ST233
AY22/23 Sem 2
github.com/SeekSaveServe

Chapter 1: Basic Concepts of Probability

DEFINITION 1 (EXPERIMENT, SAMPLE SPACE, EVENT)

A statistical experiment is any procedure that produces data or observations.

The sample space, denoted by S, is the set of all possible outcomes of a statistical experiment. The sample space depends on the problem of interest!

A sample point is an outcome (element) in the sample space.

An event is a subset of the sample space.

Union: $A \cup B = x : x \in A \vee x \in B$
Union of n events: $\cup_{i=1}^{n} = x : x \in A_1 \vee x \in A_2 ...$
Union: $A \cap B = x : x \in A \wedge x \in B$
Interection of n events: $\cap_{i=1}^{n} = x : x \in A_1 \wedge x \in A_2 ...$
Complement: $A' = x : x \in S \wedge x \notin A$
Mutually Exclusive: $A \cap B = \emptyset$

MORE EVENT OPERATIONS

(a) $A \cap A' = \emptyset$ (b) $A \cap \emptyset = \emptyset$
(c) $A \cup A' = S$ (d) $(A')' = A$
(e) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
(f) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
(g) $A \cup B = A \cup (B \cap A')$
(h) $A = (A \cap B) \cup (A \cap B')$

DE MORGAN'S LAW

For any n events $A_1, A_2, \ldots, A_n$,

(i) $(A_1 \cup A_2 \cup \ldots \cup A_n)' = A_1' \cap A_2' \cap \ldots \cap A_n'$.
A special case: $(A \cup B)' = A' \cap B'$.
(j) $(A_1 \cap A_2 \cap \ldots \cap A_n)' = A_1' \cup A_2' \cup \ldots \cup A_n'$.
A special case: $(A \cap B)' = A' \cup B'$.

Multiplication	**Addition**
Sequential experiments $n_1 * n_2 * ... * n_r$ outcomes	Independent experiments $n_1 + n_2 + ... + n_r$ outcomess

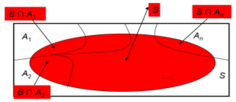
Permutation: $P_r^n = \frac{n!}{(n-r)!} = n(n-1)...(n-(r-1))$
Selection and arrangement of r objects out of n. Order is considered.
Combination: $C_r^n = \frac{n!}{r!(n-r)!} = \frac{n(n-1)...(n-(r-1))}{r!}$
Selection of r objects out of n. Order is not considered.
$C_r^n = C_{n-r}^n$
Axioms and Properties of Probability

1 $0 \leq P(A) \leq 1$
2 P(S)=1
3 $P(A \cup B) = P(A) + P(B)$ if A and B are **mutex** (not to be confused with independence)
4 P($\emptyset$)=0
5 P($A_1 \cup A_2 \cup ... A_n$)=$\sum_{i=1}^{n} A_i$
6 P(A')=$1 - P(A)$
7 $P(A \cup B) = P(A) + P(B) - P(AB)$
8 P(A)=$P(A \cap B) + P(A \cap B')$
9 $A \subset B \rightarrow P(A) \leq P(B)$
10 $P(A_1) = P(A_2) = ... = P(A_k) \rightarrow for B \subset S, P(B) = \frac{|B|}{|S|}$

Conditional Probability (B occurs given A):
$P(B|A) = \frac{P(A \cap B)}{P(A)}$
Multiplication rules
$P(A \cap B) = P(A)P(B|A)$ if $P(A)! = 0$
$P(A \cap B) = P(B)P(A|B)$ if $P(B)! = 0$
Inverse Probability Formula: $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$
Indepence $A \perp B iff P(A \cap B) = P(A)P(B)$ This implies $P(A|B) = P(A)$ and $P(B|A) = P(B)$
Partition $A_1, A_2 ... A_r$ are mutually exclusive and $\sum_{i=1}^{n} A_i = S$

THEOREM 11 (LAW OF TOTAL PROBABILITY)
Suppose $A_1, A_2, \ldots, A_n$ is a partition of S. Then for any event B, we have

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(A_i)P(B|A_i).$$



$P(B) = P(A)P(B|A) + P(A')P(B|A')$
Bayes' theorem
$P(A_k|B = \frac{P(B|A)P(A_k)}{P(A)P(B|A) + P(A')P(B|A')})$
Bayes' Theorem (n=2): $P(A|B) = \frac{P(B|A)P(A)}{\sum_{i=1}^{n} P(A_i)P(B|A_k)}$
Denom is P(B)

L2:Random Variables

Random variable X is a function from S to R
Uppercase letters denote random variables
Lowercase letters denote observed values of random variables

DEFINITION 3 (PROBABILITY MASS FUNCTION)
For a discrete random variable X, define

$$f(x) = \begin{cases} P(X = x), & \text{for } x \in R_X; \\ 0, & \text{for } x \notin R_X. \end{cases}$$

*Then $f(x)$ is known as the **probability function (pf)**, or **probability mass function (pmf)** of X.*

*The collection of pairs $(x_i, f(x_i)), i = 1, 2, 3, \ldots$, is called the **probability distribution** of X.*

PROPERTIES OF THE PROBABILITY MASS FUNCTION
The probability mass function $f(x)$ of a discrete random variable **must** satisfy:

(1) $f(x_i) \geq 0$ for all $x_i \in R_X$;
(2) $f(x) = 0$ for all $x \notin R_X$;
(3) $\sum_{i=1}^{\infty} f(x_i) = 1$, or $\sum_{x_i \in R_X} f(x_i) = 1$.

For any set $B \subset \mathbb{R}$, we have

$$P(X \in B) = \sum_{x_i \in B \cap R_X} f(x_i).$$

DEFINITION 4 (PROBABILITY DENSITY FUNCTION)
*The **probability density function** of a continuous random variable X, denoted by $f(x)$, is a function that satisfies:*

(1) $f(x) \geq 0$ for all $x \in R_X$; and $f(x) = 0$ for $x \notin R_X$;
(2) $\int_{R_X} f(x)\,dx = 1$;
(3) For any a and b such that $a \leq b$,

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx.$$

note: $P(a < X < b) = P(a \leq X \leq B)$
To check that a function is a PDF, chcek conditions 1 and 2
Cumulative Distribution Function: $F(x) = P(X \leq x)$ and
$P(a \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a-)$
Deriving CDF from probability distribution

x	0	1	2
$f(x)$	1/4	1/2	1/4

We have

$$F(0) = f(0) = 1/4, F(1) = f(0) + f(1) = 3/4, F(2) = f(0) + f(1) + f(2) = 1.$$

D ... bution from cdf

$$F(x) = \begin{cases} 0, & x < 0 \\ 1/4, & 0 \leq x < 1 \\ 3/4, & 1 \leq x < 2 \\ 1, & 2 \leq x \end{cases}.$$

- f(1)=P(X=0)=F(0)-F(0-)=$\frac{1}{4}$-0=$\frac{1}{4}$
- f(2)=P(X=1)=F(1)-F(1-)=$\frac{3}{4}$-$\frac{1}{4}$=$\frac{1}{2}$
- f(3)=P(X=2)=F(2)-F(2-)=1-$\frac{3}{4}$=$\frac{1}{4}$

CDF - Continuous Random Variable

- F(X) assumes different values in R_x when F(X)!=0
- F(X) is always non-decreasing (for discrete varaible too) $x_1 < x_2 \rightarrow F(x_1) < F(x_2)$
- Probability function and CDF is one-to-one and uniquely determined
- $0 < F(x) < 1$
- (discrete), $0 < f(x) < 1$
- (continuous), $f(x) \geq 0$ but $f(x)$ is not necessarily ≤ 1. While the PDF itself cannot exceed 1, there are points within its range that are greater than 1

CDF: CONTINUOUS RANDOM VARIABLE
If X is a continuous random variable,

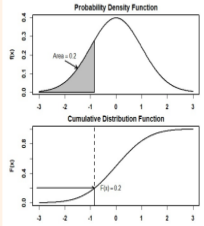
$$F(x) = \int_{-\infty}^{x} f(t)dt,$$

and

$$f(x) = \frac{dF(x)}{dx}.$$

Further

$$P(a \leq X \leq b) = P(a < X < b) = F(b) - F(a).$$



Find CDF of a continuous random variable

$$f(x) = \begin{cases} \frac{x}{\theta^2} e^{-x^2/(2\theta^2)}, & \text{for } x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \int_0^x \frac{t}{\theta^2} e^{-t^2/(2\theta^2)}\,dt$$

$$= \left[-e^{-t^2/(2\theta^2)}\right]_{t=0}^{x} = 1 - e^{-x^2/(2\theta^2)}.$$

Discrete: Expectation (mean) and Variance
- $E(X) = \sum_{x \in R_x} x_i f(x_i) = \mu_x$

Random: Expectation (mean) and Variance
- $E(X) = \mu_x = \int_{-\infty}^{\infty} x f(x) dx = \infty_{x \in R_x} x f(x) dx$
- The mean of X is not necessarily a possible value

Properties of Expectation
1 $E(aX + b) = aE(x) + b$
2 $E(X + Y) = E(X) + E(Y)$
3 (discrete) $E|g(x)| = \sum_{x \in R_x} g(x) f(x)$
3 (continuous) $E|g(x)| = \int_{x \in R_x} g(x) f(x) dx$
- Using properties 1, 2 we have $E(a_1 X_1 + a_2 X_2 + ... + a_k X + k) = a_1 E(X_1) + ... + a_k E(X_k)$

Variance
- $Var(X) = \sigma_x^2 = E(X - \mu_x)^2 = E(X^2) - E(X)^2$
- (discrete) $Var(X) = \sum_{x \in R_x} (x - \mu_x)^2 f(x)$
- (continuous) $Var(X) = \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x) dx$
- V(X) le 0 for any X, V(X)=0 when X is a constant
-