Mathematics 3 (SM 211): Probability and Statistics

Amit Chattopadhyay

IIIT-Bangalore

Ch. 1: The Concept of Probability

Syllabus: Outline

Probability:

1. The Concept of Probability
2. Compound or Joint Experiment
3. Probability Distributions-I
4. Mathematical Expectation-I
5. Probability Distributions-II
6. Mathematical Expectation-II
7. Some Important Continuous Univariate Distributions
8. Convergence of a Sequence of Random Variables and Limit Theorems

Statistics:

1. Random Samples
2. Sampling Distributions
3. Estimation of Parameters
4. Testing of Hypothesis

Reference Books

1. Mathematical Probability by *A. Banerjee, S.K. De and S. Sen*
2. Mathematical Statistics by *S.K. De and S. Sen*
3. Groundwork of Mathematical Probability and Statistics by *Amritava Gupta*
4. Introduction to Probability and Statistics for Engineers and Scientists by *S.M. Ross*
5. Introduction to Probability Models, by *S.M. Ross*
6. Probability and Statistics, (Schaum's Outlines) by *Murray R Spiegel, John J Schiller and R Alu Srinivasan*

The Concept of Probability

Objective

- Classical definition
- Frequency definition
- Axiomatic definition

Introduction

Probability $\Leftrightarrow$ Synonymous with the word 'chance' (originated from 'random' behavior of certain incident)

Probability theory $\Leftrightarrow$ Mathematical modelling of 'randomness'

Probability of What?

Probability of an "Event" related to a "Random Experiment"

Examples: Informal usage

- the probability that it will rain tomorrow is 70%
- the probability of getting a head in tossing a coin is 40%

(degree of belief on happening of some events)

Experiment

Definition: An act that has some outcome

A. Deterministic Experiment

B. Non-deterministic/ Random Experiment

Examples

1. Measuring boiling point of water (we know the outcome beforehand)
2. Throwing a die
3. Tossing a coin
4. Drawing a card from a pack of 52 cards at random
5. Choosing a point from an interval (1,2) at random

An experiment E satisfying:

(i) all possible outcomes of E are known in advance

(ii) it is impossible to predict which outcome will occur at a particular performance of E

(iii) E can be repeated (at least conceptually) under identical conditions infinite number of times.

Examples

Terminologies

1. **Trial:** Any particular performance of an experiment

2. **Event Space/ Sample Space (S):** All possible outcomes of a random experiment E

3. **Event:** Informally, any subset of the sample space S

S : Certain event; $\emptyset$: Impossible event

Examples:

1. E : Throwing a die; $S = \{1, 2, 3, 4, 5, 6\}$ Event: Getting an even face.
$\{2, 4, 6\}$

2. E : Throwing a die 3 times;
$S = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$ $= \{(1,1,1), (1,2,1), \ldots\}$
Event: Sum of the outcomes is an even number.

$\{(x, y, z) : x+y+z = \text{even number}, (x, y, z) \in S\}$

Terminologies

4. **Simple Event:** If an event A contains exactly one element of S
5. **Composite Event:** If an event A contains more than one element of S

Examples:

E : throwing a coin;

Simple events: $\{H\}$ and $\{T\}$

Composite event: $\{H, T\}$

4. **Mutually Exclusive Events:** Two events A, B connected to a random experiment E are mutually exclusive if

$$AB = \emptyset.$$

(A and B can never happen simultaneously in any performance of E)

Examples:

E : throwing a die;
$A =$ even face, $B =$ odd face.

$AB = \phi$

5. **Exhaustive Set of Events:** A collection of events $\{A_\alpha : \alpha \in I\}$ ← index set
connected to a random experiment E is exhaustive if and only if

$$\sum_{\alpha \in I} A_\alpha = S.$$

(At any performance of E at least one event of the collection is sure to occur)

Examples:

E : throwing a die;
A = even face, B = odd face.

$\{A, B\}$

6. **Equally Likely Events:** A collection of events $\{A_\alpha : \alpha \in I\}$ connected to a random experiment E are equally likely if there's no reason to believe any one of the events to occur rather than any other.

7. **Equally Likely Sample Points:** If the elementary events of a sample space are equally likely events.

Classical Definition (Laplace, 19th Century)

Let E be a random experiment with sample space S.
If S contains finite number (say, n) of equally likely sample points, then the probability of an event $A \subseteq S$ is defined as

$$P(A) = \frac{m}{n}$$

where A contains m sample points.

Classical Definition (Laplace, 19th Century)

Let E be a random experiment with sample space S.
If S contains finite number (say, n) of equally likely sample points, then the probability of an event $A \subseteq S$ is defined as

$$P(A) = \frac{m}{n}$$

where A contains m sample points.

Defects:

1. The definition can be applied to a limited number of random experiments whose sample space is finite.
2. The definition uses the concept of *equally likely* or *equally probable* sample points. Thus we are defining probability using probability.

PS-1, P1: What is the probability of an odd sum when two dice are thrown?

$S = \{(1,1), (1,2), \ldots, (6,6)\}, \; |S| = 36$

$A = \{(1,2), (2,1), \ldots, (6,5)\}$

$|A| = 18$

$P(A) = \frac{|A|}{|S|} = \frac{18}{36} = \frac{1}{2}$

PS-1, P3: Two urns contain respectively 3 white, 7 red, 15 black balls and 10 white, 6 red and 9 black balls. One ball is drawn from each urn. Find the probability that both the balls are of same colour.

3W	10W
7R	6R
15B	9B

$$P(A) = \frac{3 \times 10 + 7 \times 6 + 15 \times 9}{25 \times 25}$$

PS-1, P5: From an urn containing n balls any number of balls are drawn. Show that the probability of drawing an even number of balls is $\frac{2^{n-1}-1}{2^n-1}$.

$$|S| = \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n - 1$$

$$|A| = \binom{n}{2} + \binom{n}{4} + \cdots = 2^{n-1} - 1 .$$

Statistical Regularity

E : be a random experiment
S : sample space
A : an event.

Let E be repeated (under identical conditions) N times and A occurs N_A times. Then the *frequency ratio* of the event A is given as:

$$\boxed{f_N(A) = \frac{N_A}{N}.}$$

Now if E is repeated very large number of times, $f_N(A)$ gradually stabilises to a constant number.
This tendency of stability of frequency ratio is called *statistical regularity*.

(**Empirical/ experimental fact**)

$f_N(A) = \frac{600}{1000} \sim \frac{1}{2}$

$\underbrace{\text{HHTHTH, TH} \cdots}_{1000}$

$H \to 600$

Frequency Definition of Probability

On the basis of statistical regularity, we assume $\boxed{\lim_{N\to\infty} f_N(A)}$ exists finitely and the value of this limit is called the probability of the event A, i.e.

$$\boxed{P(A) = \lim_{N\to\infty} f_N(A).}$$

Frequency Definition of Probability

On the basis of statistical regularity, we assume $\lim_{N\to\infty} f_N(A)$ exists finitely and the value of this limit is called the probability of the event A, i.e.

$$P(A) = \lim_{N\to\infty} f_N(A).$$

Defects:

The statistical regularity is an empirical/ experimental fact whereas the limit is a rigorous mathematical concept. We cannot mix them together and it is unwise to build the theory of probability based on this definition.

Deductions: Using Classical and Frequency Definitions

1. (a) $0 \leq P(A) \leq 1$
   (b) $P(S) = 1$
   (c) $P(\emptyset) = 0$
   (d) $P(\bar{A}) = 1 - P(A)$.

Theorem of Total Probability

If $A_1, A_2, \ldots, A_k$ are pairwise mutually exclusive events, then

$$P(A_1 \overset{\cup}{+} A_2 + \ldots + A_k) = P(A_1) + P(A_2) + \ldots + P(A_k).$$

$|S| = n$

$A_1 \rightarrow n_1$ pts of S

$A_2 \rightarrow n_2 \cdots\cdots$

$$P(A_1 + \cdots + A_k) = \frac{n_1 + n_2 + \cdots + n_k}{n} = \frac{n_1}{n} + \cdots + \frac{n_k}{n} = P(A_1) + \cdots + P(A_k).$$

Event (In modern probability theory)

σ-Algebra/σ-Field/Borel-field: A class Δ of subsets of S satisfying:

(i) $S \in \Delta$

(ii) If $A \in \Delta$, then $\bar{A} \in \Delta$

(iii) If $A_1, A_2, \ldots, A_k, \ldots \in \Delta$, then $\sum_{i=1}^{\infty} A_i \in \Delta$.

Event: Any member of Δ is called an *event*.

Axiomatic Definition of Probability (Kolmogorov, 1933)

Event (In modern probability theory)

σ-Algebra/σ-Field/Borel-field: A class Δ of subsets of S satisfying:

(i) $S \in \Delta$

(ii) If $A \in \Delta$, then $\bar{A} \in \Delta$

(iii) If $A_1, A_2, \ldots, A_k, \ldots \in \Delta$, then $\sum_{i=1}^{\infty} A_i \in \Delta$.

Event: Any member of Δ is called an *event*.

Examples of σ-Algebra

1. $C_1 = \{\emptyset, S\}$ (trivial σ-field)
2. $C_2 = \{$ all subsets of $S\}$ (discrete σ-field)
3. $C_3 = \{\emptyset, S, A, \bar{A}\}$
4. $C_4 = \{$all subsets of S which are countable or whose complements are countable$\}$

Axiomatic Definition of Probability (Kolmogorov, 1933)

set of real numbers

E : Random experiment, S : Sample space, Δ : σ-Algebra
A mapping $P : \Delta \to \mathbb{R}$ is called a probability function and the unique number $P(A)$ corresponding to an event $A \in \Delta$ is called the probability of the event A if

Axiom (i): $P(A) \geq 0$ for any $A \in \Delta$

Axiom (ii): $P(S) = 1$

Axiom (iii): If $A_1, A_2, \ldots, A_k, \ldots$ be countably infinite number of mutually exclusive events then

$$P(A_1 + A_2 + \ldots + A_k + \ldots) = P(A_1) + P(A_2) + \ldots + P(A_k) + \ldots$$

(S, Δ, P) : **Probability Space**

Deductions

1. $P(\bar{A}) = 1 - P(A)$
2. $P(\emptyset) = 0$
3. $P(A) \leq 1$
4. If $A \subseteq B$, then $P(A) \leq P(B)$
5. $P(A+B) = P(A) + P(B) - P(AB)$
6. Deduction of Classical Definition
7. **Continuity theorems of probability**:
   If $\{A_n\}$ be a monotonic sequence of events, then

$$\boxed{P(\lim A_n) = \lim P(A_n).}$$

$A + \bar{A} = S$

$\Rightarrow P(A+\bar{A}) = P(S) = 1$ (Axiom (ii))

$\Rightarrow P(A) + P(\bar{A}) = 1$ (Axiom (iii))

$\Rightarrow P(\bar{A}) = 1 - P(A)$

(2) $S = S + \phi$

$\Rightarrow P(S) = P(S+\phi)$

$\Rightarrow 1 = P(S) + P(\phi)$

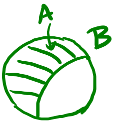
$\Rightarrow 1 = 1 + P(\phi)$

$\Rightarrow P(\phi) = 0.$

$\lim_{x \to x_0} f(x) = f(x_0)$

for any seq. $\{a_n\} \to x_0$

$$\boxed{\lim_{n\to\infty} f(a_n) = f(\lim_{n\to\infty} a_n)}$$



$B = A + (B \setminus A) \Rightarrow P(B) = P(A) + \underbrace{P(B \setminus A)}_{\geq 0}$ (Ax (iii))

$\Rightarrow P(A) \leq P(B).$

(6) Let S has n number of equally likely sample points.

Let, $U_1, U_2, \dots, U_n$ be corresponding elementary events.

$$S = U_1 + U_2 + \dots + U_n$$

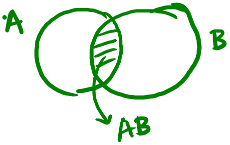
$$1 = P(S) = P(U_1) + \dots + P(U_n) \quad \text{(Axiom (ii) \& (iii))}$$

$$P(U_1) = P(U_2) = \dots = P(U_n) = \frac{1}{n}$$

$$A = U_{i_1} + U_{i_2} + \dots + U_{i_m}$$

$$P(A) = P(U_{i_1}) + P(U_{i_2}) + \dots + P(U_{i_m}) \quad \text{(Axiom (iii))}$$

$$= \frac{m}{n}.$$



$$P(A_1+A_2+\cdots+A_n)$$
$$= P(A_1)+\cdots+P(A_n)$$
$$-\overbrace{P(A_1A_2)-P(A_2A_3)-\cdots-P(A_{n-1}A_n)}^{\sum_{1\le i<j\le n} P(A_iA_j)}$$
$$+\cdots$$
$$+(-1)^{n-1}P(A_1A_2\cdots A_n)$$

$$A+B=(A\setminus AB)+AB+(B\setminus AB)$$

$$P(A+B)=P(A\setminus AB)+P(AB)+P(B\setminus AB) \quad (\text{Axiom iii})$$

$$A=(A\setminus AB)+AB$$

$$\Rightarrow P(A)=P(A\setminus AB)+P(AB) \quad (\text{Axiom iii})$$

$$\Rightarrow P(A\setminus AB)=P(A)-P(AB)$$

$$P(A+B)=P(A)-\cancel{P(AB)}+\cancel{P(AB)}+P(B)-P(AB)$$
$$=P(A)+P(B)-P(AB).$$

Monotonic Sequence of Events and Properties

Defn
1. $\{A_n\}$ is a monotonic increasing of events if: $\boxed{A_n \subseteq A_{n+1}}$ $\forall n \in \mathbb{N}$

Defn
2. $\{A_n\}$ is a monotonic decreasing of events if: $\boxed{A_n \supseteq A_{n+1}}$ $\forall n \in \mathbb{N}$
3. For a monotonic increasing sequence of events $\{A_n\}$:

$$\lim A_n = \bigcup_{n=1}^{\infty} A_n = \sum_{n=1}^{\infty} A_n$$

4. For a monotonic decreasing sequence of events $\{A_n\}$:

$$\lim A_n = \bigcap_{n=1}^{\infty} A_n = \prod_{n=1}^{\infty} A_n$$

Conditional Probability

The conditional probability of the event A on the hypothesis that event B has occurred, denoted by $P(A|B)$, is defined as

$$P(A|B) = \frac{P(AB)}{P(B)}, \text{ provided } P(B) \neq 0.$$

If $P(B) = 0$, $P(A|B)$ is not defined.

The conditional probability of the event A on the hypothesis that event B has occurred, denoted by $P(A|B)$, is defined as

$$P(A|B) = \frac{P(AB)}{P(B)}, \text{ provided } P(B) \neq 0.$$

If $P(B) = 0$, $P(A|B)$ is not defined.

classical Def^n

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{\frac{|AB|}{|S|}}{\frac{|B|}{|S|}} = \frac{|AB|}{|B|}$$

Example:

$E =$ throwing a fair die; $S = \{1, 2, \ldots, 6\}$
$A =$ even face; $A = \{2, 4, 6\}$
$B =$ multiple of 3. $B = \{3, 6\}$
Compute $P(A|B)$ and $P(B|A)$.

$$\frac{P(AB)}{P(B)} = \frac{\frac{1}{6}}{\frac{2}{6}} = \frac{1}{2}$$

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}.$$

1. Show that the conditional probability satisfies all the axioms of probability.

(i) for any two events A, B

$$P(A|B) = \frac{P(AB)}{P(B)} \geq 0$$

(ii)

$$P(S|B) = \frac{P(SB)}{P(B)} = \frac{P(B)}{P(B)} = 1 .$$

(iii) $\{A_n\}$: seq. of pairwise mutually exclusive events

$$P\left(\sum A_i \,|\, B\right) = \frac{P\left(\sum A_i B\right)}{P(B)} = \frac{\sum P(A_i B)}{P(B)} = \sum P(A_i|B)$$

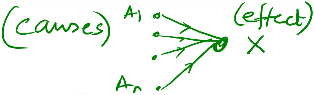
Multiplication Rule

2. (i) If $P(A), P(B) \neq 0$,
$P(AB) = P(A)P(B|A) = P(B)P(A|B)$

(ii) $P(ABC) = P(A)P(B|A)P(C|AB)$

(iii) $P(A_1A_2\ldots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\ldots P(A_n|A_1A_2\ldots A_{n-1})$

$$\begin{aligned} RHS &= P(A)\,P(B|A) \\ &= P(A)\,\frac{P(AB)}{P(A)} \\ &= P(AB) \\ &= LHS. \end{aligned}$$



If $A_1, A_2, \ldots, A_n$ be a set of n

(i) pairwise exclusive i.e. $A_iA_j = \emptyset$ $(i \neq j; i,j = 1,2,\ldots,n)$ and

P

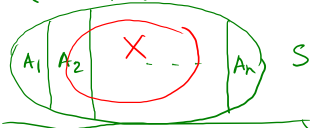
(ii) exhaustive set of events, i.e. $A_1 + A_2 + \ldots + A_n = S$

then for any arbitrary event X

(I) $P(X) = P(A_1)P(X|A_1) + P(A_2)P(X|A_2) + \ldots + P(A_n)P(X|A_n)$

(II) If $P(X) \neq 0$,

$$\underbrace{P(A_i|X)}_{P(\text{cause}\,|\,\text{effect})} = \frac{P(A_i)P(X|A_i)}{P(A_1)P(X|A_1)+P(A_2)P(X|A_2)+\ldots+P(A_n)P(X|A_n)}$$



$$X = XS = X(A_1 + A_2 + \cdots + A_n)$$
$$= XA_1 + XA_2 + \cdots + XA_n$$
$$(XA_i)(XA_j) = \Phi \quad \text{for } i \neq j$$

$$P(A_i|X) = \frac{P(A_iX)}{P(X)} = \frac{P(A_i)P(X|A_i)}{}$$

$$P(X) = P(XA_1) + \cdots + P(XA_n)$$
$$= P(A_1)P(X|A_1) + \cdots + P(A_n)P(X|A_n)$$

Independence of Events

(stochastically)

Two events A, B are said to be independent if $P(AB) = P(A)P(B)$.

$$P(A|B) = \frac{P(AB)}{P(B)}$$

if A is independent of B : $P(A|B) = P(A)$

$$\Rightarrow \quad P(A) = \frac{P(AB)}{P(B)}$$

$$\Rightarrow \quad P(AB) = P(A)P(B).$$

Independence of Events

Two events A, B are said to be independent if $P(AB) = P(A)P(B)$.

Pairwise independent: Events A, B, C are pairwise independent if $P(AB) = P(A)P(B), P(BC) = P(B)P(C), P(CA) = P(C)P(A)$.

Mutually independent: A, B, C are mutually independent if

$$P(AB) = P(A)P(B)$$
$$P(BC) = P(B)P(C)$$
$$P(CA) = P(C)P(A)$$
$$P(ABC) = P(A)P(B)P(C)$$

Assume that neither A nor B has zero probability

1. If A and B are mutually exclusive can they be always independent?
2. If A and B are independent can they be always mutually exclusive?

$AB = \phi$

$\underline{P(AB)} = 0$

$P(A)P(B|A)$

$P(B|A) = 0$

1 No. $E \rightarrow$ two coins are tossed

$A \rightarrow$ both are heads $= \{HH\}$

$B \rightarrow$ both are tails $= \{TT\}$

$S = \{HH, TT, HT, TH\}$

$AB = \phi$, $P(A) = \frac{1}{4}$, $P(B) = \frac{1}{4}$

$P(AB) = 0$

$P(AB) \neq P(A)P(B)$

2 No. $A \rightarrow$ 'H' in the first coin

$B \rightarrow$ 'H' - - - second coin.

Problems

PS-2, P1: Let A, B be two independent events. Prove that (i) $A, \bar{B}$ are independent, (ii) $\bar{A}, \bar{B}$ are independent.

Sol.

$$P(AB) = P(A)P(B)$$

$$A\bar{B} + AB = A$$

$$\Rightarrow P(A\bar{B}) + P(AB) = P(A)$$

$$\Rightarrow P(A\bar{B}) = P(A) - P(AB)$$

$$= P(A) - P(A)P(B)$$

$$= P(A)(1 - P(B))$$

$$= P(A)P(\bar{B})$$

Problems

PS-2, P3: An urn contains 4 white and 6 black balls. Two balls are drawn successively without replacement. If the first ball is seen to be white, what is the probability that the second ball is also white?

Problems

PS-2, P5: There are two identical urns containing 4 white and 3 red balls; 3 white and 7 red balls. An urn is chosen at random and a ball is drawn from it. Find the probability that the ball is white. If the ball drawn is white, what is the probability that it is from the first urn?