Marmara University, **2021**

Probability and Statistics

Subject 4
Probability and Probability Distributions

Mujdat Soyturk, Ph.D.
Associate Professor

Aim

- You have learned to describe a data set,
  - How can you use sample data to draw conclusions about the sampled populations?
  - The technique involves a statistical tool called *probability*.
- The variables that we measured can now be redefined as *random variables*, whose values depend on the chance selection of the elements in the sample.
- Using the probability as a tool, you can create *probability distributions* that serve as models for discrete random variables.

Contents

- Events and the Sample Space
- Calculating Probabilities
- Counting Rules
- Event Relations and Probability Rules
- Independence and Conditional Probability
- Baye's Rule
- Discrete Random Variables

Most parts of the slides are derived from the textbook: "Mendenhall, Beaver, Beaver, *Introduction to Probability and Statistics*, 14th Ed., Brooks/Cole, Cengage Learning, 2013"

What is Probability?

- In Subject 2 and 3, we used graphs and numerical measures to describe data sets which were usually **samples.**
- We measured "how often" using

Relative frequency = *f/n*

Sample → **Population**

"How often" = Relative frequency → **Probability**

What is Probability?

- When the population is ***known***, probability is used to describe the likelihood of observing a particular sample outcome.
  - (subject of next 2-3 weeks)
- When the population is ***unknown*** and only a sample from that population is available, probability is used in making statements about the makeup of the population, that is, in making statistical inferences.
  - (subject of last 5 weeks)

Basic Concepts

- An **experiment** is the process by which an observation (or measurement) is obtained.
- **Experiment: Record an age**
- **Experiment: Toss a die and observe the result**
- **Experiment: Record an opinion (yes, no)**
- **Experiment: Toss two coins and observe the results**

Basic Concepts

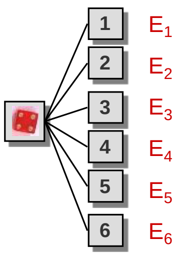
- A **simple event** is the outcome that is observed on a single repetition of the experiment.
  - The basic element to which probability is applied.
  - One and only one simple event can occur when the experiment is performed.
- A **simple event** is denoted by *E* with a subscript.

Basic Concepts

- Each simple event will be assigned a probability, measuring “how often” it occurs.
- The set of all simple events of an experiment is called the **sample space, S.**

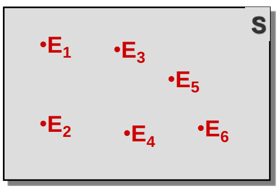
Example

- **The die toss:**
- Simple events:



Sample space:

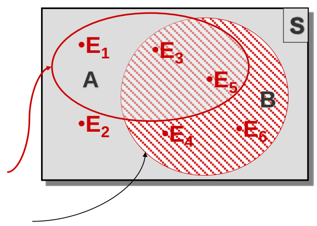
$S = \{E_1, E_2, E_3, E_4, E_5, E_6\}$



Basic Concepts

- An **event** is a collection of one or more **simple events.**

•**The die toss:**
- A: an odd number
- B: a number > 2



$A = \{E_1, E_3, E_5\}$

$B = \{E_3, E_4, E_5, E_6\}$

Basic Concepts

- Two events are **mutually exclusive** if, when one event occurs, the other cannot, and vice versa.

•**Experiment: Toss a die**

- A: observe an odd number
- B: observe a number greater than 2

Not Mutually Exclusive

- C: observe a 6
- D: observe a 3

Mutually Exclusive

B and C?
B and D?

The Probability of an Event

- The probability of an event A measures “how often” we think ***A*** will occur. We write ***P(A)***.
- Suppose that an experiment is performed *n* times. The relative frequency for an event A is

$$\frac{\text{Number of times A occurs}}{n} = \frac{f}{n}$$

•If we let *n* get infinitely large,

$$P(A) = \lim_{n\to\infty} \frac{f}{n}$$

The Probability of an Event

- *P(A)* must be between 0 and 1.
  - If event *A* can never occur, *P(A)* = 0. If event *A* always occurs when the experiment is performed, *P(A)* =1.
- The sum of the probabilities for all simple events in S equals 1.

•The **probability of an event *A*** is found by adding the probabilities of all the simple events contained in ***A***.

Finding Probabilities

- Probabilities can be found using
  - Estimates from empirical studies
  - Common sense estimates based on equally likely events.

•**Examples:**

–Toss a fair coin. P(Head) = 1/2

–10% of the U.S. population has red hair. Select a person at random. P(Red hair) = .10

Finding Probabilities

How to Calculate the Probability of an Event

1. List all the simple events in the sample space.
2. Assign an appropriate probability to each simple event.
3. Determine which simple events result in the event of interest.
4. Sum the probabilities of the simple events that result in the event of interest.

Example

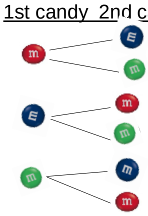
- Toss a fair coin twice. What is the probability of observing at least one head?

1st Coin	2nd Coin	E_i	$P(E_i)$
H	H	HH	1/4
	T	HT	1/4
T	H	TH	1/4
	T	TT	1/4

P(at least 1 head)
$= P(E_1) + P(E_2) + P(E_3)$
= 1/4 + 1/4 + 1/4 = 3/4

Example

- A bowl contains three candies, one red, one blue and one green. A child selects two candies at random. What is the probability that at least one is red?



1st candy 2nd candy	E_i	$P(E_i)$
	RB	1/6
	RG	1/6
	BR	1/6
	BG	1/6
	GB	1/6
	GR	1/6

P(at least 1 red)
= P(RB) + P(BR)+ P(RG) + P(GR)
= 4/6 = 2/3

Counting Rules

- If the simple events in an experiment are **equally likely,** you can calculate

$$P(A) = \frac{n_A}{N} = \frac{\text{number of simple events in A}}{\text{total number of simple events}}$$

- You can use **counting rules** to find n_A and N.

The *mn* Rule

- If an experiment is performed in two stages, with ***m*** ways to accomplish the first stage and ***n*** ways to accomplish the second stage, then there are ***mn*** ways to accomplish the experiment.
- This rule is easily extended to ***k*** stages, with the number of ways equal to

$$n_1 n_2 n_3 \dots n_k$$

Example: Toss two coins. The total number of simple events is:

$2 \times 2 = 4$

Examples

Example: Toss three coins. The total number of simple events is: $2 \times 2 \times 2 = 8$

Example: Toss two dice. The total number of simple events is: $6 \times 6 = 36$

Example: Two candies are drawn from a dish containing two red and two blue candies. The total number of simple events is: $4 \times 3 = 12$

Permutations

- The number of ways you can arrange ***n*** distinct objects, taking them ***r*** at a time is

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

where $n! = n(n-1)(n-2)...(2)(1)$ and $0! \equiv 1$.

Example: How many 3-digit lock combinations can we make from the numbers 1, 2, 3, and 4?

The order of the choice is important!

$$P_3^4 = \frac{4!}{1!} = 4(3)(2) = 24$$

Examples

Example: A lock consists of five parts and can be assembled in any order. A quality control engineer wants to test each order for efficiency of assembly. How many orders are there?

The order of the choice is important!

$$P_5^5 = \frac{5!}{0!} = 5(4)(3)(2)(1) = 120$$

Combinations

- Sometimes the ordering or arrangements of the objects is not important, but only the objects that are chosen.
- In this case, you can use a counting rule for **combinations.**

Combinations

- The number of distinct combinations of ***n*** distinct objects that can be formed, taking them ***r*** at a time is

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

Example: Three members of a 5-person committee must be chosen to form a subcommittee. How many different subcommittees could be formed?

The order of the choice is not important!

$$C_3^5 = \frac{5!}{3!(5-3)!} = \frac{5(4)(3)(2)1}{3(2)(1)(2)1} = \frac{5(4)}{(2)1} = 10$$

Example

- A box contains six candies, four red and two green. A child selects two candies at random. What is the probability that exactly one is red?

The order of the choice is not important!

$C_2^6 = \frac{6!}{2!4!} = \frac{6(5)}{2(1)} = 15$ ways to choose 2 candies.

$C_1^2 = \frac{2!}{1!1!} = 2$ ways to choose 1 green candy.

$C_1^4 = \frac{4!}{1!3!} = 4$ ways to choose 1 red candy.

4 × 2 =8 ways to choose 1 red and 1 green candy.

P(exactly one red) = 8/15

Ex 4.15, pp 136

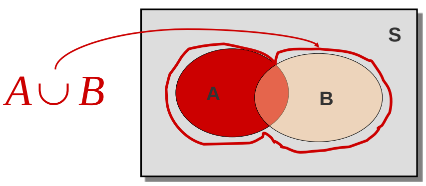
Event Relations

- Sometimes the event of interest can be formed as a combination of several other events.
- Then, we have to look at the relationships between events.

Event Relations

- The **union** of two events, A and B, is the event that either A **or** B **or both** occur when the experiment is performed. We write

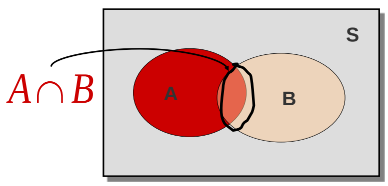
A ∪ B



Event Relations

- The **intersection** of two events, **A** and **B**, is the event that both A **and** B occur when the experiment is performed.

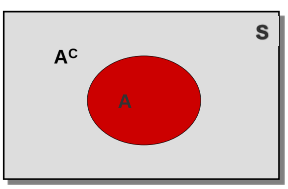
We write **A ∩ B.**



- If two events A and B are **mutually exclusive,** then P(**A ∩ B) = 0.**

Event Relations

- The **complement** of an event **A** consists of all outcomes of the experiment that do not result in event A. We write **A^C.**



Example

- Select a student from the classroom and record his/her **hair color** and **gender.**
  - **A:** student has brown hair
  - **B:** student is female
  - **C:** student is male

Mutually exclusive; B = C^C

- What is the relationship between events **B** and **C**?
- **A^C:** Student does not have brown hair
- **B∩C:** Student is both male and female = ∅
- **B∪C:** Student is either male and female = all students = S

Ex 4.16, pp 140

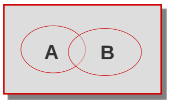
Calculating Probabilities for Unions and Complements

- There are special rules that will allow you to calculate probabilities for composite events.

The Additive Rule for Unions:

- For any two events, **A** and **B**, the probability of their union, **P(A ∪ B)**, is

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



Example: Additive Rule

Example: Suppose that there were 120 students in the classroom, and that they could be classified as follows:

A: brown hair
P(A) = 50/120
B: female
P(B) = 60/120

	Brown	Not Brown
Male	20	40
Female	30	30

P(A∪B) = P(A) + P(B) – P(A∩B)
= 50/120 + 60/120 - 30/120
= 80/120 = 2/3

Check: P(A∪B)
= (20 + 30 + 30)/120

A Special Case

When two events A and B are **mutually exclusive,**
P(A∩B) = 0
and **P(A∪B) = P(A) + P(B).**

A: male with brown hair
P(A) = 20/120
B: female with brown hair
P(B) = 30/120

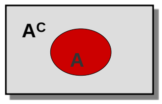
	Brown	Not Brown
Male	20	40
Female	30	30

A and B are mutually exclusive, so that

P(A∪B) = P(A) + P(B)
= 20/120 + 30/120
= 50/120

Calculating Probabilities for Complements

- We know that for any event **A:**
  - **P(A ∩ A^C) = 0**
- Since either **A** or **A^C** must occur,
  P(A ∪ A^C) =1
- so that **P(A ∪ A^C) = P(A)+ P(A^C) = 1**



P(A^C) = 1 – P(A)

Example

Select a student at random from the classroom.

Define:

A: male
P(A) = 60/120
B: female

	Brown	Not Brown
Male	20	40
Female	30	30

A and B are complementary, so that

P(B) = 1- P(A)
= 1- 60/120 = 60/120

Calculating Probabilities for Intersections

- In the previous example, we found P(A ∩ B) directly from the table. Sometimes this is impractical or impossible. The rule for calculating P(A ∩ B) depends on the idea of **independent and dependent events.**

Two events, **A** and **B**, are said to be **independent** if and only if the probability that event **A** occurs does not change, depending on whether or not event **B** has occurred.

Conditional Probabilities

- The probability that ***A*** occurs, given that event ***B*** has occurred is called the **conditional probability** of ***A*** given ***B*** and is defined as

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

"given"

Example 1

- Toss a fair coin twice. Define
  - **A**: head on second toss
  - **B**: head on first toss

HH	1/4
HT	1/4
TH	1/4
TT	1/4

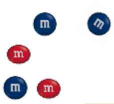
P(A|B) = ½

P(A|not B) = ½

P(A) does not change, whether B happens or not… → A and B are independent!

Example 2

- A bowl contains five candies, two red and three blue. Randomly select two candies, and define
  - **A**: first candy is blue.
  - **B**: second candy is red.



P(B|A) =P(2nd red|1st blue)= 2/4 = 1/2

P(B|not A) = P(2nd red|1st red) = 1/4

P(B) does change, depending on whether A happens or not… → A and B are dependent!

Defining Independence

- We can redefine independence in terms of conditional probabilities:

Two events A and B are **independent** if and only if

P(A|B) = P(A) or **P(B|A) = P(B)**

Otherwise, they are **dependent**.

- Once you've decided whether or not two events are independent, you can use the following rule to calculate their intersection.

The Multiplicative Rule for Intersections

- For any two events, A and B, the probability that both A and B occur is

P(A ∩ B) = P(A) P(B given that A occurred)
= P(A)P(B|A)

- If the events **A** and **B** are independent, then the probability that both **A** and **B** occur is

P(A ∩ B) = P(A) P(B)

Example 1

In a certain population, 10% of the people can be classified as being high risk for a heart attack. Three people are randomly selected from this population. What is the probability that **exactly** one of the three are high risk?

Define ***H***: high risk ***N***: not high risk

P(exactly one high risk) = P(HNN) + P(NHN) + P(NNH)

= P(H)P(N)P(N) + P(N)P(H)P(N) + P(N)P(N)P(H)

$= (.1)(.9)(.9) + (.9)(.1)(.9) + (.9)(.9)(.1)= 3(.1)(.9)^2 = .243$

Example 2

Suppose we have additional information in the previous example. We know that only 49% of the population are female. Also, of the female patients, 8% are high risk. A single person is selected at random. What is the probability that it is a high risk female?

Define ***H***: high risk ***N***: not high risk

From the example, P(F) = .49 and P(H|F) = 0.08. Use the Multiplicative Rule:

P(high risk female) = P(H∩F)

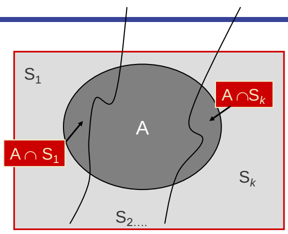
= P(F)P(H|F) =.49(.08) = .0392

The Law of Total Probability

- Let S_1 , S_2 , S_3 ,..., S_k be mutually exclusive and exhaustive events (that is, one and only one must happen). Then the probability of another event A can be written as

P(A) = P(A ∩ S_1) + P(A ∩ S_2) + … + P(A ∩ S_k)
= P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + … + P(S_k)P(A|S_k)

The Law of Total Probability



P(A) = P(A ∩ S_1) + P(A ∩ S_2) + … + P(A ∩ S_k)
= P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + … + P(S_k)P(A|S_k)

Bayes' Rule

- Let S_1 , S_2 , S_3 ,..., S_k be mutually exclusive and exhaustive events with prior probabilities P(S_1), P(S_2),…,P(S_k).
- If an event A occurs, the posterior probability of S_i, given that A occurred is

$$P(S_i \mid A) = \frac{P(S_i)P(A \mid S_i)}{\sum P(S_i)P(A \mid S_i)} \text{ for } i \ = \ 1, 2,...k$$

Example

From a previous example, we know that 49% of the population are female. Of the female patients, 8% are high risk for heart attack, while 12% of the male patients are high risk. A single person is selected at random and found to be high risk. What is the probability that it is a male?

Define H: high risk F: female M: male

We know:	
P(F) =	.49
P(M) =	.51
P(H|F) =	.08
P(H|M) =	.12

$$P(M \mid H) = \frac{P(M)P(H \mid M)}{P(M)P(H \mid M) + P(F)P(H \mid F)}$$

$$= \frac{.51(.12)}{.51(.12) + .49(.08)} = .61$$

Random Variables

- A quantitative variable *x* is a **random variable** if the value that it assumes, corresponding to the outcome of an experiment is a chance or random event.
- Random variables can be **discrete** or **continuous.**
- **Examples:**
  - *x* = SAT score for a randomly selected student
  - *x* = number of people in a room at a randomly selected time of day
  - *x* = number on the upper face of a randomly tossed die

Probability Distributions for Discrete Random Variables

- The **probability distribution** *for a random variable x* is the ***relative frequency distribution*** constructed for the entire population of measurements. It is a ***graph, table or formula*** that gives the possible values of ***x*** and the probability ***p(x)*** associated with each value.

We must have

$$0 \le p(x) \le 1 \quad \text{and} \quad \sum p(x) = 1$$

Example

- Toss a fair coin three times and define ***x* = number of heads.**

		x
HHH	1/8	3
HHT	1/8	2
HTH	1/8	2
THH	1/8	2
HTT	1/8	1
THT	1/8	1
TTH	1/8	1
TTT	1/8	0

P(*x* = 0) = 1/8
P(*x* = 1) = 3/8
P(*x* = 2) = 3/8
P(*x* = 3) = 1/8

Probability distribution for ***x***

x	*p(x)*
0	1/8
1	3/8
2	3/8
3	1/8

Probability Histogram for *x*

Probability Distributions

- Probability distributions can be used to describe the population, just as we described samples in Ch.1.
  - **Shape:** Symmetric, skewed, mound-shaped…
  - **Outliers:** unusual or unlikely measurements
  - **Center and spread:** mean and standard deviation. A population mean is called μ and a population standard deviation is called σ.

The Mean and Standard Deviation

- Let *x* be a discrete random variable with probability distribution *p(x)*. Then the mean, variance and standard deviation of *x* are given as

$$\text{Mean}: \mu = \sum x p(x)$$

Expected Value

$$\text{Variance}: \sigma^2 = \sum (x-\mu)^2 p(x)$$

$$\text{Standard deviation}: \sigma = \sqrt{\sigma^2}$$

The Mean and Standard Deviation

- The ***population mean***, which measures the average value of ***x*** in the population, is also called the ***expected value*** of the random variable ***x*** is written as ***E(x)***.

- It is the value that you would ***expect*** to observe on ***average*** *if the experiment is repeated over and over again*.

Example

- Toss a fair coin 3 times and record *x* the number of heads.

x	*p(x)*	*xp(x)*	$(x-\mu)^2 p(x)$
0	1/8	0	$(-1.5)^2(1/8)$
1	3/8	3/8	$(-0.5)^2(3/8)$
2	3/8	6/8	$(0.5)^2(3/8)$
3	1/8	3/8	$(1.5)^2(1/8)$

$$\mu = \sum xp(x) = \frac{12}{8} = 1.5$$

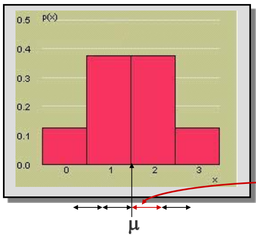
$$\sigma^2 = \sum (x-\mu)^2 p(x)$$

$$\sigma^2 = .28125 + .09375 + .09375 + .28125 = .75$$

$$\sigma = \sqrt{.75} = .688$$

Example

- The probability distribution for *x* the number of heads in tossing 3 fair coins.



- Shape? Symmetric; mound-shaped
- Outliers? None
- Center? $\mu = 1.5$
- Spread? $\sigma = .688$

Key Concepts

I. Experiments and the Sample Space

1. Experiments, events, mutually exclusive events, simple events
2. The sample space
3. Venn diagrams, tree diagrams, probability tables

II. Probabilities

1. Relative frequency definition of probability
2. Properties of probabilities
   a. Each probability lies between 0 and 1.
   b. Sum of all simple-event probabilities equals 1.
3. *P(A)*, the sum of the probabilities for all simple events in *A*

Key Concepts

III. Counting Rules

1. *mn* Rule; extended *mn* Rule
2. Permutations: $P_r^n = \dfrac{n!}{(n-r)!}$
3. Combinations: $C_r^n = \dfrac{n!}{r!(n-r)!}$

IV. Event Relations

1. Unions and intersections
2. Events
   a. Disjoint or mutually exclusive: $P(A \cap B) = 0$
   b. Complementary: $P(A) = 1 - P(A^C)$

Key Concepts

3. Conditional probability: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$
4. Independent and dependent events
5. Additive Rule of Probability:
   $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
6. Multiplicative Rule of Probability:
   $P(A \cap B) = P(A)P(B|A)$
7. Law of Total Probability
8. Bayes' Rule

Key Concepts

V. Discrete Random Variables and Probability Distributions

1. Random variables, discrete and continuous
2. Properties of probability distributions
   $0 \le p(x) \le 1 \text{ and } \sum p(x) = 1$
3. Mean or expected value of a discrete random variable:
   $\text{Mean}: \mu = \sum x p(x)$
4. Variance and standard deviation of a discrete random variable:
   $\text{Variance}: \sigma^2 = \sum (x-\mu)^2 p(x)$
   $\text{Standard deviation}: \sigma = \sqrt{\sigma^2}$