Lecture 21. Independence and Conditioning

The following material on independence is from Chapter 8, section 2.

Definition. Let X and Y be discrete random variables on the same probability space. Suppose the joint pmf is $p(x,y)$ and the individual pmfs (i.e., the marginals) are $p_X(x)$ and $p_Y(y)$. We say X and Y are *independent* if

$$p(x,y) = p_X(x) \cdot p_Y(y).$$

Comment. This definition should remind you of what we said about recognizing independence in a table.

Problem. Suppose X and Y take the values $1,2,3$. Make up and depict in a table a joint pmf for X and Y such that X and Y are: i) independent, ii) not independent.

Definition. Let X and Y be continuous random variables on the same probability space. Suppose the joint pdf is $f(x,y)$ and the individual pdfs (i.e., the marginals) are $f_X(x)$ and $f_Y(y)$. We say X and Y are independent if

$$f(x,y) = f_X(x)\, f_Y(y).$$

Fact. The following are equivalent:

a) X and Y are independent.

b) For any measurable sets $A, B \subseteq \mathbb{R}$, $P(X \in A \,\&\, Y \in B) = P(X \in A) \cdot P(Y \in B)$

Problem. Suppose X and Y are continuous random variables whose joint pdf is constant on a region in the plane. If X and Y are independent, what properties must the shape of the region have? (Could it be a triangle? A rectangle? A disk?)

Problem. Buffon Needle. See Mathematica demo.

Fact (8.5). If X and Y are independent, then so are $g(X)$ and $h(Y)$ for any (measurable) functions $g, h : \mathbb{R} \to \mathbb{R}$.

Fact (8.6). Suppose X and Y are independent, and $g, h : \mathbb{R} \to \mathbb{R}$ are (measurable) functions. Then

$$\mathrm{E}(g(X)\, h(Y)) = \mathrm{E}(g(X))\, \mathrm{E}(h(Y)).$$

Example. We applied 8.5 and 8.6 when dealing with moment generating functions. Suppose X and Y are **independent** random variables on the same sample space. By definition, $M_X(t) = \mathrm{E}(e^{tX})$. Thus

$$\begin{aligned}
M_{X+Y}(t) &= \mathrm{E}(e^{t(X+Y)}) \\
&= \mathrm{E}(e^{tX} \cdot e^{tY}) \\
&= \mathrm{E}(e^{tX}) \cdot \mathrm{E}(e^{tY}) \\
&= M_X(t) \cdot M_Y(t)
\end{aligned}$$

The following material on conditioning is from Chapter 8, section 3.

According to the definition of conditional probability, if Ω is a probability space and $A, B \subseteq \Omega$ are events, then

$$P(A \mid B) := \frac{P(A \,\&\, B)}{P(A)}$$

If X is a random variable on a discrete probability space Ω and $B \subseteq \Omega$, then

$$P(X(\omega) = x \mid \omega \in B) := \frac{P(X(\omega) = x \,\&\, \omega \in B)}{P(\omega \in B)}$$

This is more commonly written:

$$P(X = x \mid B) = \frac{P(X = x \,\&\, B)}{P(B)}.$$

If Y is another random variable on the same space,

$$P(X = x \mid Y = y) = \frac{P(X = x \,\&\, Y = y)}{P(Y = y)} = \frac{p(x,y)}{p_Y(y)}.$$

The left hand side of this equation is called the *conditional pmf of X given $Y = y$*. It is denoted $p_{X|Y}(x \mid y)$.

Similarly, in the continuous case, we define the conditional pdf as follows:

$$f_{X|Y}(x \mid y) := \frac{f(x,y)}{f_Y(y)} \quad \text{provided } f_Y(y) \neq 0.$$

I find it easiest to understand conditioning by viewing it in terms of reduction of the sample space.