Worksheet 19 (Solutions)

1. Let $X \sim N(\mu, \sigma^2)$, with $\sigma^2 > 0$ a fixed and known constant. (a) Compute the Fisher Information $\mathcal{I}(\mu)$. (b) The MLE for μ is equal to X (generally it's the mean, but in the one-observation case the mean is equal to X). Find the efficency of the MLE.

Solution: (a) We have the following for the first derivative the of the log likelihood:

$$\begin{aligned}\frac{\partial}{\partial\mu}\log(f(\mu;x)) &= \frac{\partial}{\partial\mu}\left[\frac{-1}{2\sigma^2}(x-\mu)^2\right] \\ &= \frac{+2}{2\sigma^2}(x-\mu) \\ &= \frac{1}{\sigma^2}(x-\mu)\end{aligned}$$

And for the second derivative:

$$\begin{aligned}\frac{\partial^2}{\partial^2\mu}\log(f(\mu;x)) &= \frac{\partial}{\partial\mu}\left[\frac{1}{\sigma^2}(x-\mu)\right] \\ &= \frac{-1}{\sigma^2}.\end{aligned}$$

Then, the Fisher information is:

$$\begin{aligned}\mathcal{I}(\mu) &= -\mathbb{E}\left[\frac{\partial^2}{\partial\mu^2}\log f(\mu;x)\right] \\ &= \mathbb{E}\left[\frac{1}{\sigma^2}\right] \\ &= \frac{1}{\sigma^2}.\end{aligned}$$

(b) The variance of the MLE is equal to:

$$Var(\hat{\mu}) = Var(X) = \sigma^2.$$

And therefore the efficency is:

$$e(\hat{\mu}) = \frac{\mathcal{I}(\theta)^{-1}}{Var(\widehat{\theta})} = \frac{\sigma^2}{\sigma^2} = 1.$$

So, the MLE is optimally efficent. It does as well as any unbiased estimator can do in terms of predicting the value of μ from the data.

2. Let $X \sim Poisson(\lambda)$. (a) Compute the Fisher Information $\mathcal{I}(\lambda)$. (b) The MLE for λ is equal to X (generally it's the mean, but in the one-observation case the mean is equal to X). Find the efficency of the MLE.

Solution: (a) We have the following for the first derivative the of the log likelihood:

$$\begin{aligned}\frac{\partial}{\partial\lambda}\log(f(\lambda;x)) &= \frac{\partial}{\partial\lambda}\left[x\log(\lambda) - \lambda + \log(x!)\right] \\ &= \frac{x}{\lambda} - 1.\end{aligned}$$

And for the second derivative:

$$\begin{aligned}\frac{\partial^2}{\partial^2\lambda}\log(f(\lambda;x)) &= \frac{\partial}{\partial\lambda}\left[\frac{x}{\lambda} - 1\right] \\ &= \frac{-x}{\lambda^2}.\end{aligned}$$

Then, the Fisher information is:

$$\begin{aligned}\mathcal{I}(\lambda) &= -\mathbb{E}\left[\frac{\partial^2}{\partial\lambda^2}\log f(\lambda;x)\right] \\ &= \mathbb{E}\left[\frac{x}{\lambda^2}\right] \\ &= \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}.\end{aligned}$$

(b) The variance of the MLE is equal to:

$$Var(\hat{\lambda}) = Var(X) = \lambda.$$

And therefore the efficency is:

$$e(\hat{\lambda}) = \frac{\mathcal{I}(\theta)^{-1}}{Var(\widehat{\theta})} = \frac{\lambda}{\lambda} = 1.$$

So, the MLE is optimally efficent. It does as well as any unbiased estimator can do in terms of predicting the value of λ from the data.

3. Let $X \sim Binomial(n, p)$ with $n > 0$ a fixed and known constant. (a) Compute the Fisher Information $\mathcal{I}(p)$.[1] (b) The MLE for p is equal to X/n. Find the efficency of the MLE.

[1] Try to simplify this as much as possible. You should be able to get something that has a denominator equal to $p(1-p)$.

Solution: (a) We have the following for the first derivative the of the log likelihood:

$$\begin{aligned}\frac{\partial}{\partial p}\log(f(p;x)) &= \frac{\partial}{\partial p}\left[\log(\binom{n}{x}) + x\cdot\log(p) + (n-x)\cdot\log(1-p)\right] \\ &= \frac{x}{p} - \frac{n-x}{1-p}.\end{aligned}$$

And for the second derivative:

$$\begin{aligned}\frac{\partial^2}{\partial^2 p}\log(f(p;x)) &= \frac{\partial}{\partial p}\left[\frac{x}{p} - \frac{n-x}{1-p}\right] \\ &= \frac{-x}{p^2} + \frac{n-x}{(1-p)^2}\end{aligned}$$

Then, the Fisher information is:

$$
\begin{aligned}
\mathcal{I}(p) &= -\mathbb{E}\left[\frac{\partial^2}{\partial p^2}\log f(p;x)\right] \\
&= \mathbb{E}\left[\frac{x}{p^2} - \frac{n-x}{(1-p)^2}\right] \\
&= \frac{np}{p^2} - \frac{n-np}{(1-p)^2} \\
&= \frac{n}{p} - \frac{n(1-p)}{(1-p)^2} \\
&= \frac{n}{p} - \frac{n}{(1-p)} \\
&= n \cdot \left[\frac{1}{p} - \frac{1}{1-p}\right] \\
&= n \cdot \left[\frac{(1-p)+p}{p(1-p)}\right] \\
&= n \cdot \left[\frac{1}{p(1-p)}\right] \\
&= \frac{n}{p(1-p)}
\end{aligned}
$$

(b) The variance of the MLE is equal to:

$$Var(\hat{p}) = Var(X/n) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

And therefore the efficency is:

$$e(\hat{p}) = \frac{\mathcal{I}(\theta)^{-1}}{Var(\widehat{\theta})} = \frac{\frac{p(1-p)}{n}}{\frac{p(1-p)}{n}} = 1.$$

So, the MLE is optimally efficent. It does as well as any unbiased estimator can do in terms of predicting the value of p from the data.