Probability Theory

Section 7.2

Section Summary

- Assigning Probabilities
- Probabilities of Complements and Unions of Events
- Conditional Probability
- Independence
- Bernoulli Trials and the Binomial Distribution
- Random Variables

Assigning Probabilities

Laplace's definition assumes that all outcomes are equally likely. Now we introduce a more general definition of probabilities that avoids this restriction.

- Let S be a sample space of an experiment with a finite number of outcomes. We assign a probability $p(s)$ to each outcome s, so that:
    - *i.* $0 \leq p(s) \leq 1$ for each $s \in S$
    - *ii.* $$\sum_{s \in S} p(s) = 1$$
- The function p from the set of all outcomes of the sample space S is called a *probability distribution*.

Assigning Probabilities

Example: A trick coin is biased so that when flipped, the heads come up twice as often as tails. What probabilities should we assign to the outcomes H (heads) and T (tails) when the biased coin is flipped?

Solution: We are given that $p(H) = 2p(T)$

Because $p(H) + p(T) = 1$, it follows that

$$2p(T) + p(T) = 3p(T) = 1.$$

Hence, $p(T) = 1/3$ and $p(H) = 2/3$.

Uniform Distribution

Definition: Suppose that S is a set with n elements. The *uniform distribution* assigns the probability $1/n$ to each element of S. (Note that we could have used Laplace's definition here.)

Example: Consider again the coin flipping example, but with a fair coin. Now $p(H) = p(T) = 1/2$.

Probability of an Event

Definition: The *probability* of the event *E* is the sum of the probabilities of the outcomes in *E*.

$$p(E) = \sum_{s \in E} p(s)$$

- Note that now no assumption is being made about the distribution.

Example

Example: Suppose that a 6-sided die is biased so that 3 appears twice as often as each other number, but that the other five outcomes are equally likely. What is the probability that an odd number appears when we roll this die?

Solution: We want the probability of the event $E = \{1,3,5\}$. We have $p(3) = 2/7$ and

$$p(1) = p(2) = p(4) = p(5) = p(6) = 1/7.$$

Hence, $p(E) = p(1) + p(3) + p(5) =$

$$1/7 + 2/7 + 1/7 = 4/7.$$

Probabilities of Complements and Unions of Events

- Complements: $p(\overline{E}) = 1 - p(E)$ still holds. Since each outcome is in either E or $\overline{E}$, but not both,

$$\sum_{s \in S} p(s) = 1 = p(E) + p(\overline{E}).$$

- Unions: $p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$

  also still holds under the new definition.

Combinations of Events

Theorem: If E_1, E_2, … is a sequence of pairwise disjoint events in a sample space S, then

$$p\left(\bigcup_i E_i\right) = \sum_i p(E_i)$$

see Exercises 36 *and* 37 *for the proof*

Conditional Probability

Events can be dependent, which means they can be affected by previous events

Definition: Let E and F be events with $p(F) > 0$. The conditional probability of E given F, denoted by $P(E|F)$, is defined as:

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$

Conditional Probability

Example: A bit string of length four is generated at random so that each of the 16 bit strings of length 4 is equally likely. What is the probability that it contains at least two consecutive 0s, given that its first bit is a 0?

Solution: Let E be the event that the bit string contains at least two consecutive 0s, and F be the event that the first bit is a 0.

- Since $E \cap F = \{0000, 0001, 0010, 0011, 0100\}$, $p(E \cap F) = 5/16$.
- Because 8 bit strings of length 4 start with a 0, p(F) = 8/16= ½.

Hence,
$$p(E|F) = \frac{p(E \cap F)}{p(F)} = \frac{5/16}{1/2} = \frac{5}{8}.$$

Conditional Probability

Example: What is the conditional probability that a family with two children has two boys, given that they have at least one boy. Assume that each of the possibilities *BB*, *BG*, *GB*, and *GG* is equally likely (where *B* represents a boy and *G* represents a girl).

Solution: Let E be the event that the family has two boys and let F be the event that the family has at least one boy.

- Then $E = \{BB\}$, $F = \{BB, BG, GB\}$,
- $E \cap F = \{BB\}$.
- It follows that $p(F) = 3/4$ and $p(E \cap F) = 1/4$.

Hence,
$$p(E|F) = \frac{p(E \cap F)}{p(F)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Independence

Events can be independent, which means the occurrence of one event gives no information about the probability of another event. That is,

- $p(E|F) = p(E)$
- $p(F)$ has no impact on $p(E|F)$

Definition: The events E and F are independent if and only if

$$p(E \cap F) = p(E)p(F).$$

Independence

Example: Suppose E is the event that a randomly generated bit string of length four begins with a 1 and F is the event that this bit string contains an even number of 1s. Are E and F independent if the 16 bit strings of length four are equally likely?

Solution: There are eight bit strings of length four that begin with a 1, and eight bit strings of length four that contain an even number of 1s.

- Since the number of bit strings of length 4 is 16,

$$p(E) = p(F) = 8/16 = ½.$$

- Since $E \cap F = \{1111, 1100, 1010, 1001\}$, $p(E \cap F) = 4/16 = 1/4$.

We conclude that E and F are independent, because

$$p(E \cap F) = 1/4 = (½)\ (½) = p(E)\ p(F)$$

Independence

Example: Assume (as in the previous example) that each of the four ways a family can have two children (*BB*, *GG*, *BG*, *GB*) is equally likely. Are the events *E*, that a family with two children has two boys, and *F*, that a family with two children has at least one boy, independent?

Solution:

- $E = \{BB\}$, so $p(E) = 1/4$.
- We saw previously that that $p(F) = 3/4$ and $p(E \cap F) = 1/4$.

The events *E* and *F* are not independent since

$$p(E)\,p(F) = 3/16 \neq 1/4 = p(E \cap F) .$$

Pairwise and Mutual Independence

Definition: The events E_1, E_2, …, E_n are *pairwise independent* if and only if $p(E_i \cap E_j) = p(E_i)\,p(E_j)$ for all pairs i and j with $i \leq j \leq n$.

- Any 2 pairs of events are independent.

Definition: The events are *mutually independent* if

$$p(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_m}) = p(E_{i_1})p(E_{i_2}) \cdots p(E_{i_m})$$

whenever i_j, $j = 1,2,\ldots, m$ are integers with

$$1 \leq i_1 < i_2 < \cdots < i_m \leq n \quad \text{and} \quad m \geq 2.$$

- Any m events are independent.

James Bernoulli
(1854 – 1705)

Bernoulli Trials

Definition: Suppose an experiment can have only two possible outcomes, *e*.*g*., the flipping of a coin or the random generation of a bit.

- Each performance of the experiment is called a *Bernoulli trial*.
- One outcome is called a *success* and the other a *failure*.
- If p is the probability of success and q the probability of failure, then $p + q = 1$.

- Many problems involve determining the probability of k successes when an experiment consists of n mutually independent Bernoulli trials.

Bernoulli Trials

Example: A fair coin is flipped 3 times. P(heads) = ½ = P(tails). What is the probability that we get three, two, one, or no heads?

Solution: There are $2^3 = 8$ possible outcomes.

P(three heads) = C(3,3) / 8 = 1/8

P(two heads) = C(3,2) / 8 = 3/8

P(one head) = C(3, 1) / 8 = 3/8

P(zero heads) = C(3, 0) / 8 = 1/8

HHH
HHT
HTH
HTT
THH
THT
TTH
TTT

Bernoulli Trials

Example: A fair coin (P(heads) = P(tails) = ½) is flipped 5 times. What is the probability that we get five, four, three, two, one, or no heads?

Solution: There are $2^5 = 32$ possible outcomes.

P(five heads) $= C(5, 5) / 32 = 1/32$

P(four heads) $= C(5, 4) / 32 = 5/32$

P(three heads) $= C(5, 3) / 32 = 10/32$

P(two heads) $= C(5, 2) / 32 = 10/32$

P(one head) $= C(5, 1) / 32 = 5/32$

P(zero heads) $= C(5, 0) / 32 = 1/32$

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1

Bernoulli Trials

Example: A coin is biased so that the probability of heads is 2/3. What is the probability that exactly four heads occur when the coin is flipped seven times?

Solution: There are $2^7 = 128$ possible outcomes.

- The number of ways four of the seven flips can be heads is $C(7,4)$.
- The probability of each of the outcomes is $(2/3)^4(1/3)^3$ since the seven flips are independent.
- Hence, the probability that exactly four heads occur is

$$C(7,4)\,(2/3)^4(1/3)^3 = 560/\,2187.$$

Probability of k Successes in n Independent Bernoulli Trials.

Theorem 2: The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$, is

$$C(n,k)p^k q^{n-k}.$$

Proof:

- The outcome of n Bernoulli trials is an n-tuple $(t_1,t_2,\ldots,t_n)$, where each is t_i either S (success) or F (failure).
- The probability of each outcome of n trials consisting of k successes and $n - k$ failures (in any order) is $p^k q^{n-k}$.
- Because there are $C(n,k)$ n-tuples of S's and F's that contain exactly k S's, the probability of k successes is $C(n,k)p^k q^{n-k}$.

Probability of *k* Successes in *n* Independent Bernoulli Trials.

Theorem 2: The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$, is

$$C(n,k)p^k q^{n-k}.$$

- We denote by $b(k{:}n,p)$ the probability of k successes in n independent Bernoulli trials with p the probability of success. Viewed as a function of k, $b(k{:}n,p)$ is the *binomial distribution*. By Theorem 2,

$$b(k{:}n,p) = C(n,k)p^k q^{n-k}.$$

.

Random Variables

Definition: A *random variable* is a function from the sample space of an experiment to the set of real numbers. That is, a random variable assigns a real number to each possible outcome.

- A random variable is a function. It is not a variable, and it is not random!
- In the late 1940s W. Feller and J.L. Doob flipped a coin to see whether both would use “random variable” or the more fitting “chance variable.” Unfortunately, Feller won and the term “random variable” has been used ever since.

Random Variables

Definition: The *distribution* of a random variable X on a sample space S is the set of pairs $(r, p(X = r))$ for all $r \in X(S)$, where $p(X = r)$ is the probability that X takes the value r.

Example: Suppose that a coin is flipped three times. Let $X(t)$ be the random variable that equals the number of heads that appear when t is the outcome. Then $X(t)$ takes on the following values:

$X(HHH) = 3,$
$X(HHT) = X(HTH) = X(THH) = 2,$
$X(TTH) = X(THT) = X(HTT) = 1$
$X(TTT) = 0.$

Each of the eight possible outcomes has probability 1/8. So, the distribution of $X(t)$ is $p(X = 3) = 1/8$, $p(X = 2) = 3/8$, $p(X = 1) = 3/8$, and $p(X = 0) = 1/8$.