Q Assume $X_1, X_2, \cdots X_n$ are i.i.d. $N(\mu, \sigma^2)$ where σ^2 is known. We know that $\bar{X}$ is the MLE for μ. Derive the Fisher Information for μ, and use it to approximate the standard error and distribution of $\bar{X}$.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \rightarrow \log f(x) = -\frac{(x-\mu)^2}{2\sigma^2} + \text{constant}$$

$$\frac{\partial \log f(x)}{\partial \mu} = \frac{x-\mu}{\sigma^2} \qquad \frac{\partial^2 \log f(x)}{\partial \mu^2} = -\frac{\mu}{\sigma^2}$$

thus $n I(\mu) = E\left[-\frac{\partial^2 l(\mu)}{\partial \mu^2}\right] = E\left[\frac{n}{\sigma^2}\right] = \frac{n}{\sigma^2}$

The asymptotic normality of MLE tells us that $\bar{X}$ is approx normal with mean μ_0 and variance $\frac{1}{nI(\mu)} = \frac{\sigma^2}{n}$

[this result is the same as CLT, which stating that $\bar{x} \sim (\mu, \frac{\sigma^2}{n})$]

Q. Assume $X_1, X_2, \cdots X_n$ are i.i.d. Exponential(λ). Derive the Fisher Information for λ and the approximate distribution of its MLE

We already know that $\hat{\lambda}_{MLE} = \frac{1}{\bar{X}}$, $f(x;\lambda) = \lambda e^{-\lambda x}$ $x>0$

$$\log f(x;\lambda) = \log \lambda - \lambda x$$

$$\Rightarrow \frac{\partial^2 \log f(x;\lambda)}{\partial \lambda^2} = -\frac{1}{\lambda^2} \qquad \text{So } I(\lambda) = E_\lambda\left[-\frac{\partial^2 \log f(x;\lambda)}{\partial \lambda^2}\right] = \frac{1}{\lambda^2}$$

thus, $\hat{\lambda}_{MLE}$ is approximately normal with mean λ_0 and Variance $1/nI(\lambda_0) = \frac{\lambda_0^2}{n}$. In practice, we will use $\frac{\hat{\lambda}^2}{n}$ to approximate the variance.

Q. Assume that $X_1, X_2, \cdots X_n$ are i.i.d. Poisson(λ). Derive the Fisher Information for λ and the approximate distribution of its MLE.

$$f(x;\lambda) = \frac{e^{-\lambda}\lambda^x}{k!} \rightarrow \log f(x;\lambda) = x\log\lambda - \log x! - \lambda$$

$$\frac{\partial \log f(x;\lambda)}{\partial\lambda} = \frac{x}{\lambda} - 1 \qquad \frac{\partial^2 \log f(x;\lambda)}{\partial\lambda^2} = -\frac{x}{\lambda^2}$$

$$\text{so} \quad I(\lambda) = E_\lambda\left[-\frac{\partial^2 \log f(x;\lambda)}{\partial\lambda^2}\right] = \underbrace{E_\lambda\left[\frac{x}{\lambda^2}\right]}_{\text{given } \lambda,\ E[X]=\lambda} = \frac{1}{\lambda}$$

thus $\hat{\lambda}_{MLE}$ is approximately normal with mean $\bar{X}$ and Variance $\frac{\hat{\lambda}}{n}$.

Multi-dimensional

Q. Suppose that $X_1, X_2, \cdots X_n$ are i.i.d. $N(\mu, \sigma^2)$. Derive the joint distribution of the MLE for $\theta = (\underset{\theta_1}{\mu}, \underset{\theta_2}{\sigma^2})$

$$f(x;\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$\frac{\partial^2}{\partial\mu^2}\log f(x;\theta) = -\frac{1}{\sigma^2} \qquad \frac{\partial^2}{\partial(\sigma^2)^2}\log f(x;\theta) = \frac{1}{2\sigma^4} - \frac{(x-\mu)^2}{\sigma^4}$$

$$\frac{\partial^2}{\partial\mu\,\partial\sigma^2}\log f(x;\theta) = -\frac{x-\mu}{\sigma^4} = \frac{\partial^2}{\partial\sigma^2\,\partial\mu}\log f(x;\theta)$$

So.

$$I(\theta) = E\begin{bmatrix} 1/\sigma^2 & \frac{x-\mu}{\sigma^4} \\ \frac{x-\mu}{\sigma^4} & \frac{(x-\mu)^2}{\sigma^4} - \frac{1}{2\sigma^4} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{bmatrix}$$

note $E[(x-\mu)^2] = \sigma^2$
(the definition of σ^2)

Since the MLE for θ is $\begin{bmatrix} \bar{x} \\ \frac{n}{n-1}s^2 \end{bmatrix}$, we can now state that the MLE for θ is approximately bivariate normal with covariance matrix

$$I^{-1}(\theta_0)/n = \begin{bmatrix} \frac{\sigma_0^2}{n} & 0 \\ 0 & \frac{2\sigma_0^4}{n} \end{bmatrix}$$

In practice, we can replace σ_0^2 with $\hat{\sigma}^2$.