1.9 Two-Dimensional Random Variables

Definition 1.14. *Let Ω be a sample space and X_1, X_2 be functions, each assigning a real number $X_1(\omega), X_2(\omega)$ to every outcome $\omega \in \Omega$, that is $X_1 : \Omega \to \mathcal{X}_1 \subset \mathbb{R}$ and $X_2 : \Omega \to \mathcal{X}_2 \subset \mathbb{R}$. Then the pair $\boldsymbol{X} = (X_1, X_2)$ is called a two-dimensional random variable. The induces sample space (range) of the two-dimensional random variable is*

$$\mathcal{X} = \{(x_1, x_2) : x_1 \in \mathcal{X}_1,\ x_2 \in \mathcal{X}_2\} \subset \mathbb{R}^2.$$

□

We will denote two-dimensional random variables by bold capital letters.

Definition 1.15. *The cumulative distribution function of a two-dimensional rv $\boldsymbol{X} = (X_1, X_2)$ is*

$$F_{\boldsymbol{X}}(x_1, x_2) = P(X_1 \le x_1, X_2 \le x_2) \tag{1.10}$$

□

1.9.1 Discrete Two-Dimensional Random Variables

If all values of $\boldsymbol{X} = (X_1, X_2)$ are countable, i.e., the values are in the range

$$\mathcal{X} = \{(x_{1i}, x_{2j}),\ i = 1, 2, \ldots,\ j = 1, 2, \ldots\}$$

then the variable is discrete. The cdf of a discrete rv $\boldsymbol{X} = (X_1, X_2)$ is

$$F_{\boldsymbol{X}}(x_1, x_2) = \sum_{x_{2j} \le x_2} \sum_{x_{1i} \le x_1} p(x_{1i}, x_{2j})$$

where $p(x_{1i}, x_{2j})$ denotes the *joint probability mass function* and

$$p(x_{1i}, x_{2j}) = P(X_1 = x_{1i}, X_2 = x_{2j}).$$

As in the univariate case, the joint pmf satisfies the following conditions.

1. $p(x_{1i}, x_{2j}) \ge 0$, for all i, j
2. $\sum_{\mathcal{X}_2} \sum_{\mathcal{X}_1} p(x_{1i}, x_{2j}) = 1$

Example 1.18. Consider an experiment of tossing two fair dice and noting the outcome on each die. The whole sample space consists of 36 elements, i.e.,

$$\Omega = \{(i, j) : i, j = 1, \dots, 6\}.$$

Now, with each of these 36 elements associate values of two random variables, X_1 and X_2, such that

$$\begin{aligned} X_1 &\equiv sum\ of\ the\ outcomes\ on\ the\ two\ dice, \\ X_2 &\equiv |\, difference\ of\ the\ outcomes\ on\ the\ two\ dice\,|. \end{aligned}$$

That is,

$$\boldsymbol{X}(i, j) = (i + i, |i - j|) \quad i, j = 1, 2, \dots, 6.$$

Then, the bivariate rv $\boldsymbol{X} = (X_1, X_2)$ has the following joint probability mass function (empty cells mean that the pmf is equal to zero at the relevant values of the rvs).

						x_1						
		2	3	4	5	6	7	8	9	10	11	12
x_2	0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
	1		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$	
	2			$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		
	3				$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$			
	4					$\frac{1}{18}$		$\frac{1}{18}$				
	5						$\frac{1}{18}$					

$\square$

Expectations of functions of bivariate random variables are calculated the same way as of the univariate rvs. Let $g(x_1, x_2)$ be a real valued function defined on $\mathcal{X}$. Then $g(\boldsymbol{X}) = g(X_1, X_2)$ is a rv and its expectation is

$$\mathrm{E}[g(\boldsymbol{X})] = \sum_{\mathcal{X}} g(x_1, x_2) p(x_1, x_2).$$

Example 1.19. Let X_1 and X_2 be random variables as defined in Example 1.18. Then, for $g(X_1, X_2) = X_1 X_2$ we obtain

$$\mathrm{E}[g(\boldsymbol{X})] = 2 \times 0 \times \frac{1}{36} + \dots + 7 \times 5 \times \frac{1}{18} = \frac{245}{18}.$$

$\square$

Marginal pmfs

Each of the components of the two-dimensional rv is a random variable and so we may be interested in calculating its probabilities, for example $P(X_1 = x_1)$. Such a uni-variate pmf is then derived in a context of the distribution of the other random variable. We call it the *marginal pmf*.

Theorem 1.12. *Let $\boldsymbol{X} = (X_1, X_2)$ be a discrete bivariate random variable with joint pmf $p(x_1, x_2)$. Then the marginal pmfs of X_1 and X_2, p_{X_1} and p_{X_2}, are given respectively by*

$$p_{X_1}(x_1) = P(X_1 = x_1) = \sum_{\mathcal{X}_2} p(x_1, x_2) \quad \textit{and}$$
$$p_{X_2}(x_2) = P(X_2 = x_2) = \sum_{\mathcal{X}_1} p(x_1, x_2).$$

Proof. For X_1:
Let us denote by $A_{x_1} = \{(x_1, x_2) : -\infty < x_2 < \infty\}$. Then, for any $x_1 \in \mathcal{X}_1$ we may write

$$\begin{aligned} P(X_1 = x_1) &= P(X_1 = x_1, -\infty < x_2 < \infty) \\ &= P\big((X_1, X_2) \subseteq A_{x_1}\big) \\ &= \sum_{(x_1,x_2)\in A_{x_1}} P(X_1 = x_1, X_2 = x_2) \\ &= \sum_{\mathcal{X}_2} p(x_1, x_2). \end{aligned}$$

For X_2 the proof is similar. □

Example 1.20. The marginal distributions of the variables X_1 and X_2 defined in Example 1.18 are following.

x_1	2	3	4	5	6	7	8	9	10	11	12
$P(X_1 = x_1)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

x_2	0	1	2	3	4	5
$P(X_2 = x_2)$	$\frac{1}{6}$	$\frac{5}{18}$	$\frac{2}{9}$	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$

□

Exercise 1.14. Students in a class of 100 were classified according to gender (G) and smoking (S) as follows:

		S			
		s	q	n	
G	male	20	32	8	60
	female	10	5	25	40
		30	37	33	100

where s, q and n denote the smoking status: "now smokes", "did smoke but quit" and "never smoked", respectively. Find the probability that a randomly selected student

1. is a male;
2. is a male smoker;
3. is either a smoker or did smoke but quit;
4. is a female who is a smoker or did smoke but quit.

1.9.2 Continuous Two-Dimensional Random Variables

If the values of $\boldsymbol{X} = (X_1, X_2)$ are elements of an uncountable set in the Euclidean plane, then the variable is continuous. For example the values might be in the range

$$\mathcal{X} = \{(x_1, x_2) : a \leq x_1 \leq b, c \leq x_2 \leq d\}$$

for some real a, b, c, d.

The cdf of a continuous rv $\boldsymbol{X} = (X_1, X_2)$ is defined as

$$F_{\boldsymbol{X}}(x_1, x_2) = \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f(t_1, t_2) dt_1 dt_2, \tag{1.11}$$

where $f(x_1, x_2)$ is the probability density function such that

1. $f(x_1, x_2) \geq 0$ for all $(x_1, x_2) \in R^2$
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 dx_2 = 1$.

The equation (1.11) implies that

$$\frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2} = f(x_1, x_2). \tag{1.12}$$

Also

$$P(a \le X_1 \le b, c \le X_2 \le d) = \int_c^d \int_a^b f(x_1, x_2) dx_1 dx_2.$$

The marginal pdfs of X_1 and X_2 are defined similarly as in the discrete case, here using integrals.

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_2, \quad \text{for} \quad -\infty < x_1 < \infty,$$

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_1, \quad \text{for} \quad -\infty < x_2 < \infty.$$

Example 1.21. Calculate $P\big(\boldsymbol{X} \subseteq A\big)$, where $A = \{(x_1, x_2) : x_1 + x_2 \ge 1\}$ and the joint pdf of $\boldsymbol{X} = (X_1, X_2)$ is defined by

$$f_{\boldsymbol{X}}(x_1, x_2) = \begin{cases} 6x_1x_2^2 & \text{for } 0 < x_1 < 1,\ 0 < x_2 < 1, \\ 0 & \text{otherwise.} \end{cases}$$

The probability is a double integral of the pdf over the region A. The region is however limited by the domain in which the pdf is positive.

We can write

$$\begin{aligned} A &= \{(x_1, x_2) : x_1 + x_2 \ge 1,\ 0 < x_1 < 1,\ 0 < x_2 < 1\} \\ &= \{(x_1, x_2) : x_1 \ge 1 - x_2,\ 0 < x_1 < 1,\ 0 < x_2 < 1\} \\ &= \{(x_1, x_2) : 1 - x_2 < x_1 < 1,\ 0 < x_2 < 1\}. \end{aligned}$$

Hence, the probability is

$$P(\boldsymbol{X} \subseteq A) = \int\int_A f(x_1, x_2) dx_1 dx_2 = \int_0^1 \int_{1-x_2}^1 6x_1x_2^2 dx_1 dx_2 = 0.9$$

Also, we can calculate marginal pdfs.

$$f_{X_1}(x_1) = \int_0^1 6x_1x_2^2 dx_2 = 2x_1x_2^3 \mid_0^1 = 2x_1,$$

$$f_{X_2}(x_2) = \int_0^1 6x_1x_2^2 dx_1 = 3x_1^2x_2^2 \mid_0^1 = 3x_2^2.$$

These functions allow us to calculate probabilities involving only one variable. For example

$$P\left(\frac{1}{4} < X_1 < \frac{1}{2}\right) = \int_{\frac{1}{4}}^{\frac{1}{2}} 2x_1 dx_1 = \frac{3}{16}.$$

□

Similarly to the case of a univariate rv the following linear property for the expectation holds.

$$\mathrm{E}[ag(\boldsymbol{X}) + bh(\boldsymbol{X}) + c] = a\,\mathrm{E}[g(\boldsymbol{X})] + b\,\mathrm{E}[h(\boldsymbol{X})] + c, \tag{1.13}$$

where a, b and c are constants and g and h are some functions of the bivariate rv $\boldsymbol{X} = (X_1, X_2)$.

1.9.3 Conditional Distributions and Independence

Definition 1.16. *Let $\boldsymbol{X} = (X_1, X_2)$ denote a discrete bivariate rv with joint pmf $p_{\boldsymbol{X}}(x_1, x_2)$ and marginal pmfs $p_{X_1}(x_1)$ and $p_{X_2}(x_2)$. For any x_1 such that $p_{X_1}(x_1) > 0$, the conditional pmf of X_2 given that $X_1 = x_1$ is the function of x_2 defined by*

$$p_{X_2}(x_2|x_1) = \frac{p_{\boldsymbol{X}}(x_1, x_2)}{p_{X_1}(x_1)}.$$

Analogously, we define the conditional pmf of X_1 given $X_2 = x_2$

$$p_{X_1}(x_1|x_2) = \frac{p_{\boldsymbol{X}}(x_1, x_2)}{p_{X_2}(x_2)}.$$

□

It is easy to check that these functions are indeed pdfs. For example,

$$\sum_{\mathcal{X}_2} p_{X_2}(x_2|x_1) = \sum_{\mathcal{X}_2} \frac{p_{\boldsymbol{X}}(x_1, x_2)}{p_{X_1}(x_1)} = \frac{\sum_{\mathcal{X}_2} p_{\boldsymbol{X}}(x_1, x_2)}{p_{X_1}(x_1)} = \frac{p_{X_1}(x_1)}{p_{X_1}(x_1)} = 1.$$

Example 1.22. Let X_1 and X_2 be defined as in Example 1.18. The conditional pmf of X_2 given $X_1 = 5$, is

x_2	0	1	2	3	4	5
$p_{X_2}(x_2\|5)$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0

□

Exercise 1.15. Let S and G denote the *smoking status* an *gender* as defined in Exercise 1.14. Calculate the probability that a randomly selected student is

1. a smoker given that he is a male;
2. female, given that the student smokes.

Analogously to the conditional distribution for discrete rvs, we define the conditional distribution for continuous rvs.

Definition 1.17. *Let* $\boldsymbol{X} = (X_1, X_2)$ *denote a continuous bivariate rv with joint pdf* $f_{\boldsymbol{X}}(x_1, x_2)$ *and marginal pdfs* $f_{X_1}(x_1)$ *and* $f_{X_2}(x_2)$*. For any* x_1 *such that* $f_{X_1}(x_1) > 0$*, the conditional pdf of* X_2 *given that* $X_1 = x_1$ *is the function of* x_2 *defined by*

$$f_{X_2}(x_2|x_1) = \frac{f_{\mathbf{X}}(x_1, x_2)}{f_{X_1}(x_1)}.$$

Analogously, we define the conditional p.d.f. of X_1 *given* $X_2 = x_2$

$$f_{X_1}(x_1|x_2) = \frac{f_{\mathbf{X}}(x_1, x_2)}{f_{X_2}(x_2)}.$$

□

Here too, it is easy to verify that these functions are pdfs. For example,

$$\begin{aligned}\int_{\mathcal{X}_2} f_{X_2}(x_2|x_1)dx_2 &= \int_{\mathcal{X}_2} \frac{f_{\boldsymbol{X}}(x_1, x_2)}{f_{X_1}(x_1)}dx_2 \\ &= \frac{\int_{\mathcal{X}_2} f_{\boldsymbol{X}}(x_1, x_2)dx_2}{f_{X_1}(x_1)} \\ &= \frac{f_{X_1}(x_1)}{f_{X_1}(x_1)} = 1.\end{aligned}$$

Example 1.23. For the random variables defined in Example 1.21 the conditional pdfs are

$$f_{X_1}(x_1|x_2) = \frac{f_{\boldsymbol{X}}(x_1, x_2)}{f_{X_2}(x_2)} = \frac{6x_1x_2^2}{3x_2^2} = 2x_1$$

and

$$f_{X_2}(x_2|x_1) = \frac{f_{\boldsymbol{X}}(x_1, x_2)}{f_{X_1}(x_1)} = \frac{6x_1x_2^2}{2x_1} = 3x_2^2.$$

□

The conditional pdfs allow us to calculate conditional expectations. The conditional expected value of a function $g(X_2)$ given that $X_1 = x_1$ is defined by

$$\mathrm{E}[g(X_2)|x_1] = \begin{cases} \displaystyle\sum_{\mathcal{X}_2} g(x_2)p_{X_2}(x_2|x_1) & \text{for a discrete r.v.,} \\ \displaystyle\int_{\mathcal{X}_2} g(x_2)f_{X_2}(x_2|x_1)dx_2 & \text{for a continuous r.v..} \end{cases} \tag{1.14}$$

Example 1.24. The conditional mean and variance of the X_2 given a value of X_1, for the variables defined in Example 1.21 are

$$\mu_{X_2|x_1} = \mathrm{E}(X_2|x_1) = \int_0^1 x_2 3x_2^2 dx_2 = \frac{3}{4},$$

and

$$\sigma^2_{X_2|x_1} = \mathrm{var}(X_2|x_1) = \mathrm{E}(X_2^2|x_1) - [\mathrm{E}(X_2|x_1)]^2 = \int_0^1 x_2^2 3x_2^2 dx_2 - \left(\frac{3}{4}\right)^2 = \frac{3}{80}.$$

$\square$

Definition 1.18. *Let* $\boldsymbol{X} = (X_1, X_2)$ *denote a continuous bivariate rv with joint pdf* $f_{\boldsymbol{X}}(x_1, x_2)$ *and marginal pdfs* $f_{X_1}(x_1)$ *and* $f_{X_2}(x_2)$*. Then* X_1 *and* X_2 *are called* ***independent random variables*** *if, for every* $x_1 \in \mathcal{X}_1$ *and* $x_2 \in \mathcal{X}_2$

$$f_{\boldsymbol{X}}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2). \tag{1.15}$$

$\square$

We define independent discrete random variables analogously.

If X_1 and X_2 are independent, then the conditional pdf of X_2 given $X_1 = x_1$ is

$$f_{X_2}(x_2|x_1) = \frac{f_{\boldsymbol{X}}(x_1, x_2)}{f_{X_1}(x_1)} = \frac{f_{X_1}(x_1) f_{X_2}(x_2)}{f_{X_1}(x_1)} = f_{X_2}(x_2)$$

regardless of the value of x_1. Analogous property holds for the conditional pdf of X_1 given $X_2 = x_2$.

Example 1.25. It is easy to notice that for the variables defined in Example 1.21 we have

$$f_{\boldsymbol{X}}(x_1, x_2) = 6x_1 x_2^2 = 2x_1 3x_2^2 = f_{X_1}(x_1) f_{X_2}(x_2).$$

So, the variables X_1 and X_2 are independent.

$\square$

In fact, two rvs are independent if and only if there exist functions $g(x_1)$ and $h(x_2)$ such that for every $x_1 \in \mathcal{X}_1$ and $x_2 \in \mathcal{X}_2$,

$$f_{\boldsymbol{X}}(x_1, x_2) = g(x_1) h(x_2).$$

Theorem 1.13. *Let X_1 and X_2 be independent random variables. Then*

1. *For any $A \subset \mathbb{R}$ and $B \subset \mathbb{R}$*

$$P(X_1 \subseteq A, X_2 \subseteq B) = P(X_1 \subseteq A)P(X_2 \subseteq B),$$

   that is, $\{X_1 \subseteq A\}$ and $\{X_2 \subseteq B\}$ are independent events.

2. *For $g(X_1)$, a function of X_1 only, and for $h(X_2)$, a function of X_2 only, we have*

$$\mathrm{E}[g(X_1)h(X_2)] = \mathrm{E}[g(X_1)]\,\mathrm{E}[h(X_2)].$$

Proof. Assume that X_1 and X_2 are continuous random variables. To prove the theorem for discrete rvs we follow the same steps with sums instead of integrals.

1. We have

$$\begin{aligned}
P(X_1 \subseteq A, X_2 \subseteq B) &= \int_B \int_A f_{\boldsymbol{X}}(x_1, x_2)dx_1dx_2 \\
&= \int_B \int_A f_{X_1}(x_1)f_{X_2}(x_2)dx_1dx_2 \\
&= \int_B \left(\int_A f_{X_1}(x_1)dx_1\right) f_{X_2}(x_2)dx_2 \\
&= \int_A f_{X_1}(x_1)dx_1 \int_B f_{X_2}(x_2)dx_2 \\
&= P(X_1 \subseteq A)P(X_2 \subseteq B).
\end{aligned}$$

2. Similar arguments as in Part 1 give

$$\begin{aligned}
\mathrm{E}[g(X_1)h(X_2)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1)h(x_2)f_{\boldsymbol{X}}(x_1,x_2)dx_1dx_2 \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1)h(x_2)f_{X_1}(x_1)f_{X_2}(x_2)dx_1dx_2 \\
&= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} g(x_1)f_{X_1}(x_1)dx_1\right) h(x_2)f_{X_2}(x_2)dx_2 \\
&= \left(\int_{-\infty}^{\infty} g(x_1)f_{X_1}(x_1)dx_1\right)\left(\int_{-\infty}^{\infty} h(x_2)f_{X_2}(x_2)dx_2\right) \\
&= \mathrm{E}[g(X_1)]\,\mathrm{E}[h(X_2)].
\end{aligned}$$

□

In the following theorem we will apply this result for the moment generating function of a sum of independent random variables.

Theorem 1.14. *Let X_1 and X_2 be independent random variables with moment generating functions $M_{X_1}(t)$ and $M_{X_2}(t)$, respectively. Then the moment generating function of the sum $Y = X_1 + X_2$ is given by*

$$M_Y(t) = M_{X_1}(t)M_{X_2}(t).$$

Proof. By the definition of the mgf and by Theorem 1.13, part 2, we have

$$M_Y(t) = \mathrm{E}\, e^{tY} = \mathrm{E}\, e^{t(X_1+X_2)} = \mathrm{E}\left(e^{tX_1}e^{tX_2}\right) = \mathrm{E}\left(e^{tX_1}\right)\mathrm{E}\left(e^{tX_2}\right) = M_{X_1}(t)M_{X_2}(t).$$

□

Note that this result can be easily extended to a sum of any number of mutually independent random variables.

Example 1.26. Let $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$. What is the distribution of $Y = X_1 + X_2$?

Using Theorem 1.14 we can write

$$\begin{aligned} M_Y(t) &= M_{X_1}(t)M_{X_2}(t) \\ &= \exp\{\mu_1 t + \sigma_1^2 t^2/2\} \exp\{\mu_2 t + \sigma_2^2 t^2/2\} \\ &= \exp\{(\mu_1 + \mu_2)t + (\sigma_1^2 + \sigma_2^2)t^2/2\}. \end{aligned}$$

This is the mgf of a normal rv with $\mathrm{E}(Y) = \mu_1 + \mu_2$ and $\mathrm{var}(Y) = \sigma_1^2 + \sigma_2^2$.

Exercise 1.16. A part of an electronic system has two types of components in joint operation. Denote by X_1 and X_2 the random length of life (measured in hundreds of hours) of component of type I and of type II, respectively. Assume that the joint density function of two rvs is given by

$$f_{\boldsymbol{X}}(x_1, x_2) = \frac{1}{8} x_1 \exp\left\{-\frac{x_1 + x_2}{2}\right\} I_{\mathcal{X}},$$

where $\mathcal{X} = \{(x_1, x_2) : x_1 > 0, x_2 > 0\}$.

1. Calculate the probability that both components will have a life length longer than 100 hours, that is, find $P(X_1 > 1, X_2 > 1)$.

2. Calculate the probability that a component of type II will have a life length longer than 200 hours, that is, find $P(X_2 > 2)$.

3. Are X_1 and X_2 independent? Justify your answer.

4. Calculate the expected value of so called relative efficiency of the two components, which is expressed by

$$\mathrm{E}\left(\frac{X_2}{X_1}\right).$$