Math Statistics
Semiweekly HW 13

Ozaner Hansha

October 30, 2020

Question 1

Problem: Consider a Gamma distribution with shape $k > 0$ and scale $\theta > 0$. Find a maximum likelihood estimator for $\frac{1}{\mu}$, where $\mu = k\theta$ is the mean. Is this a biased estimator?

Solution: Before we continue, note that:

- Each observation X_i in our sample has distribution $\text{Gamma}(k, \theta)$.
- The pdf of X_i given parameters k, θ is denoted by $f_{X_i}(x_i; k, \theta)$.
- The sample X is comprised of $n > 0$ i.i.d. observations of X_i.

Now, let us compute the MLE of the parameter μ. Recall that the MLE of a parameter is the statistic that minimizes its likelihood:

$$\begin{aligned}
\hat{\mu}_{\text{MLE}} &= \arg\max_{\mu} f_X(\mathbf{x}; k, \theta) && \text{(def. of MLE)} \\
&= \arg\max_{\mu} \prod_{i=1}^{n} f_{X_i}(x_i; k, \theta) && \text{(independent observations)} \\
&= \arg\max_{\mu} \prod_{i=1}^{n} f_{X_i}(x_i; k, \mu/k) && \text{(mean of gamma distribution)} \\
&= \arg\max_{\mu} \prod_{i=1}^{n} \frac{x_i^{k-1}}{\Gamma(k)(\mu/k)^k} e^{-\frac{x_i}{\mu/k}} && \text{(pdf of gamma distribution)} \\
&= \arg\max_{\mu} \prod_{i=1}^{n} \frac{k^k x_i^{k-1}}{\Gamma(k)\mu^k} e^{-\frac{kx_i}{\mu}} \\
&= \arg\max_{\mu} \left(\frac{k^k}{\Gamma(k)\mu^k}\right)^n \prod_{i=1}^{n} x_i^{k-1} e^{-\frac{kx_i}{\mu}} \\
&= \arg\max_{\mu} \left(\frac{k^k}{\Gamma(k)\mu^k}\right)^n \left(\prod_{i=1}^{n} x_i\right)^{k-1} \exp\left(-\frac{k}{\mu}\sum_{i=1}^{n} x_i\right) \\
&= \arg\max_{\mu} \frac{1}{\mu^{kn}} \left(\prod_{i=1}^{n} x_i\right)^{k-1} \exp\left(-\frac{k}{\mu}\sum_{i=1}^{n} x_i\right) && \left(\left(\tfrac{k^k}{\Gamma(k)}\right)^n > 0 \text{ and independent of } \mu\right) \\
&= \arg\max_{\mu} \log\left(\frac{1}{\mu^{kn}} \left(\prod_{i=1}^{n} x_i\right)^{k-1} \exp\left(-\frac{k}{\mu}\sum_{i=1}^{n} x_i\right)\right) && \text{(log is monotone increasing)} \\
&= \arg\max_{\mu} -kn\log\mu + (k-1)\sum_{i=1}^{n}\log(x_i) - \frac{k}{\mu}\sum_{i=1}^{n} x_i
\end{aligned}$$

Now we take the derivative of this function w.r.t. μ and set it equal to 0. solving for μ results in a value that produces a local extremum:

$$\begin{aligned}
0 &= \frac{\partial}{\partial \mu}\left(-kn\log\mu + (k-1)\sum_{i=1}^{n}\log(x_i) - \frac{k}{\mu}\sum_{i=1}^{n}x_i\right) && \text{(first order condition)}\\
&= -kn\frac{\partial}{\partial \mu}\log\mu + \frac{\partial}{\partial \mu}(k-1)\sum_{i=1}^{n}\log(x_i) - \frac{\partial}{\partial \mu}\frac{k}{\mu}\sum_{i=1}^{n}x_i\\
&= \frac{-kn}{\mu} + \frac{k}{\mu^2}\sum_{i=1}^{n}x_i\\
&= \frac{-n}{\mu} + \frac{1}{\mu^2}\sum_{i=1}^{n}x_i\\
\frac{n}{\mu} &= +\frac{1}{\mu^2}\sum_{i=1}^{n}x_i\\
n\mu &= \sum_{i=1}^{n}x_i\\
\mu &= \frac{1}{n}\sum_{i=1}^{n}x_i\\
\mu &= \bar{X}
\end{aligned}$$

Now we will show that this extremum is indeed a local maximum via the second derivative test:

$$\begin{aligned}
\frac{\partial^2}{\partial \mu^2}\left(-kn\log\mu + (k-1)\sum_{i=1}^{n}\log(x_i) - \frac{k}{\mu}\sum_{i=1}^{n}x_i\right)\Bigg|_{\mu=\bar{X}} &= \frac{\partial}{\partial \mu}\left(\frac{-kn}{\mu} + \frac{k}{\mu^2}\sum_{i=1}^{n}x_i\right)\Bigg|_{\mu=\bar{X}}\\
&= \left(\frac{kn}{\mu^2} - \frac{2k}{\mu^3}\sum_{i=1}^{n}x_i\right)\Bigg|_{\mu=\bar{X}}\\
&= \frac{kn^3}{(\sum_{i=1}^{n}x_i)^2} - \frac{2kn^3}{(\sum_{i=1}^{n}x_i)^2}\\
&= -\frac{kn^3}{(\sum_{i=1}^{n}x_i)^2}\\
&< 0 && (x_i > 0,\ k > 0,\ n > 0)
\end{aligned}$$

Recall that the second derivative test states that if $f''(\mu) < 0$ at some critical point μ, then $f(\mu)$ is a local maximum. And so $\mu = \bar{X}$ is a local maximum.

And so we finally have that $\hat{\mu}_{\text{MLE}} = \bar{X}$. Now recall that, due to the invariance of MLEs. We have that for any function g and parameter θ:

$$g(\hat{\theta}_{\text{MLE}}) = \widehat{g(\theta)}_{\text{MLE}}$$

And so for $g(\mu) = \frac{1}{\mu}$, we have that:

$$\begin{aligned}
\widehat{\left(\frac{1}{\mu}\right)}_{\text{MLE}} &= \frac{1}{\hat{\mu}_{\text{MLE}}} && \text{(invariance of MLEs)}\\
&= \frac{1}{\bar{X}} && \text{(MLE of } \mu\text{)}
\end{aligned}$$

Now that we have our MLE, we just need to check if it is biased. Note that:

$$(\forall i)\, X_i > 0 \implies \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{X} > 0 \implies \frac{1}{\bar{X}} > 0$$

Since $\frac{1}{x}$ is a convex function over $\mathbb{R}^+$ (i.e. any value $\frac{1}{\bar{X}}$ can take), we can apply Jenson's inequality:

$$\begin{aligned}
E\left[\widehat{\left(\frac{1}{\mu}\right)}_{\text{MLE}}\right] &= E\left[\frac{1}{\bar{X}}\right] \\
&> \frac{1}{E[\bar{X}]} && \text{(Jenson's inequality)} \\
&= \frac{1}{\mu} && \text{(mean of sample mean)}
\end{aligned}$$

You'll notice that we used a strict inequality. That is because equality between the LHS and RHS is only achieved when the RV in question has 0 variance (i.e. is constant). But since $n > 0$ this cannot happen.

This means:

$$\widehat{\left(\frac{1}{\mu}\right)}_{\text{MLE}} \neq \frac{1}{\mu}$$

Thus, our MLE is indeed biased.