ST2334 Midterm Cheatsheet

Basic probability concepts

Oberservation

Any recording of information, whether it is numerical or categorical.

Statistical Experiment

Any procedure that generates a set of data (observations).

Sample Space

The set of all possible outcomes of a statistical experiment is called the sample space and it is represented by the symbol S.

Sample Point

Every outcome in a sample space is called an element of the sample space or simply a sample point.

Event

An event is a subset of a sample space.

Simple Event

An event is said to be simple if it consists of exactly one outcome (i.e. one sample point)

Compound Event

An event is said to be compound if it consists of more than one outcomes (or sample points).

1. The sample space is itself an event and is usually called a sure event.
2. A subset of S that contains no elements at all is the empty set, denoted by $\emptyset$, and is usually called a null event.

Operations of Events

Union

The Union of two events A and B, denoted by $A \cup B$, is the event containing all the elements that belong to A or B or to both. That is,

$$A \cup B = \{x : x \in A \text{ or } x \in B\}$$

Intersection

The intersection of two events A and B, denoted by $A \cap B$ or simply AB, is the event containing all elements that are common to A and B. That is

$$A \cap B = \{x : x \in A \text{ and } x \in B\}$$

Complement

The complement of event A with respect to S, denoted by A' or A^C, is the set of all elements of S that are not in A. That is

$$A' = \{x : x \in S \text{ and } x \notin A\}$$

Mutually Exclusive Events

Two events A and B are said to be mutually exclusive or mutually disjoint if $A \cap B = \emptyset$, that is, if A and B have no elements in common.

Union of n Events

The union of n events $A_1, A_2, \cdots, A_n$, denoted by

$$A_1 \cup A_2 \cup ... \cup A_n$$

is the event containing all the elements that belong to one or more of the events A_1, A_2, or ..., or A_n. That is

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n = \{x : x \in A_1 \text{ or } \cdots \text{ or } x \in A_n\}$$

Intersection of n Events

The intersection of n events $A_1, A_2, \cdots, A_n$, denoted by

$$A_1 \cap A_2 \cap ... \cap A_n$$

is the event containing all the elements that are common to all of the events A_1, A_2, or ..., or A_n. That is

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n = \{x : x \in A_1 \text{ and } \cdots \text{ and } x \in A_n\}$$

Counting

Permutation

A permutation is an arrangement of r objects from a set of n objects, where $r \leq n$. (Note that the order is taken into consideration in permutation.)

$${}_nP_r = n(n-1)(n-2)\cdots(n-(r-1)) = n!/(n-r)!$$

When not all objects are distinct,

$$nP_{n_1, n_2, \cdots, n_k} = \frac{n!}{n_1! n_2! \cdots n_k!}$$

Combination

the number of ways of selecting r objects from n objects without regard to the order.

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Axioms of Probability

Axiom 1

$0 \leq Pr(A) \leq 1$

Axiom 2

$Pr(S) = 1$

Axiom 3

If $A_1, A_2, \cdots$ are mutually exclusive (disjoint) events, then

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i)$$

Inclusion-Exclusion Principle

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Pr(A_i \cap A_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} \Pr(A_i \cap A_j \cap A_k) - \cdots\cdots + (-1)^{n+1} \Pr(A_1 \cap A_2 \cap \cdots \cap A_n)$$

Conditional Probability

The conditional probability of B given A is defined as

$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)}, \quad \text{if } \Pr(A) \neq 0$$

Multiplication Rule of Probability

In general,

$$\Pr(A_1 \cap \cdots \cap A_n) = \Pr(A_1) \Pr(A_2|A_1) \Pr(A_3|A_1 \cap A_2) \cdots \Pr(A_n|A_1 \cap \cdots \cap A_{n-1})$$
providing that $\Pr(A_1 \cap \cdots \cap A_{n-1}) > 0$

The Law of Total Probability

Let $A_1, A_2, \cdots, A_n$ be a partition of the sample space S. That is $A_1, A_2, \cdots, A_n$ are mutually exclusive and exhaustive events such that $A_i \cap A_j = \emptyset$ for $i \neq j$ and $\bigcup_{i=1}^{n} A_i = S$.
Then for any event B

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

Baye's Theorem

Let $A_1, A_2, \cdots, A_n$ be a partition of the sample space S. Then

$$\Pr(A_k|B) = \frac{\Pr(A_k) \Pr(B|A_k)}{\sum_{i=1}^{n} \Pr(A_i) \Pr(B|A_i)}$$

for $k = 1, \cdots, n$. Or

$$\Pr(A_k|B) = \frac{\Pr(A_k) \Pr(B|A_k)}{\Pr(B)}$$

Independent Events

Two events A and B are independent iff

$$\Pr(A \cap B) = \Pr(A) \Pr(B)$$

Pairwise Independence

A set of events $A_1, A_2, \cdots, A_n$ are pairwise independent iff

$$\Pr(A_i \cap A_j) = \Pr(A_i) \Pr(A_j)$$

for $i \neq j$ and $i, j = 1, \cdots, n$

Mutual Independence

A set of events $A_1, A_2, \cdots, A_n$ are pairwise independent iff for any subset $\{A_{i1}, A_{i2}, \cdots, A_{ik}\}$ of $A_1, A_2, \cdots, A_n$,

$$\Pr\left(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\right) = \Pr(A_{i_1}) \Pr(A_{i_2}) \cdots \Pr(A_{i_k})$$

Concepts of Random Variables

Random Variable

Let S be a sample space associated with the experiment, E. A function X, which assigns a number to every element $s \in S$, is called a random variable.

Discrete Random Variable

If the number of possible values of X(i.e.,R_X,the range space) is finite or countable infinite, we call X a discrete random variable.

Probability (Mass) Function

The probability of $X = x_i$ denoted by $f(x_i)$ (i.e. $f(x_i) = \mathrm{pr}\,(X = x_i)$, must satisfy the following two conditions.
(1) $f(x_i) \geq 0$ for all x_i.
(2)$\sum_{i=1}^{\infty} f(x_i) = 1$

Continuous Random Variable

The range space R_x is an interval or a range of intervals.

Probability Density Function

Let X be a continuous random variable.

1. $f(x) \geq 0$ for all $x \in R_X$
2. $\int_{R_X} f(x)dx = 1$ or $\int_{-\infty}^{\infty} f(x)dx = 1$
   since $f(x) = 0$ for x not in R_X
3. For any c and d such that $c < d$, (i.e. $(c, d) \subset \boldsymbol{R}_X$) ,
   $\Pr(c \leq X \leq d) = \int_c^d f(x)dx$

Cumulative Distribution Function

We define $F(x)$ to be the cumulative distribution function of the random variable X (abbreviated as c.d.f.) where

$$F(x) = \Pr(X \leq x)$$

If X is a discrete random variable, then its c.d.f is a step function.

$$\begin{aligned} F(x) &= \sum_{t \leq x} f(t) \\ &= \sum_{t \leq x} \Pr(X = t) \end{aligned}$$

If X is a continuous random variable, then

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

For a continuous random variable X,

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

if the derivative exists.

Mean

If X is a discrete random variable, taking on values $x_1, x_2, \cdots$ with probability function $f(x)$, then the mean or expected value of X, denoted by $E(X)$, is defined by

$$\mu_X = E(X) = \sum_i x_i f(x_i) = \sum_x x f(x)$$

If X is a continuous random variable with probability density function $f(x)$, then the mean is defined by

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x)dx$$

For any function $g(X)$,
(a) $E[g(X)] = \sum_x g(x) f_X(x)$
(b) $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx$
Properties:

$$E(aX + b) = aE(X) + b$$

In general,

$$\begin{aligned} &E[a_1 g_1(X) + a_2 g_2(X) + \cdots + a_k g_k(X)] \\ &= a_1 E[g_1(X)] + a_2 E[g_2(X)] + \cdots + a_k E[g_k(X)] \end{aligned}$$

Variance

$$\sigma_X^2 = V(X) = E\left[(X - \mu_X)^2\right]$$

$$= \begin{cases} \sum_x (x - \mu_X)^2 f_X(x), & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx, & \text{if } X \text{ is continuous.} \end{cases}$$

Remarks:
(a) $V(X) \geq 0$
(b) $V(X) = E(X^2) - [E(X)]^2$

Properties:

$$V(aX + b) = a^2 V(X)$$

Standard Deviation

The positive square root of the variance.

Moment

The k-th moment of X is defined by $E(X^k)$.

Chebyshev's Inequality

Let X be a random variable (discrete or continuous) with $E(X) = \mu$ and $V(X) = \sigma^2$. For any positive number k,

$$\Pr(|X - \mu| \geq k\sigma) \leq 1/k^2$$

That is, the probability that the value of X lies at least k standard deviation from its mean is at most $\frac{1}{k^2}$. Alternatively,

$$\Pr(|X - \mu| < k\sigma) \geq 1 - 1/k^2$$

This is true for **all** distributions with finite mean and variance.

Two-dimensional Random Variables

Definition of 2D RV

Let E be an experiment and S a sample space associated with E. Let X and Y be two functions each assigning a real number to each $s \in S$.
We call (X, Y) a two-dimensional random variable. (Sometimes called a random vector).
The above definition can be extended to n random variables.

Range Space

$$R_{X,Y} = (x, y) | x \in X(s), y \in Y(s), s \in S$$

The above definition can be extended to more than two random variables.

Discrete vs. Continuous

Discrete: (X, Y) is a two-dimensional discrete random variable if the possible values of $(X(s), Y(s))$ are finite or countably infinite.
Continuous: (X, Y) is a two-dimensional continuous random variable if the possible values of $(X(s), Y(s))$ can assume all values in some region of the Euclidean plane $\mathbb{R}^2$.

Joint Probability Function

Let (X, Y) be a 2-dimensional discrete random variable defined on the sample space of an experiment. With each possible value (x_i, y_i), we associate a number $f^{X,Y}(x_i, y_i)$ representing $\Pr(X = x_i, Y = y_i)$ and satisfying the following conditions:
1. $f_{X,Y}(x_i, y_j) \geq 0$ for all $(x_i, y_j) \in R_{X,Y}$.
2.

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f_{X,Y}(x_i, y_j) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \Pr(X = x_i, Y = y_j) = 1$$

The function $f_{X,Y}(x, y)$ is called the joint probability function for (X, Y).

$$\Pr((X, Y) \in A) = \underbrace{\sum\sum}_{(x, y) \in A} f_{X,Y}(x, y)$$

Joint Probability Density Function

Let (X, Y) be a 2-dimensional continuous random variable assuming all values in some region R of the Euclidean plane $\mathbb{R}^2$. $f^{X,Y}(x, y)$ is called a joint probability density function if it satisfies the following conditions:

1. $f_{X,Y}(x, y) \geq 0$ for all $(x, y) \in R_{X,Y}$
2. $\iint_{(x,y) \in R_{X,Y}} f_{X,Y}(x, y)dxdy = 1$
   or
   $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)dxdy = 1$

Prepared by Zechu, AY2019/2020 Semester 1