

2. Probability

Probability'

- To provide a model for the repetition of an experiment
- To provide a model for the population frequency distributions

Definitions and Concepts

Probability: The measure of one's belief in the occurrence of a future event.

Random or Stochastic Event: The event that can not be predicted with certainty but the relative frequency with which they occur in a long series of trial remains stable.

Set: If the elements in the set A are a_1, a_2, a_3 , we write set

$$A = \{a_1, a_2, a_3\}.$$

Universal Set: S is the universal set that contains all elements under consideration.

Subset: For any two sets A and B we will say that A is subset of B or A is contained in B , denoted by $A \subset B$, if every point in A is also in B .

Null or Empty Set: A set consisting of no points and denoted by ϕ . Thus, ϕ is a subset of every set.

Venn Diagrams: Portray the sets and relationships between sets.

Union of Sets: The union of sets A and B denoted by $A \cup B$ is the set of all points in A or B or *both*. That is, the union of A and B contains all points that are in at least one of the sets. key word for expressing the union of two sets is “*or*” (meaning A or B or *both*).

Intersection of Sets: The intersection of A and B denoted by $A \cap B$ or by AB is the set of all points in both A and B . That is the intersection of A and B consists of the points where the two sets overlap. The key word for expressing intersections is “*and*” (meaning A and B *simultaneously*).

Complement of a Set: If A is a subset of S then the complement of A denoted by \bar{A} is the set of points that are in S but not in A . Note that $A \cup \bar{A} = S$.

Disjoint or Mutually Exclusive Sets: Two sets A and B are disjoint or mutually exclusive if they have no points in common, i.e., $A \cap B = \phi$.

Note: A and \bar{A} are mutually exclusive.

Exercise 2.1: Suppose a family contains two children and we are interested in the gender of these children.

Let F =female, M =male and FM denote that the older child is female and the younger male. Then the universal set

$$S = \{FF, FM, MF, MM\}$$

Let A be the subset containing no males, B containing two males, and C containing at least one male.

Then, the elements of:

$$A = \{FF\}, B = \{MM\}, C = \{MF, FM, MM\},$$

$$A \cap B = \phi, A \cup B = \{FF, MM\}, A \cap C = \phi,$$

$$A \cup C = \{FF, FM, MF, MM\} = S, B \cap C = \{MM\}$$

$$B \cup C = \{MF, FM, MM\} = C, C \cap \bar{B} = \{MF, FM\}$$

A probabilistic Model for an Experiment: The Discrete Case

Experiment: The process of making an observation.

The term experiment will be used to include observations from completely uncontrollable situations as well as those made under controlled laboratory conditions.

Events: One or more outcomes on an experiment denoted by capital letters.

Example 1 : Consider an experiment of single toss of a balanced die and some events associated with it:

A : Observe an odd number

B : Observe a number less than 5

C : Observe a 2 or a 3

E_1 : Observe a 1

E_2 : Observe a 2

E_3 : Observe a 3

E_4 : Observe a 4

E_5 : Observe a 5

E_6 : Observe a 6

Example 2: Suppose the experiment consists of counting the number of bacteria in a portion of food. Some of the events could be:

A : Exactly 110 bacteria

B : More than 200 bacteria

C : The number of bacteria present is between 100 and 300.

Simple Event: Corresponds to one and only one sample point. A simple event or corresponding sample point will be denoted by E .

Example 1: E_1, \dots, E_6 are simple events.

Compound Event: An event which can be decomposed into other events.

Example 1: $A, B,$ and C are compound events.

Compound events can be viewed as collections (sets) of sample points or as unions of the sets of single sample points corresponding to the appropriate simple events.

$$\begin{aligned}\text{Example 1: } A &= \{E_1, E_3, E_5\} \\ &= E_1 \cup E_3 \cup E_5.\end{aligned}$$

Sample Space: The set consisting of all possible sample points, denoted by S .

Discrete Sample Space: The sample space that contains either a finite or a countable number of distinct sample points.

$$\text{Example 1: } S = \{E_1, \dots, E_6\}.$$

$$\text{Example 2: } S = \{E_0, E_1, E_2, E_3, \dots\}.$$

Relative Frequency definition of Probability: Suppose that an experiment has generated the sample space S . To every event A in S , we assign a number $P(A)$ called the *PROBABILITY* of A so that the following axioms hold:

$$\text{Axiom 1: } P(A) \geq 0.$$

$$\text{Axiom 2: } P(S) = 1.$$

Axiom 3: If A_1, A_2, A_3, \dots form a sequence of pairwise mutually exclusive events in S , then

$$P(A_1, A_2, A_3, \dots) = \sum_{i=1}^{\infty} P(A_i).$$

Exercise 2.8: A vehicle arriving at an intersection can

turn right (R), turn left (L) or continue straight ahead (S). The experiment consists of observing the movement of a single vehicle through the intersection.

a. Sample Space = $\{L, R, S\}$.

b. $P(\text{vehicle turns}) = P(L \text{ or } R) = P(L \cup R)$
 $= P(L) + P(R) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$

Exercise 2.10: Volunteers coming into a blood center are such that 1 in 3 have O^+ blood, 1 in 15 have O^- , 1 in 3 have A^+ , and 1 in 16 have A^- . The name of a donor is selected from the records.

a. $P(O^+) = \frac{1}{3}$

b. $P(O) = P(O^+ \text{ or } O^-) = P(O^+ + O^-)$
 $= \frac{1}{3} + \frac{1}{15} = \frac{6}{15} = \frac{2}{5}$

c. $P(A) = P(A^+ \text{ or } A^-) = P(A^+ + A^-)$
 $= \frac{1}{3} + \frac{1}{16} = \frac{19}{48}$

d. $P(A \cup B)^c = 1 - P(A \cup B)$
 $= 1 - \left(\frac{2}{5} + \frac{19}{48}\right) = \frac{49}{240}$

Calculating the Probability:

The sample Point Method

- Define the experiment.
- List the simple events and define the sample space S .
- Assign probabilities to the sample points such that

$$P(E_i) \geq 0 \text{ and } \sum P(E_i) = 1.$$

- Define the event of interest A as a specific collection of sample points.
- Find $P(A)$ by summing the probabilities of the sample points in A .

Note: The complication occurs while incorrectly diagnosing the nature of a simple event and failing to list all of the sample points in S .

Exercise 2.16: Four equally qualified people apply for two identical positions. One and only one applicant is from a minority group. The positions are filled by choosing two of the applicants at random.

Denote four candidates by

A_1, A_2, A_3 and M (minority).

a-b. Since order is unimportant in selection, there are six

possible outcomes and we assign $\frac{1}{6}$ probability to each outcome:

$$\begin{array}{lll} E_1 : (A_1, A_2) & E_2 : (A_1, A_3) & E_3 : (A_1, M) \\ E_4 : (A_2, A_3) & E_5 : (A_2, M) & E_6 : (A_3, M) \end{array}$$

c. $P(M) = P(E_3) + P(E_5) + P(E_6)$
$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}.$$

Exercise 2.19: The Bureau of census reports that half of all families had incomes exceeding \$35,353 in 1991 and half had incomes equal to or below this amount. Suppose four families are surveyed and that each reveals whether their income exceeded \$35,353 in 1991

Define the events:

E=income exceeding \$35,353

N=income not exceeding \$35,353

a. The sample space:

$E_1 : (E, E, E, E)$	$E_7 : (E, N, E, N)$	$E_{13} : (N, E, N, E)$
$E_2 : (E, E, E, N)$	$E_8 : (N, E, E, N)$	$E_{14} : (N, N, E, E)$
$E_3 : (E, E, N, E)$	$E_9 : (E, N, N, E)$	$E_{15} : (N, N, N, E)$
$E_4 : (E, N, E, E)$	$E_{10} : (N, E, N, E)$	$E_{16} : (N, N, N, N)$
$E_5 : (N, E, E, E)$	$E_{11} : (N, N, E, E)$	
$E_6 : (E, E, N, N)$	$E_{12} : (E, N, N, N)$	

b. $A : \{E_1, E_2, \dots, E_{11}\}$

$B : \{E_6, E_7, \dots, E_{11}\}$

$C : \{E_2, E_3, E_4, E_5\}$

c. Since $P(E) = P(N)$, each simple event is equally likely and thus $P(E_i) = \frac{1}{16}$.

$$P(A) = \frac{11}{16} \quad P(B) = \frac{6}{16} \quad P(C) = \frac{4}{16} = \frac{1}{4}.$$

Counting Rules

The mn Rule

With m elements a_1, a_2, \dots, a_m and n elements b_1, b_2, \dots, b_n , it is possible to form mn pairs containing one element from each group.

This mn rule can be extended to any number of sets.

Exercise 2.34: An experimenter intends to investigate the effects of three settings each for temperature and pressure and two types of catalysts on the yield in a refining process. How many experimental runs have to be conducted if he wishes to run all possible combinations.

The mn rule is used. the total number of experiments will be $= (3)(3)(2) = 18$.

Permutation:

An ordered arrangement of r distinct objects is called a *permutation*.

The number of ways of ordering n distinct objects taken r at a time:

$$P_r^n = n(n - 1)(n - 2)\dots(n - r + 1) = \frac{n!}{(n - r)!}$$

where $n! = n(n - 1)(n - 2)\dots(2).(1)$

and $0! = 1$.

Exercise 2.26: A personnel director has hired 10 new engineers. If three (distinctly different) positions are open at a new plant, in how many ways can she fill the

positions?

Since the order of selection is important as well as the different combinations of engineers that may be selected, the number of ways in which the positions can be filled

$$P_3^{10} = \frac{10!}{(10-3)!} = \frac{10 \cdot 9 \cdot 8 \cdot 7!}{7!} = 10 \cdot 9 \cdot 8 = 720.$$

Multinomial Distribution

The number of ways of dividing n distinct objects into k groups containing n_1, n_2, \dots, n_k objects is:

$$N = \frac{n!}{n_1! n_2! \dots n_k!}, \text{ where } n = \sum n_i.$$

Exercise 2.38. A manufacturer has 9 distinct motors in stock, 2 of which came from a particular supplier. The motors must be randomly divided among 3 production lines with 3 motors going to each line.

a. The total number of ways to divide 9 motors into 3 groups of size 3 = $\frac{9!}{3!3!3!}$.

b. If both motors from a particular supplier are assigned

to the first line, there are 7 motors left to be assigned such that 1 to line 1 and 3 each to lines 2 and 3. This can be done in $= \frac{7!}{1!3!3!}$ ways.

c. Hence the probability that both motors from the particular supplier are assigned to line 1 =

$$\frac{\frac{7!}{1!3!3!}}{\frac{9!}{3!3!3!}} = \frac{1}{12}.$$

Combinations

In many situations the sample points are identified by an array of symbols in which the order of symbols is unimportant.

The number of combinations of n objects taken r at a time :

$$C_r^n \text{ or } \binom{n}{r} = \frac{n!}{(n-r)!}.$$

The terms $\binom{n}{r}$ are referred to as the binomial coefficients because they occur in the binomial expansion $(x + y)^n$.

Exercise 2.33. A local fraternity is conducting a raffle where 50 tickets (1 per person) are to be sold. There are 3 prizes. The 4 organizers of the raffle also buy the

tickets.

The total number of ways to choose 3 winners = $\binom{50}{3} = 19,600$. Since the choice is at random, each sample point is equally likely.

a. The number of ways for the 4 organizers to win all 3 prizes = $\binom{4}{3} = 4$

and hence the probability = $\frac{4}{19,600}$.

b. The organizers can win exactly 2 of the prizes if 1 of the other 46 people wins 1 prize. The number of ways this to occur = $\binom{4}{2} \binom{46}{1} = 276$. hence the desired probability = $\frac{276}{19,600}$.

c. The 4 organizers win exactly 1 prize

$$= \frac{\binom{4}{1} \binom{46}{2}}{19,600} = \frac{4140}{19,600}.$$

d. The 4 organizers win none of the prizes

$$= \frac{\binom{4}{0} \binom{46}{3}}{19,600} = \frac{15180}{19,600}.$$

Exercise 2.31. A student can select from 130 major areas of study at a University. In the registrar's records, a student's major is identified with a two or three letter

code. Some students opt for a double major. The registrar was asked to consider assigning these double majors a distinct two or three letter code so that they could be identified through the student records system.

a. The maximum number of possible double majors

$$\text{available} = \binom{130}{2} = 8385$$

b. If any two- or three-letter codes are available to identify majors or double majors, the number of major codes available:

$$\text{Two-letter codes} = 26 \cdot 26 = 676$$

$$\text{Three-letter codes} = 26 \cdot 26 \cdot 26 = 17,576$$

$$\text{Total codes} = 18,252.$$

c. The number of major codes required to identify students who have either a single major or a double major = $8385 + 130 = 8515$.

Conditional Probability

The conditional probability of an event is the probability of the event given that one or more events have already occurred.

The *conditional* probability of an event A given that event B has occurred is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) > 0.$$

Independence of Events

If the occurrence of an event A is *unaffected* by the occurrence or nonoccurrence of event B , we say that the event A is *independent* of B .

Two events A and B are said to be independent if any one of the following holds:

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

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

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

Otherwise, the events are said to be *dependent*.

The Multiplicative Law of Probability

The probability of the *intersection* of two events is

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B).$$

If A and B are independent, then

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

Note: The multiplicative law can be extended to find the

probability of any number of events. For example:

$$\begin{aligned}P(A \cap B \cap C) &= P[(A \cap B) \cap C] \\ &= P(A \cap B)P(C|A \cap B)] \\ &= P(A)P(B|A)P(C|A \cap B).\end{aligned}$$

The Additive law of Probability

The probability of the *union* of two events A and B is:

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

If A and B are independent, then

$$P(A \cup B) = P(A) + P(B).$$

This rule can be used to find

$$\begin{aligned}P(A \cup B \cup C) &= P[A \cup (B \cup C)] \\ &= P(A) + P(B \cup C) - P[A \cap (B \cup C)] \\ &= P(A) + P(B) + P(C) - P(A \cap B) \\ &\quad - P(A \cap C) - P(B \cup C) + P(A \cap B \cap C)\end{aligned}$$

Note: Relationship between $P(\bar{A})$ and $P(A)$:

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

The law of Total Probability

Assume that

$$S = B_1 \cup B_2 \cup \dots \cup B_k, P(B_i) > 0,$$

and $B_i \cap B_j = \phi$ for $i \neq j$. Then for any event A

$$P(A) = \sum_{i=1}^k P(B_i)P(A|B_i).$$

Baye's Rule

Assume that

$$S = B_1 \cup B_2 \cup \dots \cup B_k, P(B_i) > 0,$$

and $B_i \cap B_j = \phi$ for $i \neq j$. Then

$$\begin{aligned} P(B_j|A) &= \frac{P(A \cap B_j)}{P(A)} \\ &= \frac{P(B_j)P(A|B_j)}{\sum_{i=1}^k P(B_i)P(A|B_i)}. \end{aligned}$$

Random Variable

Events identified by numbers are called *numerical* events.

Suppose that Y denote a variable to be measured in an experiment. Because the value of Y will vary depending on the random outcomes of the experiment, it is called a *random* variable.

Random Sampling

A statistical experiment involves the observation of a *sample*, say of size n , selected from a larger body of data, existing or conceptual, called a *population*, of size N .

If the sampling procedure allows the same element to be selected more than once in the sample, it is called sampling *with replacement*.

And if same element can not be selected more than once, sampling is *without replacement*.

Total number of *with replacement* samples (each of size n from a population of size N) = N^n , and *without replacement* samples = $\binom{N}{n}$.

If the sampling is conducted in such a way that each of

the $\binom{N}{n}$ samples has an equal probability of being selected, the sampling is said to be *simple random sampling* and the sample is said to be a *random sample*.

Exercise 2.49. To gain some insight into the risks that employees face when taking a lie detector test, suppose that the probability is 0.05 that a particular lie detector concludes that a person is lying who in fact is telling the truth and suppose that any pair of tests are independent.

Let event F = lie detector concludes that a person is lying who in fact is telling the truth

a. Probability (a machine will conclude that each of 3 employees is lying when all are telling the truth) = $P(F_1 \text{ and } F_2 \text{ and } F_3)$

$$= P(F_1 \cap F_2 \cap F_3) = P(F_1)P(F_2)P(F_3)$$

[Because tests are independent]

$$= (.05)(.05)(.05) = .000125.$$

b. P (A machine will conclude that at least 1 of the 3 employees is lying when all are telling the truth)

$$= 1 - P(\text{a machine making no mistake in 3 tests}) = 1 - (0.95)(0.95)(0.95) = .143.$$

Exercise 2.50. A survey showed that 10% households were dissatisfied with the plumbing job in their homes. half complaints dealt with plumber A who does 40% of the plumbing jobs in the town.

	A	A ^c	Total
U	0.05	0.05	0.10
U ^c	0.35	0.55	0.90
Total	0.40	0.60	1.00

$$a. P(U|A) = \frac{P(U \cap A)}{P(A)} = \frac{.05}{.4} = .125$$

$$b. P(U^c|A) = \frac{P(U^c \cap A)}{P(A)} = \frac{.35}{.4} = .875$$

Exercise 2.65. When a defective article comes through the inspection line, the probability that it gets by the first inspector is 0.1. The second inspector will miss 5 out of 10 of the defective items that get past the first inspector.

Let event A= defective item gets past through first inspector and B= defective item gets past through second inspector

$$P(A)=0.1 \text{ and } P(B|A)=\frac{5}{10} = 0.5.$$

$$P(\text{defective item gets by both inspectors, i.e., A and B}) \\ = P(A \cap B) = P(A)P(B|A) = (.1)(.5) = .05.$$

Exercise 2.72. Three radar sets operating independently are set to detect any aircraft flying through a certain area.

Each set has a probability of 0.02 of failing to detect a plane in the area.

Let event F = a radar set fails to detect

a. $P(\text{an aircraft goes undetected})$

$$= P(F_1 \text{ and } F_2 \text{ and } F_3) = P(F)P(F)P(F)$$

$$= (.02)^3 = .0008$$

b. $P(\text{an aircraft is detected by all 3 radars}) = P(F_1^c \text{ and } F_2^c \text{ and } F_3^c) = P(F^c)P(F^c)P(F^c) = (.98)^3$

Exercise 2.73. Consider one of the radar sets of Ex 2.72.

$P(\text{it will correctly detect exactly 3 aircrafts before it fails to detect one})$

$$= P(F_1^c \text{ and } F_2^c \text{ and } F_3^c \text{ and } F)$$

$$= P(F^c)P(F^c)P(F^c)P(F) = (.98)^3 (.02)$$

Exercise 2.85. It has been observed that 70% of the females react positively to a given set of circumstances whereas only 40% of males react positively. A group of 20 people, 15 female and 5 male was subjected to these circumstances and their reactions were recorded. A response picked at random from 20 was negative. Find the probability that it was that of a male.

Let P=positive response, M= male respondent,
F=female respondent.

$$P(P|F)=.70, P(P|M)=.40,$$

$$P(M)=\frac{5}{20} = \frac{1}{4}, P(F)=\frac{15}{20} = \frac{3}{4}.$$

Using Baye's rule:

$$\begin{aligned} P(M|P^c) &= \frac{P(M)P(P^c|M)}{P(M)P(P^c|M) + P(F)P(P^c|F)} \\ &= \frac{(\frac{1}{4})(1 - .4)}{(\frac{1}{4})(1 - .4) + (\frac{3}{4})(1 - .7)} \\ &= .4 \end{aligned}$$