Conditional Probability

If we toss a coin three times, what is the probability that an odd number of tails appears (event $E$), if the first toss is a tail (event $F$) ?

If the first toss is a tail, the possible sequences are TTT, TTH, THH, and THH.

In two out of these four cases, there is an odd number of tails.

Therefore, the probability of $E$, under the condition that $F$ occurs, is 0.5.

We call this conditional probability.

Conditional Probability

Example: What is the probability of a random bit string of length four to contain at least two consecutive 0s, given that its first bit is a 0 ?

Solution:

$E$: "bit string contains at least two consecutive 0s"
$F$: "first bit of the string is a 0"

We know the formula $p(E \mid F) = p(E \cap F)/p(F)$.

$E \cap F = \{0000, 0001, 0010, 0011, 0100\}$

$p(E \cap F) = 5/16$

$p(F) = 8/16 = 1/2$

$p(E \mid F) = (5/16)/(1/2) = 10/16 = 5/8 = 0.625$

Independence

Let us return to the example of tossing a coin three times.

Does the probability of event $E$ (odd number of tails) depend on the occurrence of event $F$ (first toss is a tail) ?

In other words, is it the case that $p(E \mid F) \neq p(E)$ ?

We actually find that $p(E \mid F) = 0.5$ and $p(E) = 0.5$, so we say that $E$ and $F$ are independent events.

Independence

Example: Suppose $E$ is the event of rolling an even number with an unbiased die. $F$ is the event that the resulting number is divisible by three. Are events $E$ and $F$ independent?

Solution:

$p(E) = 1/2$, $p(F) = 1/3$.

$|E \cap F| = 1$ (only 6 is divisible by both 2 and 3)

$p(E \cap F) = 1/6$

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

Conclusion: $E$ and $F$ are independent.
Bernoulli Trials

Suppose an experiment with **two possible outcomes**, such as tossing a coin.

Each performance of such an experiment is called a **Bernoulli trial**.

We will call the two possible outcomes a **success** or a **failure**, respectively.

If \( p \) is the probability of a success and \( q \) is the probability of a failure, it is obvious that \( p + q = 1 \).

Bernoulli Trials

Often we are interested in the probability of **exactly \( k \) successes** when an experiment consists of \( n \) independent Bernoulli trials.

**Example:**
A coin is biased so that the probability of head is \( \frac{2}{3} \). What is the probability of exactly four heads to come up when the coin is tossed seven times?

**Solution:**
There are \( 2^7 = 128 \) possible outcomes.

The number of possibilities for four heads among the seven trials is \( \text{C}(7, 4) \).

The seven trials are independent, so the probability of each of these outcomes is \( (\frac{2}{3})^4(\frac{1}{3})^3 \).

Consequently, the probability of exactly four heads to appear is
\[
\text{C}(7, 4)(\frac{2}{3})^4(\frac{1}{3})^3 = \frac{560}{2187} = 25.61\%
\]

Bernoulli Trials

**Illustration:**
Let us denote a success by ‘S’ and a failure by ‘F’. As before, we have a probability of success \( p \) and probability of failure \( q = 1 - p \).

What is the probability of **two successes in five** independent Bernoulli trials?

Let us look at a possible sequence:

SSFFF

What is the probability that we will generate exactly this sequence?

\[
\text{Probability: } \frac{p^2q^3}{2^5} = \frac{p^2q^3}{32}
\]

Another possible sequence:

FSFFF

**Probability:**

\[
\frac{p^2q^3}{2^5} = \frac{p^2q^3}{32}
\]

Each sequence with two successes in five trials occurs with probability \( p^2q^3 \).

Bernoulli Trials

And how many possible sequences are there?
In other words, how many ways are there to pick two items from a list of five?

We know that there are \( \text{C}(5, 2) = 10 \) ways to do this, so there are 10 possible sequences, each of which occurs with a probability of \( p^2q^3 \).

Therefore, the probability of **any** such sequence to occur when performing five Bernoulli trials is \( \text{C}(5, 2) p^2q^3 \).

In general, for \( k \) successes in \( n \) Bernoulli trials we have a probability of \( \text{C}(n,k)p^kq^{n-k} \).
Random Variables

In some experiments, we would like to assign a numerical value to each possible outcome in order to facilitate a mathematical analysis of the experiment.

For this purpose, we introduce 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.

Note: Random variables are functions, not variables, and they are not random, but map random results from experiments onto real numbers in a well-defined manner.

Random Variables

**Example:**

Let $X$ be the result of a rock-paper-scissors game. If player A chooses symbol $a$ and player B chooses symbol $b$, then

$$X(a, b) =
\begin{cases}
  1, & \text{if player A wins,} \\
  0, & \text{if A and B choose the same symbol,} \\
  -1, & \text{if player B wins.}
\end{cases}$$

Random Variables

<table>
<thead>
<tr>
<th>$X(\text{rock, rock})$</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X(\text{rock, paper})$</td>
<td>-1</td>
</tr>
<tr>
<td>$X(\text{rock, scissors})$</td>
<td>1</td>
</tr>
<tr>
<td>$X(\text{paper, rock})$</td>
<td>1</td>
</tr>
<tr>
<td>$X(\text{paper, paper})$</td>
<td>0</td>
</tr>
<tr>
<td>$X(\text{paper, scissors})$</td>
<td>-1</td>
</tr>
<tr>
<td>$X(\text{scissors, rock})$</td>
<td>-1</td>
</tr>
<tr>
<td>$X(\text{scissors, paper})$</td>
<td>1</td>
</tr>
<tr>
<td>$X(\text{scissors, scissors})$</td>
<td>0</td>
</tr>
</tbody>
</table>

Expected Values

Once we have defined a random variable for our experiment, we can statistically analyze the outcomes of the experiment.

For example, we can ask: What is the average value (called the expected value) of a random variable when the experiment is carried out a large number of times?

Can we just calculate the arithmetic mean across all possible values of the random variable?

No, we cannot, since it is possible that some outcomes are more likely than others.

For example, assume the possible outcomes of an experiment are 1 and 2 with probabilities of 0.1 and 0.9, respectively.

Is the average value 1.5?

No, since 2 is much more likely to occur than 1, the average must be larger than 1.5.

Expected Values

Instead, we have to calculate the weighted sum of all possible outcomes, that is, each value of the random variable has to be multiplied with its probability before being added to the sum.

In our example, the average value is given by $0.1 \cdot 1 + 0.9 \cdot 2 = 0.1 + 1.8 = 1.9$.

**Definition:** The expected value (or expectation) of the random variable $X(s)$ on the sample space $S$ is equal to:

$$E(X) = \sum_{s \in S} p(s)X(s).$$
Expected Values

**Example:** Let $X$ be the random variable equal to the sum of the numbers that appear when a pair of dice is rolled.

There are **36 outcomes** (= pairs of numbers from 1 to 6).

The range of $X$ is \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}.

Are the 36 outcomes equally likely?
Yes, if the dice are not biased.

Are the 11 values of $X$ equally likely to occur?
No, the probabilities vary across values.

<table>
<thead>
<tr>
<th>$X$</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$\frac{1}{36}$</td>
</tr>
<tr>
<td>3</td>
<td>$\frac{2}{36} = \frac{1}{18}$</td>
</tr>
<tr>
<td>4</td>
<td>$\frac{3}{36} = \frac{1}{12}$</td>
</tr>
<tr>
<td>5</td>
<td>$\frac{4}{36} = \frac{1}{9}$</td>
</tr>
<tr>
<td>6</td>
<td>$\frac{5}{36}$</td>
</tr>
<tr>
<td>7</td>
<td>$\frac{6}{36} = \frac{1}{6}$</td>
</tr>
<tr>
<td>8</td>
<td>$\frac{5}{36}$</td>
</tr>
<tr>
<td>9</td>
<td>$\frac{4}{36} = \frac{1}{9}$</td>
</tr>
<tr>
<td>10</td>
<td>$\frac{3}{36} = \frac{1}{12}$</td>
</tr>
<tr>
<td>11</td>
<td>$\frac{2}{36} = \frac{1}{18}$</td>
</tr>
<tr>
<td>12</td>
<td>$\frac{1}{36}$</td>
</tr>
</tbody>
</table>