Solutions to the exercises - Chapter 6
"Models, Statistical Inference and Learning".

Marco Corneli

Exercise 1. To compute the bias we first need expectation of the estimator $\hat{\lambda}$, so

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = \lambda,$$

since the expectation of a Poisson random variable is λ. Thence, the estimator is **unbiased**. Recalling that

$$MSE(\hat{\lambda}) = \left(bias(\hat{\lambda})\right)^2 + \mathbb{V}(\hat{\lambda}),$$

since our estimator has no bias, its MSE coincides with its variance:

$$\mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{V}(X_i) = \frac{\lambda}{n},$$

where I used that the variance of a Poisson rv is still λ (and recall that the first equality in the above equation holds because of independence!).

Exercise 2. First, recall that each X_i, being uniformly distributed in the interval $[0, \theta]$, has the following pdf

$$f(x) = \frac{1}{\theta}\mathbf{1}_{]0,\theta]}(x)$$

and thus cdf

$$\mathbb{P}\left(X_i \leq x\right) = F(x) = \begin{cases} 0 & \text{if} \quad x \leq 0 \\ \frac{x}{\theta} & \text{if} \quad x \in]0, \theta] \\ 1 & \text{otherwise} \end{cases}$$

Since we need to compute expectation and variance of $\hat{\theta}$ (and we cannot move expectation inside the max function as we did for the sum, in the previous exercise) we need to know the pdf of $\hat{\theta}$. Notice that, given x in $]0, \theta]$

$$
\begin{aligned}
\mathbb{P}\left(\hat{\theta} \leq x\right) &= \mathbb{P}\left(\max\{X_1, \dots, X_n\} \leq x\right) \\
&= \mathbb{P}\left(X_1 \leq x, \dots, X_n \leq x\right) \\
&= \prod_{i=1}^{n} \mathbb{P}\left(X_i \leq x_i\right) \\
&= \left(F(x)\right)^n
\end{aligned}
$$

where the second-last equality comes from independence and the last form the identical distributions of the X_is. From the above equation we know that the estimator $\hat{\theta}$ has cdf

$$
G(x) = \begin{cases} 0 & \text{if} \quad x \leq 0 \\ \frac{x^n}{\theta^n} & \text{if} \quad x \in]0, \theta] \\ 1 & \text{otherwise} \end{cases}
$$

and thus pdf

$$
g(x) = n \frac{x^{n-1}}{\theta^n} \mathbf{1}_{]0,\theta]}(x).
$$

It is now easy to see that

$$
\mathbb{E}(\hat{\theta}) = \int_{-\infty}^{\infty} x g(x) dx = \frac{n}{n+1} \theta
$$

and thus the estimator is **biased** and its bias is $b(\hat{\theta}) = -\frac{1}{1+n}\theta$. By recalling that $\mathbb{V}(\hat{\theta}) = \mathbb{E}(\hat{\theta}^2) - (\mathbb{E}(\hat{\theta}))^2$, it can be seen that

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

and finally $MSE(\hat{\theta}) = \frac{2\theta^2}{(n+1)(n+2)}$.

Exercise 3. Given that $\mathbb{E}(X_i) = \frac{\theta}{2}$ and $\mathbb{V}(X_i) = \frac{\theta^2}{12}$ (show it!) it can easily be seen that the estimator $\hat{\theta}$ is **unbiased**. Moreover

$$
MSE(\hat{\theta}) = \mathbb{V}(\hat{\theta}) = 4 \frac{1}{n^2} \sum_{i=1}^{n} \mathbb{V}(X_i) = \frac{\theta^2}{3n}
$$

The interest of the last two exercises can be understood by looking at Figure 1. As it can be seen the MSE of the estimator in Exercise 2 is always smaller than the one of the estimator in Exercise 3, although $\hat{\theta}$ in Ex.2 is biased! (Notice, however, that the bias vanishes asymptotically). As such, that estimator in Ex.2 should be preferred.

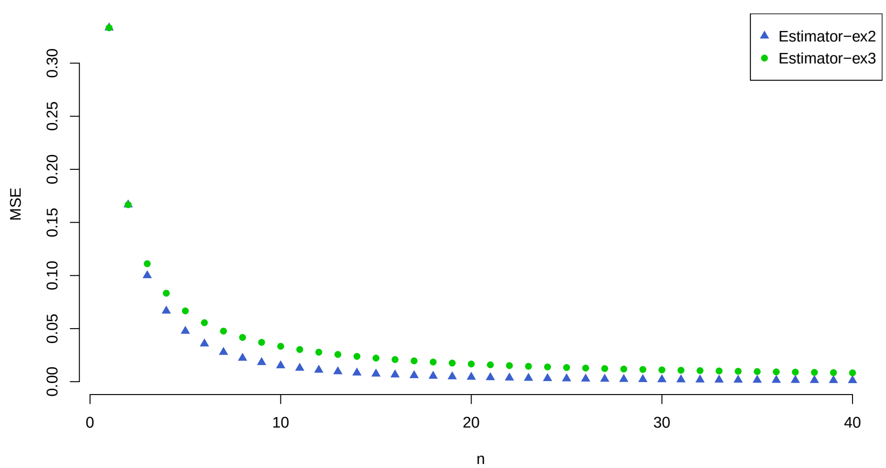

Figure 1: Mean Squared Error comparison.

$\{X_{ij}\}_{i\in 1,\dots,N, j\in 1,\dots,P}$
$X \in \{0,1\}^{N\times P}$ whose entry (i,j) is
$\gamma_k = \mathbb{P}(z_{ik}=1)$
$\mathbf{z} := \{z_{ik}\}_{i\in 1,\dots,N, k\in 1,\dots,K}$

$$X_{ij}(t) \sim \mathcal{P}\left(\int_0^t \lambda_{ij}(u)du\right)$$

where $\mathcal{P}(\lambda)$ denotes Poisson probability mass function with parameter λ. In order to cluster both rows and columns we further assume the the intensity function $\lambda_{ij}(t)$ only depends on the clusters of row i and column j. Thus, if, for instance, $z_{ik}=1$ and $w_{jl}=1$, then

$$\lambda_{ij}(t) = \lambda_{kl}(t).$$

Thus the following matrix