Statistical Inference Note

May 14, 2023

Notation

1. χ - sample space
2. X - random variable
3. $F_X(x)$ - cdf of X
4. $f_X(x)$ - pdf of X
5. $X = (X_1, \dots, X_n)$ - X is a random sample of size n

Chapter 5: Properties of a Random Sample

Definition 1 (Random Sample) *: The random variables $X_1, \dots, X_n$ are called a random sample of size n from the population $f(x)$ if $X_1, \dots, X_n$ are mutually independent variables and the marginal pdf or pmf of each X_i is the same $f(x)$. $\{X_i\}$ are called iid rv with pdf or pmf $f(x)$.*

Definition 2 (Statistics) *: Let $X_1, \dots, X_n$ be **a** random sample of size n from a population and let $T(x_1, \dots, x_n)$ be a real-valued or vector-valued function whose domain includes the sample space of $(X_1, \dots, X_n)$. Then the random variable or random vector $Y = T(X_1, \dots, X_n)$ is called a **statistics**. The probability distribution of a statistic Y is called the sampling distribution of Y.*

Definition 3 (Sample mean and variance)

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \textit{(Sample mean)}$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad \textit{(Sample variance)}$$

$\bar{x}, s^2$ denote to observed values of $\overline{X}, S^2$

Result 1 *Let $X_1, \dots, X_n$ be a random sample from a population with mean μ and variance $\sigma^2 < \infty$. Then*

1. $\mathrm{E}\bar{X} = \mu$
2. $\mathrm{Var}\bar{X} = \frac{\sigma^2}{n}$
3. $\mathrm{E}S^2 = \sigma^2$

Remark: The statistics $\bar{X}$ is unbiased estimator of μ. S^2 is unbiased estimator of σ^2 due to the $n-1$ denominator. If we use n as denominator, $\mathrm{E}S^2$ would be $\frac{n-1}{n}\sigma^2$.

Chapter 7: Point Estimation

Motivation: we want to find a good estimator for θ or $\tau\theta$ using samples from a pdf $p(x|\theta)$ since θ yields knowledge of the entire population.

Definition 4 (Point Estimator) *A point estimator is any function $W(X_1 \dots X_n)$ of a sample; that is any statistic is a point estimator.*

Remark: When a sample is taken, estimator is a function of the rv $X_1, \dots, X_n$ while an estimate is a function of realized values $X_1, \dots, X_n$.

Method of Finding Estimator

Result 2 (Method of Moments) *Let $X_1, \dots, X_n$ be a sample from population $f(x|\theta_1, \dots, \theta_n)$. The method of moments estimators are found by equating the first k sample moments ($m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$) to the corresponding k population moments ($\mu'_k = EX^k$)*

Example: Suppose $X_1, \dots, X_n$ are iid from $f(x|\theta, \sigma^2)$, we have sample moment $m_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$, $m_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ and population moment $\mu'_1 = \mathbf{E}X = \theta$, $\mu'_2 = \mathbf{E}X^2 = \theta^2 + \sigma^2$.

Then we have

$$\begin{aligned} \theta &= \bar{X} & &\rightarrow \theta = \frac{1}{n}\sum_{i=1}^{n} X_i \\ \theta^2 + \sigma^2 &= \frac{1}{n}\sum_{i=1}^{n} X_i^2 & &\rightarrow \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \theta^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 \end{aligned}$$

0.1 Maximum Likelihood Estimator

Definition 5 *(Likelihood function) If $X_1, \dots, X_n$ is an iid sample from a population $f(x|\theta_1, \dots, \theta_n)$, the likelihood function is defined by*

$$L(\theta|x) = L(\theta_1, \dots, \theta_k | x_1, \dots, x_n) = \prod_{i=1}^{n} f(x_i|\theta_1, \dots, \theta_k)$$

Definition 6 *For each sample point x, let $\hat{\theta}(x)$ be a parameter value at which $L(\theta|x)$ attains its maximum as a function of θ with x held fixed. A maximum likelihood estimator(MLE) of the parameter θ based on a sample X is $\hat{\theta}(X)$*

If the likelihood function is differentiable wrt θ_i, the candidate extrema are $\frac{\partial}{\partial\theta_i}L(\theta|x) = 0$ (Extrema can occur on boundary, we need to check those as well).

Remark: The drawbacks are:

1. The problem of actually finding the global maximum and verifying it.
2. Numerical sensitivity; how sensitive is the estimate to the change in data.