error)
3. the root average squared error loss (as an estimate of the root MSE)
4. the average absolute deviation loss (as an estimate of the L1 risk)
5. the average large deviation loss using $c = 0.1$.

(h) Comment on the table you obtained in (g).

The following R code was used to generate the table on page 21 of Lecture Note 8.

```
set.seed(1)

nn=8

x=matrix(runif(8000),ncol=8)

t1=2*apply(x,1,mean)

t2=apply(x,1,max)

t3=t2*(nn+1)/nn

b1=mean(t1-1);b2=mean(t2-1);b3=mean(t3-1);

s1=sqrt(var(t1));s2=sqrt(var(t3));s3=sqrt(var(t3));

r1=sqrt(mean((t1-1)**2));r2=sqrt(mean((t2-1)**2));r3=sqrt(mean((t3-1)**2));

rabs1=mean(abs(t1-1));rabs2=mean(abs(t2-1));rabs3=mean(abs(t3-1));

rLD1=mean(as.numeric(abs(t1-1)>0.1));

rLD2=mean(as.numeric(abs(t2-1)>0.1));

rLD3=mean(as.numeric(abs(t3-1)>0.1));

print(c(b1,s1,r1,rabs1,rLD1))
```

```
print(c(b2,s2,r2,rabs2,rLD2))

print(c(b3,s3,r3,rabs3,rLD3))
```

You can modify it for question (g). You can generate 1000 Poisson random numbers for each fixed x_i separately, then resemble your data to get 1000 sets of $(Y_1, \ldots, Y_{10})$ corresponding to $(x_1, \ldots, x_{10})$.

Note that the **apply** function is very useful. It is much faster than doing a loop. To use the **apply** function to obtain the LS estimate, you can create a function, for our specific case of $x_i = i$, $i = 1, \ldots, 10$

```
LSP=function(y){
 x=1:10
 theta=sum(x*y)/sum(x*x)
 return(theta)
}
```

Then you can use

```
t1=apply(y,1,LSP)
```

2. Let (x_i, Y_i) $(i = 1, \ldots, n)$ be independent following the **Poisson regression** model (without intercept)

$$Y_i \sim Poisson(\lambda x_i)$$

Treat $x_1, \ldots, x_n$ as given constants $(x_i > 0)$.

(a) Construct an α level likelihood ratio test of testing $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda \neq \lambda_0$ using the approximate χ^2 test.

(b) Construct an α level Wald test of testing $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda \neq \lambda_0$, using the asymptotic normality result you obtained in Homework #10.

(c) Use the test you obtained in (I.a) to construct a $1-\alpha$ confidence interval of λ.

(d) Construct a $1-\alpha$ confidence interval of λ using Wald statistic

(e) Based on the likelihood estimator of λ, find the likelihood estimator of $\eta = 1/\lambda$ and its asymptotic distribution. Construct a $1-\alpha$ confidence interval of η using Wald statistic.

3. Let $X_1, \ldots, X_m$ i.i.d $N(\theta_1, \sigma^2)$ and $Y_1, \ldots, Y_n$ i.i.d $N(\theta_2, \sigma^2)$. All X_is and Y_js are independent. σ is a known constant. θ_1 and θ_2 are the unknown parameters.

(a) Show $Q = \bar{X} - \bar{Y} - (\theta_1 - \theta_2)$ is a pivot statistic and obtain its distribution.

(b) Construct a $1-\alpha$ confidence interval for $\theta_1 - \theta_2$ using the pivot statistic.