Parametric Statistics

Sofia Triantafillou

sof.triantafillou@gmail.com

University of Crete
Department of Mathematics and Applied Mathematics

November 20, 2023

Lecture Summary

7.7 Sufficient Statistics

7.8 Jointly Sufficient Statistics

8.1 Sampling Distributions

Recap: Estimators

- Let $X_1, \ldots, X_n$ be a random sample from $Expo(\lambda)$. (e.g., the lifetimes of electronic components).
- Find $\hat{\lambda}_{MLE}$.
- Assuming a $Gamma(1, 1)$ prior for λ, find the posterior for λ.
- Is the MLE unbiased?
- Is the Bayes Estimator for the squared error loss unbiased?
- Can we improve the estimation?

Sufficient Statistics

- Both estimators use the statistic $T = \sum_i X_i$.
- Can we do better if we use something else?

Sufficient Statistic

- Let $X_1, \ldots, X_n$ be a random sample from a distribution indexed by a parameter θ, and T be a statistic.
- Suppose that, for each t, the conditional distribution of $X_1, \ldots, X_n$ given $T = t$ and θ is the same for all θ.
- Then we say that T is a **sufficient statistic** for the parameter θ.

Sufficient Statistics: Factorization Theorem

- Let $X_1, \dots, X_n$ form a random sample from either a continuous distribution or a discrete distribution for which the p.d.f. or the p.f. is $f(x \mid \theta)$, where the value of θ is unknown and belongs to a given parameter space Ω.
- A statistic $T = r(X_1, \dots, X_n)$ is a sufficient statistic for θ if and only if the joint p.d.f. or the joint p.f. $f_n(x \mid \theta)$ of $X_1, \dots, X_n$ can be factored as follows for all values of $\boldsymbol{x} = (x_1, \dots, x_n) \in R^n$ and all values of $\theta \in \Omega$:

$$f_n(\boldsymbol{x} \mid \theta) = u(\boldsymbol{x}) v[r(\boldsymbol{x}), \theta].$$

u, v are nonnegative, the function u may depend on x but does not depend on θ, and the function v will depend on θ but depends on the observed value $\boldsymbol{x}$ only through the value of the statistic $r(x)$.

Jointly Sufficient Statistics.

Suppose that for each θ and each possible value $(t_1, \ldots, t_k)$ of $(T_1, \ldots, T_k)$, the conditional joint distribution of $(X_1, \ldots, X_n)$ given $(T_1, \ldots, T_k) = (t_1, \ldots, t_k)$ does not depend on θ. Then $T_1, \ldots, T_k$ are called **jointly sufficient statistics** for θ.

Factorization Criterion for Jointly Sufficient Statistics.

Let $r_1, \ldots, r_k$ be functions of n real variables. The statistics $T_i = r_i(\boldsymbol{X}), i = 1, \ldots, k$, are jointly sufficient statistics for θ if and only if the joint p.d.f. or the joint p.f. $f_n(x \mid \theta)$ can be factored as follows for all values of $x \in \boldsymbol{R}^n$ and all values of $\theta \in \Omega$:

$$f_n(\boldsymbol{x} \mid \theta) = u(\boldsymbol{x}) v \left[r_1(\boldsymbol{x}), \ldots, r_k(\boldsymbol{x}), \theta \right].$$

u, v are nonnegative, the function u may depend on x but does not depend on θ, and the function v will depend on θ but depends on x only through the k functions $r_1(\boldsymbol{x}), \ldots, r_k(\boldsymbol{x})$.

Order Statistics

Suppose that $X_1, \ldots, X_n$ form a random sample from some distribution. Let Y_1 denote the smallest value in the random sample, let Y_2 denote the next smallest value, let Y_3 denote the third smallest value, and so on. In this way, Y_n denotes the largest value in the sample, and Y_{n-1} denotes the next largest value. The random variables $Y_1, \ldots, Y_n$ are called the order statistics of the sample.

Order Statistics Are Sufficient in Random Samples.

Let $X_1, \ldots, X_n$ form a random sample from a distribution for which the p.d.f. or the p.f. is $f(x \mid \theta)$. Then the order statistics $Y_1, \ldots, Y_n$ are jointly sufficient for θ.

Minimal Sufficient Statistic

- A statistic T is a minimal sufficient statistic if T is sufficient and is a function of every other sufficient statistic.
- Let $T = r(\boldsymbol{X})$ be a sufficient statistic for θ.
- The M.L.E. $\hat{\theta}$ of θ depends on the observations $X_1, \ldots, X_n$ only through the statistic T. Furthermore, if $\hat{\theta}$ is itself sufficient, then it is minimal sufficient.
- Every Bayes estimator $\hat{\theta}$ of θ depends on the observations $X_1, \ldots, X_n$ only through the statistic T. Furthermore, if $\hat{\theta}$ is itself sufficient, then it is minimal sufficient.

Sampling Distribution

- $\hat{\lambda}_{MLE}$ is a function of $X_1, \ldots, X_n$.
- It has a sampling distribution that depends on the value of λ.
- Suppose that the random variables $\boldsymbol{X} = (X_1, \ldots, X_n)$ form a random sample from a distribution involving a parameter θ whose value is unknown.
- Let T be a function of $\boldsymbol{X}$ and possibly θ. That is, $T = r(X_1, \ldots, X_n, \theta)$.
- The distribution of T (given θ) is called the **sampling distribution** of T.
- We will use the notation $E_\theta(T)$ to denote the mean of T calculated from its sampling distribution.

Sampling Distribution

- Assume that you have a urn with three numbers: 1,2,3.
- You draw two numbers (with replacement).
- Statistic 1 (S_1): Mean of the two samples.
- Statistic 2 (S_2): Maximum of two samples.
- Compute the sampling distributions of S_1, S_2.

x_1, x_2	$s_1(x_1, x_2)$	$s_2(x_1, x_2)$

Sampling Distribution

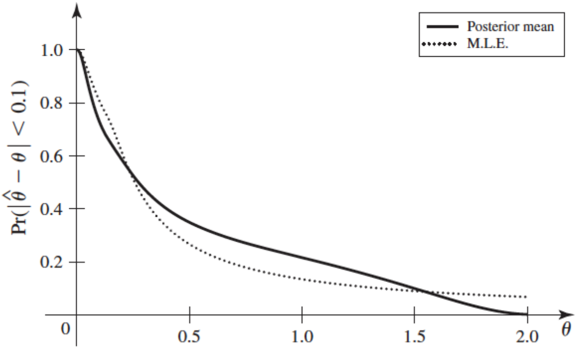
x_1, x_2	$s_1(x_1, x_2)$	$s_2(x_1, x_2)$
1,1	1	1
1,2	1.5	2
1,3	2	3
2,1	1.5	2
2,2	2	2
2,3	2.5	3
3,1	2	3
3,2	2.5	3
3,3	3	3

Why do we care about the sampling distribution.

- Assume that you want to estimate how probable it is that you will make a mistake of more than 0.1..
- You want to estimate $P(|\hat{\lambda}_{MLE} - \lambda|\lambda)$
- For every possible λ, you can compute $P(|\hat{\lambda}_{MLE} - \lambda|\lambda)$ based on the sampling distribution of $\sum_{i=1}^{n} X_i$.
- Let's say $n = 3$

Why do we care about the sampling distribution.

- Assume that you want to estimate how probable it is that you will make a mistake of more than 0.1..
- You want to estimate $P(|\hat{\lambda}_{MLE} - \lambda|\lambda)$
- For every possible λ, you can compute $P(|\hat{\lambda}_{MLE} - \lambda|\lambda)$ based on the sampling distribution of $\sum_{i=1}^{n} X_i$.
- Let's say $n = 3$



Recap

- A statistic is sufficient if no other statistic that can be calculated from the same sample provides any additional information as to the value of the parameter.
- Factorization theorem helps us find sufficient statistics.
- A sufficient statistic is minimal sufficient if it can be represented as a function of any other sufficient statistic.
- If MLE estimator is sufficient, it is minimal sufficient.
- If a Bayes estimator is sufficient, it is minimal sufficient.
- Estimators have their own distributions, known as the sampling distribution.
- We will talk more about specific sampling distributions in the next lectures, particularly for estimators of the Normal Distribution.

MLE for μ, σ^2

Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$, both unknown.
Find the MLE for μ, σ^2.

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