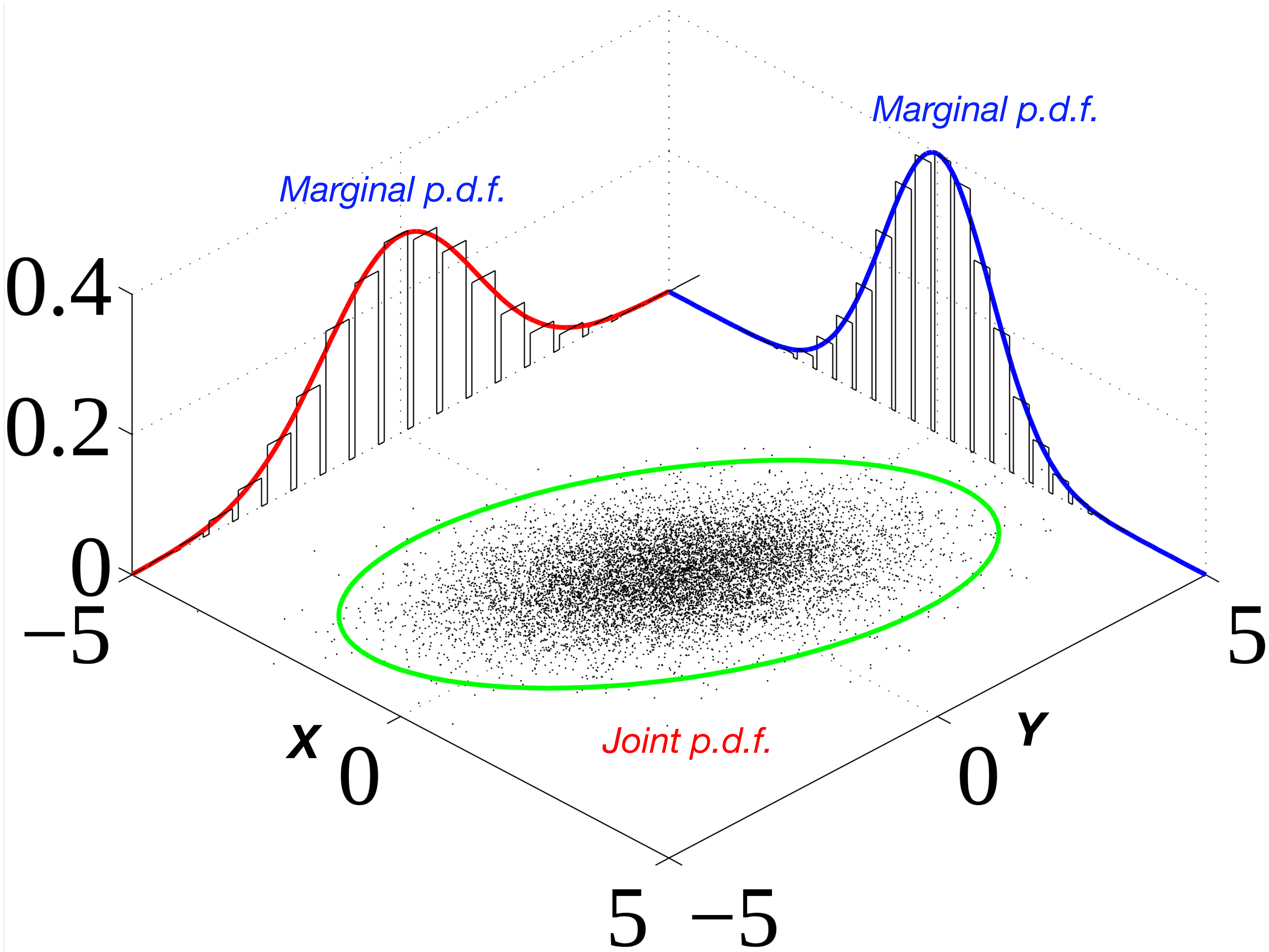
6. Jointly Distributed Random Variables



6.1 Joint Distribution Functions

Very often we are interested in two or more random variables at the same time. For example,

(1) In the population of HKUST students (our sample space), we are interested in the following characteristics of a student, his/her age (A), gender (G), major (M), and year (Y) of studies in HKUST.

Here, for each student, (A, G, M, Y) denotes a student's age, ..., etc.

(2) For any particular day, we are interested in the number of vehicle accidents, how many deaths in these accidents, and how many major injuries on the road.

Here we can use (A, D, I) to denote the 3 quantities we are interested.

6.1 Joint Distribution Functions

Definitions *For any two random variables X and Y defined on the same sample space, we define the* ***joint distribution function of X and Y*** *(we abbreviate it to joint d.f. and denote it by $F_{X,Y}$) by*

$$F_{X,Y}(x,y) = P(X \le x, Y \le y) \quad \textit{for } x,y \in \mathbb{R}.$$

$X: S \to \mathbb{R}$

$Y: S \to \mathbb{R}$.

Remark Notation

$$\begin{aligned}\{X \le x; Y \le y\} &= \{s \in S : X(s) \le x \text{ and } Y(s) \le y\} \\ &= \{s \in S : X(s) \le x\} \cap \{s \in S : Y(s) \le y\} \\ &= \{X \le x\} \cap \{Y \le y\}.\end{aligned}$$

6.1 Joint Distribution Functions

Marginal distribution function

The distribution function of X can be obtained from the joint density function of X and Y in the following way:

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x, y).$$

We call F_X the **marginal distribution function of X.**

Similarly,

$$F_Y(y) = \lim_{x \to \infty} F_{X,Y}(x, y)$$

and F_Y is called the **marginal distribution function of Y.**

6.1 Joint Distribution Functions

Some useful calculations

These formulas are not only useful in some calculations, but the derivations are just as important. Let $a, b, a_1 < a_2, b_1 < b_2$ be real numbers, then

$$P(X > a, Y > b) = 1 - F_X(a) - F_Y(b) + F_{X,Y}(a,b),$$

$$\begin{aligned} P(a_1 < X \leq a_2,\ b_1 < Y \leq b_2) = {} & F_{X,Y}(a_2,b_2) - F_{X,Y}(a_1,b_2) \\ & + F_{X,Y}(a_1,b_1) - F_{X,Y}(a_2,b_1). \end{aligned}$$

Proof Write $A = \{X \leq a\}$ and $B = \{Y \leq b\}$. Then

$$\{X > a, Y > b\} = A^c B^c = (A \cup B)^c.$$

Therefore,

$$\begin{aligned} P(X > a,\ Y > b) &= P((A \cup B)^c) = 1 - P(A \cup B) \\ &= 1 - P(A) - P(B) + P(AB) = 1 - F_X(a) - F_Y(b) + F_{X,Y}(a,b). \end{aligned}$$

The second formula is left as an exercise.

6.1.1 Joint Discrete Random Variables

In the case when both X and Y are discrete random variables, we define the **joint probability mass function of X and Y** (abbreviated to joint p.m.f. and denoted by $p_{X,Y}$) as :

$$p_{X,Y}(x,y) = P(X = x, Y = y).$$

As in the distribution functions situation, we can recover the probability mass function of X and Y in the following manner:

$$p_X(x) = P(X = x) = \sum_{y \in \mathbb{R}} p_{X,Y}(x,y),$$
$$p_Y(y) = P(Y = y) = \sum_{x \in \mathbb{R}} p_{X,Y}(x,y).$$

We call p_X the **marginal probability mass function of X** and p_Y the **marginal probability mass function of Y**.

$$\{X=x\} \cap S = \{X=x\} \cap \{\bigcup_y \{Y=y\}\}$$

$$= \bigcup_y \{Y=y\} \cap \{X=x\}$$

$$p(X=x) = \sum_y p(X=x, Y=y)$$

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

6.1.1 Joint Discrete Random Variables

Example Suppose that 3 balls are randomly selected from an urn containing 3 red, 4 white and 5 blue balls. If we let R and W denote the number of red and white balls chosen, then the joint probability mass function of R and W is

r \ w	0	1	2	3	$P(R=r)$
0	$10/220$	$40/220$	$30/220$	$4/220$	$84/220$
1	$30/220$	$60/220$	$18/220$	0	$108/220$
2	$15/220$	$12/220$	0	0	$27/220$
3	$1/220$	0	0	0	$1/220$
$P(W=w)$	$56/220$	$112/220$	$48/220$	$4/220$	

We illustrate how some of the entries are computed:

(i) $p(0,0) = \binom{5}{3}/\binom{12}{3} = \frac{10}{220}$.

(ii) $p(2,2) = 0$ as the event of getting 2 red balls and 2 white balls are impossible.

6.1.1 Joint Discrete Random Variables

Some useful calculations

(i)

$$P(a_1 < X \leq a_2,\ b_1 < Y \leq b_2) = \sum_{a_1 < x \leq a_2} \sum_{b_1 < y \leq b_2} p_{X,Y}(x,y),$$

(ii)

$$F_{X,Y}(a,b) = P(X \leq a, Y \leq b) = \sum_{x \leq a} \sum_{y \leq b} p_{X,Y}(x,y),$$

(iii)

$$P(X > a, Y > b) = \sum_{x > a} \sum_{y > b} p_{X,Y}(x,y).$$

6.1.1 Joint Discrete Random Variables

Example Suppose that 15% of the families in a certain community have no children 20% have 1, 35% have 2, and 30% have 3. We suppose further that, in each family, each child is equally likely (independently) to be a boy or a girl. If a family is chosen at random from this community, then B, the number of boys; and G, the number of girls, in this family will have the joint probability mass function as shown in the following table.

Table for $P(B = i, G = j)$.

i \ j	0	1	2	3	$P(B = i)$
0	0.1500	0.1000	0.0875	0.0375	0.3750
1	0.1000	0.1750	0.1125	0	0.3875
2	0.0875	0.1125	0	0	0.2000
3	0.0375	0	0	0	0.0375
$P(G = j)$	0.3750	0.3875	0.2000	0.0375	

Some of these probabilities are calculated as follows: $P(B = 0, G = 0) = P(\text{no children}) = 0.15.$

$$P(B = 0, G = 1) = P(1 \text{ girl and total of children is } 1) = P(1 \text{ child})P(1 \text{ girl}|1 \text{ child}) = 0.20 \times \frac{1}{2} = 0.10.$$

$$P(B = 0, G = 2) = P(2 \text{ girls and total of children is } 2) = P(2 \text{ children})P(2 \text{ girls}|2 \text{ children}) = 0.35 \times \frac{1}{2^2} = 0.0875.$$

6.1.2 Joint Continuous Random Variables

We say that X and Y are **jointly continuous random variables** if there exists a function (which is denoted by $f_{X,Y}$, called the **joint probability density function of X and Y**) if for every set $C \subset \mathbb{R}^2$, we have

$$P((X,Y) \in C) = \int\int_{(x,y)\in C} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.$$

The marginal probability density function of X is given by

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y\,.$$

Similarly, the marginal probability density function of Y is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x\,.$$

if X, Y have joint pdf $f_{X,Y}(x,y)$, what is the marginal cdf of X.

$$F_X(x) = P(X \le x)$$

$$= P(\{X \le x\} \cap S)$$

$$= P(\{X \le x\} \cap \{Y \le \infty\})$$

$$= \iint_{\substack{\{u \le x\} \\ \{v \le y\}}} f(u,v)\, dv\, du$$

$$F_X(x) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f(u,v)\, dv\, du$$

$$f_X(x) = F_X'(x) = \int_{-\infty}^{+\infty} f(x,v)\, dv$$

$$= \int_{-\infty}^{+\infty} f(x,y)\, dy.$$

6.1.2 Joint Continuous Random Variables

Some useful calculations

(i) Let $A, B \subset \mathbb{R}$, take $C = A \times B$ above

$$P(X \in A, Y \in B) = \int_A \int_B f_{X,Y}(x,y)\,\mathrm{d}y\,\mathrm{d}x\,.$$

(ii) In particular, Let $a_1, a_2, b_1, b_2 \in \mathbb{R}$ where $a_1 < a_2$ and $b_1 < b_2$, we have

$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f_{X,Y}(x,y)\,\mathrm{d}y\,\mathrm{d}x\,.$$

(iii) Let $a, b \in \mathbb{R}$, we have

$$F_{X,Y}(a,b) = P(X \leq a, Y \leq b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f_{X,Y}(x,y)\,\mathrm{d}y\,\mathrm{d}x\,.$$

As a result of this, $f_{X,Y}(x,y) = \dfrac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$.

6.1.2 Joint Continuous Random Variables

Example The joint probability density function of X and Y is given by

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y}, & 0 < x < \infty, 0 < y < \infty \\ 0, & \text{otherwise} \end{cases} .$$

Compute

(i) $P(X > 1, Y < 1)$, (ii) $P(X < Y)$, (iii) $P(X \le x)$,

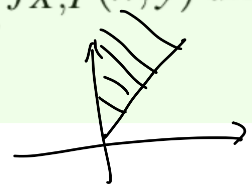
(iv) the marginal probability density function of X,

(v) the marginal distribution function of Y.

Solution

(i) $$P(X > 1, Y < 1) = \int_0^1 \int_1^\infty f_{X,Y}(x,y)\,dx\,dy = \int_0^1 \int_1^\infty 2e^{-x}e^{-2y}\,dx\,dy = e^{-1}(1-e^{-2}).$$

(ii) $$P(X < Y) = \iint_{x<y} f_{X,Y}(x,y)\,dx\,dy = \int_0^\infty \int_0^y 2e^{-x}e^{-2y}\,dx\,dy = \int_0^\infty [2e^{-2y} - 2e^{-3y}]\,dy = 1/3.$$



6.1.2 Joint Continuous Random Variables

Solution (cont.)

(iv) We will work on (iv) first.
For $x \le 0$, $f_X(x) = 0$. For $x > 0$, $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{dy} = \int_0^{\infty} 2e^{-x}e^{-2y}\,\mathrm{dy} = e^{-x}$,

Hence $$f_X(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & x \le 0 \end{cases}.$$

(iii) Marginal distribution function of X is: For $x > 0$,

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{dt} = \int_0^x e^{-t}\,\mathrm{dt} = 1 - e^{-x}.$$

And for $x \le 0$, $F_X(x) = 0$. Hence, marginal distribution function of X is

$$F_X(x) = \begin{cases} 1 - e^{-x}, & x > 0 \\ 0, & x \le 0 \end{cases}.$$

(v) Marginal distribution function of Y.
For $y \le 0$, $F_Y(y) = 0$. For $y > 0$,

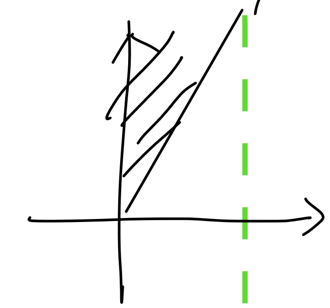
$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{y} f_Y(t)\,\mathrm{dt} = \int_0^y \int_{-\infty}^{\infty} f_{X,Y}(x,t)\,\mathrm{dx}\,\mathrm{dt} \\ &= \int_0^y \int_0^{\infty} 2e^{-x}e^{-2t}\,\mathrm{dx}\,\mathrm{dt} = 1 - e^{-2y}. \end{aligned}$$

Hence, marginal distribution function of Y $$F_Y(y) = \begin{cases} 1 - e^{-2y}, & y \ge 0 \\ 0, & y < 0 \end{cases}.$$

6.1.2 Joint Continuous Random Variables

Example The joint density of X and Y is given by

$y = ax$



$$f(x,y) = \begin{cases} e^{-(x+y)}, & 0 < x < \infty, 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}.$$

$x \leq ay$

Find the density function of the random variable X/Y.

Solution We start by computing the distribution function of X/Y. For $a > 0$,

$$F_{X/Y}(a) = P\left(\frac{X}{Y} \leq a\right) \quad = \iint_{x/y \leq a} e^{-(x+y)}\,\mathrm{d}x\,\mathrm{d}y \quad = \int_0^\infty \int_0^{ay} e^{-(x+y)}\,\mathrm{d}x\,\mathrm{d}y$$

$$= \int_0^\infty (1 - e^{-ay})e^{-y}\,\mathrm{d}y \quad = \left[-e^{-y} + \frac{e^{-(a+1)y}}{a+1}\right]_0^\infty$$

$$= 1 - \frac{1}{a+1}.$$

Differentiation yields that the density function of X/Y is given by

$$f_{X/Y}(a) = 1/(a+1)^2, \quad 0 < a < \infty.$$

$$\int_0^{\infty}\int_0^{\alpha y} e^{-(x+y)}\,dx\,dy$$

$$= \int_0^{\infty}\int_0^{\alpha y} e^{-x-y}\,dx\,dy$$

$$= \int_0^{\infty}\int_0^{\alpha y} e^{-x}\cdot e^{-y}\,dx\,dy$$

$$= \int_0^{\infty} e^{-y}\left[-e^{-x}\right]_0^{\alpha y} dy$$

$$= \int_0^{\infty} e^{-y}\left(1-e^{-\alpha y}\right) dy$$

$$= \int_0^{\infty} e^{-y} - e^{-y(\alpha+1)}\,dy$$

$$= \left[-e^{-y}\right]_0^{\infty} - \left[-\frac{1}{\alpha+1}e^{-(\alpha+1)y}\right]_0^{\infty}$$

$$= 1 - \frac{1}{\alpha+1}$$

6.2 Independent Random Variables

Definition Two random variables X and Y are said to be **independent** if

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) \qquad \text{for any } A, B \subset \mathbb{R}.$$

Random variables that are **not** independent are said to be **dependent.**

Theorem (For jointly discrete random variables). *The following three statements are equivalent:*

(i) Random variables X and Y are independent.

(ii) For all $x, y \in \mathbb{R}$, we have

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

(iii) For all $x, y \in \mathbb{R}$, we have

$$F_{X,Y}(x,y) = F_X(x)F_Y(y).$$

6.2 Independent Random Variables

Theorem (For jointly continuous random variables). *The following three statements are equivalent:*

(i) Random variables X and Y are independent.

(ii) For all $x, y \in \mathbb{R}$, we have $f_{X,Y}(x,y) = f_X(x)f_Y(y)$.

(iii) For all $x, y \in \mathbb{R}$, we have

$$F_{X,Y}(x,y) = F_X(x)F_Y(y).$$

Example Suppose that 3 balls are randomly selected from an urn containing 3 red, 4 white and 5 blue balls. If we let R and W denote the number of red and white balls chosen, *are R and W independent?*

$P(R=i, W=j) =$ (R,W): $W=0$ $W=1$ $W=2$ $W=3$ $W=4$

$R=0$

$R=1$

$R=2$

$R=3$

3 ball selected :

3R 4W 5B.

$P(R=i, W=j) =$

	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$.
$i=0$	10	40	30	4	$\frac{150}{220}$
$i=1$	30	120	0		
$i=2$	15	12	0		
$i=3$	1	0	0		
		$\frac{172}{220}$			

分母 : 220

$P_{X,Y}(1,1) = P_X(1) P_Y(1)$ No.

6.2 Independent Random Variables

Example The joint probability density function of X and Y is given by

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y}, & 0 < x < \infty, 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}.$$

One can verify that

$$f_X(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{otherwise} \end{cases} \quad ; \quad f_Y(y) = \begin{cases} 2e^{-2y}, & y > 0 \\ 0, & \text{otherwise} \end{cases}.$$

Then, we can check that $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for every $x, y \in \mathbb{R}$. Hence, X and Y are independent.

6.2 Independent Random Variables

Example Suppose that $n+m$ independent trials, having a common success probability p, are performed. If X is the number of successes in the first n trials, and Y is the number of successes in the final m trials, then X and Y are independent, since knowing the number of successes in the first n trials does not affect the distribution of the number of successes in the final m trials (by the assumption of independent trials). In fact, for integral x and y,

$$\begin{aligned} P(X=x,Y=y) &= \binom{n}{x} p^x (1-p)^{n-x} \binom{m}{y} p^y (1-p)^{m-y} \\ &= P(X=x)P(Y=y), \quad 0 \le x \le n, 0 \le y \le m. \end{aligned}$$

On the other hand, X and Z will be dependent, where $Z = X+Y$ is the total number of successes in the $n+m$ trials. To see this, consider the fact that

$$P(Z=z) = \binom{m+n}{z} p^z (1-p)^{m+n-z}, \quad 0 \le z \le m+n.$$

However, $$\begin{aligned} P(X=x,Z=z) &= P(X=x,X+Y=z) \\ &= P(X=x,Y=z-x) = \binom{n}{x} p^x (1-p)^{n-x} \binom{m}{z-x} p^{z-x} (1-p)^{m-z+x}, \end{aligned}$$

for $0 \le x \le n, x \le z \le m+x$.

So we see that $$P(X=x,Z=z) \neq P(X=x)P(Z=z),$$

indicating that X and Z are not independent.

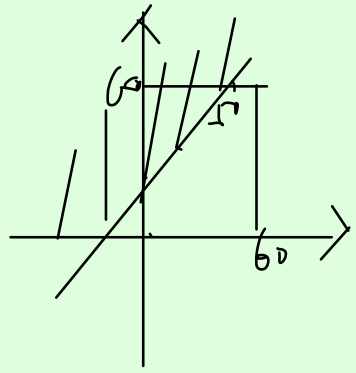
6.2 Independent Random Variables

Example A man and a woman decide to meet at a certain location. If each person independently arrives at a time uniformly distributed between 12 noon and 1 pm, find the probability that the first to arrive has to wait longer than 10 minutes.

Solution

If we let X and Y denote, respectively, the time past 12 that the man and the woman arrive, then X and Y are independent random variables, each of which is uniformly distributed over $(0, 60)$. The desired probability, $P(X+10<Y)+P(Y+10<X)$, which by symmetry equals $2P(X+10<Y)$, is obtained as follows:

$y = x+10$

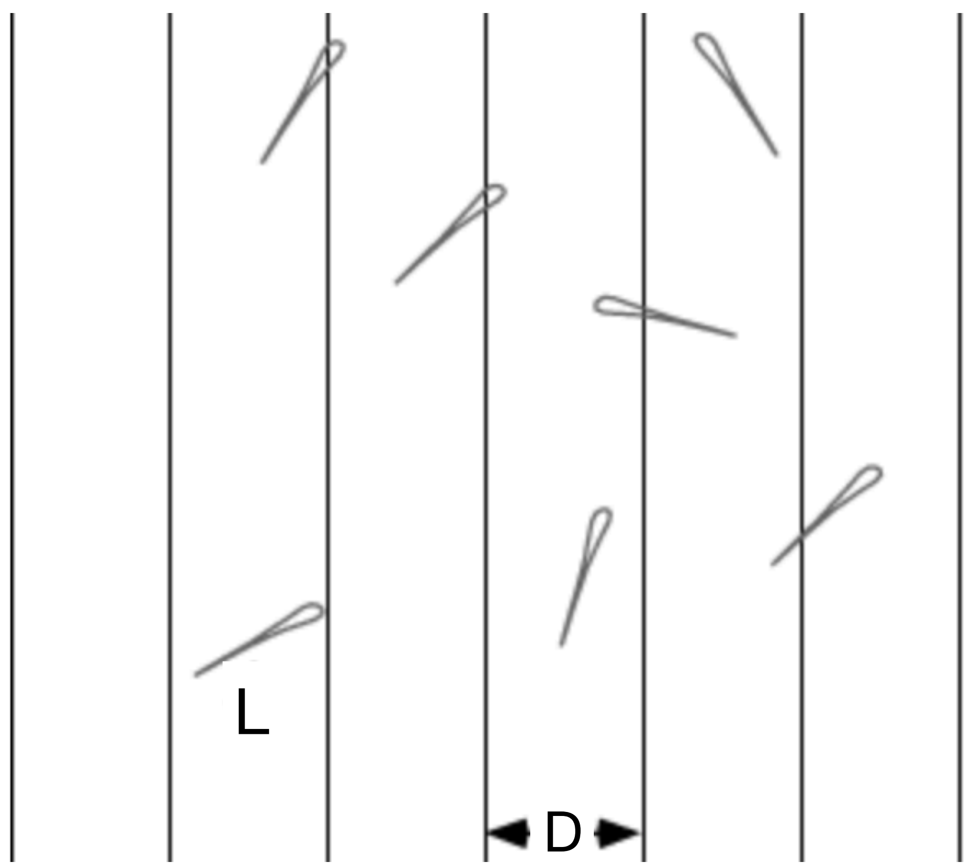


$$2P(X+10<Y) = 2\iint_{x+10<y} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 2\iint_{x+10<y} f_X(x)f_Y(y)\,\mathrm{d}x\,\mathrm{d}y$$

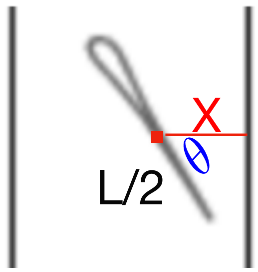
$$= 2\int_{10}^{60}\int_{0}^{y-10}\left(\frac{1}{60}\right)^2 \mathrm{d}x\,\mathrm{d}y = \frac{2}{60^2}\int_{10}^{60}(y-10)\,\mathrm{d}y = \frac{25}{36}.$$

6.2 Independent Random Variables

Example (Buffon's needle problem). A table is ruled with equidistant parallel lines a distance D apart. A needle of length L, where $L \leq D$, is randomly thrown on the table. What is the probability that the needle will intersect one of the lines (the other possibility being that the needle will be completely contained in the strip between two lines)?



6.2 Independent Random Variables



Solution Let us determine the position of the needle by specifying the distance X from the middle point of the needle to the nearest parallel line, and the angle θ between the needle and the projected line of length X. The needle will intersect a line if the hypotenuse of the right triangle is less than $L/2$, that is, if

$$\frac{X}{\cos\theta} < \frac{L}{2} \text{ or } X < \frac{L}{2}\cos\theta.$$

As X varies between 0 and $D/2$ and θ between 0 and $\pi/2$, it is reasonable to assume that they are independent, uniformly distributed random variables over these respective ranges. Hence

$$\begin{aligned} P\left(X < \frac{L}{2}\cos\theta\right) &= \iint_{x<L/2\cos\theta} f_X(x) f_\Theta(\theta)\, \mathrm{d}x\, \mathrm{d}\theta \\ &= \frac{4}{\pi D}\int_0^{\pi/2}\int_0^{L/2\cos\theta} \mathrm{d}x\, \mathrm{d}\theta \quad = \frac{4}{\pi D}\int_0^{\pi/2} \frac{L}{2}\cos\theta\, \mathrm{d}\theta \quad = \frac{2L}{\pi D}. \end{aligned}$$

6.2 Independent Random Variables

Theorem *Random variables X and Y are independent if and only if there exist functions $g, h : \mathbb{R} \longrightarrow \mathbb{R}$ such that for all $x, y \in \mathbb{R}$, we have*

$$f_{X,Y}(x,y) = g(x)h(y).$$

Proof Let us give the proof in the continuous case. First note that independence implies that the joint density is the product of the marginal densities of X and Y, so the preceding factorization will hold when the random variables are independent. Now, suppose that $f_{X,Y}(x,y) = h(x)g(y)$.

Then

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{-\infty}^{\infty} h(x)\,\mathrm{d}x \int_{-\infty}^{\infty} g(y)\,\mathrm{d}y = C_1C_2.$$

where $C_1 = \int_{-\infty}^{\infty} h(x)\,\mathrm{d}x$ and $C_2 = \int_{-\infty}^{\infty} g(y)\,\mathrm{d}y$. Also,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y = C_2h(x),$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x = C_1g(y).$$

Since $C_1C_2 = 1$, we thus see that $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ and the proof is complete.

6.2 Independent Random Variables

Example In a previous example, we have

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y}, & 0 < x < \infty, 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}.$$

We can take

$$g(x) = \begin{cases} 4e^{-x}, & 0 < x < \infty \\ 0, & \text{elsewhere} \end{cases}$$

and

$$h(y) = \begin{cases} 1/2e^{-2y}, & 0 < y < \infty \\ 0, & \text{elsewhere} \end{cases}.$$

Then it is easily verified that

$$f_{X,Y}(x,y) = g(x)h(y) \quad \text{for all } x, y \in \mathbb{R}.$$

Therefore, X and Y are independent.

6.2 Independent Random Variables

Example If the joint density function of X and Y is

$$f(x,y) = 24xy, \quad 0 < x < 1, 0 < y < 1, 0 < x + y < 1$$

and is equal to 0 otherwise, are the random variables independent?

Solution The region in which the joint density is nonzero cannot be expressed in the form $x \in A, y \in .B$. This means that the joint density does not factor and so the random variables are not independent.
More formally, we can define

$$I(x,y) = \begin{cases} 1, & \text{if } 0 < x < 1, 0 < y < 1, 0 < x + y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Then $f(x,y)$ can be written as

$$f(x,y) = 24xyI(x,y),$$

which clearly does not factor into a part depending only on x and another depending only on y.

6.2 Independent Random Variables

Example Suppose there are N traffic accidents in a day, where N is assumed to Poisson distributed with parameter λ. Each traffic accident is classified as major and minor. Suppose that given a traffic accident, it is a major accident with probability $p \in (0,1)$. Let X and Y denote the numbers of major and minor accidents respectively.

(i) Find the joint probability mass function of X and Y.

(ii) Are X and Y independent?

(iii) Can you identify the distributions of X and Y?

Solution

(i) For $0 \le i, j < \infty$,
$$P(X=i,Y=j) = P(X=i,N=i+j) = P(N=i+j)P(X=i|N=i+j)$$

$$= \frac{e^{-\lambda}\lambda^{i+j}}{(i+j)!}\binom{i+j}{i}p^i q^j = \frac{e^{-\lambda}\lambda^{i+j}p^i q^j}{i!\,j!} = \frac{e^{-\lambda}(\lambda p)^i}{i!}\frac{(\lambda q)^j}{j!}$$

$$= \frac{e^{-\lambda p}(\lambda p)^i}{i!}\frac{e^{-\lambda q}(\lambda q)^j}{j!}.$$

6.2 Independent Random Variables

Solution (cont.)

(ii) Define

$$g(x) = \begin{cases} \frac{e^{-\lambda p}(\lambda p)^x}{x!}, & \text{for } x \text{ integer} \geq 0 \\ 0, & \text{otherwise} \end{cases};$$

and

$$h(y) = \begin{cases} \frac{e^{-\lambda q}(\lambda q)^y}{y!}, & \text{for } y \text{ integer} \geq 0 \\ 0, & \text{otherwise} \end{cases}.$$

Hence, $p_{X,Y}(x,y) = g(x)h(y)$ for all $x, y \in \mathbb{R}$. And so X and Y are independent.

(iii) From the previous part, we see that

$$X \sim \text{Poisson}(\lambda p) \text{ and } Y \sim \text{Poisson}(\lambda q).$$

6.2 Independent Random Variables

Example Suppose that X and Y are independent standard normal distribution. Find the joint probability density function of X and Y.

Solution Recall that X and Y have the common probability density function given by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}.$$

Therefore the joint probability density function of X, Y is

$$\begin{aligned} f_{X,Y}(x,y) &= f_X(x) f_Y(y) \quad \text{(by independence)} \\ &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \\ &= \frac{1}{2\pi} e^{-[x^2+y^2]/2}, \quad \text{for } -\infty < x < \infty, -\infty < y < \infty. \end{aligned}$$

6.2 Independent Random Variables

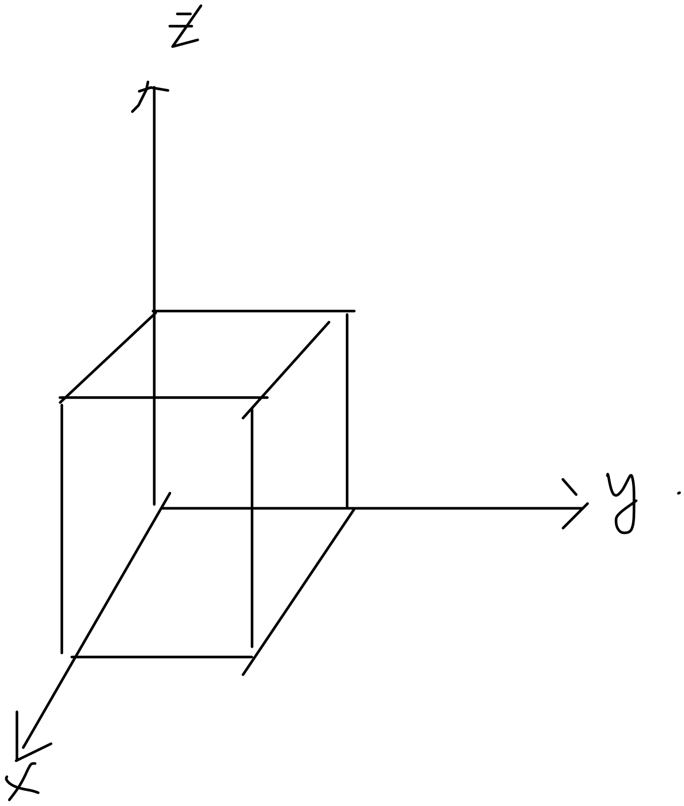
Example Let X, Y, Z be independent and uniformly distributed over $(0,1)$. Compute $P(X \geq YZ)$.

Solution Since

$$f_{X,Y,Z}(x,y,z) = f_X(x) f_Y(y) f_Z(z) = 1, \quad 0 \leq x,y,z \leq 1,$$

we have

$$\begin{aligned} P(X \geq YZ) &= \iiint_{x \geq yz} f_{X,Y,Z}(x,y,z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \\ &= \int_0^1 \int_0^1 \int_{yz}^1 \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \\ &= \int_0^1 \int_0^1 (1 - yz) \, \mathrm{d}y \, \mathrm{d}z \\ &= \int_0^1 (1 - z/2) \, \mathrm{d}z \\ &= \frac{3}{4}. \end{aligned}$$



$x = yz$

6.2 Independent Random Variables

Remark (Independence is a symmetric relation). To say that X is independent of Y is equivalent to saying that Y is independent of X, or just that X and Y are independent.

As a result, in considering whether X is independent of Y in situations where it is not at all intuitive that knowing the value of Y will not change the probabilities concerning X, it can be beneficial to interchange the roles of X and Y and ask instead whether Y is independent of X. The next example illustrates this point.

6.2 Independent Random Variables

Example If the initial throw of the dice in the game of craps results in the sum of the dice equaling 4, then the player will continue to throw the dice until the sum is either 4 or 7. If this sum is 4, then the player wins, and if it is 7, then the player loses. Let N denote the number of throws needed until either 4 or 7 appears, and let X denote the value (either 4 or 7) of the final throw. Is N independent of X?

Solution

The answer to this question is not intuitively obvious. However, suppose that we tum it around and ask whether X is independent of N. That is, does knowing how many throws it takes to obtain a sum of either 4 or 7 affect the probability that that sum is equal to 4?

Clearly not, since the fact that none of the first $n - 1$ throws were either 4 or 7 does not change the probabilities for the nth throw. Thus, we can conclude that X is independent of N, or equivalently, that N is independent of X.

6.3 Sums of Independent Random Variables

Very often we are interested in the sum of independent random variables. For example,

(1) Two dice are rolled, we are interested (as in many games) in the sum of the two numbers.

(2) In a data set, each datum collected is rounded off to the nearest integer. Let X_i denote the error in rounding the ith datum. Suppose we want to compute the total of this data set.

One quantity that we are interested is: sum of the errors due to rounding off.

6.3.1 X and Y are continuous and independent

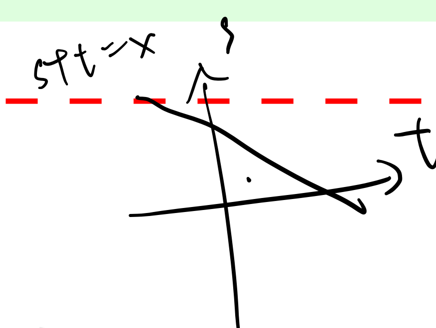
Theorem Under the assumption of independence of X and Y, we have

$$f_{X,Y}(x,y) = f_X(x) f_Y(y), \qquad \text{for } x, y \in \mathbb{R}.$$

Then it follows that

$$F_{X+Y}(x) = \int_{-\infty}^{\infty} F_X(x-t) f_Y(t)\, \mathrm{dt} = \int_{-\infty}^{\infty} F_Y(x-t) f_X(t)\, \mathrm{dt}.$$

$$f_{X+Y}(x) = \int_{-\infty}^{\infty} f_X(x-t) f_Y(t)\, \mathrm{dt} = \int_{-\infty}^{\infty} f_X(t) f_Y(x-t)\, \mathrm{dt}.$$



Proof

$$\begin{aligned}
F_{X+Y}(x) &= P(X+Y \le x) \\
&= \iint_{s+t \le x} f_{X,Y}(s,t)\, \mathrm{ds}\, \mathrm{dt} \\
&= \iint_{s+t \le x} f_X(s) f_Y(t)\, \mathrm{ds}\, \mathrm{dt} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{x-t} f_X(s) f_Y(t)\, \mathrm{ds}\, \mathrm{dt} \\
&= \int_{-\infty}^{\infty} F_X(x-t) f_Y(t)\, \mathrm{dt}.
\end{aligned}$$

$$\begin{aligned}
f_{X+Y}(x) &= \frac{d}{\mathrm{dx}} F_{X+Y}(x) \\
&= \int_{-\infty}^{\infty} \frac{d}{\mathrm{dx}} F_X(x-t) f_Y(t)\, \mathrm{dt} \\
&= \int_{-\infty}^{\infty} f_X(x-t) f_Y(t)\, \mathrm{dt}.
\end{aligned}$$

6.3.1 X and Y are continuous and independent

Example (Sum of 2 Independent Uniform Random Variables). Suppose that X and Y are independent with a common uniform distribution over $(0,1)$. Find the probability density function of $X+Y$.

Solution Recall that

$$f_X(t) = f_Y(t) = \begin{cases} 1, & \text{if } 0 \leq t \leq 1 \\ 0, & \text{otherwise} \end{cases}.$$

Also note that $X+Y$ takes values in $(0,2)$. For $x \leq 0$ or $x \geq 2$, it then follows that $f_{X+Y}(x) = 0$.

Next, for $0 < x < 2$,

$$\begin{aligned} f_{X+Y}(x) &= \int_{-\infty}^{\infty} f_X(x-t) f_Y(t)\,\mathrm{dt} \\ &= \int_0^1 f_X(x-t) \times 1\,\mathrm{dt} \quad = \quad \int_0^1 f_X(x-t)\,\mathrm{dt}. \end{aligned}$$

We see that $f_X(x-t) > 0$ (in this case, it is 1) if and only if $0 < x-t < 1$ (remember that x is fixed, and t varies in $(0,1)$). To proceed, we need to separate the range of x into 2 cases:

6.3.1 X and Y are continuous and independent

Solution (cont.)

Case 1: for $0 < x \le 1$.

$$\begin{aligned} f_{X+Y}(x) &= \int_0^1 f_X(x-t)\,\mathrm{dt} \\ &= \int_0^x f_X(x-t)\,\mathrm{dt} + \int_x^1 f_X(x-t)\,\mathrm{dt} \\ &= \int_0^x 1\,\mathrm{dt} + 0 \\ &= x. \end{aligned}$$

Case 2: for $1 < x < 2$.

$$\begin{aligned} f_{X+Y}(x) &= \int_0^1 f_X(x-t)\,\mathrm{dt} \\ &= \int_0^{x-1} f_X(x-t)\,\mathrm{dt} + \int_{x-1}^1 f_X(x-t)\,\mathrm{dt} \\ &= \int_{x-1}^1 1\,\mathrm{dt} \\ &= 2-x. \end{aligned}$$

In the second last equality, we use the fact that $0 < x-t < 1$ is equivalent to $x-1 < t < x$; also we need $0 < t < 1$.
Summing up:

$$f_{X+Y}(x) = \begin{cases} x, & 0 < x \le 1 \\ 2-x, & 1 < x < 2 \\ 0, & \text{elsewhere} \end{cases}.$$

The density function has the shape of a triangle, and so the random variable $X+Y$ is sometimes known as the **triangular distribution**.

6.3.1 X and Y are continuous and independent

Theorem (Sum of 2 Independent Gamma Random Variables). *Assume that $X \sim \Gamma(\alpha, \lambda)$ and $Y \sim \Gamma(\beta, \lambda)$, and X and Y are mutually independent. Then,*

$$X + Y \sim \Gamma(\alpha + \beta,\ \lambda).$$

Note that both X and Y must have the same second parameter.

Proof

Note that for $w > 0$, $f_{X+Y}(w) = \int_{-\infty}^{\infty} f_X(w-y) f_Y(y)\,\mathrm{dy}$. And so, for $w > 0$,

$$\begin{aligned}
f_{X+Y}(w) &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\lambda^\beta}{\Gamma(\beta)} \int_0^w (w-y)^{\alpha-1} e^{-\lambda(w-y)} y^{\beta-1} e^{-\lambda y}\,\mathrm{dy} \\
&= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda w} \int_0^w (w-y)^{\alpha-1} y^{\beta-1}\,\mathrm{dy} \\
&= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda w} w^{\alpha+\beta-1} \int_0^1 (1-u)^{\alpha-1} u^{\beta-1}\,\mathrm{du} \quad (\text{letting } u = y/w) \\
&= \left(\frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 (1-u)^{\alpha-1} u^{\beta-1}\,\mathrm{du} \right) w^{\alpha+\beta-1} e^{-\lambda w}.
\end{aligned}$$

By the definition of Beta function in Chapter 5, $B(\alpha,\ \beta) = \int_0^1 (1-u)^{\alpha-1} u^{\beta-1}\,\mathrm{du} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$

Then, $$f_{X+Y}(w) = \frac{\lambda^{\alpha+\beta} w^{\alpha+\beta-1} e^{-\lambda w}}{\Gamma(\alpha+\beta)}$$

6.3.1 X and Y are continuous and independent

Theorem Let $X_1, X_2, \ldots, X_n$ be n independent exponential random variables each having parameter λ. Then, as an exponential random variable with parameter λ is the same as a Gamma random variable with parameters $(1,\ \lambda)$, we see from *last theorem* that $X_1 + X_2 + \cdots + X_n$ is a Gamma random variable with parameters $(n,\ \lambda)$.

Theorem (Sum of Independent Normal Random Variables). *If X_i, $i = 1, \ldots, n$ are independent random variables that are normally distributed with respective parameters μ_i, σ_i^2, $i = 1, \ldots, n$, then $\sum_{i=1}^n X_i$ is normally distributed with parameters $\sum_{i=1}^n \mu_i$ and $\sum_{i=1}^n \sigma_i^2$.*

Proof To begin, let X and Y be independent normal random variables, with X having mean 0 and variance σ^2, and Y having mean 0 and variance 1. We will determine the density function of $X + Y$. Now, with

$$c = \frac{1}{2\sigma^2} + \frac{1}{2} = \frac{1 + \sigma^2}{2\sigma^2}$$

6.3.1 X and Y are continuous and independent

Proof (cont.) we have
$$f_X(a-y)f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(a-y)^2}{2\sigma^2}\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)$$
$$= \frac{1}{2\pi\sigma}\exp\left(-\frac{a^2}{2\sigma^2}\right)\exp\left(-c\left(y^2-2y\frac{a}{1+\sigma^2}\right)\right).$$

Hence,
$$f_{X+Y}(a) = \frac{1}{2\pi\sigma}\exp\left(-\frac{a^2}{2\sigma^2}\right)\exp\left(\frac{a^2}{2\sigma^2(1+\sigma^2)}\right)\int_{-\infty}^{\infty}\exp\left(-c\left(y-\frac{a}{1+\sigma^2}\right)^2\right)\mathrm{d}y$$
$$= \frac{1}{2\pi\sigma}\exp\left(-\frac{a^2}{2(1+\sigma^2)}\right)\int_{-\infty}^{\infty}\exp\left(-cx^2\right)\mathrm{d}x$$
$$= C\exp\left(-\frac{a^2}{2(1+\sigma^2)}\right)$$

where C doesn't depend on a. But this implies that $X+Y$ is normal with mean 0 and variance $1+\sigma^2$. Now, suppose that X_1 and X_2 are independent normal random variables, with X_i having mean μ_i and variance σ_i^2, $i=1,2$. Then
$$X_1+X_2 = \sigma_2\left(\frac{X_1-\mu_1}{\sigma_2}+\frac{X_2-\mu_2}{\sigma_2}\right)+\mu_1+\mu_2.$$

But since $(X_1-\mu_1)/\sigma_2$ is normal with mean 0 and variance σ_1^2/σ_2^2, and $(X_2-\mu_2)/\sigma_2$ is normal with mean 0 and variance 1, it follows from our previous result that $(X_1-\mu_1)/\sigma_2+(X_2-\mu_2)/\sigma_2$ is normal with mean 0 and variance $1+\sigma_1^2/\sigma_2^2$, implying that X_1+X_2 is normal with mean $\mu_1+\mu_2$ and variance $\sigma_2^2(1+\sigma_1^2/\sigma_2^2)=\sigma_1^2+\sigma_2^2$. Thus, this proposition is established when $n=2$.

The general case now follows by induction.

6.3.1 X and Y are continuous and independent

Example A basketball team will play a 44-game season. Twenty six of these games are against class A teams and 18 are against class B teams. Suppose that the team will win each game against a class A team with probability 0.4 and will win each game against a class B team with probability 0.7. Suppose also that the results of the different games are independent. Approximate the probability that

(a) the team wins 25 games or more;

(b) the team wins more games against class A teams that it does against class B teams.

Solution

(a) Let X_A and X_B denote respectively the number of games the team wins against class A and against class B teams. Note that X_A and X_B are independent binomial random variables and

$$E(X_A) = 26(0.4) = 10.4, \quad \mathrm{var}(X_A) = 26(0.4)(0.6) = 6.24,$$
$$E(X_B) = 18(0.7) = 12.6, \quad \mathrm{var}(X_B) = 18(0.7)(0.3) = 3.78.$$

By the normal approximation to the binomial, X_A and X_B will have approximately the same distribution as would independent normal random variables with the given parameters.

6.3.1 X and Y are continuous and independent

Solution (cont.)

So $X_A + X_B$ will have approximately a normal distribution with mean 23 and variance 10.02.

$$P(X_A + X_B \geq 25) = P(X_A + X_B \geq 24.5)$$
$$= P\left(\frac{X_A + X_B - 23}{\sqrt{10.2}} \geq \frac{24.5 - 23}{\sqrt{10.2}}\right) \approx P\left(Z \geq \frac{1.5}{\sqrt{10.2}}\right) \approx 1 - P(Z < 0.4739) \approx 0.3178.$$

(b) We note that $X_A - X_B$ will have approximately a normal distribution with mean -2.2 and variance 10.02. So we have

$$P(X_A - X_B \geq 1) = P(X_A - X_B \geq 0.5)$$
$$= P\left(\frac{X_A - X_B + 2.2}{\sqrt{10.2}} \geq \frac{0.5 + 2.2}{\sqrt{10.2}}\right) \approx P\left(Z \geq \frac{2.7}{\sqrt{10.2}}\right) \approx 1 - P(Z < 0.8530)$$
$$\approx 0.1968.$$

So there is approximately a 31.78 percent chance that the team will win at least 25 games and approximately a 19.68 percent chance that the team will win more games against class A teams than against class B teams.

6.3.2. X and Y are discrete and independent

Example (Sum of 2 Independent Poisson Random Variables).

$$X \sim \text{Poisson}(\lambda), \quad Y \sim \text{Poisson}(\mu), \quad X, Y \text{ independent.}$$

Find probability mass function of $X + Y$.

First note that $X+Y$ takes values $0, 1, 2, \ldots$. For $n = 0, 1, \ldots,$

Solution

$$\begin{aligned} P(X+Y=n) &= \sum_{k=0}^{n} P(X=k,\ Y=n-k) \\ &= \sum_{k=0}^{n} P(X=k)P(Y=n-k) \quad \text{by independence} \\ &= \sum_{k=0}^{n} \frac{e^{-\lambda}\lambda^k}{k!} \times \frac{e^{-\mu}\mu^{n-k}}{(n-k)!} = e^{-(\lambda+\mu)} \sum_{k=0}^{n} \frac{\lambda^k \mu^{n-k}}{(n-k)!k!} = \frac{e^{-(\lambda+\mu)}}{n!} \sum_{k=0}^{n} \binom{n}{k} \lambda^k \mu^{n-k} \\ &= \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^n}{n!}. \end{aligned}$$

Conclusion: Sum of 2 independent Poisson random variables is still Poisson. The mean is the sum of the means.

6.3.2. X and Y are discrete and independent

Example (Sum of 2 Independent Binomial Random Variables).

$$X \sim \text{Bin}(n, p), \quad Y \sim \text{Bin}(m, p), \quad X, Y \text{ independent.}$$

Find probability mass function of $X+Y$.

Solution

First note that $X+Y$ takes values $0, 1, 2, \ldots, n+m$. For $k = 0, 1, \ldots, n+m$,

$$\begin{aligned}
P(X+Y=k) &= \sum_{i=0}^{k} P(X=i,\, Y=k-i) \\
&= \sum_{i=0}^{k} P(X=i)P(Y=k-i) \quad \text{by independence} \\
&= \sum_{i=0}^{k} \binom{n}{i} p^i q^{n-i} \times \binom{m}{k-i} p^{k-i} q^{m-(k-i)} \quad = \quad p^k q^{n+m-k} \sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i} \\
&= \binom{n+m}{k} p^k q^{n+m-k}.
\end{aligned}$$

Conclusion: Sum of 2 independent Binomial random variables with the same success probability is still Binomial with parameters $n+m$ and p.

6.3.2. X and Y are discrete and independent

Example (Sum of 2 Independent Geometric Random Variables). Let $X \sim \text{Geom}(p)$ and $Y \sim \text{Geom}(p)$, and assume that X and Y are independent. Determine the distribution of $X+Y$.

Solution

First of all, observe that $X+Y$ takes values $2,3,\ldots$. For $k \geq 2$,

$$\begin{aligned} f_{X+Y}(k) &= \sum_{i=1}^{k-1} P(X=i,\ Y=k-i) \\ &= \sum_{i=1}^{k-1} P(X=i)P(Y=k-i) \quad \text{by independence} \\ &= \sum_{i=1}^{k-1} pq^{i-1} \cdot pq^{k-i-1} \quad = \sum_{i=1}^{k-1} p^2 q^{k-2} = (k-1)p^2q^{k-2} \quad = \binom{k-1}{1} p^2 q^{k-2}, \end{aligned}$$

which is the probability mass function of a negative binomial random variable with parameters $(2,p)$.

Conclusion: For independent $X \sim \text{Geom}(p), Y \sim \text{Geom}(p)$,

$$X+Y \sim NB(2,p).$$

6.4. Conditional distributions: discrete case

Recall that $$P(B|A) := \frac{P(AB)}{P(A)} \quad \text{if } P(A) > 0.$$

The **conditional probability mass function** of X given that $Y = y$ is defined by

$$\begin{aligned} p_{X|Y}(x|y) &:= P(X = x|Y = y) \\ &= \frac{P(X = x,\ Y = y)}{P(Y = y)} \\ &= \frac{p_{X,Y}(x,y)}{p_Y(y)} \end{aligned}$$

for all values of y such that $p_Y(y) > 0$.

Similarly, the **conditional distribution function** of X given that $Y = y$ is defined by

$$F_{X|Y}(x|y) = P(X \leq x|Y = y) \quad \text{for } y \text{ such that } p_Y(y) > 0.$$

6.4. Conditional distributions: discrete case

It now follows that

$$\begin{aligned} F_{X|Y}(x|y) &:= \frac{P(X \le x,\, Y = y)}{P(Y = y)} \\ &= \frac{\sum_{a \le x} p_{X,Y}(a,y)}{p_Y(y)} = \sum_{a \le x} \frac{p_{X,Y}(a,y)}{p_Y(y)} = \sum_{a \le x} p_{X|Y}(a|y). \end{aligned}$$

Note that the definitions are exactly parallel to the unconditional case.

Theorem *If X is independent of Y, then the conditional probability mass function of X given $Y = y$ is the same as the marginal probability mass function of X for every y such that $p_Y(y) > 0$.*

Proof For y such that $p_Y(y) > 0$,

$$\begin{aligned} p_{X|Y}(x|y) &:= \frac{p_{X,Y}(x,y)}{p_Y(y)} \\ &= \frac{p_X(x)p_Y(y)}{p_Y(y)} = p_X(x). \end{aligned}$$

6.4. Conditional distributions: discrete case

Example Suppose that $p(x,y)$, the joint probability mass function of X and Y, is given by $p(0,0)=0.4,\ p(0,1)=0.2,\ p(1,0)=0.1,\ p(1,1)=0.3.$

(i) Calculate the conditional probability mass function of X given that $Y=1$.

(ii) Calculate the conditional probability mass function of X given that $Y=0$.

Solution

First note that

$$p_Y(1)=p(0,1)+p(1,1)=0.5 \quad \text{and} \quad p_Y(0)=p(0,0)+p(1,0)=0.5.$$

(i)

$$\begin{aligned} p_{X|Y}(0|1) &= \frac{p(0,1)}{p_Y(1)}=\frac{0.2}{0.5}=\frac{2}{5}; \\ p_{X|Y}(1|1) &= \frac{p(1,1)}{p_Y(1)}=\frac{0.3}{0.5}=\frac{3}{5}. \end{aligned}$$

(ii)

$$\begin{aligned} p_{X|Y}(0|0) &= \frac{p(0,0)}{p_Y(0)}=\frac{0.4}{0.5}=\frac{4}{5}; \\ p_{X|Y}(1|0) &= \frac{p(1,0)}{p_Y(0)}=\frac{0.1}{0.5}=\frac{1}{5}. \end{aligned}$$

6.4. Conditional distributions: discrete case

Example If X and Y are independent Poisson random variables with respective parameters λ and μ, calculate the conditional distribution of X given that $X+Y=n$. That is, find

$$P(X \leq k|X+Y=n) \quad \text{for all values of } k.$$

Solution

As $$P(X \leq k|X+Y=n) = \sum_{j=0}^{k} P(X=j|X+Y=n),$$

we first calculate

$$P(X=j|X+Y=n) = \frac{P(X=j,\, X+Y=n)}{P(X+Y=n)} = \frac{P(X=j,\, Y=n-j)}{P(X+Y=n)} = \frac{P(X=j)P(Y=n-j)}{P(X+Y=n)}$$

$$= \frac{\frac{e^{-\lambda}\lambda^j}{j!} \times \frac{e^{-\mu}\mu^{n-j}}{(n-j)!}}{\frac{e^{-(\lambda+\mu)}(\lambda+\mu)^n}{n!}} = \binom{n}{j}\left(\frac{\lambda}{\lambda+\mu}\right)^j \left(1-\frac{\lambda}{\lambda+\mu}\right)^{n-j}.$$

This indicates that X conditioned on $X+Y=n$ is $\text{Bin}(n, \lambda/(\lambda+\mu))$. Hence

$$P(X \leq k|X+Y=n)$$
$$= \sum_{j=0}^{k} \binom{n}{j}\left(\frac{\lambda}{\lambda+\mu}\right)^j \left(1-\frac{\lambda}{\lambda+\mu}\right)^{n-j}.$$

6.5. Conditional distributions: continuous case

Suppose that X and Y are jointly continuous random variables. We define the **conditional probability density function** of X given that $Y = y$ as

$$f_{X|Y}(x|y) := \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

for all y such that $f_Y(y) > 0$.

The use of conditional densities allows us to define conditional probabilities of events associated with one random variable when we are given the value of a second random variable.
That is, for $A \subset \mathbb{R}$ and y such that $f_Y(y) > 0$, $\quad P(X \in A|Y = y) = \int_A f_{X|Y}(x|y)\,\mathrm{d}x\,.$

In particular, the **conditional distribution function** of X given that $Y = y$ is defined by

$$F_{X|Y}(x|y) = P(X \leq x|Y = y) = \int_{-\infty}^{x} f_{X|Y}(t|y)\,\mathrm{d}t\,.$$

6.5. Conditional distributions: continuous case

Theorem *If X is independent of Y, then the conditional probability density function of X given $Y = y$ is the same as the marginal probability density function of X for every y such that $f_Y(y) > 0$.*

Proof

For y such that $f_Y(y) > 0$,

$$\begin{aligned} f_{X|Y}(x|y) &:= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\ &= \frac{f_X(x)f_Y(y)}{f_Y(y)} \\ &= f_X(x). \end{aligned}$$

6.5. Conditional distributions: continuous case

Example Suppose that the joint probability density function of X and Y is given by

$$f(x,y) = \begin{cases} \frac{15}{2}x(2-x-y), & 0<x<1,\ 0<y<1 \\ 0, & \text{otherwise} \end{cases}.$$

Compute the conditional probability density function of X given that $Y = y$ where $0 < y < 1$.

Solution

For $0 < y < 1$, we have

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

$$= \int_0^1 \frac{15}{2}x(2-x-y)\,\mathrm{dx} = \frac{15}{2}\left[\frac{2}{3} - \frac{y}{2}\right].$$

Therefore, for $0 < y < 1$

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} \frac{6x(2-x-y)}{4-3y}, & \text{for } 0<x<1 \\ 0, & \text{otherwise} \end{cases}.$$

6.5. Conditional distributions: continuous case

Example Suppose that the joint probability density function of X and Y is given by

$$f(x,y) = \begin{cases} \frac{e^{-x/y}e^{-y}}{y}, & 0 < x < \infty,\ 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}.$$

Find $P(X > 1|Y = y)$.

Solution

For $y \leq 0$, $f_Y(y) = 0$ and $P(X > 1|Y = y)$ is not defined.

For $y > 0$,

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)\,\mathrm{dx} \\ &= \int_0^{\infty} \frac{e^{-x/y}e^{-y}}{y}\,\mathrm{dx} \quad = \quad e^{-y}\int_0^{\infty} \frac{1}{y}e^{-x/y}\,\mathrm{dx} \quad = \quad e^{-y}. \end{aligned}$$

Therefore, for $y > 0$,

$$f_{X|Y}(x|y) = \begin{cases} \frac{1}{y}e^{-x/y}, & x > 0 \\ 0, & x \leq 0 \end{cases}.$$

Hence,

$$\begin{aligned} P(X > 1|Y = y) &= \int_1^{\infty} f_{X|Y}(x|y)\,\mathrm{dx} \\ &= \frac{1}{y}\int_1^{\infty} e^{-x/y}\,\mathrm{dx} \\ &= e^{-1/y}. \end{aligned}$$

6.6 Joint Probability Distribution Function of Functions of Random Variables

Let X and Y be jointly distributed random variables with joint probability density function $f_{X,Y}$. It is sometimes necessary to obtain the joint distribution of the random variables U and V, which arise as functions of X and Y.

Specifically, suppose that

$$U = g(X,Y) \quad \text{and} \quad V = h(X,Y),$$

for some functions g and h.

We want to find the joint probability density function of U and V in terms of the joint probability density function $f_{X,Y}$, g and h.

For example, X and Y are independent exponentially distributed random variables, and we are interested to know the joint probability density function of $U = X+Y$ and $V = X/(X+Y)$. In this case,

$$g(x,y) = x+y \quad \text{and} \quad h(x,y) = x/(x+y).$$

6.6 Joint Probability Distribution Function of Functions of Random Variables

Assume that the following conditions are satisfied:

(i) Let X and Y be jointly continuously distributed random variables with known joint probability density function.

(ii) Let U and V be given functions of X and Y in the form:

$$U = g(X,Y), \quad V = h(X,Y).$$

And we can uniquely solve X and Y in terms of U and V, say $x = a(u,v)$ and $y = b(u,v)$.

(iii) The functions g and h have continuous partial derivatives and

Jacobian determinant

$$J(x,y) := \begin{vmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{vmatrix} = \frac{\partial g}{\partial x}\frac{\partial h}{\partial y} - \frac{\partial g}{\partial y}\frac{\partial h}{\partial x} \neq 0.$$

Conclusion: The joint probability density function of U and V is given by

$$f_{U,V}(u,v) = f_{X,Y}(x,y)|J(x,y)|^{-1},$$

where $x = a(u,v)$ and $y = b(u,v)$.

6.6 Joint Probability Distribution Function of Functions of Random Variables

Example Let X_1 and X_2 be jointly continuous random variables with probability density function f_{X_1,X_2}. Let $Y_1 = X_1 + X_2, Y_2 = X_1 - X_2$. Find the joint density function of Y_1 and Y_2 in terms of f_{X_1,X_2}.

Solution

Let $g_1(x_1,x_2) = x_1 + x_2$ and $g_2(x_1,x_2) = x_1 - x_2$. Then $J(x_1,\ x_2) = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2.$

For instance, if X_1 and X_2 are independent, uniform $(0,1)$ random vari ables, then

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

If X_1 and X_2 were independent, exponential random variables with respective parameters λ_1 and λ_2, then

$$f_{Y_1,Y_2}(y_1,y_2)$$
$$= \begin{cases} \frac{\lambda_1\lambda_2}{2} \exp\left(-\lambda_1\left(\frac{y_1+y_2}{2}\right) - \lambda_2\left(\frac{y_1-y_2}{2}\right)\right), & y_1+y_2 \ge 0, y_1-y_2 \ge 0 \\ 0, & \text{otherwise} \end{cases}.$$

6.6 Joint Probability Distribution Function of Functions of Random Variables

Example

Suppose that X and Y are independent standard normal random variables. Show that X+Y and X-Y are independent normal random variables.

Solution

6.6 Joint Probability Distribution Function of Functions of Random Variables

Example Let X and Y be independent standard normal.

Let random variables R and Θ denote the polar coordinates of the point (x, y). That is,

$$R = \sqrt{X^2 + Y^2} \quad ; \quad \Theta = \tan^{-1}\left(\frac{Y}{X}\right).$$

(i) Find the joint probability density function of R and Θ.

(ii) Show that R and Θ are independent.

Solution

(i) To find the joint probability density function of R and Θ, consider

Step 1. Random variables R and Θ takes values in $(0, \infty)$ and in $(0, 2\pi)$.

Step 2. Now $r = g(x, y) := \sqrt{x^2 + y^2}$ and $\theta = h(x, y) := \tan^{-1}\left(\frac{y}{x}\right)$, so $x = r\cos\theta$ and $y = r\sin\theta$.

The transformed region is $0 < r < \infty$ and $0 < \theta < 2\pi$.

Step 3.

$$J(x, y) := \begin{vmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{vmatrix} = (x^2 + y^2)^{-1/2}.$$

Step 4. Therefore, for $0 < r < \infty$ and $0 < \theta < 2\pi$,

$$\begin{aligned} f_{R,\Theta}(r, \theta) &= f_{X,Y}(x, y)/|J(x, y)| \\ &= \sqrt{x^2 + y^2} f_{X,Y}(x, y) = \sqrt{x^2 + y^2}\frac{1}{2\pi}e^{-(x^2+y^2)/2} = r\frac{1}{2\pi}e^{-r^2/2}. \end{aligned}$$

6.6 Joint Probability Distribution Function of Functions of Random Variables

Solution (cont.)

(ii) Notice that

$$f_{R,\Theta}(r,\theta) = \left[re^{-r^2/2}\right] \times \frac{1}{2\pi}$$

for all r and θ, R and Θ are independent.

Indeed,

$$f_R(r) = re^{-r^2/2}, \quad \text{for } 0 < r < \infty$$

and

$$f_\Theta(\theta) = \frac{1}{2\pi}, \quad \text{for } 0 < \theta < 2\pi.$$

R is said to have the **Rayleigh** distribution.

6.6 Joint Probability Distribution Function of Functions of Random Variables

Example If X and Y are independent Gamma random variables with parameters (α, λ) and (β, λ), respectively, compute the joint density of $U = X + Y$ and $V = X/(X+Y)$.

Solution

The joint density of X and Y is given by

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

Now, if $g_1(x,y) = x + y$, $g_2(x,y) = x/(x+y)$, then

$$J(x,y) = \begin{vmatrix} 1 & 1 \\ \frac{y}{(x+y)^2} & \frac{-x}{(x+y)^2} \end{vmatrix} = -\frac{1}{x+y}.$$

Finally, as the equations $u = x + y$, $v = x/(x+y)$ have as their solutions $x = uv$, $y = u(1-v)$, we see that

$$f_{U,V}(u,v) = f_{X,Y}[uv, u(1-v)] \times u$$
$$= \frac{\lambda e^{-\lambda u}(\lambda u)^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)} \times \frac{v^{\alpha-1}(1-v)^{\beta-1}\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}.$$

Hence $X + Y$ and $X/(X+Y)$ are independent, with $X + Y$ having a Gamma distribution with parameters $(\alpha+\beta, \lambda)$ and $X/(X+Y)$ is said to have a Beta distribution with parameters (α, β). The above also shows that $B(\alpha, \beta)$, the

6.6 Joint Probability Distribution Function of Functions of Random Variables

When the joint density function of the n random variables $X_1, X_2, \ldots, X_n$ is given and we want to compute the joint density function of $Y_1, Y_2, \ldots, Y_n$, where $Y_1 = g_1(X_1, \ldots, X_n)$, $Y_2 = g_2(X_1, \ldots, X_n)$, $\ldots$, $Y_n = g_n(X_1, \ldots, X_n)$, the approach is the same.

Namely, we assume that the functions g_j have continuous partial derivatives and that the Jacobian determinant $J(x_1, \ldots, x_n) \neq 0$ at all points $(x_1, \ldots, x_n)$, where

$$J(x_1, \ldots, x_n) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial g_n}{\partial x_n} \end{vmatrix}.$$

Furthermore, we suppose that the equations $y_1 = g_1(x_1, \ldots, x_n)$, $y_2 = g_2(x_1, \ldots, x_n)$, $\ldots$, $y_n = g_n(x_1, \ldots, x_n)$ have a unique solution, say, $x_1 = h_1(y_1, \ldots, y_n)$, $\ldots$, $x_n = h_n(y_1, \ldots, y_n)$. Under these assumptions, the joint density function of the random variables Y_i is given by

$$f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) = f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)|J(x_1, \ldots, x_n|^{-1},$$

where $x_i = h_i(y_1, \ldots, y_n)$, $i = 1, 2, \ldots, n$.

6.6 Joint Probability Distribution Function of Functions of Random Variables

Example Let X_1, X_2, and X_3 be independent standard normal random variables. If $Y_1 = X_1 + X_2 + X_3$, $Y_2 = X_1 - X_2$, $Y_3 = X_1 - X_3$, compute the joint density function of Y_1, Y_2, Y_3.

Solution

Letting $Y_1 = X_1 + X_2 + X_3$, $Y_2 = X_1 - X_2$, $Y_3 = X_1 - X_3$, the Jacobian of these transformations is given by

$$J = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{vmatrix} = 3.$$

As the transformations above yield that $X_1 = \frac{Y_1+Y_2+Y_3}{3}$, $X_2 = \frac{Y_1-2Y_2+Y_3}{3}$, $X_3 = \frac{Y_1+Y_2-2Y_3}{3}$, we see that

$$f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3) = \frac{1}{3} f_{X_1,X_2,X_3}\left(\frac{y_1+y_2+y_3}{3}, \frac{y_1-2y_2+y_3}{3}, \frac{y_1+y_2-2y_3}{3}\right).$$

Hence, as $$f_{X_1,X_2,X_3}(x_1,x_2,x_3) = \frac{1}{(2\pi)^{3/2}} e^{-\sum_{i=1}^{3} x_i^2/2}$$

we see that $$f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3) = \frac{1}{3(2\pi)^{3/2}} e^{-Q(y_1,y_2,y_3)/2}$$ where

$$Q(y_1,y_2,y_3) = \left(\frac{y_1+y_2+y_3}{3}\right)^2 + \left(\frac{y_1-2y_2+y_3}{3}\right)^2 + \left(\frac{y_1+y_2-2y_3}{3}\right)^2 = \frac{y_1^2}{3} + \frac{2}{3}y_2^2 + \frac{2}{3}y_3^2 - \frac{2}{3}y_2y_3.$$

6.7 Jointly distributed random variables: n > 2

We will illustrate the results for 3 jointly distributed random variables, called X, Y and Z. We will assume they are jointly continuous random variables.

$$F_{X,Y,Z}(x,y,z) := P(X \le x,\ Y \le y,\ Z \le z).$$

There are a number of marginal distribution functions, namely

$$\begin{aligned} F_{X,Y}(x,y) &:= \lim_{z\to\infty} F_{X,Y,Z}(x,y,z); & F_X(x) &:= \lim_{y\to\infty,\ z\to\infty} F_{X,Y,Z}(x,y,z); \\ F_{X,Z}(x,z) &:= \lim_{y\to\infty} F_{X,Y,Z}(x,y,z); & F_Y(y) &:= \lim_{x\to\infty,\ z\to\infty} F_{X,Y,Z}(x,y,z); \\ F_{Y,Z}(y,z) &:= \lim_{x\to\infty} F_{X,Y,Z}(x,y,z); & F_Z(z) &:= \lim_{x\to\infty,\ y\to\infty} F_{X,Y,Z}(x,y,z). \end{aligned}$$

6.7 Jointly distributed random variables: n > 2

Joint probability density function of X, Y and Z: $f_{X,Y,Z}(x,y,z)$

For any $D \subset \mathbb{R}^3$, we have

$$P((X,Y,Z) \in D) = \int\int\int_{(x,y,z)\in D} f_{X,Y,Z}(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,.$$

Let $A, B, C \subset \mathbb{R}$, take $D = A \times B \times C$ above

$$P(X \in A, Y \in B, Z \in C) = \int_C \int_B \int_A f_{X,Y,Z}(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,.$$

6.7 Jointly distributed random variables: n > 2

Marginal probability density function of X, Y and Z

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z)\,\mathrm{dy}\,\mathrm{dz};\\
f_Y(y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z)\,\mathrm{dx}\,\mathrm{dz};\\
f_Z(z) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z)\,\mathrm{dx}\,\mathrm{dy};\\
f_{X,Y}(x,y) &= \int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z)\,\mathrm{dz};\\
f_{Y,Z}(y,z) &= \int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z)\,\mathrm{dx};\\
f_{X,Z}(x,z) &= \int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z)\,\mathrm{dy}\,.
\end{aligned}
$$

6.7 Jointly distributed random variables: n > 2

Independent random variables

Theorem *For jointly continuous random variables, the following three statements are equivalent:*

(i) Random variables X, Y and Z are independent.

(ii) For all $x, y, z \in \mathbb{R}$, we have

$$f_{X,Y,Z}(x,y,z) = f_X(x) f_Y(y) f_Z(z).$$

(iii) For all $x, y, z \in \mathbb{R}$, we have

$$F_{X,Y,Z}(x,y,z) = F_X(x) F_Y(y) F_Z(z).$$

6.7 Jointly distributed random variables: n > 2

Checking for independence

Theorem *Random variables X, Y and Z are independent if and only if there exist functions $g_1, g_2, g_3 : \mathbb{R} \longrightarrow \mathbb{R}$ such that for all $x, y, z \in \mathbb{R}$, we have*

$$f_{X,Y,Z}(x,y,z) = g_1(x)g_2(y)g_3(z).$$

Conditional distributions

There are a number of conditional probability density functions, to name a few,

$$\begin{aligned} f_{X,Y|Z}(x,y|z) &:= f_{X,Y,Z}(x,y,z)/f_Z(z); \\ f_{X|Y,Z}(x|y,z) &:= f_{X,Y,Z}(x,y,z)/f_{Y,Z}(y,z) \end{aligned}$$

and so on ...