Lecture 7. More on Independence. Random Variables

More on Independence.

Comment. On the quiz last Thursday, more than half the class asserted incorrectly that"independent events are events that have no outcomes in common." So, let us review this idea.

Two events A and B are said to be independent if the occurrence of one has no effect on the probability of the other. In terms of conditional probability, we may interpret the independence of A and B to mean $P(A) = P(A|B)$: the probability of A is the same as the probability of A given B. But we might also interpret it to mean $P(B) = P(B|A)$. Which is the right way? Fortunately, if $P(A)$ and $P(B)$ are both non-zero:

$$\begin{aligned} P(B) = P(B|A) &\iff P(B) = \frac{P(B \cap A)}{P(A)} \\ &\iff P(A)P(B) = P(B \cap A) \\ &\iff P(A) = \frac{P(A \cap B)}{P(B)} \\ &\iff P(A) = P(A|B) \end{aligned}$$

Thus, either of these conditions implies the other. To make the symmetry of the independence relation obvious, we use the condition $P(A)P(B) = P(B \cap A)$ as the definition. Let us repeat it:

Definition. Events A and B are said to be independent if $P(A)P(B) = P(B \cap A)$.

This definition has the advantage that it does not mention conditional probabilities, though of course it implies that the conditional probabilities are equal to the unconditioned ones, which was the idea that motivated the concept of independence. It has the advantage that we don't need to worry if the probability of one of the events happens to be zero. The definition implies that an event of probability zero is independent of any other event.

Caution. Two events are said to be *mutually exclusive* if they have no outcomes in common. If A and B are mutually exclusive, then $A \cap B$ is the empty event, and therefore $P(A \cap B) = 0$. Accordingly, if A and B have non-zero probability and are mutually exclusive, then they are **not** independent. It is a common for learners to assume erroneously that the two concepts are the same.

Example. Example 1.12 of our textbook reads: "An elevator with two passengers stops at the second, third and fourth floors. If it is equally likely that a passenger gets off at any of the three floors, what is the probability that the passengers get off at different floors."

Comment. Let's say the passengers are Mary and Sarah. The problem stipulates that Mary has a $1/3$ probability of getting off at the 2^{nd} floor, a $1/3$ probability of getting off at the 3^{rd} floor and a $1/3$ probability of getting off at the 4^{th} floor. The same is true of Sarah. The sample space consists of all the ways the two of them may get off. The outcomes, therefore, correspond to the cells in the following table.

Sarah

Mary		2	3	4	+
	2				1/3
	3				1/3
	4				1/3
	+	1/3	1/3	1/3	1

Here, the cells in the first row are the elements in the event "Mary gets of on the second floor." The problem states that this event has probability 1/3, and this is indicated in the margin, at the end of the row. The same is true of the other floors at which Mary may exit. Similarly, the columns correspond to the floors at which Sarah may exit, and the event of her exiting at any one of the three floors is the same. The event of them exiting at the same floor consists of the cells on the diagonal from the upper left to the lower right. We must figure out how to assign probability masses to these cells, and then add them up. Unfortunately, the problem does not give us enough information to do this, as we see from the following 3 different "solutions," each based on a different assumption that is perfectly consistent with the way the problem was stated.

Solution 1. Mary's last name is Landrieu, and Sarah's is Palin. No way will they get off at the same floor. We must fill the chart in as follows:

Sarah

Mary		2	3	4	+
	2	0	1/6	1/6	1/3
	3	1/6	0	1/6	1/3
	4	1/6	1/6	0	1/3
	+	1/3	1/3	1/3	1

Note that the events "Mary gets off at 2 and "Sarah gets off at 2" have no outcomes of non-zero probability in common, so they are essentially mutually exclusive.

Solution 2. Mary's last name is Smith, and so is Sarah's. Sarah, in fact, is Mary's infant daughter. Since Mary is carrying Sarah in her arms, the two are certain to get off at the same floor. We must fill the chart in as follows:

Sarah

Mary		2	3	4	+
	2	1/3	0	0	1/3
	3	0	1/3	0	1/3
	4	0	0	1/3	1/3
	+	1/3	1/3	1/3	1

Note that the events "Mary gets off at 2 and "Sarah gets off at 2" contain exactly the same outcomes of non-zero probability, so they are essentially the same event.

Solution 3. Mary and Sarah are absolute strangers. The event of Mary getting off at her floor is independent of the event of Sarah getting off at hers. Using the multiplication rule in the definition of independence, we must fill in the table as follows:

		Sarah			
		2	3	4	+
Mary	2	1/9	1/9	1/9	1/3
	3	1/9	1/9	1/9	1/3
	4	1/9	1/9	1/9	1/3
	+	1/3	1/3	1/3	1

The event of them getting off at the same floor has probability $1/9 + 1/9 + 1/9 = 1/3$.

Random Variables: Definition, comments, examples and problems

Definition. A *real-valued random variable on a discrete probability space* is a function from that space to the real numbers.

It is common to use the symbol X to denote a random variable. When one says, "X is a random variable," one has a probability space Ω in mind and a real-valued function, named X, on that space. To state in symbols that X is a function from Ω to $\mathbb{R}$, one writes $X : \Omega \to \mathbb{R}$.

As with the terms "experiment," "outcome," "sample space" and "event," the idea of a random variable is intended to transpose an intuitive notion into a rigid mathematical framework, where one can work with meanings that are precise, stable and agreed upon by all competent users of the theory of probability. As with these other concepts, there is an intuitive notion at the origin. Consider a randomly varying quantity, such as the number of white cells in a microliter of blood, the number of defective items coming off an assembly line in an hour or the weight of a truck arriving at a particular weigh station. Such a quantity varies randomly from occasion to occasion, but over many measurements certain regularities are perceived. Certain counts or weights appear more frequently, others less, and for each number x, there is, in the long run, a certain proportion of measurements that are less than or equal to x. The variation can be modeled by assuming that the number x is the value of a function on a probability space. The relative frequency with which numbers less than or equal to x occur is modeled as the probability of the event

$$\{\omega \in \Omega \mid X(\omega) \le x\}.$$

We mentioned previously that in a probability space that is *not* discrete, some subsets may fail to have probability measures. For this reason, an additional stipulation is needed to define the concept of random variable in the non-discrete case: we require that for any real numbers $a < b$, the set $\{ \omega \in \Omega \mid a < X(\omega) < b \}$ must have a probability measure. In many basic applications, this requirement is automatically met. But attention to this detail is critical in the mathematical foundations of the theory. For the time being, we will be speaking only of discrete random variables, so we will postpone further discussion of this issue.

Example 1.

An experiment consists in rolling a die 5 times and recording the 5 numbers rolled. The outcomes are the numbers that appear, in the order in which they appear. For example, 63435 denotes the outcome in which the first roll yielded a 6, the second a 3, etc. For each outcome ω, let $X(\omega)$ be the sum of all the numbers rolled. Then X is a real-valued function defined on the sample space of this the experiment, and hence it is a random variable. An equation such as $X = 21$ defines an event:

$$\{ \omega \in \Omega \mid X(\omega) = 21 \}.$$

This event contains 63435, since $X(63435) = 21$, but it does not contain 64433, since $X(64433) = 20$. The probability of this event is denoted $P(X = 21)$. In this example, there are $6^5 = 7776$ equally-likely outcomes, and 540 of them (see Mathematica Notebook 7) sum to 21. Thus,

$$P(X = 21) = 540/7776 \approx 0.0694444.$$

Problems 1.

Examine the computations related to this problem in the Notebook 7. The notebook contains a table showing the values of $P(X = n)$, as well as a graph of these probabilities as a function of n.

a) How would you graph the function X itself? What does the graph look like?
b) Make a table showing the number of outcomes in the event $X = n$ for each $n \in \{5, 6, \ldots, 30\}$. How is this related to the graph of $f(n) := P(X = n)$? How is it related to the graph of X?
c) Suppose the problem is modified by rolling the die 6 times instead of 5. make a table of the values of $P(X = n)$ for $n \in \{6, 7, \ldots, 36\}$, and graph these probabilities as a function of n.

Example 2.

An experiment consists rolling a die over and over until a 3 is rolled. The sample space Ω of this experiment is the set of all sequences $\omega = a_1 a_2 \cdots a_{n-1} 3$ in which each a_i, $i < n$, is in the set $\{1, 2, 4, 5, 6\}$. (Note the missing 3.) Let $X(\omega)$ denote the length of ω:

$$\begin{aligned} X(3) &= 1; \\ X(a_1 3) &= 2 \ \ (a_1 \in \{1, 2, 4, 5, 6\}); \\ X(a_1 a_2 3) &= 3 \ \ (a_1, a_2 \in \{1, 2, 4, 5, 6\}); \\ &\mathit{etc.} \end{aligned}$$

In this example, $P(X=1)=1/6$, since there is a probability of $1/6$ that the first roll will result in a 3. $P(X=2)=(5/6)(1/6)=5/36$, since there is a probability of $5/6$ that the first roll will NOT result in a 3 and a probability of $(1/6)$ that the second roll will. In general,

$$P(X=n)=(5/6)^{n-1}(1/6).$$

Problems 2.

a) The outcomes in this probability space are not equally likely. Letting $f:\Omega\to[0,1]$ denote the pmf, explain why $f(\omega)=(1/6)^{X(\omega)}$.
b) Let $g(n):=P(X=n)$. Show that g is a pmf on the set $\{1,2,3,\ldots\}$.

Example 3.

An urn contains n beads, and s of them are special. An experiment consists in taking k beads. Each outcomes ω is a k-subset of the n-set of all beads, so there are $\binom{n}{k}$ outcomes, all equally likely. Let $X(\omega)$ denote the number of special beads in ω. Let x be an integer in $[0,k]$. We have previously calculated $P(X=x)$, using the following reasoning: there are $\binom{s}{x}$ ways to draw x of the s special beads and $\binom{n-s}{k-x}$ ways to draw $k-x$ of the $n-s$ non-special beads. We multiply these two numbers to count the number of ways of drawing k beads of which x are special. Then we form the ratio to the total number of outcomes to make this a probability. Thus:

$$P(X=x)=\frac{\binom{s}{x}\binom{n-s}{k-x}}{\binom{n}{k}}.$$

Problems 3.

a) In a blank Mathematica notebook, define a Mathematica function that calculate $f(x,n,s,k)$, the probability of getting x special beads, given that k beads are taken from a jar containing n beads, of which s are special. Note that in this problem, we are dealing with a family of different probability spaces, which vary with the **parameters** n, s, and k. (Hint: Begin: `f[x_,n_,s_,k_]:=`)
b) Using the function you have defined, find $P(30<X\leq 40)$ if X is the number of special beads in a sample of size 100 from a jar containing 1000 beads of which 250 are special.