1.3 Continuous Joint Distribution

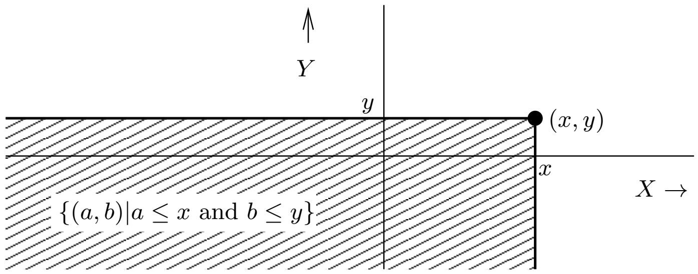
Joint distribution function

Before we calculated probabilities concerning 2 or more discrete random variables, defined on the same probability space, with the use of the joint proba-bility function. For 2 or more random variables, we define the joint distribution function (which in principle is usable for discrete random variables):

Definition **1.3.1** The joint distribution function $F_{X,Y}$of the random variables X and $Y$is defined as

$$F_{X,Y}(x,y) = P(X \leq x \text{ and } Y \leq y).$$

And analogous for $X_1, X_2, \ldots, X_n$:
$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = P(X_1 \leq x_1 \text{ and } \ldots \text{ and } X_n \leq x_n)$.
So the distribution function of X and Y in the point (x,y) is the probability that $\{X \leq x \text{ and } Y \leq y\}$ occurs. This means that the pair (X,Y) takes values in the bottom-left "quadrant" of point (x,y) in the flat plane.



If we in this figure let x increase with a constant y value, the half-surface $\{Y \leq y\}$ develops in the limit. So

$$\lim_{x\to\infty} F_{X,Y}(x,y) = \lim_{x\to\infty} P(X \leq x \text{ and } Y \leq y) \stackrel{*}{=} P(Y \leq y) = F_Y(y).$$

(The limit transition * is intuitively right, but actually has to be verified with measure theory.)***

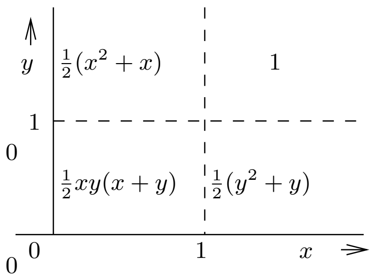
As well, $F_X(x) = \lim_{y \to \infty} F_{X,Y}(x, y)$ applies. In this respect the functions F_X and F_Y are called the marginal distribution function of X and Y respectively. Furthermore

$$\lim_{\substack{x \to \infty \\ y \to \infty}} F_{X,Y}(x, y) = 1; \lim_{x \to -\infty} F_{X,Y}(x, y) = \lim_{y \to -\infty} F_{X,Y}(x, y) = 0$$

applies. The not-falling and semi-continuity of the marginal distribution function F_X also applies for both arguments of the joint distribution function $F_{X,Y}$, as described in the following mathematical example.

Example 1.3.2 The continuous random variables X and Y have the following joint distribution function $F_{X,Y}$(the defined area are indicated schematically):

$$F_{X,Y}(x, y) = \begin{cases} \frac{1}{2}xy(x+y) & \text{if } x, y \in [0, 1] \\ \frac{1}{2}(x^2 + x) & \text{if } x \in [0, 1] \text{ and } y > 1 \\ \frac{1}{2}(y^2 + y) & \text{if } y \in [0, 1] \text{ and } x > 1 \\ 1 & \text{if } x > 1 \text{ and } y > 1 \\ 0 & \text{else} \end{cases}$$



From this we can easily derive:

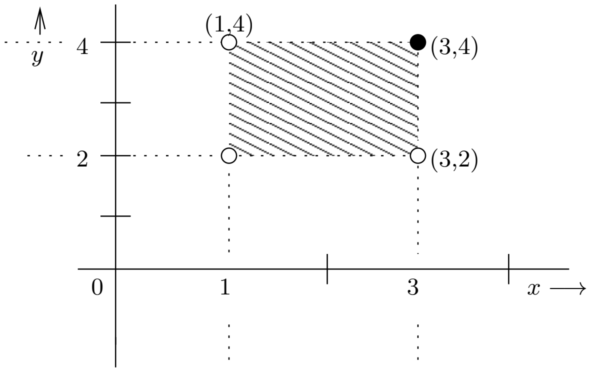
$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x, y) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{2}(x^2 + x) & \text{if } x \in [0, 1] \\ 1 & \text{if } x > 1 \end{cases}$$

Out of symmetry arguments $F_Y(y) = F_X(y)$ applies. These are continuous functions of continuous random variables with probability densities of $f_X(x) = f_Y(x) = x + \frac{1}{2}$ for $x \in [0, 1]$. □

Example 1.3.3 Calculate the probability $P(1 < X \leq 3 \text{ and } 2 < Y \leq 4)$ using the joint distribution function of X and Y, given by:

$$F(x, y) = \begin{cases} 1 - e^{-2x} - e^{-y} + e^{-2x-y} & \text{for } x \geq 0 \text{ and } y \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

$F(3, 4)$ calculates the probability of the event $\{X \leq 3 \text{ and } Y \leq 4\}$, that encompasses the event $\{1 < X \leq 3 \text{ and } 2 < Y \leq 4\}$ (see figure)



Because $\{1 < X \leq 3$ and $2 < Y \leq 4\}$ with $\{X \leq 1$ and $Y \leq 4\}$ and $\{1 < X \leq 3$ and $Y \leq 2\}$ form a partition with $\{X \leq 3$ and $Y \leq 4\}$, we can say that

$$\begin{aligned} F(3,4) &= P(X \leq 3 \text{ and } Y \leq 4) \\ &= P(1 < X \leq 3 \text{ and } 2 < Y \leq 4) + P(X \leq 1 \text{ and } Y \leq 4) \\ &\quad + P(1 < X \leq 3 \text{ and } Y \leq 2). \end{aligned}$$

If we also apply

$$P(1 < X \leq 3 \text{ and } Y \leq 2) = F(3,2) - F(1,2),$$

we find

$$\begin{aligned} &P(1 < X \leq 3 \text{ and } 2 < Y \leq 4) = F(3,4) - F(1,4) - F(3,2) + F(1,2) \\ &= (1 - e^{-6} - e^{-4} + e^{-10}) - (1 - e^{-2} - e^{-4} + e^{-6}) - (1 - e^{-6} - e^{-2} + e^{-8}) \\ &\quad + (1 - e^{-2} - e^{-2} + e^{-4}) \\ &= e^{-10} - e^{-6} - e^{-8} + e^{-4} (\approx 1.55\%). \end{aligned}$$

As well, we find $P(0.5 < X \leq 1.5$ and $0.5 < Y \leq 1.5) \approx 0.122$. □

In this example, we saw the calculation of the probability that the pair (X, Y) takes on values in a rectangular area in the xy-plane, using the joint distribution function. Now, we will generalise this property and summarize the most important properties of joint distribution functions.

Property 1.3.4 If X and Y have a joint distribution function $F_{X,Y}$, then:

a. $P(X \in < a, b]$ and $Y \in < c, d]) = F_{X,Y}(b,d) - F_{X,Y}(a,d) - F_{X,Y}(b,c) + F_{X,Y}(a,c)$

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

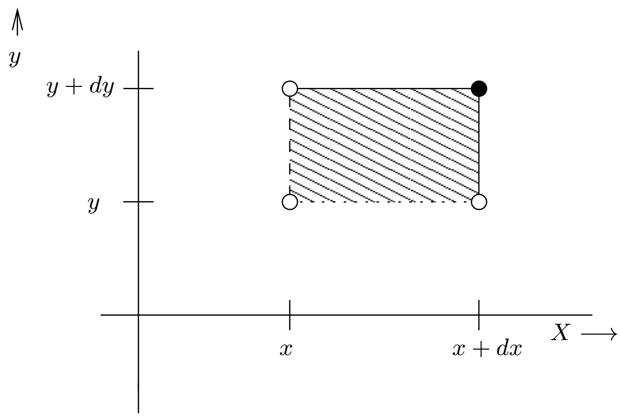
Equal properties can be described for three or more random variables. If X and Y are continuous random variables, the intervals $]a,b]$ and $]c,d]$ can be either open or closed, after all: $P(X = x) = 0$ and $P(Y = y) = 0$ for all x and y.
In example 7.1.2, the functions of X and Y that were derived from $F_{X,Y}$ appeared to be continuous. $F_{X,Y}$ is continuous in both arguments as well. Because $P(X = x) = 0$ applies for all $x \in \mathbb{R}$, $P(X = x \text{ and } Y = y) = 0$ holds true as well. (X,Y)

Joint probability density

The marginal distribution functions derived from $F_{X,Y}$ of X and Y, appeared to be both continuous in example 1.3.2. $F_{X,Y}$ as well is continuous in both arguments. Because $P(X = x) = 0$ for all $x \in \mathbb{R}$, we can definitely say that $P(X = x \text{ and } Y = y) = 0$. The pair (X,Y) can solely with a positive probability become a subspace B of $\mathbb{R}^2$, if B does not consist of either line segments or points, but consists of a positive area.

Definition 1.3.5 A pair of random variables (X,Y) are called **continuous** if the joint distribution function $F_{X,Y}$is continuous in both arguments.

If we consider the probability $X \in]x, x+dx]$ and $Y \in]y, y+dy]$ (see figure), rectangular property 7.1.4 holds true.



If the distribution function F of the pair of continuous random variables (X,Y) is differentiable in both arguments, we can find the following (heuristic) equation:

$$P(x < X \leq x+dx \text{ and } y < Y \leq y+dy)$$
$$= (F(x+dx, y+dy) - F(x+dx, y)) - (F(x, y+dy) - F(x,y))$$

$$\approx \frac{\partial F(x+dx,y)}{\partial y} \cdot dy - \frac{\partial F(x,y)}{\partial y} \cdot dy$$
$$= \frac{\partial}{\partial y}[F(x+dx,y) - F(x,y)] \cdot dy$$
$$\approx \frac{\partial}{\partial y}\frac{\partial}{\partial x}[F(x,y)] \cdot dxdy.$$

Analogous to the situation with one-dimensional distribution functions as in chapter VI, we assume here too that the partial derivatives indeed exist, except for cases where the area is zero. Therefore we define:

Definition 1.3.6 The **joint probability density** $f_{X,Y}$ of a pair of continuous random variables (X, Y) is defined by

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y).$$

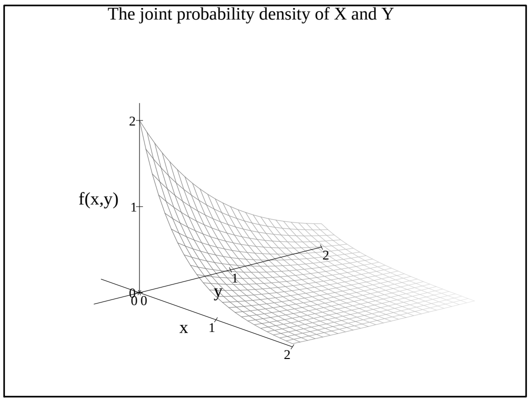
(Sometimes we omit the subscripts X, Y)
The function $F_{X,Y}(x,y)$ should be continuously differentiable.
So $P(X \in]x, x+dx]$ and $Y \in]y, y+dy])$ is approximately given by the probability density $f(x,y)$ times the area $dxdy$ of rectangle: $f(x,y)dxdy$. This probability is bigger or equal to 0, and therefore also $f(x,y) \geq 0$.

Example 1.3.7 With the help of definition 1.3.2 we find for the continuous F of example 1.3.3:

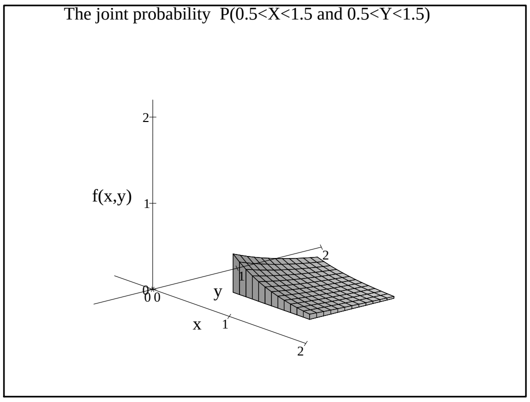
$$f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y) = \begin{cases} 2e^{-2x-y} & \text{if } x > 0 \text{ and } y > 0 \\ 0 & \text{if } x < 0 \text{ or } y < 0 \end{cases}$$

The graph of f can be plotted in a three-dimensional figure by putting the values of $f(x,y)$ on the Z-axis.

The joint probability density of X and Y



The joint probability P(0.5<X<1.5 and 0.5<Y<1.5)



We can leave f for $x = 0$ or $y = 0$ undefined, but usually we give f an obvious (non-negative) value, so $f(x,y) \geq 0$ for all pairs (x,y).

Furthermore we can see that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = \int_{0}^{\infty}\int_{0}^{\infty} 2e^{-2x-y}dxdy$$
$$= \int_{0}^{\infty} 2e^{-2x}dx \int_{0}^{\infty} e^{-y}dy = 1.$$

Graphically, the calculation above means that the encompassed volume in the graph of f and the XY-plane is exactly 1. De probability $P(0.5 < X \leq 1.5$ and $0.5 < Y \leq 1.5)$ can also be calculated using the probability density $f(x,y)$. Therefore we define the area $A = \{(a,b)|0.5 < a \leq 1.5 \text{ and } 0.5 < b \leq 1.5\}$ in a finite amount of rectangles of form $]x, x+dx]\times]y, y+dy]$.

Then we can say that:

$$P(0.5 < X \leq 1.5 \text{ and } 0.5 < Y \leq 1.5) \approx \sum_{A} f(x,y)dxdy,$$

which gives for infinitesimal dx and dy the following integral:

$$P(0.5 < X \leq 1.5 \text{ and } 0.5 < Y \leq 1.5) = \int_{0.5}^{1.5}\int_{0.5}^{1.5} f(x,y)dxdy = \int_{0.5}^{1.5}\int_{0.5}^{1.5} 2e^{2x-y}dxdy$$
$$= \int_{0.5}^{1.5} 2e^{-2x}dx \int_{0.5}^{1.5} e^{-y}dy = e^{-2x}\Big|_{0.5}^{1.5} \cdot e^{-y}\Big|_{0.5}^{1.5}$$
$$= (e^{-3} - e^{-1})(e^{-1.5} - e^{-0.5}) \approx 0.122.$$

the probability can be graphically interpreted as the volume below the plot of f and above rectangle A in the XY -plane (see above figure). Compare this calculation with that of example 1.3.3.

If we choose a quarter-plane at the left bottom of point (x,y) for the previously given derivation for A, we can derive $F_{X,Y}$ from f again:

$$F_{X,Y}(x,y) = P(X \leq x \text{ and } Y \leq y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)dudv.$$

Voor $u < 0$ of $v < 0$ is $f(u,v) = 0$ so $F_{X,Y}(x,y) = 0$ if $x < 0$ of $y < 0$. For $x \geq 0$ and $y \geq 0$ we find

$$F_{X,Y}(x,y) = \int_{0}^{y}\int_{0}^{x} 2e^{-2u-v}dudv = (1 - e^{-2x})(1 - e^{-y}).$$

□

Property 1.3.8 For two continuous random variables (X, Y) with joint probability density $f_{X,Y}$ the following applies:

a. $f_{X,Y}(x, y) \geq 0$ for all pairs (x, y).

b. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$.

c. Voor $A \subset \mathbb{R}^2$ the following applies: $P((X, Y) \in A) = \int\int_A f_{X,Y}(x, y) dx dy$.

d. $F_{X,Y}(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(u, v) du dv$.

e. $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$ and $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$.

f. $P(X = x) = P(Y = y) = P(X = x \text{ and } Y = y) = 0$.

Properties 1.3.8a and 1.3.8c are by definition 1.3.6 and respectively example 1.3.7 made plausible. Property 1.3.8b is deduced from 1.3.8c by choosing $A = \mathbb{R}^2$. Likewise men can deduce the properties 1.3.8d and 1.3.8f from 1.3.8c. Property 1.3.8e, determining the marginal probability densities out of the joint probability density, can be proven by using properties 1.3.8b and 1.3.8d:

$$\begin{aligned} f_X(x) &= \frac{d}{dx} F_X(x) = \frac{d}{dx}\left[\lim_{y\to\infty} F_{X,Y}(x, y)\right] \\ &= \frac{d}{dx}\left[\int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{X,Y}(u, v) dv du\right] \\ &= \int_{-\infty}^{\infty} f_{X,Y}(x, v) dv. \end{aligned}$$

In case a two-dimensional function ***f* satisfies the properties** 1.3.8**a and** 1.3.8**b**, we can speak of a **joint probability density** again. Property 1.3.8c can also be used in case the area A is not rectangular.

Example 1.3.9 After differentiation, The joint distribution function in ex. 1.3.2 leads to the joint probability density

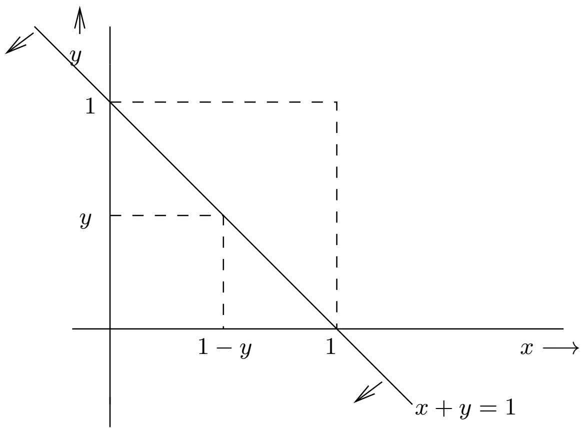
$$f_{X,Y}(x, y) = \begin{cases} x + y & \text{if } 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

The probability density of X can be derived with the property 7.2.4e out of $f_{X,Y}$ as well:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_0^1 (x+y)dy = xy + \frac{1}{2}y^2 \Big|_{y=0}^{y=1} = x + \frac{1}{2},$$

for $0 \leq x \leq 1$ and $f_X(x) = 0$ elsewhere.

Calculation of the probability $P(X+Y \leq 1)$ is not possible with the distribution function $F_{X,Y}$, because the half-plane $A = \{(x,y)|x+y \leq 1\}$ is not rectangular (see figure).



This IS possible with the joint probability density (Property 1.3.8c)

$$\begin{aligned} P(X+Y \leq 1) &= \int\int_A f_{XY}(x,y)dxdy = \int_0^1 \int_0^{1-y} (x+y)dxdy \\ &= \int_0^1 \frac{1}{2}(1-y^2)dy = \frac{1}{3}\,. \end{aligned}$$

□

Remark
Analogous to the discrete case, we can also look at continuous conditional distributions. Without discussing this further, we will give the definition of the conditional probability density and conditional expected value (notice the similarities to the conditional probability function!)

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

and

$$E(X|Y=y) = \int\limits_{-\infty}^{\infty} x f_X(x|Y=y)dx.$$

In example 1.3.9, as long as value y of Y is known, the following applies:

$$f_X(x|Y=y) = \frac{f_{X,Y}(x,y)}{f_y(y)} = \frac{x+y}{\frac{1}{2}+y} \text{ voor } 0 \leq x \leq 1$$
$$E(X|Y=y) = \int\limits_{-\infty}^{\infty} x \cdot f_X(x|Y=y)dx = \int\limits_{0}^{1} x \cdot \frac{x+y}{\frac{1}{2}+y} dx = \frac{2+3y}{3+6y}.$$

The conditional expectation, given $Y = y$, is dependent on the value of y, so between $\frac{5}{9}$ and $\frac{2}{3}$. □

Independence

In example 1.3.2 we can see t hat events {X ≤ 1/2} and {Y ≤ 1/2} are not independent:

$P(X \leq \frac{1}{2}$ and $Y \leq \frac{1}{2}) = F_{X,Y}(\frac{1}{2}, \frac{1}{2}) = \frac{1}{8}$ and

$P(X \leq \frac{1}{2}) \cdot P(Y \leq \frac{1}{2}) = F_X(\frac{1}{2}) \cdot F_Y(\frac{1}{2}) = \frac{3}{8} \cdot \frac{3}{8}$ are not the same.

X and Y will be called dependent.

Definition 1.3.10 X and Y are (mutually) independent random variables if or all $B \subset \mathbb{R}$ and $C \subset \mathbb{R}$

$$P(X \in B \text{ and } Y \in C) = P(X \in B) \cdot P(Y \in C).$$

(An analogous definition applies for $X_1, X_2, \ldots, X_n$.)
In an exercise it is assessed if the definition 5.3.1 on independence of discrete random variables is equal to this general definition. For continuous random variables the independence can be assessed with the following property.

Property 1.3.11 For continuous random variables X and Y the following applies:

a. X and Y are independent. $\Longleftrightarrow F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y)$ for all $(x,y) \in \mathbb{R}^2$

b. X and Y are independent. $\Longleftrightarrow f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$ for all $(x, y) \in \mathbb{R}^2$.

A part of this proof is asked in the exercises

Example 1.3.11 We can check if the random variables X and Y with joint distribution function F in example 1.3.3 independent, by the following two methods (according to 1.3.10).

1. First calculate the marginal distribution function.

$$F_X(x) = \lim_{y\to\infty} F(x, y) = \begin{cases} 1 - e^{-2x} & \text{for } x \geq 0 \\ 0 & \text{elsewhere..} \end{cases}$$

$$F_Y(y) = \lim_{x\to\infty} F(x, y) = \begin{cases} 1 - e^{-y} & \text{for } y \geq 0 \\ 0 & \text{elsewhere.} \end{cases}$$

For $x \geq 0$ and $y \geq 0$ now applies:

$$F(x, y) = 1 - e^{-2x} - e^{-y} + e^{-2x-y} = (1 - e^{-2x})(1 - e^{-y}) = F_X(x) \cdot F_Y(y)$$

and for other values of x and y the following applies:

$$F(x, y) = 0 = F_X(x) \cdot F_Y(y).$$

So according to property 1.3.10a, X and Y are independent.

2. The same conclusion can me made for property 7.3.2b:
For $x \geq 0$ and $y \geq 0$

$$f(x, y) = 2e^{-2x-y} = 2e^{-2x} \cdot e^{-y} = f_X(x) \cdot f_Y(y)$$

and

$$f(x, y) = 0 = f_X(x) \cdot f_Y(y) \text{ elsewhere.}$$

The marginal probability densities f_X and f_Y can be derived from $f_{X,Y}$ or F_X and F_Y.

Lastly we notice that X and Y both contain an exponential distribution with parameter $\lambda = 2$ resp. $\lambda = 1$. □

Usually we will not check the independence, but if it is implicitly or explicitly given, then: If $X_1, X_2, \ldots, X_n$ relate to an equal amount of independent experiments and the marginal distributions of $X_1, X_2, \ldots, X_n$ are known, the joint distribution of $X_1, X_2, \ldots, X_n$ follows. Generalisation of the property 1.3.10 with independence gives us:

$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)$$

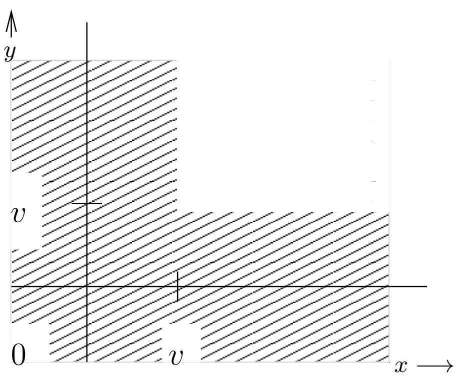
and

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

Functions of continuous random variables; correlation

Example 1.3.12 A certain hospital has two backup power supplies of the same type. It is presumed that the working time X and Y of these supplies at the point of power outage are distributed mutually independent and exponentially with known parameter λ. In case one single supply is capable of powering all of the hospital on its own, the other one can be powered up if the first one fails. We want to know the distribution of $Z = X + Y$. If both supplies have to be powered up at the moment of power failure we want to know the distribution of $V = \min(X, Y)$ as well as that of $W = \max(X, Y)$: the probability distribution of complete resp. partial power provision.

For defining the distribution $V = \min(X, Y)$ we firstly try to express the distribution function of V in that of X and Y.



$$\begin{aligned} F_V(v) &= P(\min(X,Y) \leq v) \\ &= 1 - P(\min(X,Y) > v) \\ &= 1 - P(X > v \text{ and } Y > v) \\ &= 1 - P(X > v) \cdot P(Y > v), \text{because } X \text{ and } Y \text{ are o.o.} \\ &= 1 - [1 - F_X(v)] \cdot [1 - F_Y(v)]. \end{aligned}$$

Because both X and Y are exponentially distributed with parameter λ, we can say that:

$$F_V(v) = 1 - (1 - F_X(v))^2 = \begin{cases} 1 - e^{-2\lambda v} & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases}$$

and

$$f_V(v) = \frac{d}{dv} F_V(v) = \begin{cases} 2\lambda e^{-2\lambda v} & \text{if } v \geq 0 \\ 0 & \text{als } v < 0. \end{cases}$$

$V = \min(X,Y)$ is apparently exponentially distributed as well, with parameter 2λ and thus $E(V) = \frac{1}{2\lambda} = \frac{1}{2}E(X)$.
For the distribution of $W = \max(X,Y)$ we refer to problem 9.
We later come back to the distribution of $X + Y$ in example 7.4.4. □

In example 7.4.1 found the expected value of $V = g(X,Y)$ by firstly deriving the probability density of V from that of X and Y. We can (analogous to the discrete case) calculate $Eg(X,Y)$ as well by weighing the values of $g(x,y)$ with the "probability" $f_{X,Y}(x,y)dxdy$.

Property 1.3.13 Without proof for two random variables X and Y and a real function g applies:

$$Eg(X,Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)dxdy.$$

Analogous to the discrete case, we can derive from this property (which is easily generalised to n random variables) that:

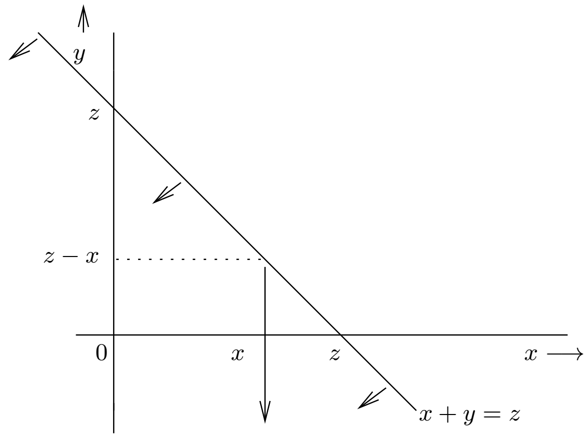
$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} EX_i.$$

But if we are interested in the probability distribution $X_1 + \ldots + X_n$, we must derive that out of the joint probability density. Again, we look at two random variables X and Y and their sum $Z = X + Y$.

Then we can firstly derive the equation of Z als follows:

$$F_Z(z) = P(Z \leq z) = P(X + Y \leq z).$$

The area (in $\mathbb{R}^2$) where (X, Y) takes on values, so $X + Y \leq z$, the half-plane is defined $\{(x, y)|x + y \leq z\}$, as plotted in the following figure:



We can interpret $F_Z(z)$ as the volume under the graph of $f_{X,Y}$ above this half-plane, so:

$$\begin{aligned} F_Z(z) &= \iint\limits_{x+y\leq z} f_{X,Y}(x, y)dydx \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{z-x} f_{X,Y}(x, y)dy \right] dx. \end{aligned}$$

With the use of this expression, in principle we can determine the distribution Z and thus also determine $E(Z)$, if $f_{X,Y}$ is known (see example 1.3.10). We can however derive an expression for the probability density of Z.
Therefore we substitute, in $\int_{-\infty}^{z-x} f_{X,Y}(x, y)dy, \quad y = t - x$,
so $dy = dt$ and $t = y + x$ goes from $-\infty$ to z:

$$\begin{aligned} F_Z(z) &= \int_{-\infty}^{\infty} \int_{-\infty}^{z} f_{X,Y}(x, t - x)dtdx \\ &= \int_{-\infty}^{z} \int_{-\infty}^{\infty} f_{X,Y}(x, t - x)dxdt \end{aligned}$$

(by changing integration order)

Where this integral has the form of $F_Z(z) = \int_{-\infty}^{z} g(t)dt$ and therefore $f_Z(z) = g(z)$ applies:

$$f_Z(z) = \frac{d}{dz}F_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z - x)dx.$$

If X and Y are also mutually independent then:

$$f_{X,Y}(x, z-x) = f_X(x) \cdot f_Y(z-x).$$

In that case we find an integral, which is the continuous equivalent of a convolution sum (compare with property 1.3.11).

Property 1.3.14 (the **convolution integral)**
If X and Y are independent continuous random variables, for $Z = X+Y$ applies that:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx.$$

(f_Z is also written as $f_X * f_Y$.)
Example 1.3.15 If Z is the total working time $X + Y$ of the two backup power supplies in example 7.4.1, it follows from property 7.4.3 that:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \int_0^z \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda(z-x)}dx \\ &= \int_0^z \lambda^2 e^{-\lambda z}dx = \lambda^2 z e^{-\lambda z} \text{ als } z > 0 \end{aligned}$$

for $z < 0$, $f_Z(z) = 0$ applies, because $f_X(x) = 0$ is of $f_Y(z-x) = 0$.

This means that $E(Z) = \int_0^\infty z \cdot \lambda^2 z e^{-\lambda z} dz = \ldots = \frac{2}{\lambda}$, a value that follows immediately from $E(Z) = E(X+Y) = E(X) + E(Y) = \frac{1}{\lambda} + \frac{1}{\lambda}$. □

With the use of property 7.4.2, $\text{cov}(X,Y) = E(X-\mu_X)(Y-\mu_Y)$ and $\rho(X,Y)$ are also defined in the continuous case. All properties for covariance and correlation still apply (where needed, the prove only needs to change $P(X = x \text{ and } Y = y)$ for $f_{X,Y}(x,y)dxdy$ and the summations for integrals). Also, the weak law of large numbers and properties concerning independence, which are discussed in chapter V, still apply.

Example 1.3.16 Example 1.3.9 shows the joint probability density $f_{X,Y}(x,y) = x + y$, for $0 \leq x \leq 1$ and $0 \leq y \leq 1$, marginal probability density of X and Y derived:

$$f_X(x) = f_Y(x) = \begin{cases} \frac{1}{2} + x & \text{for } 0 \leq x \leq 1 \\ 0 & \text{elsewhere.} \end{cases}$$

From that we calculate:

$$E(Y) = E(X) = \int_{-\infty}^{\infty} x f_X(x)dx = \int_0^1 x(\frac{1}{2} + x)dx = \frac{7}{12}$$

$$E(Y^2) = E(X^2) = \int_0^1 x^2 \cdot (\frac{1}{2} + x)dx = \frac{5}{12}$$

$$\text{var}(Y) = \text{var}(X) = E(X^2) - E(X)^2 = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{11}{144}$$

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf_{X,Y}(x,y)dxdy = \int_0^1 \int_0^1 xy(x+y)dxdy = \frac{1}{3}$$

$$\text{cov}(X,Y) = E(XY) - \mu_X\mu_Y = \frac{1}{3} - \frac{7}{12} \cdot \frac{7}{12} = -\frac{1}{144}$$

$$\rho(X,Y) = \frac{\text{cov}(X,Y)}{\sigma_X\sigma_Y} = \frac{-1/144}{11/144} = -\frac{1}{11}$$

$$\text{var}(X+Y) = \text{var}(X)+\text{var}(Y)+2\text{cov}(X,Y) = \frac{11}{144}+\frac{11}{144}-2 \cdot \frac{1}{144} = \frac{20}{144}\,.$$

We notice that $\text{cov}(X,Y) \neq 0$, so X and Y are dependent random variables, which also follows from the probability densities of the distribution functions. □