ECS315 2012/1 Part III.3 Dr.Prapun

10 Multiple Random Variables

One is often interested not only in individual random variables, but also in relationships between two or more random variables. Furthermore, one often wishes to make inferences about one random variable on the basis of observations of other random variables.

Example 10.1. If the experiment is the testing of a new medicine, the researcher might be interested in cholesterol level, blood pressure, and the glucose level of a test person.

10.1 A Pair of Random Variables

In this section, we consider two random variables, say X and Y, simultaneously.

10.2. The analysis are different from Section 9.2 in two main aspects. First, there may be *no* deterministic relationship (such as $Y = g(X)$) between the two random variables. Second, we want to look at both random variables as a whole, not just X alone or Y alone.

Example 10.3. Communication engineers may be interested in the input and output of a channel.

Example 10.4. Of course, to rigorously define (any) random variables, we need to go back to the sample space Ω. Recall Example 7.4 where we considered several random variables defined on the sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$ where the outcomes are equally likely. In that example, we define $X(\omega) = \omega$ and $Y(\omega) = (\omega - 3)^2$.

Example 10.5. Consider the scores of 20 students below:

$$\underbrace{10, 9, 10, 9, 9, 10, 9, 10, 10, 9}_{\text{Room \#1}}, \underbrace{1, 3, 4, 6, 5, 5, 3, 3, 1, 3.}_{\text{Room \#2}}$$

The first ten scores are from (ten) students in room #1. The last 10 scores are from (ten) students in room #2.

Suppose we have the a score report card for each student. Then, in total, we have 20 report cards.

Figure 6: In Example 10.5, we pick a report card randomly from a pile of cards.

I pick one report card up randomly. Let X be the score on that card.

- What is the chance that $X > 5$? (Ans: $P[X > 5] = 11/20$.)
- What is the chance that $X = 10$? (Ans: $p_X(10) = P[X = 10] = 5/20 = 1/4$.)

Now, let the random variable Y denote the room# of the student whose report card is picked up.

- What is the probability that $X = 10$ and $Y = 2$?

- What is the probability that $X = 10$ and $Y = 1$?

- What is the probability that $X > 5$ and $Y = 1$?

- What is the probability that $X > 5$ and $Y = 2$?

Now suppose someone informs me that the report card that I picked up is from a student in room #1. (He may be able to tell this by the color of the report card of which I have no knowledge.) I now have an extra information that $Y = 1$.

- What is the probability that $X > 5$ given that $Y = 1$?

- What is the probability that $X = 10$ given that $Y = 1$?

Definition 10.6. If X and Y are random variables, we use the shorthand[30]

$$[X \in B, Y \in C] = [X \in B \text{ and } Y \in C] = [X \in B] \cap [Y \in C]$$

Observe that the "," in $[X \in B, Y \in B]$ means "and". Consequently,

$$\begin{aligned} P[X \in B, Y \in C] &= P[X \in B \text{ and } Y \in C] \\ &= P([X \in B] \cap [Y \in C]). \end{aligned}$$

Similarly, the concept of conditional probability can be straightforwardly applied to random variables via

$$\begin{aligned} P[X \in B | Y \in C] = P([X \in B] | [Y \in C]) &= \frac{P([X \in B] \cap [Y \in C])}{P([Y \in C])} \\ &= \frac{P[X \in B, Y \in C]}{P[Y \in C]}. \end{aligned}$$

Example 10.7. We also have

$$\begin{aligned} P[X = x, Y = y] &= P[X = x \text{ and } Y = y], \\ P[X = x | Y = y] &= \frac{P[X = x \text{ and } Y = y]}{P[Y = y]}, \end{aligned}$$

and

$$\begin{aligned} P[3 \le X < 4, Y < 1] &= P[3 \le X < 4 \text{ and } Y < 1] \\ &= P[X \in [3,4) \text{ and } Y \in (-\infty, 1)]. \\ P[3 \le X < 4 | Y < 1] &= \frac{P[3 \le X < 4 \text{ and } Y < 1]}{P[Y < 1]} \end{aligned}$$

[30]Linking back to the original sample space, this shorthand actually says

$$\begin{aligned} [X \in B, Y \in C] &= [X \in B \text{ and } Y \in C] \\ &= \{\omega \in \Omega : X(\omega) \in B \text{ and } Y(\omega) \in C\} \\ &= \{\omega \in \Omega : X(\omega) \in B\} \cap \{\omega \in \Omega : Y(\omega) \in C\} \\ &= [X \in B] \cap [Y \in C]. \end{aligned}$$

Definition 10.8. ***Joint pmf***: If X and Y are two discrete random variables (defined on a same sample space with probability measure P), the function $p_{X,Y}(x,y)$ defined by

$$p_{X,Y}(x,y) = P\left[X = x, Y = y\right]$$

is called the ***joint probability mass function*** of X and Y. We can then evaluate $P\left[(X,Y) \in R\right]$ by

$$P\left[(X,Y) \in R\right] = \sum_{(x,y):(x,y)\in R} p_{X,Y}(x,y).$$

Definition 10.9. The ***conditional pmf*** of X given Y is defined as

$$p_{X|Y}(x|y) = P\left[X = x|Y = y\right]$$

which gives

$$p_{X,Y}(x,y) = p_{X|Y}(x|y)p_Y(y) = p_{Y|X}(y|x)p_X(x).$$

Example 10.10. Consider a binary symmetric channel. Suppose the input X to the channel is bernoulli(0.3). At the output Y of this channel, the crossover (bit-flipped) probability is 0.1. Find the the joint pmf $p_{X,Y}(x,y)$ of X and Y.

In most situation, it is much more convenient to focus on the "important" part of the joint pmf. To do this, we usually present the joint pmf (and the conditional pmf) in their matrix forms:

Definition 10.11. When both X and Y take finitely many values (both have finite supports), say $S_X = \{x_1, \ldots, x_m\}$ and $S_Y = \{y_1, \ldots, y_n\}$, respectively, we can arrange the probabilities $p_{X,Y}(x_i, y_j)$ in an $m \times n$ matrix

$$\begin{bmatrix} p_{X,Y}(x_1,y_1) & p_{X,Y}(x_1,y_2) & \ldots & p_{X,Y}(x_1,y_n) \\ p_{X,Y}(x_2,y_1) & p_{X,Y}(x_2,y_2) & \ldots & p_{X,Y}(x_2,y_n) \\ \vdots & \vdots & \ddots & \vdots \\ p_{X,Y}(x_m,y_1) & p_{X,Y}(x_m,y_2) & \ldots & p_{X,Y}(x_m,y_n) \end{bmatrix}.$$

- We shall call this matrix the ***joint pmf matrix***.
- The sum of all the entries in the matrix is one.

- The sum of the entries in the ith row is[31] $p_X(x_i)$, and the sum of the entries in the jth column is $p_Y(y_j)$:

$$p_X(x_i) = \sum_{j=1}^{n} p_{X,Y}(x_i, y_j) \quad \text{and} \quad p_Y(y_j) = \sum_{i=1}^{m} p_{X,Y}(x_i, y_j)$$

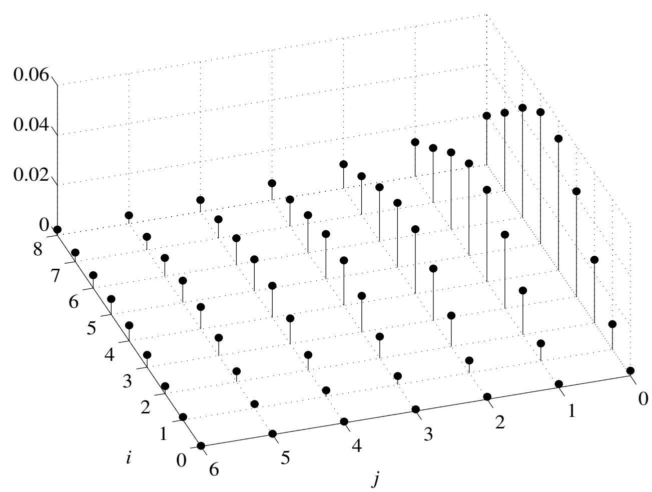

Figure 7: Example of the plot of a joint pmf. [9, Fig. 2.8]

[31] To see this, we consider $A = [X = x_i]$ and a collection defined by $B_j = [Y = y_j]$ and $B_0 = [Y \notin S_Y]$. Note that the collection $B_0, B_1, \ldots, B_n$ partitions Ω. So, $P(A) = \sum_{j=0}^{n} P(A \cap B_j)$. Of course, because the support of Y is S_Y, we have $P(A \cap B_0) = 0$. Hence, the sum can start at $j = 1$ instead of $j = 0$.

- $p_{X,Y}(x,y) = 0$ if[32] $x \notin S_X$ or $y \notin S_Y$. In other words, we don't have to consider the x and y outside the supports of X and Y, respectively.

Exercise 10.12. Toss-and-Roll Game:

Step 1 Toss a fair coin. Define X by

$$X = \begin{cases} 1, & \text{if result} = \text{H}, \\ 0, & \text{if result} = \text{T}. \end{cases}$$

Step 2 You have two dice, Dice 1 and Dice 2. Dice 1 is fair. Dice 2 is unfair with $p(1) = p(2) = p(3) = \frac{2}{9}$ and $p(4) = p(5) = p(6) = \frac{1}{9}$.

(i) If $X = 0$, roll Dice 1.

(ii) If $X = 1$, roll Dice 2.

Record the result as Y.

Find the joint pmf $p_{X,Y}(x,y)$ of X and Y.

10.13. From the joint pmf, we can find $p_X(x)$ and $p_Y(y)$ by

$$p_X(x) = \sum_y p_{X,Y}(x,y) \tag{18}$$

$$p_Y(y) = \sum_x p_{X,Y}(x,y) \tag{19}$$

In this setting, $p_X(x)$ and $p_Y(y)$ are call the ***marginal pmfs*** (to distinguish them from the joint one).

In MATLAB, if we define the joint pmf matrix as `P_XY`, then the marginal pmf (row) vectors `p_X` and `p_Y` can be found by

```
p_X = (sum(P_XY,2))'
p_Y = (sum(P_XY,1))
```

[32] To see this, note that $p_{X,Y}(x,y)$ can not exceed $p_X(x)$ because $P(A \cap B) \le P(A)$. Now, suppose at $x = a$, we have $p_X(a) = 0$. Then $p_{X,Y}(a,y)$ must also $= 0$ for any y because it can not exceed $p_X(a) = 0$. Similarly, suppose at $y = a$, we have $p_Y(a) = 0$. Then $p_{X,Y}(x,a) = 0$ for any x.

Example 10.14. Consider the following joint pmf matrix

Exercise 10.15 (F2011)**.** Random variables X and Y have the following joint pmf

$$p_{X,Y}(x,y) = \begin{cases} c\,(x+y), & x \in \{1,3\} \text{ and } y \in \{2,4\}, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Check that $c = 1/20$.

(b) Find $P\left[X^2 + Y^2 = 13\right]$.

(c) Find $p_X(x)$.

(d) Find $\mathbb{E}X$.

(e) Find $p_{Y|X}(y|1)$. Note that your answer should be of the form

$$p_{Y|X}(y|1) = \begin{cases} ?, & y = 2, \\ ?, & y = 4, \\ 0, & \text{otherwise.} \end{cases}$$

(f) Find $p_{Y|X}(y|3)$.

Definition 10.16. The ***joint cdf*** of X and Y is defined by

$$F_{X,Y}(x,y) = P\left[X \le x, Y \le y\right].$$

Definition 10.17. Two random variables X and Y are said to be ***identically distributed*** if, for every B, $P[X \in B] = P[Y \in B]$.

10.18. The following statements are equivalent:

(a) Random variables X and Y are ***identically distributed***.

(b) For every B, $P[X \in B] = P[Y \in B]$

(c) $p_X(c) = p_Y(c)$ for all c

(d) $F_X(c) = F_Y(c)$ for all c

Definition 10.19. Two random variables X and Y are said to be ***independent*** if the events $[X \in B]$ and $[Y \in C]$ are independent for all sets B and C.

10.20. The following statements are equivalent:

(a) Random variables X and Y are ***independent***.

(b) $[X \in B] \perp\!\!\!\perp [Y \in C]$ for all B, C.

(c) $P[X \in B, Y \in C] = P[X \in B] \times P[Y \in C]$ for all B, C.

(d) $p_{X,Y}(x, y) = p_X(x) \times p_Y(y)$ for all x, y.

(e) $F_{X,Y}(x, y) = F_X(x) \times F_Y(y)$ for all x, y.

Definition 10.21. Two random variables X and Y are said to be ***independent and identically distributed (i.i.d.)*** if X and Y are both independent and identically distributed.

10.22. Being identically distributed does not imply independence. Similarly, being independent, does not imply being identically distributed.

Example 10.23. Roll a dice. Let X be the result. Set $Y = X$.

Example 10.24. Suppose the pmf of a random variable X is given by

$$p_X(x) = \begin{cases} 1/4, & x = 3, \\ \alpha, & x = 4, \\ 0, & \text{otherwise.} \end{cases}$$

Let Y be another random variable. Assume that X and Y are i.i.d.

Find

(a) α,

(b) the pmf of Y, and

(c) the joint pmf of X and Y.

Example 10.25. Consider a pair of random variables X and Y whose joint pmf is given by

$$p_{X,Y}(x,y) = \begin{cases} 1/15, & x=3, y=1, \\ 2/15, & x=4, y=1, \\ 4/15, & x=3, y=3, \\ \beta, & x=4, y=3, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Are X and Y identically distributed?

(b) Are X and Y independent?

10.2 Extending the Definitions to Multiple RVs

Definition 10.26. Joint pmf:

$$p_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = P\left[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\right].$$

Joint cdf:

$$F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = P\left[X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n\right].$$

10.27. Marginal pmf:

Definition 10.28. ***Identically distributed*** random variables: The following statements are equivalent.

(a) Random variables $X_1, X_2, \ldots$ are ***identically distributed***

(b) For every B, $P\left[X_j \in B\right]$ does not depend on j.

(c) $p_{X_i}(c) = p_{X_j}(c)$ for all c, i, j.

(d) $F_{X_i}(c) = F_{X_j}(c)$ for all c, i, j.

Definition 10.29. ***Independence*** among finite number of random variables: The following statements are equivalent.

(a) $X_1, X_2, \ldots, X_n$ are ***independent***

(b) $[X_1 \in B_1], [X_2 \in B_2], \ldots, [X_n \in B_n]$ are independent, for all $B_1, B_2, \ldots, B_n$.

(c) $P\left[X_i \in B_i, \forall i\right] = \prod_{i=1}^{n} P\left[X_i \in B_i\right]$, for all $B_1, B_2, \ldots, B_n$.

(d) $p_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = \prod_{i=1}^{n} p_{X_i}(x_i)$ for all $x_1, x_2, \ldots, x_n$.

(e) $F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)$ for all $x_1, x_2, \ldots, x_n$.

Example 10.30. Toss a coin n times. For the ith toss, let

$$X_i = \begin{cases} 1, & \text{if H happens on the } i\text{th toss,} \\ 0, & \text{if T happens on the } i\text{th toss.} \end{cases}$$

We then have a collection of i.i.d. random variables $X_1, X_2, X_3, \ldots, X_n$.

Example 10.31. Roll a dice n times. Let N_i be the result of the ith roll. We then have another collection of i.i.d. random variables $N_1, N_2, N_3, \ldots, N_n$.

Example 10.32. Let X_1 be the result of tossing a coin. Set $X_2 = X_3 = \cdots = X_n = X_1$.

10.33. If $X_1, X_2, \ldots, X_n$ are independent, then so is any subcollection of them.

10.34. For i.i.d. $X_i \sim$ Bernoulli(p), $Y = X_1 + X_2 + \cdots + X_n$ is $\mathcal{B}(n, p)$.

Definition 10.35. A ***pairwise independent*** collection of random variables is a collection of random variables any two of which are independent.

(a) Any collection of (mutually) independent random variables is pairwise independent

(b) Some pairwise independent collections are not independent. See Example (10.36).

Example 10.36. Let suppose X, Y, and Z have the following joint probability distribution: $p_{X,Y,Z}(x, y, z) = \frac{1}{4}$ for $(x, y, z) \in \{(0,0,0), (0,1,1), (1,0,1), (1,1,0)\}$. This, for example, can be constructed by starting with independent X and Y that are Bernoulli-$\frac{1}{2}$. Then set $Z = X \oplus Y = X + Y \mod 2$.

(a) X, Y, Z are pairwise independent.

(b) X, Y, Z are not independent.