16/7/24

NOTES ON STATISTICAL INFERENCE

Abstract: NOTES ARE BASED ON BOOK INTRODUCTION TO STATISTICAL INFERENCE BY CASELLA BERGER

MOHIT SRIVASTAVA

TABLE OF CONTENTS

INTRODUCTION TO PROBABILITY THEORY AND RANDOM VARIABLE

Definition 1.1- Sample Space (S)

The Set of all possible outcome of a particular experiment is known as Sample Space.

For example - tossing a coin one time has two outcomes Head and Tails, so its sample space will be

$$S = \{H, T\}$$

Definition 1.2- Event (E)

A subset of Sample Spaces which represent collection of possible outcomes of an Experiment.

for Example, In tossing a coin,Heads and Tails are two outcomes, so event will be {H}, {T}, {H,T}.

Definition 1.3= Sigma Algebra

A sigma Algebra A is collection of Subsets of S satisfying these 3 properties

1. φ, S $\in$ A
2. if X $\in$ A, then $X^c \in$ A
3. If $X_1, X_2, X_3, ... \in A$ then $X_1 \cup X_2 \cup X_3 \cup .... \in A$

for example let S = $(-\infty, \infty)$ then smallest sigma algebra known as Borel sigma algebra, B contain all sets of the form

$$[a, b],\ (a, b),\ [a, b), (a, b]$$

for all real numbers a and b.

Definition 1.4= Probability

A Probability Function is a function from associated sigma algebra B of Sample Space S to real number having these properties

1. $$P(A) \geq 0 \forall A \in \text{B}$$
2. $$P(S) = 1$$
3. If $A_1, A_2, A_3, ... \in$ B are pairwise disjoint then $P(\cup A_i) = \Sigma A_i$

These 3 properties are also known as Axioms Of Probability. Any function which satisfies these 3 axioms known as Probability Function.

Theorem 1.1

Let S = $\{s_1, s_2, s_3, ..., s_n\}$ be a finite set. Let B be sigma algebra of subsets of S. Let $p_1, ..., p_n$ be non negative number that sum to 1. For any A $\in$ B, we define P(A) by

$$P(A) = \Sigma_{\{i:s_i \in A\}} P_i$$

Then P is a probability function on B. This remains true even for countable set.

Proof :

Let S be a finite set. For any A ∈ B, $P(A) = \Sigma_{\{i:s_i \in A\}} P_i \geq 0$ as $p_i \geq 0$ for all i. Thus, Axiom 1 is true.

Now for A = S,$P(S) = \Sigma_{\{i:s_i \in S\}} P_i = 1$. Let $A_1, A_2, ... \in$ B denote pairwise disjoint event, Then

$$P\left(\cup_{i=1}^{k} A_i\right) = \Sigma_{\{j:s_j \in \cup_{i=1}^{k} A_i\}} p_j = \Sigma_{i=1}^{k} \Sigma_{\{j:s_j \in A_i\}} p_j = \Sigma_{i=1}^{k} P(A_i)$$

Definition 1.5

A Collection of events $A_1, A_2, ..., A_n$ are said to be mutually independent if for any subcollection $A_{i_1}, A_{i_2}, ..., A_{i_k}$, we have

$$P\left(\cap_{j=1}^{k} A_{i_j}\right) = \prod_{i=1}^{k} \left(A_{i_j}\right)$$

Definition 1.6

A Random Variable is a function from Sample Space S into the real numbers.

Example-

1. Suppose we are tossing a coin n times, Let X = number of heads .
2. Tossing 2 dice, Let Y = sum of numbers.

Suppose we have a sample space $S = \{s_1, s_2, ..., s_n\}$ with Probability function P and we define a random variable X with $\text{range}(X) = \{x_1, x_2, ..., x_m\}$. We will observe $X = x_i$ iff outcome of an random experiment is an $s_j \in S$ such that $X(s_j) = x_j$. Thus,

$$P_x(X = x_i) = P(\{s_i \in S : X(s_i) = x_i\})$$

Example

Let S be 2^{50} strings of 50 Os and 1s, X = number of 1s and Ϗ $= \{0, 1, 2, ..., 50\}$. Suppose that each of strings is equally likely. Probability that X = 27 equals

$$P_X(X = 27) = \frac{\text{\#Strings with 27 1s}}{\text{\# Strings}} = \frac{\binom{50}{27}}{2^{50}}$$

In general for any $i \in$ Ϗ,

$$P_X(X = i) = \frac{\binom{50}{i}}{2^{50}}$$

For any set A ⊂ X,

$$P_X(X \in A) = P(\{s \in S : X(s) \in A\})$$

Definition 1.7

The cumulative Distributive Function of Random Variable X, $F_X(x)$ is defined by

$$F_X(x) = P_X\Big(X \leq x\Big) \forall x$$

F_X can be discontinuous but will always be right continuous i.e function is continuous when a point is approached from right.

Theorem 1.2

F(x) is a cdf if and only if following 3 conditions hold:
1. $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.
2. $F(x)$ is non decreasing function of x.
3. $F(x)$ is right continuous i.e for every number x_0, $\lim\limits_{x\to x_0} F(x) = F(x_0)$.

Proof:

- >

Let F(x) be cdf. Then using definition of Probability function,

$$\lim_{x\to-\infty} F_x = \lim_{x\to-\infty} P\Big(X \leq x\Big) = 0$$

and

$$\lim_{x\to\infty} F_x = \lim_{x\to\infty} P\Big(X \leq x\Big) = 1.$$

as P(S) = 1 where S is sample space hence for any subset A ∈ S $P(A) \leq 1$. Property 1 is satisfied.

$F_X(x) = P_X\Big(X \leq x\Big)$ is a non decreasing function as for $x_1 \leq x_2$

$$F_X(x_1) = P_X\Big(X \leq x_1\Big) \leq P_X\Big(X \leq x_2\Big) = F_X(x_2)$$

using property

$$\text{if A} \subset B \text{ then } P(A) \leq P(B)$$

For proving right continuity of $F_X(x)$ we can use third axiom of Probability i.e if $A_1 \subset A_2 \subset .... \subset A_N$ is countable sequence of sets then,

$$P\Big(\lim_{n\to\infty} A_n\Big) = \lim_{n\to\infty} A_n$$

Note: Probability is also a measure functions. It will follow all general properties of measure function. $A_n = \Big\{X \leq x_n + x\Big\}$ with $x_1 < x_2 < x_3 < ....$ a sequence of monotone decreasing sequence with

$$\lim_{n\to\infty} x_n = 0$$

hence

$$P(\lim_{n\to\infty} A_n) = P\Big(X \leq x\Big) = F_X(x) = \lim_{n\to\infty} P\Big(X \leq x + x_n\Big) = \lim_{n\to\infty} F_X(x + x_n)$$

Definition 1.8

A Random Variable X is continuous if $F_X(x)$ is continuous function of x. If $F_X(x)$ is step function of x then Random Variable is said to be discrete.

Definition 1.9

The Random Variables X and Y are identically distributed if for every set

$$A \in \mathrm{B}^1, \text{we have } P(X \in A) = P(Y \in A)$$

Example: Consider the experiment of tossing a fair coin 3 times. Define the Random Variable X and Y by X = no of heads observed and Y = no of tails observed

Therefore, distribution of X and Y are exactly the same. That is for each k = 0,1,2,3 ,...., we have $P(X = k) = P(Y = k)$, so X and Y are identically distributed.

Theorem 1.3

The Following two statements are equivalent.
1. The Random Variable X and Y are identically distributed.
2. $$F_X(x) = F_Y(x) \text{ for every x.}$$

Proof: i →ii

Given X and Y are identically distributed, For any A $\in \mathrm{B}^1, P(X \in A) = P(Y \in A)$.

For every x, $(-\infty, x] \in \mathrm{B}^1$,

$$F_X(x) = P(X \in (-\infty, x]) = P(Y \in (i - \infty, x]) = F_Y(x)$$

ii → i

If X and Y are probabilities agree on all sets then will also agree on all intervals. we have to prove the opposite that if X and Y probabilities agree on all interval, they agree on all sets. We will prove i → ii only for all intervals then extend it to all sets using properties of sigma algebras.
$(-\infty, a), (a, \infty), (a, b), [a, b), (-\infty, a], [a, \infty), (a, b])$

1. For $(-\infty, a)$, already proved from statement ii.

2. For $(a, \infty), P(X \in [a, \infty)) = 1 - P(X \in (-\infty, a]))$therefore it is also true for $[a, \infty)$.

1. m,

$$P(X \in [a, \infty)) = P(X \in (a, \infty)) + P(X = a), \therefore \text{ for } x \in (a, \infty) \text{ they are identically distributed, similar reasons for } (-\infty, a]$$

1.