MAP 531: Homework

Paul-Antoine GIRARD & Adrien TOULOUSE

Problem 1: Estimating parameters of a Poisson distribution

We recall that the Poisson distribution with parameter $\theta > 0$ has a pdf given by $(p(\theta, k), k \in \mathbb{N})$ w.r.t the counting measure on $\mathbb{N}$:

$$p(\theta, k) = e^{-\theta} \frac{\theta^k}{k!}$$

Question 1

The poisson distribution is a discrete distribution since it has a countable number of possible values ($\mathbb{N}$).

In statistics, we use this distribution to compute the probability of a given number of (rare) events in a time period.

For example a poisson distribution can model:

- The number of patients arriving in an emergency room between 9 and 10am.
- The number of network failures per day.
- In quality control, the number of manufacturing defect.

Question 2

We assume that $\mathbb{X}$ follows a Poisson distribution with parameter $\theta > 0$.

We will use the fact that $e^\theta = \sum_{i=0}^{\infty}(\frac{\theta^i}{i!}), \forall \theta \in \mathbb{R}$

$$\mathbb{E}[\mathbb{X}] = \sum_{i=0}^{\infty}(i * p(\theta, i)) = \sum_{i=0}^{\infty}(i * e^{-\theta}\frac{\theta^i}{i!}) = \theta * e^{-\theta}\sum_{i=1}^{\infty}(\frac{\theta^{i-1}}{(i-1)!}) = \theta * e^{-\theta}\sum_{i=0}^{\infty}(\frac{\theta^i}{i!}) = \theta * e^{-\theta} * e^{\theta} = \theta$$

$$\mathbb{E}[\mathbb{X}^2] = \sum_{i=0}^{\infty}(i^2 * p(\theta, i)) = \sum_{i=0}^{\infty}(i^2 * e^{-\theta}\frac{\theta^i}{i!}) = \theta * e^{-\theta}\sum_{i=1}^{\infty}(i\frac{\theta^{i-1}}{(i-1)!}) = \theta * e^{-\theta}\sum_{i=0}^{\infty}((i+1)\frac{\theta^i}{i!})$$

$$= \theta * e^{-\theta}[\sum_{i=0}^{\infty}(i\frac{\theta^i}{i!}) + \sum_{i=0}^{\infty}(\frac{\theta^i}{i!})] = \theta * e^{-\theta}[\theta * e^{\theta} + e^{\theta}] = \theta(\theta + 1)$$

$$\mathbb{V}(\mathbb{X}) = \mathbb{E}[\mathbb{X}^2] - \mathbb{E}[\mathbb{X}]^2 = \theta(\theta + 1) - \theta^2 = \theta$$

Question 3

We are provided with n independent observations of a Poisson random variable of parameter $\theta \in \Theta = \mathbb{R}_+^*$. Our observations are $X_k \sim Pois(\theta), \forall k \in 1, ..., n$.
The corresponding statistical model is $(\mathbb{N}, \ , \ \{p(. \mid \theta), \ \theta \in \Theta\})$ with $\mathbb{P}(x \mid \theta) = \mathbb{P}_\theta(\mathbb{X} = x)$ We are trying to estimate the parameter θ.

Question 4

The likelihood function is the function on θ that makes our n observations most likely.

Using the independance of the X_k:

$$l(\theta) = \prod_{k=1}^{n} p(\theta, x_k) = \prod_{k=1}^{n} e^{-\theta} \frac{\theta^{x_k}}{x_k!}, with\ x_k \in \mathbb{N}, \forall k \in 1, ..., n$$

$$L(\theta) = log(l(\theta)) = \sum_{k=1}^{n} (-\theta + x_k log(\theta) - log(x_k!)) = -n\theta + log(\theta) \sum_{k=1}^{n} x_k - \sum_{k=1}^{n} log(x_k!)$$

By derivating with respect to θ, we have:

$$L'(\theta) = -n + \frac{\sum_{k=1}^{n} x_k}{\theta}$$

$$L''(\theta) = -\frac{\sum_{k=1}^{n} x_k}{\theta^2} < 0$$

Since, the second derivative of the log-likelihood function is negative, the function is concave and admits a global maximum, given by:

$$L'(\theta) = 0 \Leftrightarrow -n + \frac{\sum_{k=1}^{n} X_k}{\theta} = 0 \Leftrightarrow \hat{\theta}_{MLE} = \overline{X}$$

So, the maximum likelihood estimator is:

$$\hat{\theta}_{MLE} = \overline{X}$$

Question 5

We have that:

$$\mathbb{E}[\overline{X}] = \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}[X_k] = \mathbb{E}[X_1] = \theta$$

$$\mathbb{V}(\overline{X}) = \frac{1}{n^2} \sum_{k=1}^{n} \mathbb{V}(X_k) = \frac{1}{n} \mathbb{V}[X_1] = \frac{\theta}{n}$$

Applying the central limit theorem, we have that $\sqrt{n}(\hat{\theta}_{MLE} - \theta)$ converges towards a Gaussian $\mathcal{N}(0, \theta)$.

Question 6

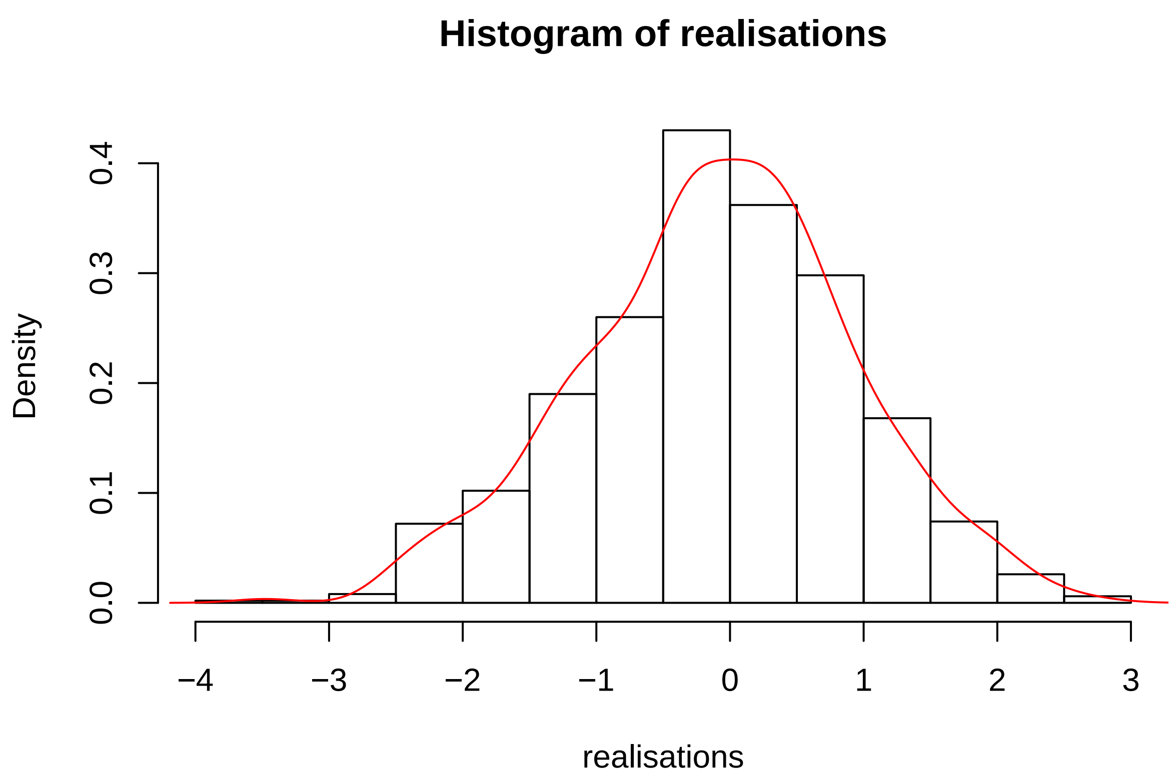
By continuous mapping, $\sqrt{\hat{\theta}_{MLE}}$ converges in probability towards $\sqrt{\theta}$. Then, by Slutsky's theorem, we have that $\sqrt{n}\frac{(\hat{\theta}_{MLE} - \theta)}{\sqrt{\hat{\theta}_{MLE}}}$ converges in law towards a gaussian $\mathcal{N}(0, 1)$.

Let's check this result in R by simulating 1000 times our random variable $\sqrt{n}\frac{(\hat{\theta}_{MLE} - \theta)}{\sqrt{\hat{\theta}_{MLE}}}$ with a sample size of 100:

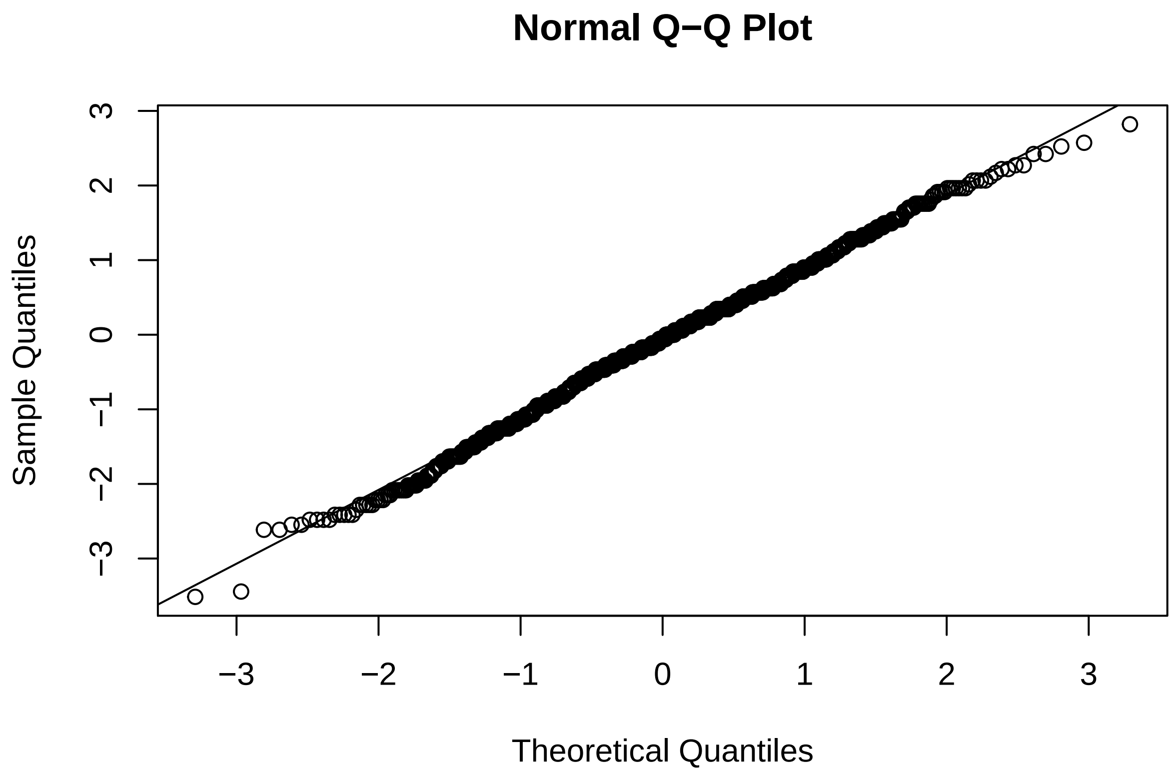
```
estim <- function(x, theta){
  n <- length(x)
  est <- sqrt(n) * (mean(x) - theta) / sqrt(mean(x))
  return(est)}
```

```r
set.seed(43)
Nattempts = 1e3
nsample = 100
theta = 3
samples <- lapply(1:Nattempts, function(i) rpois(nsample, theta))
realisations <- sapply(samples, function(x) estim(x, theta))

hist(realisations, probability = TRUE)
d = density(realisations, kernel='gaussian')
lines(d, col = 'red')
```




```r
qqnorm(realisations)
qqline(realisations)
```



This confirms what we found theoretically: the random variable follows a standard gaussian distribution.

Question 7

Let $Z_n = \sqrt{n}\frac{(\hat{\theta}_{MLE}-\theta)}{\sqrt{\hat{\theta}_{MLE}}}$ be our random variable.

Denote z_{alpha} the α-quantile for the standard Normal distribution.

$$\lim_{n\to+\infty}\mathbb{P}(-z_{1-\alpha/2} \leq Z_n \leq z_{1-\alpha/2}) \geq 1-\alpha \Leftrightarrow \lim_{n\to+\infty}\mathbb{P}(-z_{1-\alpha/2}\sqrt{\frac{\hat{\theta}_{MLE}}{n}} \leq \hat{\theta}_{MLE}-\theta \leq z_{1-\alpha/2}\sqrt{\frac{\hat{\theta}_{MLE}}{n}}) \geq 1-\alpha$$

For $\alpha \in (0,1)$, an asymptotic confidence interval for θ of level α is therefore :

$$[\hat{\theta}_{MLE} - z_{1-\alpha/2}\frac{\sqrt{\hat{\theta}_{MLE}}}{\sqrt{n}};\ \hat{\theta}_{MLE} + z_{1-\alpha/2}\frac{\sqrt{\hat{\theta}_{MLE}}}{\sqrt{n}}]$$

Question 8

We apply the δ-method with $g(x) = 2\sqrt{x}$ We have: $g'(x) = \frac{1}{\sqrt{x}}$
So,

$$\sqrt{n}(g(\hat{\theta}_{MLE}) - g(\theta)) \xrightarrow{d} \mathcal{N}(0,\ g'(\theta)^2 \times \theta) \Leftrightarrow \sqrt{n}(g(\hat{\theta}_{MLE}) - g(\theta)) \xrightarrow{d} \mathcal{N}(0,1)$$

Question 9

Let $W_n = \sqrt{n}(2\sqrt{\hat{\theta}_{MLE}} - 2\sqrt{\theta})$ be our random variable.

We know by the last question that $W_n \xrightarrow{d} \mathcal{N}(0,1)$.

$$\lim_{n\to+\infty} \mathbb{P}(-z_{1-\alpha/2} \leq W_n \leq z_{1-\alpha/2}) \geq 1-\alpha \Leftrightarrow \lim_{n\to+\infty} \mathbb{P}(-\frac{z_{1-\alpha/2}}{2\sqrt{n}} \leq \sqrt{\hat{\theta}_{MLE}} - \sqrt{\theta} \leq \frac{z_{1-\alpha/2}}{2\sqrt{n}}) \geq 1-\alpha$$

$$\Leftrightarrow \mathbb{P}(\sqrt{\hat{\theta}_{MLE}} - \frac{z_{1-\alpha/2}}{2\sqrt{n}} \leq \sqrt{\theta} \leq \sqrt{\hat{\theta}_{MLE}} + \frac{z_{1-\alpha/2}}{2\sqrt{n}}) = 1-\alpha$$

For $\alpha \in (0,1)$, an asymptotic confidence interval for θ of level α is therefore :

$$[\hat{\theta}_{MLE} - z_{1-\alpha/2}\frac{\sqrt{\hat{\theta}_{MLE}}}{\sqrt{n}};\ \hat{\theta}_{MLE} + z_{1-\alpha/2}\frac{\sqrt{\hat{\theta}_{MLE}}}{\sqrt{n}}]$$

For $\alpha \in (0,1)$, an asymptotic confidence interval for θ of level α is therefore :

$$[**]$$

Question 10

Based on the first moment of a poisson distribution, we easily have that:

$$\hat{\theta}_{MME} = \overline{X}$$

We can remark that $\hat{\theta}_{MME} = \hat{\theta}_{MLE}$

Based on the second moment of a poisson distribution, we have:

$$n^{-1}\sum_{k=1}^{n} X_k^2 = \hat{\theta}_2(\hat{\theta}_2 + 1)$$

Let's define the function $h(x) = x(x+1)$
Its inverse on $\mathbb{R}_+^*$ is $h^{-1}(x) = \frac{1}{2}[-1+\sqrt{4x+1}]$ and this gives us:

$$\hat{\theta}_2 = \frac{1}{2}[-1 + \sqrt{(4n^{-1}\sum_{k=1}^{n} X_k^2) + 1}]$$

Question 11

$\mathbb{E}(\hat{\theta}_{MLE}) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i)$ by linearity of the expectation. So,

$$\mathbb{E}(\hat{\theta}_{MLE}) = \frac{1}{n} * n\theta = \theta$$

Therefore, $\hat{\theta}_{MLE}$ is an unbiased estimator of θ, ie. $b_{\theta^*}(\hat{\theta}_{MLE}) = 0$

$\mathbb{V}(\hat{\theta}_{MLE}) = \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{V}(X_i)$ by independance of the X_k.

$$\mathbb{V}(\hat{\theta}_{MLE}) = \frac{1}{n^2} * n\theta = \frac{\theta}{n}$$

The quadratic risk Q is:

$$Q = b_{\theta^*}(\hat{\theta}_{MLE})^2 + \mathbb{V}^*(\hat{\theta}_{MLE}) = 0 + \frac{\theta}{n} = \frac{\theta}{n}$$

Question 12

$\hat{\theta}_{MLE}$ is an unbiased estimator. So the Cramer-Rao bound is given by:

$$\frac{1}{I_n(\theta^*)} = \frac{1}{\mathbb{E}[-L''(\theta^*)]}$$

By derivating the log-likelihood function with respect to θ, we have:

$$L'(\theta^*) = -n + \frac{\sum_{i=1}^{n} x_k}{\theta}$$

$$-L''(\theta^*) = \frac{\sum_{i=1}^{n} x_k}{\theta^2}$$

Therefore,

$$\mathbb{E}[-L''(\theta^*)] = \frac{\sum_{i=1}^{n} \mathbb{E}[X_k]}{\theta^2} = \frac{n}{\theta}$$

Finally,

$$\frac{1}{I_n(\theta^*)} = \frac{\theta}{n} = \mathbb{V}(\hat{\theta}_{MLE})$$

We can conclude that our estimator $\hat{\theta}_{MLE}$ is efficient.

Question 13

$$\hat{\theta}_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X_n})^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \theta + \theta - \overline{X_n})^2 = \frac{1}{n}\sum_{i=1}^{n}[(X_i - \theta)^2 + (\theta - \overline{X_n})^2 + 2(X_i - \theta)(\theta - \overline{X_n})]$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i - \theta)^2 + (\theta - \overline{X_n})^2 + \frac{2}{n}(\theta - \overline{X_n})\sum_{i=1}^{n}(X_i - \theta) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \theta)^2 + (\theta - \overline{X_n})^2 + 2(\theta - \overline{X_n})(\overline{X_n} - \theta)$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i - \theta)^2 - (\theta - \overline{X_n})^2$$

Question 14

$$\mathbb{E}(\theta - \overline{X_n})^2 = \mathbb{E}(\theta^2 - 2\theta\overline{X_n} + \overline{X_n}^2) = \theta^2 - 2\theta\mathbb{E}(\overline{X_n}) + \mathbb{E}(\overline{X_n})^2$$
$$= -\theta^2 + \mathbb{V}(\overline{X_n}) + \mathbb{E}(\overline{X_n})^2 = -\theta^2 + \frac{\theta}{n} + \theta^2 = \frac{\theta}{n}$$

$$\mathbb{E}(\hat{\theta}_2) = \mathbb{E}(\frac{1}{n}\sum_{i=1}^{n}(X_i - \theta)^2 - (\theta - \overline{X_n})^2)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i - \theta)^2 - \mathbb{E}(\theta - \overline{X_n})^2 = \frac{1}{n}\sum_{i=1}^{n}\mathbb{V}(X_i) - \frac{\theta}{n} = \theta(1 - \frac{1}{n})$$

Therefore the bias is:

$$b_{\hat{\theta}_2} = -\frac{\theta}{n}$$

We can get an unbiased estimator $\hat{\theta}_3$ by defining $\hat{\theta}_3 = (1 - \frac{1}{n})^{-1}\hat{\theta}_2$

Question 15

Using the previous questions, we know that:

$$\hat{\theta}_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \theta)^2 - (\theta - \overline{X_n})^2$$

therefore, we have:

$$\sqrt{n}(\hat{\theta}_2 - \theta) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - \theta)^2 - \sqrt{n}(\theta - \overline{X_n})^2 - \sqrt{n}\theta = \sqrt{n}(\overline{Y_n} - \theta) - \sqrt{n}(\theta - \overline{X_n})^2$$

where

$$\forall i \in [\![1, n]\!], Yi = (X_i - \theta)^2$$

$$\bar{Y_n} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

Since :

$$\mathbb{E}(Y_i) = \mathbb{V}(X_i) = \theta$$

and

$$\mathbb{V}(Y_i) = 2\theta^2 + \theta$$

Applying the central limit theorem, we have that $\sqrt{n}(\overline{Y_n} - \theta)$ converges towards a Gaussian $\mathcal{N}(0, 2\theta^2 + \theta)$.

We have:

$$\sqrt{n}(\theta - \overline{X_n})^2 = \sqrt{n}(\theta - \overline{X_n})(\theta - \overline{X_n})$$

Applying the central limit theorem, we have that $\sqrt{n}(\theta - \overline{X_n})$ converges towards a Gaussian $\mathcal{N}(0, \theta)$. On the other hand , applying the law of large numbers : $(\theta - \overline{X_n})$ converges in probability towards 0.

Applying the Slutsky theorem,$\sqrt{n}(\theta - \overline{X_n})^2$ converges in distribution towards the constant 0. Therefore it converges in probability towards 0.

Now, we can apply the Slutsky theorem to $\sqrt{n}(\overline{Y_n} - \theta) - \sqrt{n}(\theta - \overline{X_n})^2$ which gives us finally that $\sqrt{n}(\hat{\theta}_2 - \theta)$ converges in distribution towards a Gaussian $\mathcal{N}(0, 2\theta^2 + \theta)$.

Question 16

Let $s \in \mathbb{R}$. The probability generating function of the Poisson distribution is given by:

$$G_{\mathbb{X}}(s) = \mathbb{E}[exp(s\mathbb{X})] = \sum_{k=0}^{\infty} e^{ks} e^{-\theta}\frac{\theta^k}{k!} = e^{-\theta}\sum_{k=0}^{\infty}\frac{(\theta e^s)^k}{k!} = e^{-\theta}e^{\theta e^s} = e^{\theta(e^s - 1)}$$

In order to compute the first and second moment of the Poisson distribution, we can now use the moment generating function. Let's compute its first and second order derivatives.

$$G'_{\mathbb{X}}(s) = \theta e^s e^{\theta(e^s-1)}$$

$$G''_{\mathbb{X}}(s) = \theta[e^s e^{\theta(e^s-1)} + \theta e^{2s} e^{\theta(e^s-1)}] = \theta e^s[e^{\theta(e^s-1)} + \theta e^s e^{\theta(e^s-1)}]$$

Then, we have:

$$\mathbb{E}[\mathbb{X}] = G'_{\mathbb{X}}(0) = \theta$$

$$\mathbb{E}[\mathbb{X}^2] = G''_{\mathbb{X}}(0) = \theta(1+\theta)$$
$$\mathbb{V}(\mathbb{X}) = \mathbb{E}[\mathbb{X}^2] - \mathbb{E}[\mathbb{X}]^2 = \theta(1+\theta) - \theta^2 = \theta$$

We will now show that: $\mathbb{V}[(\mathbb{X}_i - \theta)^2] = 2\theta^2 + \theta$

$$G^{(3)}_{\mathbb{X}}(s) = (1 + 3\theta e^s + \theta^2 e^{2s})\theta e^{s+\theta(e^s-1)}$$
$$G^{(4)}_{\mathbb{X}}(s) = (1 + \theta^3 e^{3s} + 6\theta^2 e^{2s} + 7\theta e^s)\theta e^{s+\theta(e^s-1)}$$
$$\mathbb{V}[(\mathbb{X}_i - \theta)^2] = \mathbb{E}[(\mathbb{X} - \theta)^4] - \mathbb{E}[(\mathbb{X} - \theta)^2]^2 = ... = 2\theta^2 + \theta$$