Section 5.6 Independent random variables

Tuesday, November 15, 2016 12:30 PM

Section 5.6 Independ...

(b) To compute $P(Y_1 > 0.5|Y_2 = 0.3)$, we work with the conditional pdf $f_{Y_1|Y_2}(y_1|y_2)$, which for $y_2 = 0.3$, is given by

$$f_{Y_1|Y_2}(y_1|y_2) = \begin{cases} \left(\frac{200}{49}\right) y_1, & 0 < y_1 < 0.7 \\ 0, & \text{otherwise.} \end{cases}$$

Thus,

$$P(Y_1 > 0.5|Y_2 = 0.3) = \int_{0.5}^{0.7} \left(\frac{200}{49}\right) y_1 dy_1 \approx 0.489.$$

(c) To compute $P(Y_1 > 0.5)$, we can either use the marginal pdf $f_{Y_1}(y_1)$ or the joint pdf $f_{Y_1,Y_2}(y_1, y_2)$. Marginally, it turns out that $Y_1 \sim \text{beta}(2, 3)$ as well (verify!). Thus,

$$P(Y_1 > 0.5) = \int_{0.5}^{1} 12y_1(1 - y_1)^2 dy_1 \approx 0.313.$$

REMARK: Notice how $P(Y_1 > 0.5|Y_2 = 0.3) \neq P(Y_1 > 0.5)$; that is, knowledge of the value of Y_2 has affected the way that we assign probability to events involving Y_1. Of course, one might expect this because of the support in the joint pdf $f_{Y_1,Y_2}(y_1, y_2)$. □

5.6 Independent random variables

TERMINOLOGY: Suppose (Y_1, Y_2) is a random vector (discrete or continuous) with joint cdf $F_{Y_1,Y_2}(y_1, y_2)$, and denote the marginal cdfs of Y_1 and Y_2 by $F_{Y_1}(y_1)$ and $F_{Y_2}(y_2)$, respectively. We say that the random variables Y_1 and Y_2 are **independent** if and only if

$$P(Y_1 \leq y_1, Y_2 \leq y_2) = F_{Y_1,Y_2}(y_1, y_2) = F_{Y_1}(y_1)F_{Y_2}(y_2) \qquad P(Y_1 \leq y_1) \times P(Y_2 \leq y_2)$$

for all values of y_1 and y_2. Otherwise, we say that Y_1 and Y_2 are **dependent**.

RESULT: Suppose that (Y_1, Y_2) is a random vector (discrete or continuous) with joint pdf (pmf) $f_{Y_1,Y_2}(y_1, y_2)$, and denote the marginal pdfs (pmfs) of Y_1 and Y_2 by $f_{Y_1}(y_1)$ and $f_{Y_2}(y_2)$, respectively. Then, Y_1 and Y_2 are independent if and only if

$$f_{Y_1,Y_2}(y_1, y_2) = f_{Y_1}(y_1)f_{Y_2}(y_2)$$

for all values of y_1 and y_2. Otherwise, Y_1 and Y_2 are dependent.

Proof. Exercise. □

Example 5.7. Suppose that the pmf for the discrete random vector (Y_1, Y_2) is given by

$$p_{Y_1,Y_2}(y_1, y_2) = \begin{cases} \frac{1}{18}(y_1 + 2y_2), & y_1 = 1, 2,\ y_2 = 1, 2 \\ 0, & \text{otherwise.} \end{cases}$$

The marginal distribution of Y_1, for values of $y_1 = 1, 2$, is given by

$$p_{Y_1}(y_1) = \sum_{y_2=1}^{2} p_{Y_1,Y_2}(y_1, y_2) = \sum_{y_2=1}^{2} \frac{1}{18}(y_1 + 2y_2) = \frac{1}{18}(2y_1 + 6),$$

and $p_{Y_1}(y_1) = 0$, otherwise. Similarly, the marginal distribution of Y_2, for values of $y_2 = 1, 2$, is given by

$$p_{Y_2}(y_2) = \sum_{y_1=1}^{2} p_{Y_1,Y_2}(y_1, y_2) = \sum_{y_1=1}^{2} \frac{1}{18}(y_1 + 2y_2) = \frac{1}{18}(3 + 4y_2),$$

and $p_{Y_2}(y_2) = 0$, otherwise. Note that, for example,

$$\frac{3}{18} = p_{Y_1,Y_2}(1, 1) \neq p_{Y_1}(1)p_{Y_2}(1) = \frac{8}{18} \times \frac{7}{18} = \frac{14}{81};$$

thus, the random variables Y_1 and Y_2 are dependent. $\square$

Example 5.8. Let Y_1 and Y_2 denote the proportions of time (out of one workday) during which employees I and II, respectively, perform their assigned tasks. Suppose that the random vector (Y_1, Y_2) has joint pdf

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} y_1 + y_2, & 0 < y_1 < 1,\ 0 < y_2 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

$f_{Y_1}(y_1) = \int_0^1 f(y_1, y_2)\,dy_2$

$= \int_0^1 (y_1 + y_2)\,dy_2$

$= y_1 + \int_0^1 y_2\,dy_2$

$= y_1 + \frac{y_2^2}{2}\Big|_0^1$

$= y_1 + \frac{1}{2}$

It is straightforward to show (verify!) that

$$f_{Y_1}(y_1) = \begin{cases} y_1 + \frac{1}{2}, & 0 < y_1 < 1 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad f_{Y_2}(y_2) = \begin{cases} y_2 + \frac{1}{2}, & 0 < y_2 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Thus, since

$$f_{Y_1,Y_2}(y_1, y_2) = y_1 + y_2 \neq \left(y_1 + \frac{1}{2}\right)\left(y_2 + \frac{1}{2}\right) = f_{Y_1}(y_1)f_{Y_2}(y_2),$$

for $0 < y_1 < 1$ and $0 < y_2 < 1$, Y_1 and Y_2 are dependent. $\square$

A CONVENIENT RESULT: Let (Y_1, Y_2) be a random vector (discrete or continuous) with pdf (pmf) $f_{Y_1,Y_2}(y_1, y_2)$, If the support set R does not constrain y_1 by y_2 (or y_2 by y_1), and additionally, we can factor the joint pdf (pmf) $f_{Y_1,Y_2}(y_1, y_2)$ into two nonnegative expressions

$$f_{Y_1,Y_2}(y_1, y_2) = g(y_1)h(y_2),$$

then Y_1 and Y_2 are independent. Note that $g(y_1)$ and $h(y_2)$ are simply functions; **they need not be pdfs (pmfs)**, although they sometimes are. The only requirement is that $g(y_1)$ is a function of y_1 only, $h(y_2)$ is a function of y_2 only, and that both are nonnegative. *If the support involves a constraint, the random variables are automatically dependent.*

Example 5.9. In Example 5.6, Y_1 denoted the amount of brand 1 grain in stock and Y_2 denoted the amount of brand 2 grain in stock. Recall that the joint pdf of (Y_1, Y_2) was given by

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} 24y_1y_2, & y_1 > 0,\ y_2 > 0,\ 0 < y_1 + y_2 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Here, the support is $R = \{(y_1, y_2) : y_1 > 0,\ y_2 > 0,\ 0 < y_1 + y_2 < 1\}$. Since knowledge of y_1 affects the value of y_2, and vice versa, the support involves a constraint, and Y_1 and Y_2 are dependent. $\square$

Example 5.10. Suppose that the random vector (X, Y) has joint pdf

$$f_{X,Y}(x, y) = \begin{cases} [\Gamma(\alpha)\Gamma(\beta)]^{-1}\lambda e^{-\lambda x}(\lambda x)^{\alpha+\beta-1}y^{\alpha-1}(1-y)^{\beta-1}, & x > 0, 0 < y < 1 \\ 0, & \text{otherwise.} \end{cases}$$

for $\lambda > 0$, $\alpha > 0$, and $\beta > 0$. Since $R = \{(x, y) : x > 0,\ 0 < y < 1\}$ does not involve a constraint, it follows immediately that X and Y are independent, since we can write

$$f_{X,Y}(x, y) = \underbrace{\lambda e^{-\lambda x}(\lambda x)^{\alpha+\beta-1}}_{g(x)} \times \underbrace{\frac{y^{\alpha-1}(1-y)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}}_{h(y)},$$

where $g(x)$ and $h(y)$ are nonnegative functions. Note that we are not saying that $g(x)$ and $h(y)$ are marginal distributions of X and Y, respectively (in fact, they are not the marginal distributions, although they are proportional to the marginals). $\square$

EXTENSION: We generalize the notion of **independence** to n-variate random vectors. We use the conventional notation

$$\boldsymbol{Y} = (Y_1, Y_2, ..., Y_n)$$

and

$$\boldsymbol{y} = (y_1, y_2, ..., y_n).$$

We denote the joint cdf of $\boldsymbol{Y}$ by $F_{\boldsymbol{Y}}(\boldsymbol{y})$ and the joint pdf (pmf) of $\boldsymbol{Y}$ by $f_{\boldsymbol{Y}}(\boldsymbol{y})$.

TERMINOLOGY: Suppose that the random vector $\boldsymbol{Y} = (Y_1, Y_2, ..., Y_n)$ has joint cdf $F_{\boldsymbol{Y}}(\boldsymbol{y})$, and suppose that the random variable Y_i has cdf $F_{Y_i}(y_i)$, for $i = 1, 2, ..., n$. Then, $Y_1, Y_2, ..., Y_n$ are independent random variables if and only if

$$F_{\boldsymbol{Y}}(\boldsymbol{y}) = \prod_{i=1}^{n} F_{Y_i}(y_i);$$

that is, the joint cdf can be factored into the product of the marginal cdfs. Alternatively, $Y_1, Y_2, ..., Y_n$ are independent random variables if and only if

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \prod_{i=1}^{n} f_{Y_i}(y_i);$$

that is, the joint pdf (pmf) can be factored into the product of the marginals.

Example 5.11. In a small clinical trial, $n = 20$ patients are treated with a new drug. Suppose that the response from each patient is a measurement $Y \sim \mathcal{N}(\mu, \sigma^2)$. Denoting the 20 responses by $\boldsymbol{Y} = (Y_1, Y_2, ..., Y_{20})$, then, assuming independence, the joint distribution of the 20 responses is, for $\boldsymbol{y} \in \mathcal{R}^{20}$,

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \prod_{i=1}^{20} \underbrace{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{y_i-\mu}{\sigma}\right)^2}}_{f_{Y_i}(y_i)} = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{20} e^{-\frac{1}{2}\sum_{i=1}^{20}\left(\frac{y_i-\mu}{\sigma}\right)^2}.$$

What is the probability that at least one patient's response is greater than $\mu + 2\sigma$?

SOLUTION. Define the event

$$B = \{\text{at least one patient's response exceeds } \mu + 2\sigma\}.$$

We want to compute $P(B)$. Note that

$P(\overline{B}) = P(Y_1 < \mu+2\sigma, Y_2 < \mu+2\sigma, \vdots, Y_{20} < \mu+2\sigma)$

$$\overline{B} = \{\text{all 20 responses are less than } \mu + 2\sigma\}$$

and recall that $P(B) = 1 - P(\overline{B})$. We will compute $P(\overline{B})$ because it is easier. The probability that the first patient's response Y_1 is less than $\mu + 2\sigma$ is given by

indep. $= \prod_{i=1}^{20} P(Y_i < \mu+2\sigma)$

$$F_{Y_1}(\mu + 2\sigma) = P(Y_1 < \mu + 2\sigma) = P(Z < 2) = F_Z(2) = 0.9772,$$

where $Z \sim \mathcal{N}(0, 1)$ and $F_Z(\cdot)$ denotes the standard normal cdf. This probability is same for each patient, because each patient's response follows the same $\mathcal{N}(\mu, \sigma^2)$ distribution. Because the patients' responses are independent random variables,

$Y_i \sim N(\mu, \sigma^2)$

$P(Y_i < \mu+2\sigma)$

$$\begin{aligned} P(\overline{B}) &= P(Y_1 < \mu + 2\sigma, Y_2 < \mu + 2\sigma, ..., Y_{20} < \mu + 2\sigma) \\ &= \prod_{i=1}^{20} F_{Y_i}(\mu + 2\sigma) \\ &= [F_Z(2)]^{20} \approx 0.63. \end{aligned}$$

$P\left(\frac{Y_i-\mu}{\sigma} < \frac{\mu+2\sigma-\mu}{\sigma}\right)$

$= P(Z < 2)$

normalcdf$(-10^{99}, 2, 0, 1)$

Finally, $P(B) = 1 - P(\overline{B}) \approx 1 - 0.63 = 0.37$. $\square$

5.7 Expectations of functions of random variables

RESULT: Suppose that $\boldsymbol{Y} = (Y_1, Y_2, ..., Y_n)$ has joint pdf $f_{\boldsymbol{Y}}(\boldsymbol{y})$, or joint pmf $p_{\boldsymbol{Y}}(\boldsymbol{y})$, and suppose that $g(\boldsymbol{Y}) = g(Y_1, Y_2, ..., Y_n)$ is a real vector valued function of $Y_1, Y_2, ..., Y_n$; i.e., $g : \mathcal{R}^n \to \mathcal{R}$. Then,

- if $\boldsymbol{Y}$ is discrete,

$$E[g(\boldsymbol{Y})] = \sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} g(\boldsymbol{y}) p_{\boldsymbol{Y}}(\boldsymbol{y}),$$

- and if $\boldsymbol{Y}$ is continuous,

$$E[g(\boldsymbol{Y})] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\boldsymbol{y}) f_{\boldsymbol{Y}}(\boldsymbol{y}) d\boldsymbol{y}.$$

If these quantities are not finite, then we say that $E[g(\boldsymbol{Y})]$ does not exist.