Lecture Notes: Probability Part I

Lecturer: Pheak NEANG

Date: May 6, 2024

0.1 Concept of Probability

0.1.1 Sample Spaces, Events, and Their Probabilities

Definition 0.1.1. A **random experiment** is a mechanism that produces a definite outcome that cannot be predicted with certainty. The **sample space** associated with a random experiment is the set of all possible outcomes. An **event** is a subset of the sample space.

Definition 0.1.2. An event E is said to occur on a particular trial of the experiment if the outcome observed is an element of the set E.

Example 0.1.3. Construct a sample space for the experiment that consists of rolling a single die. Find the events that correspond to the phrases "an even number is rolled" and "a number greater than two is rolled."

Example 0.1.4. Construct a sample space that describes all three-child families according to the genders of the children with respect to birth order.

0.1.2 Complements, Intersections, and Unions

Definition 0.1.5. The **complement** of an event A in a sample space S, denoted A^c is the collection of all outcomes in S that are not elements of the set A. It corresponds to negating any description in words of the event A.

The Probability rule for Complements:

$$\Pr(A^c) = 1 - \Pr(A)$$

Example 0.1.6. Two events connected with the experiment of rolling a single die are E: "the number rolled is even" and T: "the number rolled is greater than two." Find the complement of each.

Definition 0.1.7. The intersection of events A and B, denoted $A \cap B$, is the collection of all outcomes that are elements of both of the sets A and B. It corresponds to combining descriptions of the two events using the word "and."

Events A and B are mutually exclusive if they have no elements in common.

Definition 0.1.8. Events A and B are mutually exclusive if and only if

$$\Pr(A \cap B) = 0$$

Definition 0.1.9. The union of events A and B, denoted $A \cup B$, is the collection of all outcomes that are elements of one or the other of the sets A and B, or of both of them. It corresponds to combining descriptions of the two events using the word "or."

The Additive rule of Probability

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

Example 0.1.10. Find the probability that at least one heads will appear in five tosses of a fair coin.

Example 0.1.11. A single die is rolled.

a. Suppose the die is fair. Find the probability that the number rolled is both even and greater than two.

b. Suppose the die has been "loaded" so that$P(1) = 1/12$, $P(6) = 3/12$, and the remaining four outcomes are equally likely with one another. Now find the probability that the number rolled is both even and greater than two.

0.1.3 Conditional Probability and Independent Events

Definition 0.1.12. The **conditional probability** of A given B, denoted $\Pr(A|B)$, is the probability that event A has occurred in a trial of a random experiment for which it is known that event B has definitely occurred. It may be computed by means of the following formula:

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

Example 0.1.13. A fair die is rolled.

a. Find the probability that the number rolled is a five, given that it is odd.

b. Find the probability that the number rolled is odd, given that it is a five.

Definition 0.1.14. Event A and B are independent if

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

If A and B are not independent, then they are **dependent.**

Example 0.1.15. The specificity of a diagnostic test for a disease is the probability that the test will be negative when administered to a person who does not have the disease. The higher the specificity, the lower the false positive rate. Suppose the specificity of a diagnostic procedure to test whether a person has a particular disease is 89%.

a. A person who does not have the disease is tested for it using this procedure. What is the probability that the test result will be positive?

b. A person who does not have the disease is tested for it by two independent laboratories using this procedure. What is the probability that both test results will be positive ?

Example 0.1.16. Suppose a particular species of trained dogs has a 90% chance of detecting contraband in airline luggage. If the luggage is checked three times by three different dogs independently of one another, what is the probability that contraband will be detected?

0.2 Discrete Random Variables

0.2.1 Random Variables

Definition 0.2.1. A **random variable** is a numerical quantity that is generated by a random experiment.

Definition 0.2.2. A random variable is called **discrete** if it has either a finite or a countable number of possible values. A random variable is called **continuous** if its possible values contain a whole interval of numbers.

Example 0.2.3. Classify each random variable as either discrete or continuous.

a. The number of arrivals at an emergency room between midnight and $6:00$ a.m.

b. The weight of a box of cereal labeled "18 ounces."

c. The number of kernels of popcorn in a 1-pound container.

d. The number of applicants for a job.

Example 0.2.4. Classify each random variable as either discrete or continuous.

a. The number of boys in a randomly selected three-child family.

b. The temperature of a cup of coffee served at a restaurant.

c. The number of no-shows for every 100 reservations made with a commercial airline.

d. The number of vehicles owned by a randomly selected household.

e. The average amount spent on electricity each July by a randomly selected household in a certain state.

0.2.2 Probability Distributions for Discrete Random Variables

Definition 0.2.5. The probability distribution of a discrete random variable X is a list of each possible value of X together with the probability that X takes that value in one trial of the experiment.

The probabilities in the probability distribution of a random variable X must satisfy the following two conditions:

1. Each probability $\Pr(X = x)$ must be between 0 and 1: $0 \leq \Pr(x) \leq 1$.
2. The sum of all the probabilities is 1, that is $\sum \Pr(x) = 1$.

Example 0.2.6. A fair coin is tossed twice. Let X be the number of heads that are observed.

a. Construct the probability distribution of X.

b. Find the probability that at least one head is observed.

Definition 0.2.7. The **mean** (also the expected value) of a discrete random variable X is the number.

$$\mu = E(X) = \sum x \Pr(x)$$

The mean of a random variable may be interpreted as the average of the values assumed by the random variable in repeated trials of the experiment.

Example 0.2.8. A service organization in a large town organizes a raffle each month. One thousand raffle tickets are sold for \$1 each. Each has an equal chance of winning. The first prize is \$300, the second prize is \$200, and the third prize is \$100. Let X denote the net gain from the purchase of one ticket.

a. Construct the probability distribution of X.

b. Find the probability of winning any money in the purchase of one ticket.

c. Find the expected value of X, and interpret its meaning.

Definition 0.2.9. The **variance**, σ^2, of a discrete random variable X is the number

$$\sigma^2 = \sum (x-\mu)^2 \Pr(X = x),$$

which by algebra equivalent by

$$\sigma^2 = \left[\sum x^2 \Pr(X = x)\right] - \mu^2$$

Definition 0.2.10. The **standard divation**, σ, of a discrete random variable X is the number

$$\sigma = \sqrt{\sum (x-\mu)^2 \Pr(X = x)},$$

which by algebra equivalent by

$$\sigma = \sqrt{\left[\sum x^2 \Pr(X = x)\right] - \mu^2}$$

Example 0.2.11. A discrete random variable X has the following probability distribution:

X	-1	0	1	4
$\Pr(X = x)$	0.2	0.5	a	0.1

Compute each of the following quantities.

a. a

b. $\Pr(0)$

c. $\Pr(X > 0)$

d. $\Pr(X \geq 0)$

e. $\Pr(X \leq 0)$

f. The mean μ of X

g. The Variance σ^2 of X

h. The standard deviation σ of X

0.2.3 The Binomial Distribution

Definition 0.2.12. Suppose a random experiment has the following characteristics.

1. There are n identical and independent trials of a common procedure.
2. There are exactly two possible outcomes for each trial, one termed "success" and the other "failure."
3. The probability of success on any one trial is the same number p.

Then the discrete random variable X that counts the number of successes in the n trials is the binomial random variable with parameters n and p. We also say that X has a binomial distribution with parameters n and p.

Remark 0.2.13. Probability Formula for a Binomial Random Variable. If X is a binomial random variable with parameters n and p, then

$$\Pr(x) = \frac{x!}{x!(n-x)!} p^x q^{n-x},$$

where $q = 1 - p$ and where for any counting number $m, m!$ ("m factorial) is defined by

$$0! = 1, 1! = 1, 2! = 1.2, \quad 3! = 1.2.3$$

and generally

$$m! = 1.2 \cdots (m-1).m$$

Example 0.2.14. Seventeen percent of victims of financial fraud know the perpetrator of the fraud personally.

a. Use the formula to construct the probability distribution for the number X of people in a random sample of five victims of financial fraud who knew the perpetrator personally.

b. A investigator examines five cases of financial fraud every day. Find the most frequent number of cases each day in which the victim knew the perpetrator.

c. A investigator examines five cases of financial fraud every day. Find the average number of cases per day in which the victim knew the perpetrator.

d. Find the mean and standard deviation of the random variable X.

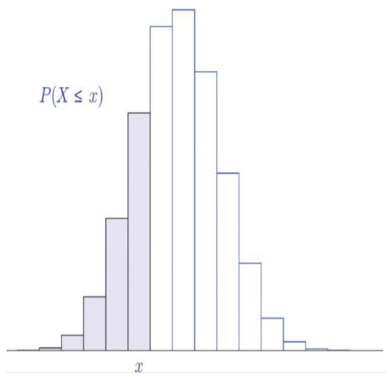
Remark 0.2.15. If X is a binomial random variable with parameters n and p, then

$$\mu = np, \quad \sigma^2 = npq, \quad \sigma = \sqrt{npq},$$

where $q = 1 - p$.

Property 0.2.16. The Cumulative Probability Distribution of a Binomial Random Variable: If X is a discrete random variable, then

$$\Pr(X \geq x) = 1 - \Pr(X \leq x - 1) \text{ and } \Pr(x) = \Pr(X \leq x) - \Pr(X \leq x - 1)$$



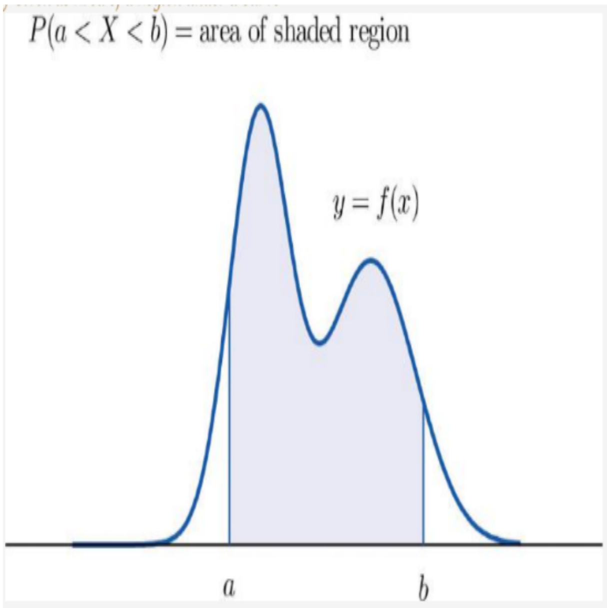
Example 0.2.17. A student takes a ten-question true/false exam.

a. Find the probability that the student gets exactly six of the questions right simply by guessing the answer on every question.

b. Find the probability that the student will obtain a passing grade of 60% or greater simply by guessing.

0.3 Continuous Random Variables

Definition 0.3.1. The probability distribution of a continuous random variable X is an assignment of probabilities to intervals of decimal numbers using a function $f(x)$, called a **density function**, in the following way: the probability that X assumes a value in the interval $[a, b]$ is equal to the area of the region that is bounded above by the graph of the equation $y = f(x)$, bounded below by the x-axis, and bounded on the left and right by the vertical lines through a and b, as illustrated in the following



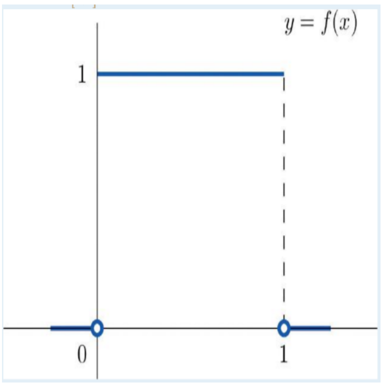
Remark 0.3.2. Every density function f (x) must satisfy the following two conditions:

1. For all numbers x, $f(x) \geq 0$, so that the graph of $y = f(x)$ never drops below the x-axis.
2. The area of the region under the graph of $y = f(x)$ and above the x-axis is 1.

For any continuous random variable X:

$$\Pr(a \leq X \leq b) = \Pr(a < X \leq b) = \Pr(a \leq X < b) = \Pr(a < X < b)$$

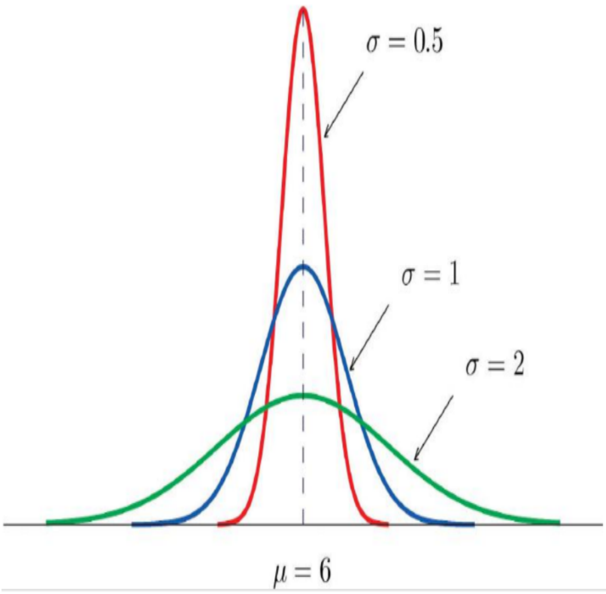
Example 0.3.3. A random variable X has the uniform distribution on the interval $[0, 1]$: the density function is $f(x) = 1$ if x is between 0 and 1 and $f(x) = 0$ for all other values of x, as shown in following



1. Find $\Pr(X > 0.75)$, the probability that X assumes a value greater than 0.75.
2. Find $\Pr(X \leq 0.2)$, the probability that X assumes a value less than or equal to 0.2.
3. Find $\Pr(0.4 < X < 0.7)$, the probability that X assumes a value between 0.4 and 0.7.

0.3.1 Normal Distributions

Definition 0.3.4. The probability distribution corresponding to the density function for the bell curve with parameters μ and σ is called the normal distribution with mean μ and standard deviation σ.

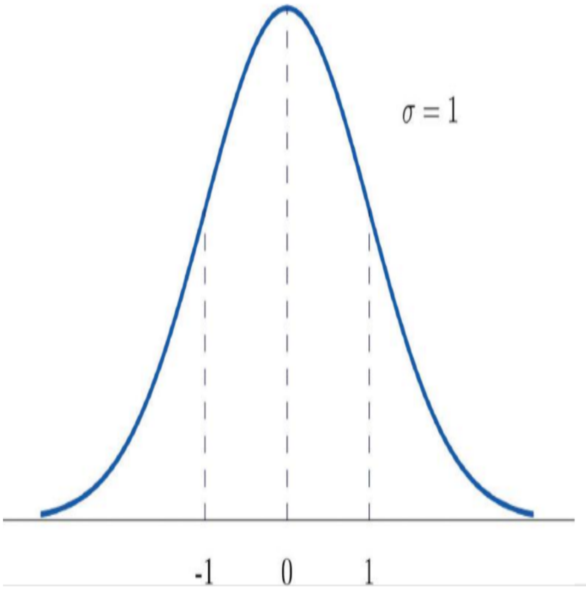


Definition 0.3.5. A continuous random variable whose probabilities are described by the normal distribution with mean μ and standard deviation σ is called a normally distributed random variable, or a normal random variable for short, with mean μ and standard deviation σ.

Example 0.3.6. Heights of 25-year-old men in a certain region have mean 69.75 inches and standard deviation 2.59 inches. These heights are approximately normally distributed. Thus the height X of a randomly selected 25-year-old man is a normal random variable with mean $\mu = 69.75$ and standard deviation $\sigma = 2.59$. Sketch a qualitatively accurate graph of the density function for X. Find the probability that a randomly selected 25-year-old man is more than 69.75 inches tall.

0.3.2 The Standard Normal Distribution

Definition 0.3.7. A standard normal random variable is a normally distributed random variable with mean $\mu = 0$ and standard deviation $\sigma = 1$. It will always be denoted by the letter Z.



Example 0.3.8. Find the probabilities indicated, where as always Z denotes a standard normal random variable.

a. $\Pr(Z < 1.48)$.

b. $\Pr(Z < -0.25)$.

c. $\Pr(Z > 1.60)$.

d. $\Pr(Z > -1.02)$.

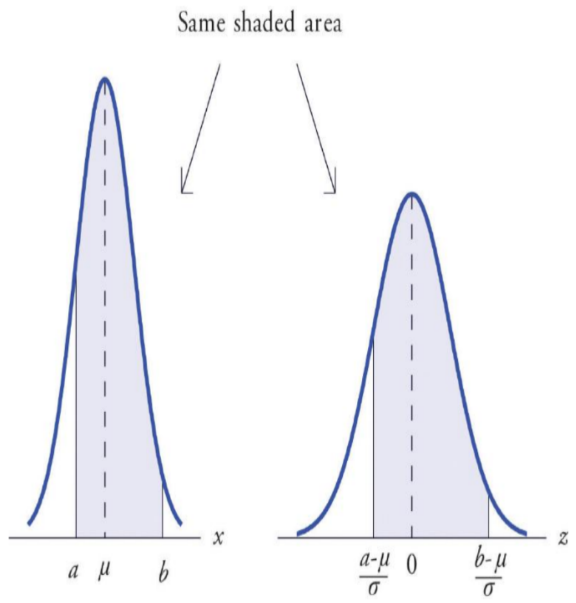
e. $\Pr(0.5 < Z < 1.57)$

0.3.3 Probability Computations for General Normal Random Variables

Definition 0.3.9. If X is a normally distributed random variable with mean μ and standard deviation σ, then

$$\Pr(a < X < b) = \Pr(\frac{a-\mu}{\sigma} < Z < \frac{a-\mu}{\sigma})$$

where Z denotes a standard normal random variable. a can be any decimal number or $-\infty$; b can be any decimal number or ∞.



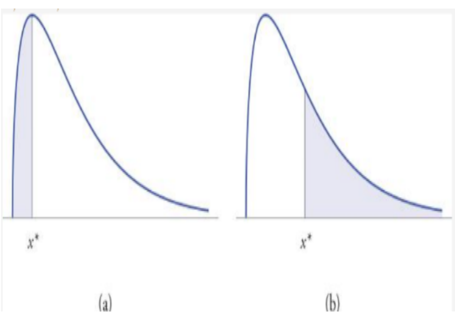
Example 0.3.10. Let X be a normal random variable with mean $\mu = 10$ and standard deviation $\sigma = 2.5$. Compute the following probabilities.

a. $\Pr(X < 14)$

b. $\Pr(8 < X < 14)$.

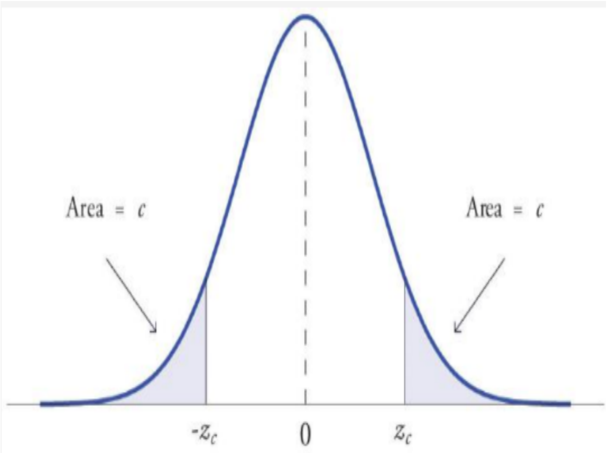
0.3.4 Areas of Tails of Distributions

Definition 0.3.11. The left tail of a density curve $y = f(x)$ of a continuous random variable X cut off by a value x^* of X is the region under the curve that is to the left of x^*, as shown by the shading in Figure "Right and Left Tails of a Distribution"(a). The right tail cut off by $x^*$is defined similarly, as indicated by the shading in Figure "Right and Left Tails of a Distribution"(b).



Example 0.3.12. Find the value z^* of Z as determined by z-table the value z^* that cuts off a left tail of area 0.0125 in the standard normal distribution. In symbols, find the number z^* such that $P(Z < z^*) = 0.0125$.

Definition 0.3.13. The value of the standard normal random variable Z that cuts off a right tail of area c is denoted z_c. By symmetry, value of Z that cuts off the left tail of area c is $-z_c$.



Example 0.3.14. Find $z_{0.1}$ and $-z_{.01}$, the values of Z that cut off right and left tails of area 0.01 in the standard normal distribution.