Section 2.4. Independent Random Variables

Note. Recall from Definition 1.4.2 that two events A and B from a sample space are independent if $P(A \cap B) = P(A)P(B)$; this implies that $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$. In this section, we use the conditional probability functions developed in the previous section to define independent random variables. Several theorems are proved showing how computations can be simplified when two random variables are independent, and how to recognize if two random variables are independent.

Note 2.4.1. With X_1 and X_2 as random variables with joint probability density function $f(x_1, x_2)$ and marginal probability density functions $f_1(x_1)$ and $f_2(x)$, we have the conditional probability density functions

$$f_{X_1|X_2}(x_1 \mid x_2) = f_{1|2}(x_1 \mid x_2) = f(x_1, x_2)/f_2(x_2),$$

$$f_{X_2|X_1}(x_2 \mid x_1) = f_{2|1}(x_2 \mid x_1) = f(x_1, x_2)/f_1(x_1).$$

If $f_{2|1}(x_2 \mid x_1)$ does not depend on x_1, then

$$f_2(x_2) = \int_{-\infty}^{\infty} f_{2|1}(x_2 \mid x_1) f_x(x_1)\, dx_1 = f_{2|1}(x_2 \mid x_1) \int_{-\infty}^{\infty} f_1(x_1)\, dx_1 = f_{2|1}(x_2 \mid x_1).$$

So if $f_{2|1}(x_2 \mid x_1)$ does not depend on x_1 then

$$f_2(x_2) = f_{2|1}(x_2 \mid x_1) \text{ and so } f(x_1, x_2) = f(x_1)f(x_2).$$

We have similar results for discrete random variables. This motivates the following definition.

Definition 2.4.1. Let the continuous random variables X_1 and X_2 have the joint probability density function $f(x_1, x_2)$ and the marginal probability density functions $f_1(x_1)$ and $f_2(x_2)$. Random variables X_1 and X_2 are *independent* if and only if $f(x_1, x_2) \equiv f_1(x_1)f(x_2)$. Random variables that are not independent are *dependent*. We have similar definitions for discrete random variables.

Note. Hogg, McKean, and Craig comment on page 118 about the expression "$f(x_1, x_2) \equiv f_1(x_1)f_2(x_2)$. The claim that there may be points $(x_1, x_2) \in \mathcal{S}_1 \times \mathcal{S}_2$ which violate this equality, but that with A representing the set of all such points we have $P(A) = 0$. In measure theory, we would say that $f(x_1, x_2) = f_1(x_1)f_2(x_2)$ almost everywhere (see my online notes for Real Analysis 1 [MATH 5210] on 2.5. Countable Additivity, Continuity, and the Borel-Cantelli Lemma).

Example 2.4.2 Let the joint probability density function of X_1 and X_2 be

$$f(x_1, x_2) = \begin{cases} x_1 + x_2 & \text{for } 0 < x_1 < 1, 0 < x_2 < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

The marginal probability density functions are

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2 = \int_0^1 (x_1 + x_2)\, dx_2$$

$$= \left(x_1x_2 + \frac{1}{2}(x_2)^2 \right) \Bigg|_{x_2=0}^{x_2=1} = \begin{cases} x_1 + 1/2 & \text{for } 0 < x_1 < 1 \\ 0 & \text{elsewhere,} \end{cases}$$

and

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1 = \int_0^1 (x_1 + x_2)\, dx_1$$

$$= \left(\frac{1}{2}(x_1)^2 + x_1 x_2 \right) \Bigg|_{x_1=0}^{x_1=1} = \begin{cases} 1/2 + x_2 & \text{for } 0 < x_2 < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

Now

$$f_1(x_1)f_2(x_2) = (x_1+1/2)(x_2+1/2) = x_1x_2+x_2/2+x_2/2+1/4 \not\equiv x+1+x_2 = f(x_1,x_2)$$

so that X_1 and X_2 are not independent; X_1 and X_2 are dependent.

Note. The following theorem classifies two independent random variables in direct terms of the joint probability density function.

Theorem 2.4.1. Let the random variables X_1 and X_2 have supports $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively, and have the joint probability density function $f(x_1, x_2)$. Then X_1 and X_2 are independent if and only if $f(x_1, x_2)$ can be written as a product of a nonnegative function of x_1 and a nonnegative function of x_2. That is, $f(x_1, x_2) \equiv g(x_1)h(x_2)$ for some $g(x_1) > 0$ for $x_1 \in \mathcal{S}_1$ and 0 elsewhere, and some $h(x_2) > 0$ for $x_2 \in \mathcal{S}_2$ and 0 elsewhere.

Note. We see in the proof of Theorem 2.4.1 that some additional hypotheses on g and h are necessary. Namely, g and h must be measurable and integrable (that is, $\int_{-\infty}^{\infty} g(x_1)\,dx_1 < \infty$ and $\int_{-\infty}^{\infty} h(x_2)\,dx_2 < \infty$).

Example 2.4.3. There is a subtle requirement on the support of $f(x_1, x_2)$ implied in Theorem 2.4.1. Since $g(x_1) > 0$ if and only if $x_1 \in \mathcal{S}_1$ and $h(x_2) > 0$ if and only if $x_2 \in \mathcal{S}_2$, then $g(x_1)h(x_2) > 0$ if and only if $(x_1, x_2) \in \mathcal{S}_1 \times \mathcal{S}_2$. That is, for X_1 and X_2 to be independent we need the support of $f(x_1, x_2)$ to be a "product space") (...almost everywhere). For example, the probability density function $f(x_1, x_2) = 8x_1x_2$ for $0 < x_1 < x_2 < 1$, 0 elsewhere, appears to imply that X_1 and X_2 are independent by Theorem 2.4.1 (with $g(x_1) = 8x_1$ and $h(x_2) = x_2$, say). However we cannot find $g(x_1)$ and $h(x_2)$ with appropriate supports such that $f(x_1, x_2) \equiv g(x_1)h(x_2)$. So for $f(x_1, x_2) = 8x_1x_2$ with support $0 < x_1 < x_2 < 1$, the random variables X_1 and X_2 are dependent.

Note. The next result allows us to discuss independence of random variables in terms of cumulative distribution functions.

Theorem 2.4.2. Let (X_1, X_2) be a random vector with joint cumulative distribution function $F(x_1, x_2)$ and let X_1 and X_2 have the marginal cumulative distribution functions $F_1(x_1)$ and $F_2(x_2)$, respectively. Then X_1 and X_2 are independent if and only if $F(x_1, x_2) = F_1(x_1)F_2(x_2)$ for all $(x_1, x_2) \in \mathbb{R}^2$.

Note. The next theorem allows us to calculate probabilities over "rectangles" for independent random variables.

Theorem 2.4.3. The random variables X_1 and X_2 are independent random variables if and only if

$$P(a < X_1 \leq b, c < X_2 \leq d) = P(a < X_1 \leq b)P(c < X_2 \leq d)$$

for every $a < b$ and $c < d$, where a, b, c, d are constants.

Note. In Example 2.4.4 (page 122) it is observed that for the probability density function of Example 2.4.2, in which X_1 and X_2 are dependent, the corresponding cumulative distribution functions violate Theorem 2.4.3. So this gives a specific example showing that independence is necessary for the probability equation in Theorem 2.4.3 to hold.

Note. The previous theorem and the next two theorems show how certain computations are simplified for independent variables.

Theorem 2.4.4. Suppose X_1 and X_2 are independent and that $E[u(X_1)]$ and $E[v(X_2)]$ exists. Then

$$E[u(X_1)v(X_2)] = E[u(X_1)]E[v(X_2)].$$

Theorem 2.4.5. Suppose the joint moment generating function $M(t_1, t_2)$ exists for the random variables X_1 and X_2. Then X_1 and X_2 are independent if and only if $M(t_1, t_2) = M(t_1, 0)M(0, t_2)$; that is, the joint moment generating function is identically equal to the product of the marginal moment generating functions.

Example. Exercise 2.4.6(a).