Discrete Probability

Example: What is the probability of a positive integer selected at random from the set of positive integers not exceeding 100 to be divisible by 2 or 5?

Solution:

E_2: "integer is divisible by 2"
E_5: "integer is divisible by 5"

$E_2 = \{2, 4, 6, \ldots, 100\}$

$|E_2| = 50$

$p(E_2) = 0.5$

Discrete Probability

$E_5 = \{5, 10, 15, \ldots, 100\}$
$|E_5| = 20$
$p(E_5) = 0.2$

$E_2 \cap E_5 = \{10, 20, 30, \ldots, 100\}$
$|E_2 \cap E_5| = 10$
$p(E_2 \cap E_5) = 0.1$

$p(E_2 \cup E_5) = p(E_2) + p(E_5) – p(E_2 \cap E_5)$
$p(E_2 \cup E_5) = 0.5 + 0.2 – 0.1 = 0.6$

Discrete Probability

What happens if the outcomes of an experiment are **not** equally likely?

In that case, we assign a probability p(s) to each outcome $s \in S$, where S is the sample space.

Two conditions have to be met:
(1): $0 \leq p(s) \leq 1$ for each $s \in S$, and
(2): $\sum_{s \in S} p(s) = 1$

This means, as we already know, that (1) each probability must be a value between 0 and 1, and (2) the probabilities must add up to 1, because one of the outcomes is **guaranteed** to occur.

Discrete Probability

How can we obtain these probabilities p(s) ?

The probability p(**s**) assigned to an outcome **s** equals the limit of the number of times **s** occurs divided by the number of times the experiment is performed.

Once we know the probabilities p(s), we can compute the **probability of an event E** as follows:

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

Discrete Probability

Example I: A die is biased so that the number 3 appears twice as often as each other number.

What are the probabilities of all possible outcomes?

Solution: There are 6 possible outcomes $s_1, \ldots, s_6$.

$p(s_1) = p(s_2) = p(s_4) = p(s_5) = p(s_6)$

$p(s_3) = 2p(s_1)$

Since the probabilities must add up to 1, we have:

$5p(s_1) + 2p(s_1) = 1$

$7p(s_1) = 1$

$p(s_1) = p(s_2) = p(s_4) = p(s_5) = p(s_6) = 1/7$, $p(s_3) = 2/7$

Discrete Probability

Example II: For the biased die from Example I, what is the probability that an odd number appears when we roll the die?

Solution:

$E_{odd} = \{s_1, s_3, s_5\}$

Remember the formula $p(E) = \sum_{s \in E} p(s)$.

$p(E_{odd}) = \sum_{s \in E_{odd}} p(s) = p(s_1) + p(s_3) + p(s_5)$
$p(E_{odd}) = 1/7 + 2/7 + 1/7 = 4/7 = 57.14\%$

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, THT, 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

If we want to compute the conditional probability of E given F, we use F as the sample space.

For any outcome of E to occur under the condition that F also occurs, this outcome must also be in $E \cap F$.

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

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

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

Because we have $p(E \mid F) = p(E \cap F)/p(F)$, $p(E \mid F) = p(E)$ if and only if $p(E \cap F) = p(E)p(F)$.

Definition: The events E and F are said to be independent if and only if $p(E \cap F) = p(E)p(F)$.

Obviously, this definition is **symmetrical** for E and F. If we have $p(E \cap F) = p(E)p(F)$, then it is also true that $p(F \mid E) = p(F)$.

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 2/3. What is the probability of exactly four heads to come up when the coin is tossed seven times?

Bernoulli Trials

Solution:

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

The number of possibilities for four heads among the seven trials is C(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

$C(7, 4)(2/3)^4(1/3)^3 = 560/2187 = 25.61\%$

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:

SSFFF

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

Bernoulli Trials

Sequence: S S F F F

Probability: $p \cdot p \cdot q \cdot q \cdot q = p^2q^3$

Another possible sequence:

Sequence: F S F S F

Probability: $q \cdot p \cdot q \cdot p \cdot q = p^2q^3$

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 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
$C(5, 2)\, p^2q^3$.

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