

## ***Basic Probability Concepts***

When you take a multiple-choice exam, the chances of guessing the correct answer are usually 1 out of 4, or 25 %. When the exam questions are of True/False type, the chances of guessing correctly are 1 out of 2, or 50%. Whenever you are talking about chances or odds of something happening, you are applying the basics of probability theory, which we are about to discuss in more detail.

### ***Probability of Occurrence***

When we said that the chances of guessing the correct answer in a multiple choice test were 1 out of 4, we calculated the probability of occurrence:

$$\text{Probability of Occurrence (P)} = \text{Number of desired outcomes} / \text{Total possible outcomes}$$

The number of desired results is one because there is only one correct answer to the question. The total number of possible outcomes is four since there are usually four choices in every multiple-choice question.

If there are 100 Baruch students in the room and they equally represent freshmen, sophomores, juniors and seniors (meaning 25 students for each group), the probability of you running into a junior is 25/100 since there are 25 juniors in the room (desired outcomes) out of 100 students in total. Each possible type of probability of occurrence is called an event. If an event has only one characteristic or condition, like specifically running into a junior in that crowded room, it is called a **simple event**. If an event has two or more characteristics, like running into a junior that lives in Queens and is an accounting major, it is called a **joint event**.

### ***Contingency Tables***

Right after Thanksgiving you decided to survey 100 people in a neighborhood shopping mall about their shopping preferences. These were the results of your survey:

Out of 100 people, 65 were females and 35 were males. Out of 65 females, 40 said that they enjoyed shopping and the rest said that they hated it around the holiday season. Out of 35 males, 20 said that they hated shopping and the rest confessed that they liked it.

You can conveniently present your findings in the table (see if you can fill in the missing data – check your answers on p. 3):

	<b>Female</b>	<b>Male</b>	<b>Total</b>
<b>Enjoy Shopping</b>	40		
<b>Hate Shopping</b>		20	
<b>Total</b>	65	35	100

The above table is called a **contingency table** and with its help, you were able to easily calculate the total number of people who liked shopping and the number of people who disliked shopping (regardless of gender). An example of a simple event is liking to shop. An example of a joint event is being a female and disliking shopping. Can you name 3 more simple events and 3 more joint events?

### *Simple and Joint Probability*

When conducting your survey, you asked the participants to indicate their gender and their attitude towards shopping on small pieces of paper. By the end of the day, you collected 100 pieces of paper, which you placed into the bowl in order to study simple and joint probability.

Simple probability is the probability of a simple event occurring. For example, a probability of picking out a piece of paper that said “female” on it (regardless of the shopping interest) would be simple probability. Numerically, the probability of randomly selecting a piece of paper with the word “female” on it would be  $65/100$ , since there are 65 desired outcomes (female participants) in your survey and 100 total possible outcomes:

$$P(\text{female}) = 65 / 100 = 0.65 = 65\%$$

Similarly, the joint probability is the probability of a joint event occurring. If you were aiming at selecting a paper that said, “male, hate shopping,” how would you calculate the probability of this happening? The number in the denominator is still 100 – the total number of possible outcomes, but the number in the numerator – the number of desired outcomes – will be 20 – for the number of males that dislike shopping:

$$P(\text{male and dislike shopping}) = 20 / 100 = 0.20 = 20\%$$

There is a special kind of joint probability that we have to discuss. For example, you would like to find the probability of randomly selecting a female **or** a person that likes to shop. In practice that would mean finding the probability of randomly selecting a female (simple probability) plus the probability of selecting a person that likes to shop (simple probability) minus the probability of picking a female that likes to shop (joint probability) to avoid counting these participants twice:

$$\begin{aligned} P(\text{female}) &= 65 / 100 \\ P(\text{enjoy shopping}) &= 55 / 100 \\ P(\text{female and enjoy shopping}) &= 40 / 100 \\ P(\text{female or enjoy shopping}) &= 65/100 + 55/100 - 40/100 = 0.8 = 80\% \end{aligned}$$

It might be even easier to use the table to understand how to calculate a joint probability of this kind:

	Female	Male	Total
Enjoy Shopping	40	15	55
Hate Shopping	25	20	45
Total	65	35	100

Everyone who enjoys shopping

All females

Females that enjoy shopping

The formula for finding this kind of probability is called a **general addition rule**:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

### ***Conditional Probability***

Let's return to the bowl with the small pieces of paper that you collected. Assume that you were now only concerned with females' shopping interests. What would you do? You would remove all the papers that said "male" on them to make sure that your total number of possible outcomes is narrowed down to females only.

Whenever there is an assumption or a limit to the total number of outcomes that you can use, you are dealing with conditional probability. The words like "suppose," "assume," or "given" are hints that the problem asks you to limit the total number of possible outcomes.

If you were asked to find the probability of selecting a person that likes shopping *given* that that person is a male, you would be dealing with conditional probability. To answer this question, you have to follow the general formula for the probability of occurrence and use the correct total number of possible outcomes. In our case, the total number of possible outcomes will become 35 since we are now dealing only with males. The number of desired outcomes will become 15 – that is, how many people indicated that they like shopping out of the 35 total respondents. Conditional probability is denoted by a separating line or slash between the simple event and the limiting condition:

$$P(\text{enjoy shopping} / \text{male}) = 15 / 35 \approx 0.43 = 43\%$$

Using the contingency table below, you are now only concerned with the third column and disregarding the rest of the information, since the condition was to observe only males:

	Female	<b>Male</b>	Total
Enjoy Shopping	40	<b>15</b>	55
Hate Shopping	25	<b>20</b>	45
Total	65	<b>35</b>	100

Number of desired outcomes

Number of possible outcomes