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 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 \( \binom{7}{4} \). The seven trials are independent, so the probability of each of these outcomes is \( (2/3)^4(1/3)^3 \).
Consequently, the probability of exactly four heads to appear is
\[
\binom{7}{4}(2/3)^4(1/3)^3 = 560/2187 = 25.61\%
\]

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:

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>S F F F F</td>
<td>( p \cdot p \cdot q \cdot q \cdot q = p^2q^3 )</td>
</tr>
</tbody>
</table>

Another possible sequence:

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>F S F S F</td>
<td>( q \cdot p \cdot q \cdot p = p^2q^3 )</td>
</tr>
</tbody>
</table>

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 \( \binom{5}{2} = 10 \) 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
\[
\binom{5}{2}p^2q^3 = 10 \cdot \frac{2}{3}^2 \cdot \frac{1}{3}^3 = \frac{40}{243} \\
\]
In general, for \( k \) successes in \( n \) Bernoulli trials we have a probability of \( \binom{n}{k}p^kq^{n-k} \).
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

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

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rock, rock</td>
<td>0</td>
</tr>
<tr>
<td>rock, paper</td>
<td>-1</td>
</tr>
<tr>
<td>rock, scissors</td>
<td>1</td>
</tr>
<tr>
<td>paper, rock</td>
<td>1</td>
</tr>
<tr>
<td>paper, paper</td>
<td>0</td>
</tr>
<tr>
<td>paper, scissors</td>
<td>-1</td>
</tr>
<tr>
<td>scissors, rock</td>
<td>-1</td>
</tr>
<tr>
<td>scissors, paper</td>
<td>1</td>
</tr>
<tr>
<td>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

\[
0.1 \times 1 + 0.9 \times 2 = 0.1 + 1.8 = 1.9.
\]

**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.

Expected Values

$$E(X) = \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{7}{36} + \frac{8}{36} + \frac{9}{36} + \frac{10}{36} + \frac{11}{36} + \frac{12}{36}$$

This means that if we roll the dice many times, sum all the numbers that appear and divide the sum by the number of trials, we expect to find a value of 7.

Expected Values

Theorem:
If $X$ and $Y$ are random variables on a sample space $S$, then $E(X + Y) = E(X) + E(Y)$.

Furthermore, if $X_1, X_2, \ldots, X_n$ are random variables on $S$, then $E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$.

Moreover, if $a$ and $b$ are real numbers, then $E(aX + b) = aE(X) + b$.

Expected Values

We can use our knowledge about expected values to compute the average-case complexity of an algorithm.

Let the sample space be the set of all possible inputs $a_1, a_2, \ldots, a_n$, and the random variable $X$ assign to each $a_i$ the number of operations that the algorithm executes for that input.

For each input $a_i$, the probability that this input occurs is given by $p(a_i)$.

The algorithm's average-case complexity then is:

$$E(X) = \sum_{i=1}^{n} p(a_i)X(a_i)$$
Expected Values

However, in order to conduct such an average-case analysis, you would need to find out:

- the number of steps that the algorithms takes for any (!) possible input, and
- the probability for each of these inputs to occur.

For most algorithms, this would be a highly complex task, so we will stick with the worst-case analysis.

Independent Random Variables

**Definition:** The random variables $X$ and $Y$ on a sample space $S$ are **independent** if

$$p(X(s) = r_1 \land Y(s) = r_2) = p(X(s) = r_1) \cdot p(Y(s) = r_2).$$

In other words, $X$ and $Y$ are independent if the probability that $X(s) = r_1 \land Y(s) = r_2$ equals the product of the probability that $X(s) = r_1$ and the probability that $Y(s) = r_2$ for all real numbers $r_1$ and $r_2$.

**Example:** Are the random variables $X_1$ and $X_2$ from the “pair of dice” example independent?

**Solution:**

$p(X_1 = i) = \frac{1}{6}$

$p(X_2 = j) = \frac{1}{6}$

$p(X_1 = i \land X_2 = j) = \frac{1}{36}$

Since $p(X_1 = i \land X_2 = j) = p(X_1 = i) \cdot p(X_2 = j)$, the random variables $X_1$ and $X_2$ are **independent**.

**Theorem:** If $X$ and $Y$ are independent random variables on a sample space $S$, then $E(XY) = E(X)E(Y)$.

**Note:**

$E(X + Y) = E(X) + E(Y)$ is true for any $X$ and $Y$, but $E(XY) = E(X)E(Y)$ needs $X$ and $Y$ to be independent.

How come?

**Example:** Let $X$ and $Y$ be random variables on some sample space, and each of them assumes the values 1 and 3 with equal probability.

Then $E(X) = E(Y) = 2$

If $X$ and $Y$ are **independent**, we have:

$E(X + Y) = \frac{1}{4}(1 + 1) + \frac{1}{4}(1 + 3) + \frac{1}{4}(3 + 1) + \frac{1}{4}(3 + 3) = 4 = E(X) + E(Y)$

$E(XY) = \frac{1}{4}(1 \cdot 1) + \frac{1}{4}(1 \cdot 3) + \frac{1}{4}(3 \cdot 1) + \frac{1}{4}(3 \cdot 3) = 4 = E(X)E(Y)$

Let us now assume that $X$ and $Y$ are **correlated** in such a way that $Y = 1$ whenever $X = 1$, and $Y = 3$ whenever $X = 3$.

Then $E(X + Y) = \frac{1}{2}(1 + 1) + \frac{1}{2}(3 + 3)$

$= 4 = E(X) + E(Y)$

$E(XY) = \frac{1}{2}(1 \cdot 1) + \frac{1}{2}(3 \cdot 3)$

$= 5 \neq E(X)E(Y)$