# Statistics, Fall 2025 - Practice Session 3
# Likelihood, Consistency, Asymptotics

TA: Stylianos Daskalinas

Email: stylianosdaskalinas@gmail.com

Office hours by appointment. Office D3, 3rd floor , Bld. B

9/10/2025

## Problem 1

Assume an IID sample $\{X_i\}_{i=1}^n$ where $X_i \sim \Gamma(a, c)$, and assume $a = 2$. Then the density becomes

$$f(x; a, c) = \frac{1}{c^2\Gamma(2)} x^{2-1} e^{-\frac{x}{c}} = \frac{1}{c^2} x e^{-\frac{x}{c}}$$

Calculate the following

- The Maximum Likelihood estimator for $c$

- The value of the Fisher Information

- The asymptotic distribution of $\hat{c}$

### Solution

First, formulate the log-Likelihood as

$$\log\Big(L(c; x_1, \ldots, x_n)\Big) = \log\Big(\prod_{i=1}^n c^{-2} x_i e^{-\frac{x_i}{c}}\Big) = -2n\log(c) + \sum_{i=1}^n \log(x_i) - \frac{1}{c}\sum_{i=1}^n x_i$$
$$= -2n\log(c) + n\overline{\log(x)} - \frac{n}{c}\overline{x}$$

The Maximum Likelihood estimate is given by

$$\frac{d}{dc}\log\Big(L(c; x_1, \ldots, x_n)\Big) = 0 \Rightarrow -\frac{2n}{c} + \frac{n\overline{x}}{c^2} = 0 \Rightarrow \hat{c}_{\text{MLE}} = \frac{\overline{x}}{2}$$

Now, for the Fisher information, we need the second derivative of the Likelihood function

$$\begin{aligned} I_n(c) &= -\mathbb{E}\Big[\frac{d^2}{dc^2}\log\Big(L(c;x_1,\dots,x_n)\Big)\Big] = -\mathbb{E}\Big[\frac{2n}{c^2}-\frac{2n\overline{x}}{c^3}\Big] = -2n\frac{1}{c^2}\mathbb{E}\Big[1-\frac{\overline{x}}{c}\Big] \\ &= -\frac{2n}{c^2}\Big[1-\frac{1}{c}\mathbb{E}[\overline{x}]\Big] = -\frac{2n}{c^2}\Big[1-\frac{1}{c}\Big[\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[x_i]\Big]\Big] = -\frac{2n}{c^2}\Big[1-\frac{1}{c}[2c]\Big] = \frac{2n}{c^2} \end{aligned}$$

For the last line, the expectation of the Gamma distribution is used.

As the sample becomes asymptotic ($n \to \infty$), the consistency of the Maximum Likelihood estimator ensures that it is distributed normally:

$$\sqrt{n}\Big(\hat{c}_{\text{MLE}}-c\Big) \sim N\Big(0, \text{Var}(\hat{c}_{\text{MLE}})\Big)$$

Here, $\text{Var}(\hat{c}_{\text{MLE}})$ gives the variance of the Maximum Likelihood estimator, which can be recovered by the reciprocal of the Fisher information, that is

$$\text{Var}(\hat{c}_{\text{MLE}}) = \frac{c^2}{2n}$$

Then, the asymptotic distribution of the Maximum Likelihood estimator is

$$\sqrt{n}\Big(\hat{c}_{\text{MLE}}-c\Big) \sim N\Big(0, \frac{c^2}{2n}\Big) \Rightarrow \;\; \hat{c}_{\text{MLE}} \sim N\Big(c, \frac{c^2}{2}\Big)$$

To ensure the validity of this result, we need to verify the 3$^{\text{rd}}$ derivative regularity property of the likelihood function, which is required for the convergence-in-distribution of the Maximum Likelihood estimate. Essentially, we need the third derivative of the logarithm of the density of the distribution to be bounded by some function $M(x)$ with a finite first moment. Notice that

$$\frac{d^3}{dc^3}\log\Big(f(x;c)\Big) = 2c^{-4}(3x-2c)$$

It is easy to come up with a bounding function for the absolute value of the third derivative, e.g.

$$M(x) = 2c^{-4}(3x+2c)$$

which has a finite first moment (in that case $16/c^3$). Therefore, the asymptotic result holds.

## Problem 2

Let $\hat{a}_1$ and $\hat{a}_2$ be two unbiased estimators of a parameter $a$, where $\text{Var}(\hat{a}_1) = 1$, $\text{Var}(\hat{a}_2) = 2$ and $\text{Cov}(\hat{a}_1, \hat{a}_2) = \frac{1}{2}$. Which $c_1, c_2 \in \mathbb{R}$ produce unbiased estimators of the form

$$c_1\hat{a}_1 + c_2\hat{a}_2$$

that have minimum variance among the estimators of this type?

**Solution**

First, since we need the estimators to be unbiased, we need

$$\mathbb{E}\Big[c_1\hat{a}_1 + c_2\hat{a}_2\Big] = a \Rightarrow c_1\mathbb{E}[\hat{a}_1] + c_2\mathbb{E}[\hat{a}_2] = a \Rightarrow (c_1 + c_2)a = a \Rightarrow c_2 = 1 - c_1$$

Then, to move towards the minimum variance estimators, we have

$$\begin{aligned} \text{Var}\Big[c_1\hat{a}_1 + c_2\hat{a}_2\Big] &= c_1^2\text{Var}(\hat{a}_1) + c_2^2\text{Var}(\hat{a}_2) + 2c_1c_2\text{Cov}(\hat{a}_1, \hat{a}_2) \\ &= c_1^2 + 2(1 - c_1)^2 + \frac{1}{2}2c_1(1 - c_1) = 2c_1^2 - 3c_1 + 2 \end{aligned}$$

Then, variance is a function of the coefficient $c_1$, and all we need to do is to find a value that minimizes the variance function. Thus, naming the function $h(c_1)$, we're looking for

$$c_1^* = \arg\min_{c_1\in\mathbb{R}} \Big\{h(c_1)\Big\} = \arg\min_{c_1\in\mathbb{R}} \Big\{2c_1^2 - 3c_1 + 2\Big\}$$

Then, the first order condition gives the extremum

$$\frac{d}{dc_1}h(c_1) = 0 \Rightarrow 4c_1 - 3 = 0 \Rightarrow c_1^* = \frac{3}{4}$$

while the second order condition

$$\frac{d^2}{dc_1^2}h(c_1) = 4 > 0$$

indicates that indeed the extremum is a minimum, since the function is convex.

Then, the coefficients $c_1$, $c_2$ that produce the minimum-variance, unbiased estimators of the form

$$c_1\hat{a}_1 + c_2\hat{a}_2$$

are

$$c_1^* = \frac{3}{4} \quad c_2^* = 1 - c_1^* = \frac{1}{4}$$

## Problem 3

Is the following estimator for variance of a random sample biased? Is it consistent?

$$\hat{s}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

Justify the answer mathematically.

**Solution**

Consider the following manipulations:

$$\mathbb{E}\hat{s}^2 = \frac{1}{n}\sum_{i=1}^{n}[x_i - \overline{x}]^2 = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[x_i - \frac{1}{n}\sum_{k=1}^{n}x_k\Big]^2 = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[\frac{1}{n}\sum_{k=1}^{n}x_i - \frac{1}{n}\sum_{k=1}^{n}x_k\Big]^2$$

$$= \frac{1}{n^3}\sum_{i=1}^{n}\mathbb{E}\Big[\sum_{k=1}^{n}[x_i - x_k]\Big]^2 = \frac{1}{n^3}\sum_{i=1}^{n}\mathbb{E}\Big[\sum_{k=1}^{n}\sum_{j=1}^{n}[x_i - x_k][x_i - x_j]\Big]$$

$$= \frac{1}{n^3}\sum_{i=1}^{n}\mathbb{E}\Big[\sum_{k=1}^{n}\sum_{j=1}^{n}[x_i^2 - x_k x_i - x_j x_i + x_j x_k]\Big] = \frac{1}{n^3}\sum_{i=1}^{n}\mathbb{E}\Big[[n^2 x_i^2 - n x_i\sum_{k=1}^{n}x_k - n x_i\sum_{j=1}^{n}x_j + \sum_{k=1}^{n}\sum_{j=1}^{n}x_j x_k\Big]$$

$$= \frac{1}{n^3}\sum_{i=1}^{n}\Big[n^2\mathbb{E}[x_i^2] - n[\mathbb{E}[x_i^2] + (n-1)(\mathbb{E}[x_i])^2] - n[\mathbb{E}[x_i^2] + (n-1)(\mathbb{E}[x_i])^2] + n\mathbb{E}[x_i^2] + (n^2-n)(\mathbb{E}[x_i])^2\Big]$$

$$= \frac{1}{n^3}\sum_{i=1}^{n}\Big[[n^2 - n]\mathbb{E}[x_1^2] - [n^2 - n](\mathbb{E}[x_1])^2\Big] = \frac{1}{n^3}nn(n-1)\mathrm{Var}(x_1) = \frac{n-1}{n}\mathrm{Var}(x_1)$$

Thus, the initial estimator for variance is biased and, instead of dividing by $n$, we'd need to divide by $n-1$ to remove bias. This is the famous Bessel correction for the estimation of variance.

However, the estimator is consistent, since the limit of

$$\frac{n-1}{n}$$

approaches 1 as $n \to \infty$, therefore we can trust this estimator with large samples. In the small sample case, we need to apply the Bessel correction.