Lecture 20. Bivariate Distributions

This lecture is related to the material in Chapter 8, section 1.

Motivation.

The iLEAP/LEAP test is a standardized test given to all Louisiana public schools every April. In each grade, there are about 40,000 tested. There are only finitely many possible scores on these tests; each score is an integer between 100 and 500.

Let Ω be the set of all the students who took the 7th-grade math iLEAP in 2009 *and* the 8th-grade math LEAP 2010. If we choose a student at random from Ω, then each has equal probability—of about 1/40,000—of being chosen. If X is the score of the random student on the 7th-grade test, then $P(X = x)$ is the proportion of students who scored x on that test. Thus, Ω can be viewed as a probability space with random variable X (the 7th-grade score). There is also the 8th-grade score. If Y is the score of a random student on the 8th-grade test, then $P(Y = y)$ is the proportion of students who scored y on that test.

Now, suppose we measure both scores at the same time. Then $p(x, y) = P(X = x \,\&\, Y = y)$ can be interpreted as the proportion of students who scored x on the 7th-grade test *and* scored y on the 8th-grade test. We not expect to be able to deduce the joint distribution from the distribution of the scores of the 7th-grade tests (the pmf $p_X(x) = P(X = x)$) and the distribution of the scores of the 8th-grade tests (the pmf $p_Y(y) = P(Y = y)$), since kids who do well in 7th grade typically also do well in 8th-grade. Even if we know the proportion of students getting each possible 7th-grade score (that is, even if we know the $p_X(x)$) *and* we know the proportion of students getting each possible 8th-grade score (that is, even if we know $p_Y(y)$), we do not know the joint pmf $p(x, y)$.

It is only in the case that we know X and Y to be *independent* that we can deduce the joint pmf $p(x, y)$ from the two, so-called, *marginal* pmfs $P_X(x)$ and $P_Y(y)$. In the case that X and Y are independent,

$$p(x, y) = p_X(x) \cdot_Y (y).$$

This equality will fails to the extent there is some influence on one score on the other.

We will return to the topic of independence in the next lecture. In the present lecture, we do little more than establish basic notation.

The Discrete Case

Definition. Let Ω be a discrete probability space. Let $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ be two real-valued random variables. The *joint probability mass function for the variables X and Y* is the function

$$p(x, y) := P(X = x \,\&\, Y = y).$$

Let A be the set of possible values of X and let B is the set of possible values of Y. Then the values of $p(x, y)$ can be depicted in a table with rows labelled by the elements of A and columns labelled by the elements of B. (See the example below.)

The row and column sums give the so-called *marginal probabilities*. If $a \in A$ and $b \in B$:

$$p_X(a) := P(X = a) = \sum_{y \in B} p(a, y);$$

$$p_Y(b) := P(Y = b) = \sum_{x \in A} p(x, b).$$

Example. Suppose $\Omega = \{\alpha, \beta, \gamma, \delta, \epsilon\}$, with $p(\alpha) = 1/15$, $p(\beta) = 2/15$, $p(\gamma) = 3/15$, $p(\delta) = 4/15$, and $p(\epsilon) = 5/15$. Suppose

$$X(\omega) = \begin{cases} 1, & \text{if } \omega \in \{\alpha, \epsilon\}; \\ 2, & \text{if } \omega \in \{\beta, \gamma, \delta\}; \end{cases}$$

and

$$Y(\omega) = \begin{cases} 1, & \text{if } \omega \in \{\alpha, \gamma\}; \\ 2, & \text{if } \omega \in \{\beta, \delta\}; \\ 3, & \text{if } \omega = \epsilon. \end{cases}$$

We can represent the joint probability mass function in a table, as follows:

	$Y = 1$	$Y = 2$	$Y = 3$	row sum
$X = 1$	$1/15$	0	$5/15$	$6/15$
$X = 2$	$3/15$	$6/15$	0	$9/15$
column sum	$4/15$	$6/13$	$5/15$	$15/15$

Facts about Expectation.

The expectations of X and Y may be computed as follows:

$$\mathrm{E}(X) = \sum_{x \in A} x\, P(X = x) = \sum_{x \in A} x \left(\sum_{y \in B} p(x, y) \right) = \sum_{x \in A} \sum_{y \in B} x\, p(x, y),$$

$$\mathrm{E}(Y) = \sum_{y \in B} y\, P(Y = y) = \sum_{y \in B} y \left(\sum_{x \in A} p(x, y) \right) = \sum_{y \in B} \sum_{x \in A} y\, p(x, y).$$

Fact. Suppose $h : \mathbb{R}^2 \to \mathbb{R}$ is a function of two variables. Let $Z = h(X, Y)$. Then

$$\mathrm{E}(Z) = \mathrm{E}(h(X, Y)) = \sum_{x \in A} \sum_{y \in B} h(x, y) p(x, y)$$

Fact. $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$. (See p. 315.)

The Continuous Case

This is a natural generalization of the discrete case. Suppose Ω be a continuous probability space. Let $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ be two real-valued random variables. The *joint probability density function for the variables* X *and* Y is the function $f(x, y)$ such that

$$P(a < X < b \,\&\, c < Y < d) = \int_a^b \int_c^d f(x, y)\, dy\, dx.$$

The densities of X and Y separately are called the *marginal densities*. They are denoted $f_X(x)$ and $f_Y(y)$ respectively, and they are related to the joint density by:

$$f_X(a) = \int_{-\infty}^{\infty} f(a, y)\, dy,$$

$$f_Y(b) = \int_{-\infty}^{\infty} f(x, b)\, dx.$$

If h is a function of two variables, then $Z = h(X, Y)$ is a random variable. Its expectation is:

$$\mathrm{E}(Z) = \mathrm{E}(h(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y)\, f(x, y)\, dy\, dx.$$