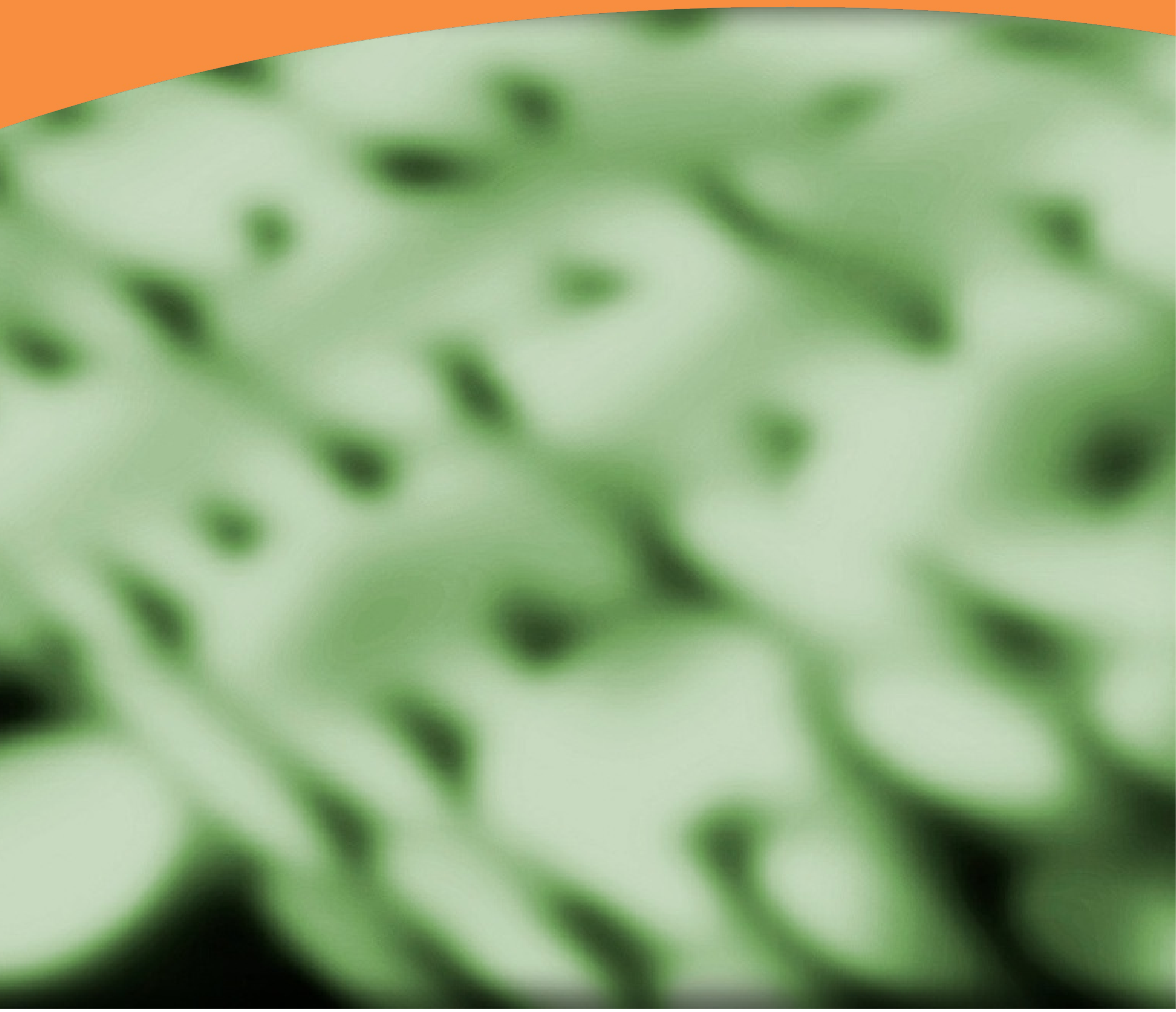
Random variables III

Probability Examples c-4

Leif Mejlbro



bookboon
The eBook company

Leif Mejlbro

Probability Examples c-2

Random variables I

Probability Examples c-2 – Random variables I

ISBN 978-87-7681-516-5

Contents

	Introduction	**3**
1	**Some theoretical results**	**4**
2	**Simple introducing examples**	**18**
3	**Frequencies and distribution functions in 1 dimension**	**20**
4	**Frequencies and distributions functions, 2 dimensions**	**33**
5	**Functions of random variables, in general**	**75**
6	**Inequalities between two random variables**	**81**
7	**Functions Y = f(X) o f random variables**	**93**
8	**Functions of two random variables, f(X; Y)**	**107**
9	**Means and moments of higher order**	**129**
10	**Mean and variance in special cases**	**143**
	Index	**157**

Introduction

This is the second book of examples from the *Theory of Probability.* This topic is not my favourite, however, thanks to my former colleague, Ole Jørsboe, I somehow managed to get an idea of what it is all about. The way I have treated the topic will often diverge from the more professional treatment. On the other hand, it will probably also be closer to the way of thinking which is more common among many readers, because I also had to start from scratch.

The topic itself, *Random Variables*, is so big that I have felt it necessary to divide it into three books, of which this is the first one. We shall here deal with the basic stuff, i.e. frequencies and distribution functions in 1 and 2 dimensions, functions of random variables and inequalities between random variables, as well as means and variances.

The prerequisites for the topics can e.g. be found in the *Ventus: Calculus 2* series, so I shall refer the reader to these books, concerning e.g. plane integrals.

Unfortunately errors cannot be avoided in a first edition of a work of this type. However, the author has tried to put them on a minimum, hoping that the reader will meet with sympathy the errors which do occur in the text.

Leif Mejlbro
25th October 2014

1 Some theoretical results

The abstract (and precise) definition of a *random variable* X is that X is a real function on Ω, where the triple $(\Omega, \mathcal{F}, P)$ is a probability field, such that

$$\{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F} \qquad \text{for every } x \in \mathbb{R}.$$

This definition leads to the concept of a *distribution function* for the random variable X, which is the function $F : \mathbb{R} \to \mathbb{R}$, which is defined by

$$F(x) = P\{X \leq x\} \qquad (= P\{\omega \in \Omega \mid X(\omega) \leq x\}),$$

where the latter expression is the mathematically precise definition which, however, for obvious reasons everywhere in the following will be replaced by the former expression.

A distribution function for a random variable X has the following properties:

$0 \leq F(x) \leq 1 \qquad$ for every $x \in \mathbb{R}$.

The function F is weakly increasing, i.e. $F(x) \leq F(y)$ for $x \leq y$.

$\lim_{x \to -\infty} F(x) = 0 \qquad$ and $\qquad \lim_{x \to +\infty} F(x) = 1$.

The function F is continuous from the right, i.e. $\lim_{h \to 0+} F(x+h) = F(x)$ for every $x \in \mathbb{R}$.

One may in some cases be interested in giving a crude description of the behaviour of the distribution function. We define a *median* of a random variable X with the distribution function $F(x)$ as a real number $a = (X) \in \mathbb{R}$, for which

$$P\{X \leq a\} \geq \frac{1}{2} \qquad \text{and} \qquad P\{X \geq a\} \geq \frac{1}{2}.$$

Expressed by means of the distribution function it follows that $a \in \mathbb{R}$ is a median, if

$$F(a) \geq \frac{1}{2} \qquad \text{and} \qquad F(a-) = \lim_{h \to 0-} F(x+h) \leq \frac{1}{2}.$$

In general we define a *p-quantile*, $p \in]0,1[$, of the random variable as a number $a_p \in \mathbb{R}$, for which

$$P\{X \leq a_p\} \geq p \qquad \text{and} \qquad P\{X \geq a_p\} \geq 1-p,$$

which can also be expressed by

$$F(a_p) \geq p \qquad \text{and} \qquad F(a_p-) \leq p.$$

If the random variable X only has a finite or a countable number of values, x_1, x_2, $\dots$, we call it *discrete*, and we say that X has a *discrete distribution*.

A very special case occurs when X only has one value. In this case we say that X is *causally distributed*, or that X is constant.

The random variable X is called *continuous*, if its distribution function $F(x)$ can be written as an integral of the form

$$F(x) = \int_{-\infty}^{x} f(u)\, du, \qquad x \in \mathbb{R},$$

where f is a nonnegative integrable *function*. In this case we also say that X has a continuous *distribution*, and the integrand $f : \mathbb{R} \to \mathbb{R}$ is called a *frequency* of the random variable X.

Let again $(\Omega, \mathcal{F}, P)$ be a given probability field. Let us consider *two* random variables X and Y, which are both defined on Ω. We may consider the pair (X, Y) as a *2-dimensional random variable*, which implies that we then shall make precise the extensions of the previous concepts for a single random variable.

We say that the *simultaneous distribution*, or just the *distribution*, of (X, Y) is known, if we know

$$P\{(X, Y) \in A\} \qquad \text{for every Borel set } A \subseteq \mathbb{R}^2.$$

When the simultaneous distribution of (X, Y) is known, we define the *marginal distributions* of X and Y by

$$P_X(B) = P\{X \in B\} := P\{(X, Y) \in B \times \mathbb{R}\}, \qquad \text{where } B \subseteq \mathbb{R} \text{ is a Borel set,}$$

$$P_Y(B) = P\{Y \in B\} := P\{(X, Y) \in \mathbb{R} \times B\}, \qquad \text{where } B \subseteq \mathbb{R} \text{ is a Borel set.}$$

Notice that we can always find the *marginal distributions* from the *simultaneous distribution*, while it is far from always possible to find the *simultaneous distribution* from the *marginal distributions*. We now introduce

The *simultaneous distribution function* of the 2-dimensional random variable (X,Y) is defined as the function $F : \mathbb{R}^2 \to \mathbb{R}$, given by

$$F(x,y) := P\{X \le x \wedge Y \le y\}.$$

We have

- If $(x,y) \in \mathbb{R}^2$, then $0 \le F(x,y) \le 1$.
- If $x \in \mathbb{R}$ is kept fixed, then $F(x,y)$ is a weakly increasing function in y, which is continuous from the right and which satisfies the condition $\lim_{y\to-\infty} F(x,y) = 0$.
- If $y \in \mathbb{R}$ is kept fixed, then $F(x,y)$ is a weakly increasing function in x, which is continuous from the right and which satisfies the condition $\lim_{x\to-\infty} F(x,y) = 0$.
- When both x and y tend towards infinity, then

  $$\lim_{x,\,y\to+\infty} F(x,y) = 1.$$

- If x_1, x_2, y_1, $y_2 \in \mathbb{R}$ satisfy $x_1 \le x_2$ and $y_1 \le y_2$, then

  $$F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_2) \ge 0.$$

Given the simultaneous distribution function $F(x,y)$ of (X,Y) we can find the distribution functions of X and Y by the formulæ

$$F_X(x) = F(x,+\infty) = \lim_{y\to+\infty} F(x,y), \qquad \text{for } x \in \mathbb{R},$$

$$F_y(x) = F(+\infty,y) = \lim_{x\to+\infty} F(x,y), \qquad \text{for } y \in \mathbb{R}.$$

The 2-dimensional random variable (X,Y) is called *discrete*, or that it has a *discrete distribution*, if both X and Y are discrete.

The 2-dimensional random variable (X,Y) is called *continuous*, or we say that it has a *continuous distribution*, if there exists a nonnegative integrable function (a *frequency*) $f : \mathbb{R}^2 \to \mathbb{R}$, such that the distribution function $F(x,y)$ can be written in the form

$$F(x,y) = \int_{-\infty}^{x} \left\{ \int_{-\infty}^{y} f(t,u)\,du \right\} dt, \qquad \text{for } (x,y) \in \mathbb{R}^2.$$

In this case we can find the function $f(x,y)$ at the differentiability points of $F(x,y)$ by the formula

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}.$$

It should now be obvious why one should know something about the theory of integration in more variables, cf. e.g. the *Ventus: Calculus 2* series.

We note that if $f(x,y)$ is a frequency of the continuous 2-dimensional random variable (X,Y), then X and Y are both continuous 1-dimensional random variables, and we get their (marginal) frequencies by

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy, \qquad \text{for } x \in \mathbb{R},$$

and

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx, \qquad \text{for } y \in \mathbb{R}.$$

It was mentioned above that one far from always can find the simultaneous distribution function from the marginal distribution function. It is, however, possible in the case when the two random variables X and Y are *independent*.

Let the two random variables X and Y be defined on the same probability field $(\Omega, \mathcal{F}, P)$. We say that X and Y are *independent*, if for all pairs of Borel sets A, $B \subseteq \mathbb{R}$,

$$P\{X \in A \wedge Y \in B\} = P\{X \in A\} \cdot P\{Y \in B\},$$

which can also be put in the simpler form

$$F(x,y) = F_X(x) \cdot F_Y(y) \qquad \text{for every } (x,y) \in \mathbb{R}^2.$$

If X and Y are *not* independent, then we of course say that they are *dependent*.

In two special cases we can obtain more information of independent random variables:

If the 2-dimensional random variable (X,Y) is *discrete*, then X and Y are independent, if

$$h_{ij} = f_i \cdot g_j \qquad \text{for every } i \text{ and } j.$$

Here, f_i denotes the probabilities of X, and g_j the probabilities of Y.

If the 2-dimensional random variable (X,Y) is *continuous*, then X and Y are independent, if their frequencies satisfy

$$f(x,y) = f_X(x) \cdot f_Y(y) \qquad \text{almost everywhere.}$$

The concept *"almost everywhere"* is rarely given a precise definition in books on applied mathematics. Roughly speaking it means that the relation above holds outside a set in $\mathbb{R}^2$ of area zero, a so-called *null set*. The common examples of null sets are either finite or countable sets. There exists, however, also non-countable null sets. Simple examples are graphs of any (piecewise) C^1-curve.

Concerning maps of random variables we have the following very important results,

Theorem 1.1 *Let X and Y be independent random variables. Let $\varphi : \mathbb{R} \to \mathbb{R}$ and $\psi : \mathbb{R} \to \mathbb{R}$ be given functions. Then $\varphi(X)$ and $\psi(Y)$ are again independent random variables.*

If X is a continuous random variable of the frequency I, then we have the following important theorem, where it should be pointed out that one always shall check *all* assumptions in order to be able to conclude that the result holds:

Theorem 1.2 *Given a continuous random variable X of frequency f.*

1) Let I be an open interval, such that $P\{X \in I\} = 1$.

2) Let $\tau : I \rightarrow J$ be a bijective *map of I onto an* open *interval J.*

3) Furthermore, assume that τ is differentiable *with a* continuous *derivative τ', which satisfies*

$$\tau'(x) \neq 0 \qquad \text{for alle } x \in I.$$

Under the assumptions above $Y := \tau(X)$ is also a continuous random variable, and its frequency $g(y)$ is given by

$$g(y) = \begin{cases} f\left(\tau^{-1}(y)\right) \cdot \left|\left(\tau^{-1}\right)'(y)\right|, & \text{for } y \in J, \\ \\ 0, & \text{otherwise.} \end{cases}$$

We note that if just one of the assumptions above is *not* fulfilled, then we shall instead find the distribution function $G(y)$ of $Y := \tau(X)$ by the general formula

$$G(y) = P\{\tau(X) \in]-\infty\,,\,y]\} = P\left\{X \in \tau^{\circ -1}(]-\infty\,,\,y])\right\},$$

where $\tau^{\circ -1} = \tau^{-1}$ denotes the inverse set map.

Note also that if the assumptions of the theorem are all satisfied, then τ is necessarily monotone.

At a first glance it may be strange that we at this early stage introduce 2-dimensional random variables. The reason is that by applying the simultaneous distribution for (X, Y) it is fairly easy to define the elementary operations of calculus between X and Y. Thus we have the following general result for a *continuous* 2-dimensional random variable.

Theorem 1.3 *Let (X, Y) be a* continuous *random variable of the frequency $h(x, y)$.*

The frequency of the sum $X + Y$ is $\qquad k_1(z) = \int_{-\infty}^{+\infty} h(x, z - x)\,dx.$

The frequency of the difference $X - Y$ is $\qquad k_2(z) = \int_{-\infty}^{+\infty} h(x, x - z)\,dx.$

The frequency of the product $X \cdot Y$ is $\qquad k_3(z) = \int_{-\infty}^{+\infty} h\left(x\,,\,\frac{z}{x}\right) \cdot \frac{1}{|x|}\,dx.$

The frequency of the quotient X/Y is $\qquad k_4(z) = \int_{-\infty}^{+\infty} h(zx\,,\,x) \cdot |x|\,dx.$

Notice that one must be very careful by computing the product and the quotient, because the corresponding integrals are improper.

If we furthermore assume that X and Y are *independent*, and $f(x)$ is a frequency of X, and $g(y)$ is a frequency of Y, then we get an even better result:

Theorem 1.4 *Let X and Y be* continuous *and* independent *random variables with the frequencies $f(x)$ and $g(y)$, resp..*

The frequency of the sum $X+Y$ is $k_1(z) = \int_{-\infty}^{+\infty} f(x)g(z-x)\,dx.$

The frequency of the difference $X-Y$ is $k_2(z) = \int_{-\infty}^{+\infty} f(x)g(x-z)\,dx.$

The frequency of the product $X \cdot Y$ is $k_3(z) = \int_{-\infty}^{+\infty} f(x)\, g\left(\frac{z}{x}\right) \cdot \frac{1}{|x|}\,dx.$

The frequency of the quotient X/Y is $k_4 = \int_{-\infty}^{+\infty} f(zx)g(x) \cdot |x|\,dx.$

Let X and Y be *independent* random variables with the distribution functions F_X and F_Y, resp.. We introduce two random variables by

$$U := \max\{X,Y\} \qquad \text{and} \qquad V := \min\{X,Y\},$$

the distribution functions of which are denoted by F_U and F_V, resp.. Then these are given by

$$F_U(u) = F_X(u) \cdot F_Y(u) \qquad \text{for } u \in \mathbb{R},$$

and

$$F_V(v) = 1 - (1 - F_X(v)) \cdot (1 - F_Y(v)) \qquad \text{for } v \in \mathbb{R}.$$

These formulæ are general, provided only that X and Y are independent.

If X and Y are *continuous and independent*, then the frequencies of U and V are given by

$$f_U(u) = F_X(u) \cdot f_Y(u) + f_X(u) \cdot F_Y(u), \qquad \text{for } u \in \mathbb{R},$$

and

$$f_V(v) = (1 - F_X(v)) \cdot f_Y(v) + f_X(v) \cdot (1 - F_y(v)), \qquad \text{for } v \in \mathbb{R},$$

where we note that we shall apply both the frequencies and the distribution functions of X and Y.

The results above can also be extended to bijective maps $\underline{\varphi} = (\varphi_1 , \varphi_2) : \mathbb{R}^2 \to \mathbb{R}^2$, or subsets of $\mathbb{R}^2$. We shall need the *Jacobian* of $\underline{\varphi}$, introduced in e.g. the *Ventus: Calculus 2* series.

It is important here to define the notation and the variables in the most convenient way. We start by assuming that D is an open domain in the $(x_1\, x_2)$ plane, and that $\tilde{D}$ is an open domain in the (y_1 , y_2) plane. Then let $\underline{\varphi} = (\varphi_1 , \varphi_2)$ be a bijective map *of* $\tilde{D}$ *onto* D with the inverse $\underline{\tau} = \underline{\varphi}^{-1}$, i.e. the opposite of what one probably would expect:

$$\underline{\varphi} = (\varphi_1 , \varphi_2) : \tilde{D} \to D, \qquad \text{with } (x_1 , x_2) = \underline{\varphi}(y_1 , y_2).$$

The corresponding *Jacobian* is defined by

$$J_{\underline{\varphi}} = \frac{\partial (x_1 , x_2)}{\partial (y_1 , y_2)} = \begin{vmatrix} \dfrac{\partial \varphi_1}{\partial y_1} & \dfrac{\partial \varphi_2}{\partial y_1} \\[2ex] \dfrac{\partial \varphi_1}{\partial y_1} & \dfrac{\partial \varphi_2}{\partial y_2} \end{vmatrix},$$

where the independent variables (y_1 , y_2) are in the "denominators". Then recall the *Theorem of transform of plane integrals*, cf. e.g. the *Ventus: Calculus 2* series: If $h : D \to \mathbb{R}$ is an integrable function, where $D \subseteq \mathbb{R}^2$ is given as above, then for every (measurable) subset $A \subseteq D$,

$$\int_A h(x_1 , x_2)\, dx_1 dx_2 = \int_{\underline{\varphi}^{-1}(A)} h(x_1 , x_2) \cdot \left| \frac{\partial (x_1 , x_2)}{\partial (y_1 , y_2)} \right| dy_1 dy_2.$$

Of course, this formula is not mathematically correct; but it shows intuitively what is going on: Roughly speaking we "delete the y-s". The correct mathematical formula is of course the well-known

$$\int_A h(x_1 , x_2)\, dx_1 dx_2 = \int_{\underline{\varphi}^{-1}(A)} (\varphi_1(y_1 , y_2) , \varphi_2(y_1 , y_2)) \cdot \left| J_{\underline{\varphi}}(y_1 , y_2) \right| dy_1 dy_2,$$

although experience shows that it in practice is more confusing then helping the reader.

Theorem 1.5 *Let (X_1, X_2) be a continuous 2-dimensional random variable with the frequency $h(x_1\,,\,x_2)$. Let $D \subseteq \mathbb{R}^2$ be an open domain, such that*

$$P\{(X_1\,,\,X_2) \in D\} = 1.$$

Let $\underline{\tau} : D \to \tilde{D}$ be a bijective map of D onto another open domain $\tilde{D}$, and let $\underline{\varphi} = (\varphi_1\,,\,\varphi_2) = \underline{\tau}^{-1}$, where we assume that φ_1 and φ_2 have continuous partial derivatives and that the corresponding Jacobian is different from 0 in all of $\tilde{D}$.
Then the 2-dimensional random variable

$$(Y_1\,,\,Y_2) = \underline{\tau}(X_1\,,\,X_2) = (\tau_1(X_1\,,\,X_2)\,,\,\tau_2(X_1\,,\,X_2))$$

has the frequency $k(y_1\,,\,y_2)$, given by

$$k(y_1\,,\,y_2) = \begin{cases} h(\varphi_1(y_1\,,\,y_2)\,,\,\varphi_2(y_1\,,\,y_2)) \cdot \left|\dfrac{\partial(x_1\,,\,x_2)}{\partial(y_1\,,\,y_2)}\right|, & \text{for } (y_1\,,\,y_2) \in \tilde{D}, \\[2ex] 0, & \text{otherwise} \end{cases}$$

We have previously introduced the concept *conditional probability*. We shall now introduce a similar concept, namely the *conditional distribution*.

If X and Y are discrete, we define the *conditional distribution of X for given $Y = y_j$* by

$$P\{X = x_i \mid Y = y_j\} = \frac{P\{X = x_i \wedge Y = y_j\}}{P\{Y = y_j\}} = \frac{h_{ij}}{g_j}.$$

It follows that for fixed j we have that $P\{X = x_i \mid Y = y_j\}$ indeed is a distribution. We note in particular that we have the *law of the total probability*

$$P\{X = x_i\} = \sum_j P\{X = x_i \mid Y = y_j\} \cdot P\{Y = y_j\}.$$

Analogously we define for two continuous random variables X and Y the *conditional distribution function of X for given $Y = y$* by

$$P\{X \le x \mid Y = y\} = \frac{\int_{-\infty}^{x} f(u,y)\,du}{f_Y(y)}, \qquad \text{forudsat, at } f_Y(y) > 0.$$

Note that the conditional distribution function is not defined at points in which $f_Y(y) = 0$.

The corresponding frequency is

$$f(x \mid y) = \frac{f(x,y)}{f_Y(y)}, \qquad \text{provided that } f_Y(y) = 0.$$

We shall use the convention that “0 times undefined = 0”. Then we get the *Law of total probability*,

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

We now introduce the *mean*, or *expectation* of a random variable, provided that it exists.

1) Let X be a *discrete* random variable with the possible values $\{x_i\}$ and the corresponding probabilities $p_i = P\{X = x_i\}$. The *mean*, or *expectation*, of X is defined by

$$E\{X\} := \sum_i x_i\, p_i,$$

provided that the series is absolutely convergent. If this is not the case, the mean does not exists.

2) Let X be a *continuous* random variable with the frequency $f(x)$. We define the *mean*, or *expectation* of X by

$$E\{X\} = \int_{-\infty}^{+\infty} x\, f(x)\, dx,$$

provided that the integral is absolutely convergent. If this is not the case, the mean does not exist.

If the random variable X only has nonnegative values, i.e. the image of X is contained in $[0, +\infty[$, and the mean exists, then the mean is given by

$$E\{X\} = \int_0^{+\infty} P\{X \geq x\}\, dx.$$

Concerning maps of random variables, means are transformed according to the theorem below, provided that the given expressions are absolutely convergent.

Theorem 1.6 *Let the random variable $Y = \varphi(X)$ be a function of X.*

1) If X is a discrete *random variable with the possible values $\{x_i\}$ of corresponding probabilities $p_i = P\{X = x_i\}$, then the* mean *of $Y = \varphi(X)$ is given by*

$$E\{\varphi(X)\} = \sum_i \varphi(x_i)\, p_i,$$

provided that the series is absolutely convergent.

2) If X is a continuous *random variable with the frequency $f(x)$, then the* mean *of $Y = \varphi(X)$ is given by*

$$E\{\varphi(X)\} = \int_{-\infty}^{+\infty} \varphi(x)\, g(x)\, dx,$$

provided that the integral is absolutely convergent.

Assume that X is a random variable of mean μ. We add the following concepts, where $k \in \mathbb{N}$:

The k-th moment,	$E\left\{X^k\right\}$.
The k-th absolute moment,	$E\left\{\lvert X\rvert^k\right\}$.
The k-th central moment,	$E\left\{(X-\mu)^k\right\}$.
The k-th absolute central moment,	$E\left\{\lvert X-\mu\rvert^k\right\}$.
The variance, i.e. the second central moment,	$V\{X\} = E\left\{(X-\mu)^2\right\}$,

provided that the defining series or integrals are absolutely convergent. In particular, the *variance* is very important. We mention

Theorem 1.7 *Let X be a random variable of mean $E\{X\} = \mu$ and variance $V\{X\}$. Then*

$$E\left\{(X-c)^2\right\} = V\{X\} + (\mu - c)^2 \qquad \textit{for every } c \in \mathbb{R},$$

$$V\{X\} = E\left\{X^2\right\} - (E\{X\})^2 \qquad \textit{for } c = 0,$$

$$E\{aX+b\} = a\,E\{X\} + b \qquad \textit{for every } a,\, b \in \mathbb{R},$$

$$V\{aX+b\} = a^2 V\{X\} \qquad \textit{for every } a,\, b \in \mathbb{R}.$$

It is not always an easy task to compute the distribution function of a random variable. We have the following result which gives an estimate of the probability that a random variable X differs more than some given $a > 0$ from the mean $E\{X\}$.

Theorem 1.8 (Čebyšev's inequality). *If the random variable X has the mean μ and the variance σ^2, then we have for every $a > 0$,*

$$P\{|X-\mu| \geq a\} \leq \frac{\sigma^2}{a^2}.$$

If we here put $a = k\sigma$, we get the equivalent statement

$$P\{\mu - k\sigma < X < \mu + k\sigma\} \geq 1 - \frac{1}{k^2}.$$

These concepts are then generalized to 2-dimensional random variables. Thus,

Theorem 1.9 *Let $Z = \varphi(X,Y)$ be a function of the 2-dimensional random variable (X,Y).*

1) If (X,Y) is discrete*, then the* mean *of $Z = \varphi(X,Y)$ is given by*

$$E\{\varphi(X,Y)\} = \sum_{i,j} \varphi\left(x_i\,,\,y_j\right) \cdot P\left\{X = x_i \,\wedge\, Y = y_j\right\},$$

provided that the series is absolutely convergent.

2) If (X,Y) is continuous*, then the* mean *of $Z = \varphi(X,Y)$ is given by*

$$E\{\varphi(X,Y)\} = \int_{\mathbb{R}^2} \varphi(x,y)\, f(x,y)\, dxdy,$$

provided that the integral is absolutely convergent.

It is easily proved that if (X,Y) is a 2-dimensional random variable, and $\varphi(x,y) = \varphi_1(x) + \varphi_2(y)$, then

$$E\left\{\varphi_1(X) + \varphi_2(Y)\right\} = E\left\{\varphi_1(X)\right\} + E\left\{\varphi_2(Y)\right\},$$

provided that $E\left\{\varphi_1(X)\right\}$ and $E\left\{\varphi_2(Y)\right\}$ exists. In particular,

$$E\{X+Y\} = E\{X\} + E\{Y\}.$$

If we furthermore assume that X and Y are independent and choose $\varphi(x,y) = \varphi_1(x) \cdot \varphi_2(y)$, then also

$$E\left\{\varphi_1(X) \cdot \varphi_2(Y)\right\} = E\left\{\varphi_1(X)\right\} \cdot E\left\{\varphi_2(Y)\right\},$$

provided that $E\left\{\varphi_1(X)\right\}$ and $E\left\{\varphi_2(Y)\right\}$ exists. In particular we get under the assumptions above that

$$E\{X \cdot Y\} = E\{X\} \cdot E\{Y\},$$

and

$$E\{(X - E\{X\}) \cdot (Y - E\{Y\})\} = 0.$$

These formulæ are easily generalized to n random variables. We have e.g.

$$E\left\{\sum_{i=1}^{n} X_i\right\} = \sum_{i=1}^{n} E\left\{X_i\right\},$$

provided that all means $E\left\{X_i\right\}$ exist.

If two random variables X and Y are not independent, we shall find a measure of how much they "depend" on each other. This measure is described by the *correlation*, which we now introduce.

Consider a 2-dimensional random variable (X,Y), where

$$E\{X\} = \mu_X, \qquad E\{Y\} = \mu_Y, \qquad V\{X\} = \sigma_X^2 > 0, \qquad V\{Y\} = \sigma_Y^2 > 0,$$

all exist. We define the *covariance* between X and Y, denoted by $\mathrm{Cov}(X,Y)$, as

$$\mathrm{Cov}(X,Y) := E\left\{(X-\mu_X)\cdot(Y-\mu_Y)\right\}.$$

We define the *correlation* between X and Y, denoted by $\varrho(X,Y)$, as

$$\varrho(X,Y) := \frac{\mathrm{Cov}(X,Y)}{\sigma_X\cdot\sigma_Y}.$$

Theorem 1.10 *Let X and Y be two random variables, where*

$$E\{X\}=\mu_X, \qquad E\{Y\}=\mu_Y, \qquad V\{X\}=\sigma_X^2>0, \qquad V\{Y\}=\sigma_Y^2>0,$$

all exist. Then

$$\mathrm{Cov}(X,Y)=0, \qquad \text{if } X \text{ and } Y \text{ are independent},$$

$$\mathrm{Cov}(X,Y)=E\{X\cdot Y\}-E\{X\}\cdot E\{Y\},$$

$$|\mathrm{Cov}(X,Y)|\le \sigma_X\cdot\sigma_y,$$

$$\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X),$$

$$V\{X+Y\}=V\{X\}+V\{Y\}+2\mathrm{Cov}(X,Y),$$

$$V\{X+Y\}=V\{X\}+V\{Y\}, \qquad \text{if } X \text{ and } Y \text{ are independent},$$

$$\varrho(X,Y)=0, \qquad \text{if } X \text{ and } Y \text{ are independent},$$

$$\varrho(X,X)=1, \qquad \varrho(X,-X)=-1, \qquad |\varrho(X,Y)|\le 1.$$

Let Z be another random variable, for which the mean and the variance both exist- Then

$$\mathrm{Cov}(aX+bY,Z)=a\,\mathrm{Cov}(X,Z)+b\,\mathrm{Cov}(Y,Z), \qquad \text{for every } a,\, b\in\mathbb{R},$$

and if $U=aX+b$ and $V=cY+d$, where $a>0$ and $c>0$, then

$$\varrho(U,V)=\varrho(aX+b,cY+d)=\varrho(X,Y).$$

Two independent random variables are always non-correlated, while two non-correlated random variables are not necessarily independent.

By the obvious generalization,

$$V\left\{\sum_{i=1}^n X_i\right\}=\sum_{i=1}^n V\left\{X_i\right\}+2\sum_{j=2}^n\sum_{i=1}^{j-1}\mathrm{Cov}\left(X_i,X_j\right).$$

If all X_1, X_2, $\dots$, X_n are independent of each other, this is of course reduced to

$$V\left\{\sum_{i=1}^n X_i\right\}=\sum_{i=1}^n V\left\{X_i\right\}.$$

Finally we mention the various types of convergence which are natural in connection with sequences of random variables. We consider a sequence X_n of random variables, defined on the same probability field $(\Omega,\mathcal{F},P)$.

1) We say that X_n *converges in probability* towards a random variable X on the probability field $(\Omega, \mathcal{F}, P)$, if

$$P\{|X_n - X| \geq \varepsilon\} \to 0 \qquad \text{for } n \to +\infty,$$

for every fixed $\varepsilon > 0$.

2) We say that X_n *converges in probability* towards a constant c, if every fixed $\varepsilon > 0$,

$$P\{|X_n - c| \geq \varepsilon\} \to 0 \qquad \text{for } n \to +\infty.$$

3) If each X_n has the distribution function F_n, and X has the distribution function F, we say that the sequence X_n of random variables *converges in distribution* towards X, if at every point of continuity x of $F(x)$,

$$\lim_{n \to +\infty} F_n(x) = F(x).$$

Finally, we mention the following theorems which are connected with these concepts of convergence. The first one resembles *Čebyšev's inequality.*

Theorem 1.11 (The weak law of large numbers). *Let X_n be a sequence of* independent *random variables, all defined on $(\Omega, \mathcal{F}, P)$, and assume that they all have the same mean and variance,*

$$E\{X_i\} = \mu \qquad \textit{and} \qquad V\{X_i\} = \sigma^2.$$

Then for every fixed $\varepsilon > 0$,

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^n X_i - \mu\right| \geq \varepsilon\right\} \to 0 \qquad \textit{for } n \to +\infty.$$

A slightly different version of the weak law of large numbers is the following

Theorem 1.12 *If X_n is a sequence of* independent identical distributed *random variables, defined on $(\Omega, \mathcal{F}, P)$ where $E\{X_i\} = \mu$, (notice that we do not assume the existence of the variance), then for every fixed $\varepsilon > 0$,*

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^n X_i - \mu\right| \geq \varepsilon\right\} \to 0 \qquad \textit{for } n \to +\infty.$$

We have concerning convergence in distribution,

Theorem 1.13 (Helly-Bray's lemma). *Assume that the sequence X_n of random variables converges in distribution towards the random variable X, and assume that there are real constants a and b, such that*

$$P\{a \leq X_n \leq b\} = 1 \qquad \textit{for every } n \in \mathbb{N}.$$

If φ is a continuous function on the interval $[a, b]$, then

$$\lim_{n \to +\infty} E\{\varphi(X_n)\} = E\{\varphi(X)\}.$$

In particular,

$$\lim_{n \to +\infty} E\{X_n\} \qquad \textit{and} \qquad \lim_{n \to +\infty} V\{X_n\} = V\{X\}.$$

Finally, the following theorem gives us the relationship between the two concepts of convergence:

Theorem 1.14 *1) If X_n converges in probability towards X, then X_n also converges in distribution towards X.*

2) If X_n converges in distribution towards a constant c, then X_n also converges in probability towards the constant c.

2 Simple introducing examples

Example 2.1 *A motorist shall pass 4 traffic lights. We assume that at each of the traffic lights there is the probability p that he must stop. There is furthermore such a long distance between the traffic signals that there is no synchronization between them. Let X be the random variable, which indicates the number of stops. Find the distribution of X. Sketch in the case $p = \frac{1}{2}$ the corresponding diagram. Let Y have the value k, if the first stop is at signal number k, $k = 1, 2, 3, 4$. Is Y a random variable?*

In this case the model is given by the binomial distribution $X \in B(4,p)$, thus

$$P\{X = k\} = \binom{4}{k} p^k (1-p)^{4-k}, \qquad k = 0, 1, 2, 3, 4.$$

We define here a "success" as a stop (what we otherwise would not consider as a success).

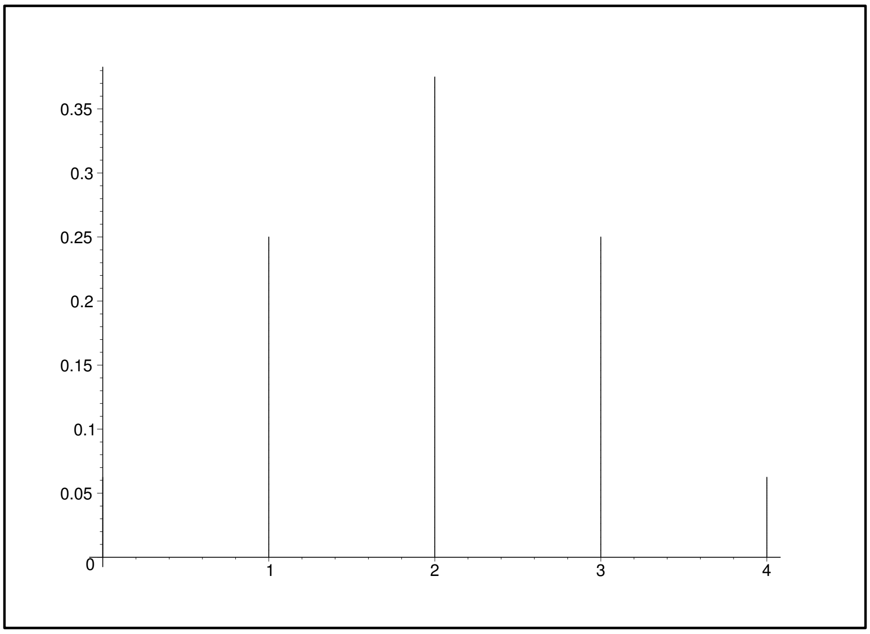


Figure 1: The diagram, for $p = \frac{1}{2}$.

We get in particular For $p = \frac{1}{2}$,

$$P\{X = k\} = \binom{4}{k} \left(\frac{1}{2}\right)^4, \qquad k = 0, 1, 2, 3, 4,$$

thus

$$p_0 = p_4 = \frac{1}{16}, \qquad p_1 = p_3 = \frac{4}{16} = \frac{1}{4}, \qquad p_2 = \frac{6}{16} = \frac{3}{8}.$$

If the first stop is at signal number k, the car has not stopped earlier, so

$$P\{Y = k\} = (1-p)^{k-1} p, \qquad k = 1, 2, 3, 4.$$

Then

$$\sum_{k=1}^{4} P\{Y=k\} = \sum_{k=1}^{4} p(1-p)^{k-1} = 1-(1-p)^4 < 1, \qquad \text{når } p<1.$$

Since the sum is *not* equal to 1, we conclude that Y is *not* a random variable.

The reason why Y is not a random variable, is that we have in the setup forgotten the possibility of "no stops at all" of the probability $(1-p)^4$. If we add this value to Y for this event (where one e.g. arbitrarily could let it correspond to the event $Y=\pi^2$), then Y becomes a random variable. A more reasonable definition would of course be $Y=5$. However, there is nothing wrong in choosing $Y=\pi^2$.

Example 2.2 *A random variable X can have the possible values 1, 2, ... of the probabilities*

$$P\{X=k\} = A\,\frac{q^k}{k}, \qquad k \in \mathbb{N} \qquad (\textit{where } q \in]0\,,\,1[).$$

Find the constant A.

We put $p=1-q$. Since

$$\sum_{k=1}^{\infty} P\{X=k\} = 1,$$

it follows from the series of logarithm that

$$1 = \sum_{k=1}^{\infty} P\{X=k\} = A\sum_{k=1}^{\infty} \frac{qk}{k} = -A\left\{-\sum_{k=1}^{\infty}\frac{q^k}{k}\right\} = -A\,\ln(1-q) = A\,\ln\frac{1}{p}.$$

From $\dfrac{1}{p} > 1$ follows that $\ln\dfrac{1}{p} > 0$, hence

$$A = \frac{1}{\ln\frac{1}{p}} = \frac{1}{|\ln p|},$$

and thus

$$P\{X=k\} = \frac{q^k}{k\,|\ln(1-q)|}, \qquad k \in \mathbb{N}.$$

3 Frequencies and distribution functions in 1 dimension

Example 3.1 *Check if the function*

$$f(x) = \begin{cases} \frac{1}{2} - kx, & x \in [0,6], \\ \\ 0 & \textit{otherwise}, \end{cases}$$

is a frequency for some k.

If $f(x)$ is a frequency, then the following two conditions must be fulfilled:

1) $f(x) \geq 0$ for every $x \in \mathbb{R}$,

2) $\int_{-\infty}^{\infty} f(x)\, dx = 1$.

By putting $x = 6$ into (1) we get $\dfrac{1}{2} - 6k \geq 0$, thus $k \leq \dfrac{1}{12}$.

A computation of (2) gives

$$1 = \int_0^6 \left\{ \frac{1}{2} - kx \right\} dx = \left[\frac{1}{2}\, x - \frac{1}{2}\, kx^2 \right]_0^6 = 3 - 18k,$$

hence

$$k = \frac{3-1}{18} = \frac{1}{9} > \frac{1}{12}.$$

The two requirements can never be satisfied simultaneously, so $f(x)$ is not a frequency for any $k \in \mathbb{R}$.

Example 3.2 *Find k, such that*

$$f(x) = \begin{cases} kx^2\left(1 - x^3\right), & x \in [0,1], \\ \\ 0, & \textit{otherwise}, \end{cases}$$

is a frequency of a random variable, and sketch the function.
Find the median of X.

Obviously, $f(x) \geq 0$. Then by an integration,

$$\int_0^1 x^2\left(1 - x^3\right)\, dx = \int_0^1 \left(x^2 - x^5\right)\, dx = \frac{1}{3} - \frac{1}{6} = \frac{1}{6}.$$

If we choose $k = 6$, then $f(x)$ becomes a frequency, thus

$$f(x) = \begin{cases} 6x^2\left(1 - x^3\right) = 6x^2 - 6x^5 & \text{for } x \in [0,1], \\ \\ 0 & \text{otherwise.} \end{cases}$$

Since $f(0) = f(1) = 0$, and $f'(x) = 12x - 30x^4 = 0$ for $x = \sqrt[3]{\frac{2}{5}} \approx 0,74$, we have a (global) maximum $\frac{18}{5} \cdot \left(\frac{2}{5}\right)^{\frac{2}{3}} \approx 1.95.$

The function f is increasing in $\left]0, \sqrt[3]{\frac{2}{5}}\right[$ and decreasing in $\left]\sqrt[3]{\frac{2}{5}}\right[$.

The distribution function $F(x)$ is in the interval $[0,1]$ given by

$$F(x) = \int_0^x f(t)\,dt = \int_0^x \left\{6t^2 - 6t^5\right\}\,dt = 2x^3 - x^6.$$

The *median* is given as the $x \in]0,1[$, for which $F(x) = \frac{1}{2}$, thus

$$2x^3 - x^6 = \frac{1}{2}, \quad \text{or} \quad x^6 - 2x^3 + 1 = \frac{1}{2}, \quad \text{i.e.} \quad \left(x^3 - 1\right)^2 = \frac{1}{2}.$$

We get $x^3 = 1 \pm \frac{\sqrt{2}}{2}$. However, since $x \in]0,1[$, we can only use the sign $-$, and the median is (here uniquely determined by)

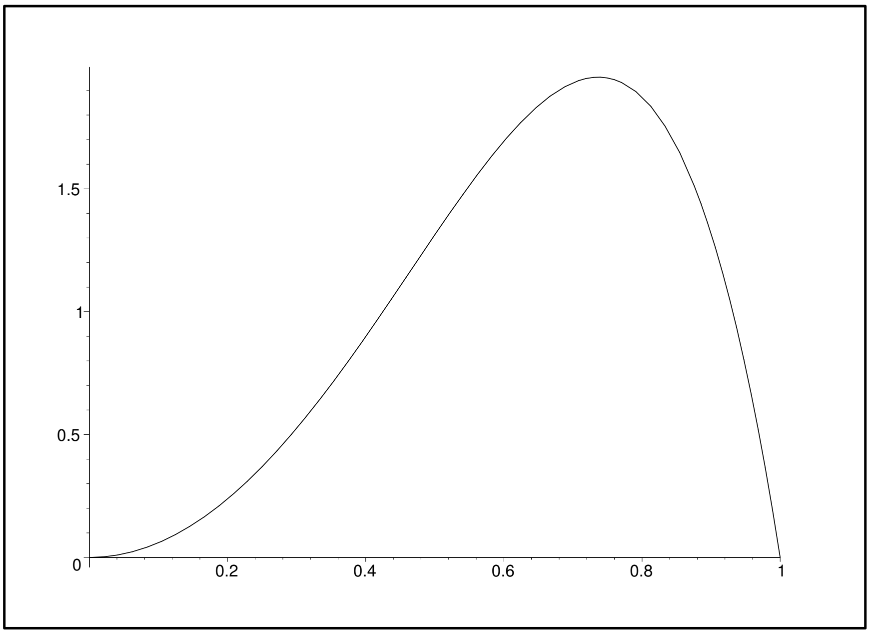
$$x_{\text{med}} = \sqrt[3]{1 - \frac{\sqrt{2}}{2}} \approx 0.66.$$

It is possible to apply MAPLE, e-g. by:

```
> f:=x->6*x^2*(1-x^3);
```

$$f := x \longrightarrow 6x^2\left(1-x^3\right)$$

```
> plot(f,0..1,color=black);
```




```
> F:=int(f(x),x=0..t);
```

$$F := -t^6 + 2t^3$$

```
> fsolve(2*t^3-t^5 = 1/2,t=0..1);
```

$$.6641045243$$

```
> solve(2*t^3-t^6 = 1/2,t);
```

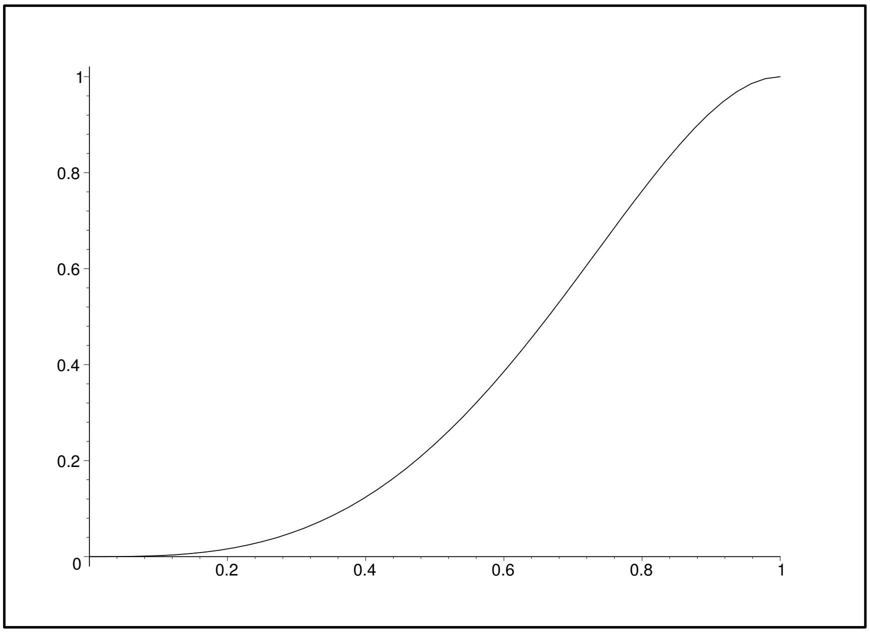
$$\frac{1}{2}(8+4\sqrt{2})^{1/3}, -\frac{1}{4}(8+4\sqrt{2})^{1/3},$$
$$-\frac{1}{4}(8+4\sqrt{2})^{1/3} - \frac{1}{4}I\sqrt{3}(8+4\sqrt{2})^{1/3}, \frac{1}{2}(8-4\sqrt{2})^{1/3},$$
$$-\frac{1}{4}(8-4\sqrt{2})^{1/3} + \frac{1}{4}I\sqrt{3}(8-4\sqrt{2})^{1/3}, -\frac{1}{4}(8-4\sqrt{2})^{1/3} - \frac{1}{4}I\sqrt{3}(8-4\sqrt{2})^{1/3}$$

```
> 1/2*(8-4*sqrt(2))^(1/3); # this is the median
```

$$\frac{1}{2}(8-4\sqrt{2})^{1/3}$$

```
> F1:=x->2*x^3-x^6;
```

$$F1 := x \longrightarrow 2x^3 - x^6$$




```
> plot(F1,0..1,color=black);
```

The former figure shows the graph of the frequency, and the latter figure shows the graph of the distribution function. Notice the difference between using `fsolve` or `solve`.

With the exception of the sketches of the graphs we see that it is easy to perform the same computations without using MAPLE. Furthermore, the MAPLE program is also less transparent then an explanation in plain words.

Example 3.3 *A random variable X has the frequency*

$$f(x)=\begin{cases} C(x-a), & a\leq x\leq \dfrac{a+b}{2},\\[2mm] C(b-x), & \dfrac{a+b}{2}\leq x\leq b,\\[2mm] 0, & \textit{otherwise}.\end{cases}$$

Find the constant C and the distribution function.
Compute

$$P\left\{X\leq \frac{a+b}{2}\right\} \qquad \textit{and} \qquad P\left\{\frac{2a+b}{3}\leq X\leq \frac{a+2b}{3}\right\}.$$

This distribution is called the triangular distribution over $]a,b[$.

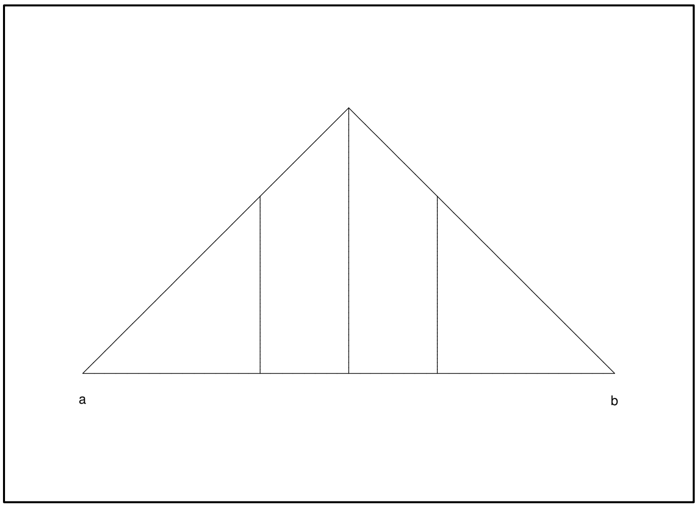


Figure 2: The graph of the frequency f.

1) By considering the graph we immediately get

$$1=\int_{-\infty}^{\infty} f(x)\,dx=\frac{1}{2}\cdot C\,\frac{b-a}{2}\cdot(b-a)=C\left\{\frac{b-a}{2}\right\}^2,$$

because the integral can be interpreted as the area of a triangle. Then

$$C=\left(\frac{2}{b-a}\right)^2.$$

2) When $a\leq x\leq \dfrac{a+b}{2}$ the distribution function is given by

$$F(x)=\int_a^x C(t-a)\,dt=C\left[\frac{(t-a)^1}{2}\right]_a^x=\frac{1}{2}\left(\frac{2}{b-a}\right)^2(x-a)^2=2\left(\frac{x-a}{b-a}\right)^2.$$

If instead $\dfrac{a+b}{2} \le x \le b$, then

$$\begin{aligned} F(x) &= \frac{1}{2} + \int_{\frac{a+b}{2}}^{x} C(b-t)\,dt = \frac{1}{2} + \left(\frac{2}{b-a}\right)^2 \left[-\frac{(b-t)^2}{2}\right]_{\frac{a+b}{2}}^{x} \\ &= \frac{1}{2} + \frac{1}{2}\left(\frac{2}{b-a}\right)^2 \left\{\left(\frac{b-a}{2}\right)^2 - (b-x)^2\right\} = \frac{1}{2} + \frac{1}{2} - 2\left(\frac{b-x}{b-a}\right)^2 \\ &= 1 - 2\left(\frac{b-x}{b-a}\right)^2 . \end{aligned}$$

Summing up we get

$$F(x) = \begin{cases} 0, & \text{for } x \le a, \\ 2\left(\dfrac{x-a}{b-a}\right)^2, & \text{for } a < x \le \dfrac{a+b}{2}, \\ 1 - 2\left(\dfrac{b-x}{b-a}\right)^2, & \text{for } \dfrac{a+b}{2} < x \le b, \\ 1, & \text{for } x > b. \end{cases}$$

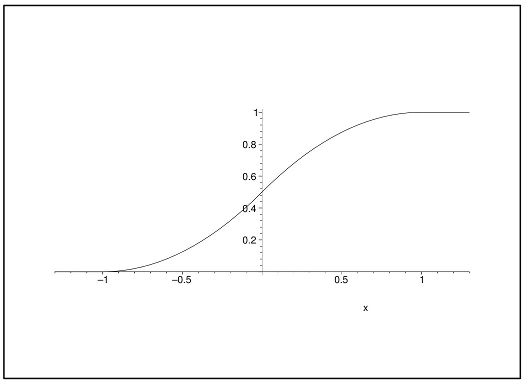


Figure 3: The distribution function for $a = -1$ and $b = 1$.

3) By considering the graph (or by insertion of $x\frac{a+b}{2}$) we get

$$P\left\{X \le \frac{a+b}{2}\right\} = F\left(\frac{a+b}{2}\right) = \frac{1}{2}.$$

Another consideration of the graph gives

$$P\left\{\frac{2a+b}{3} \le X \le \frac{a+2b}{3}\right\} = 1 - 2\,F\left(\frac{2a+b}{3}\right) = 1 - 2\cdot 2\cdot \left\{\frac{\frac{b-a}{3}}{b-a}\right\}^2 = 1 - \frac{4}{9} = \frac{5}{9}.$$

Example 3.4 *Prove for some choice of the constant k that the function*

$$f(x) = k \cdot e^{-|x-2|}, \qquad x \in \mathbb{R},$$

is the frequency of a random variable X.
Find the distribution function of X, and compute $P\{-1 \le X \le 3\}$ and $P\{X \ge 0\}$.
Find the median of X.

Obviously, $f(x)$ is continuous, and $f(x) > 0$, when $k > 0$. The remaining condition of a frequency is that $\int_{-\infty}^{\infty} f(x)\,dx = 1$. Then by a computation,

$$\int_{-\infty}^{\infty} f(x)\,dx = k\left\{\int_{-\infty}^{2} e^{x-2}\,dx + \int_{2}^{\infty} e^{-(x-2)}\,dx\right\} = 2k,$$

which is equal to 1 for $k = \frac{1}{2}$.

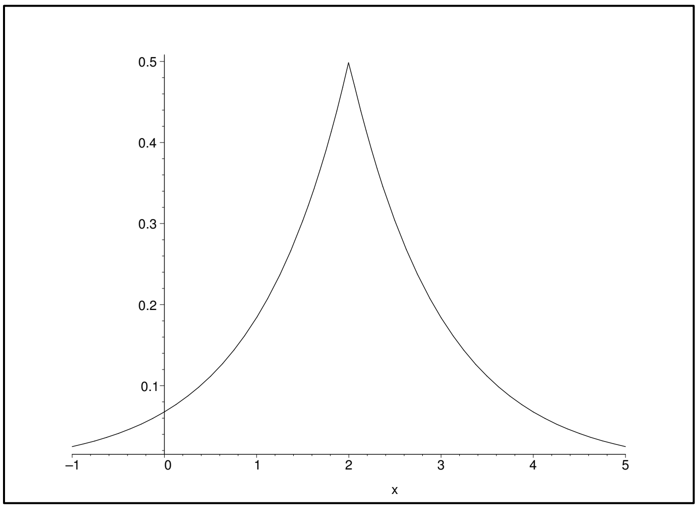


Figure 4: The graph of the frequency f. (NB Different scales on the axes).

The random variable X has the *frequency*

$$f(x) = \frac{1}{2}\,e^{-|x-2|}, \qquad x \in \mathbb{R}.$$

Now

$$\int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} \frac{1}{2}\,e^{t-2}\,dt = \frac{1}{2}\,e^{x-2} \qquad \text{for } x \le 2,$$

and

$$\begin{aligned}\int_{-\infty}^{x} f(t)\,dt &= \int_{-\infty}^{2} \frac{1}{2}\,e^{t-2}\,dt + \int_{2}^{x} e^{-(t-2)}\,dt = \frac{1}{2} + \left\{\frac{1}{2} - \frac{1}{2}\,e^{-(x-2)}\right\} \\ &= 1 - \frac{1}{2}\,e^{-(x-2)} \qquad \text{for } x > 2,\end{aligned}$$

so the *distribution function* becomes

$$F(x) = \begin{cases} \dfrac{1}{2}\,e^{x-2}, & \text{for } x \le 2, \\ 1 - \dfrac{1}{2}\,e^{-(x-2)}, & \text{for } x > 2. \end{cases}$$

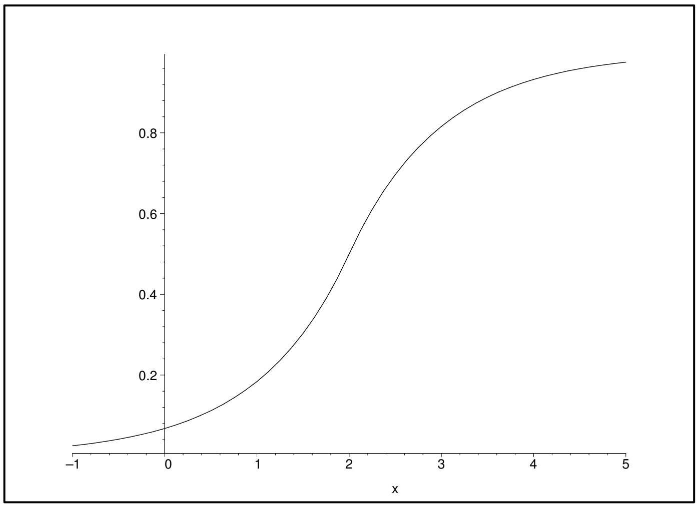


Figure 5: The graph of the distribution function F.

Finally,

$$P\{-1 \le X \le 3\} = F(3) - F(-1) = 1 - \frac{1}{2}\,e^{-1} - \frac{1}{2}\,e^{-3} \approx 0.79,$$

and

$$P\{X \ge 0\} = 1 - F(0) = 1 - \frac{1}{2}\,e^{-2} \approx 0.93.$$

Since $F(2) = \dfrac{1}{2}$, it is obvious that 2 is the median of X.

Example 3.5 *Prove that the function*

$$f(x) = \begin{cases} k\,x\,\exp\left(-\dfrac{1}{2}\,x^2\right), & x \ge 0, \\[2ex] \qquad 0, & x < 0, \end{cases}$$

for some choice of the constant k, can be considered as the frequency of a random variable X.
Find the distribution function F of X.
Sketch the graph of the function f and of the function F.
Find the median of X.

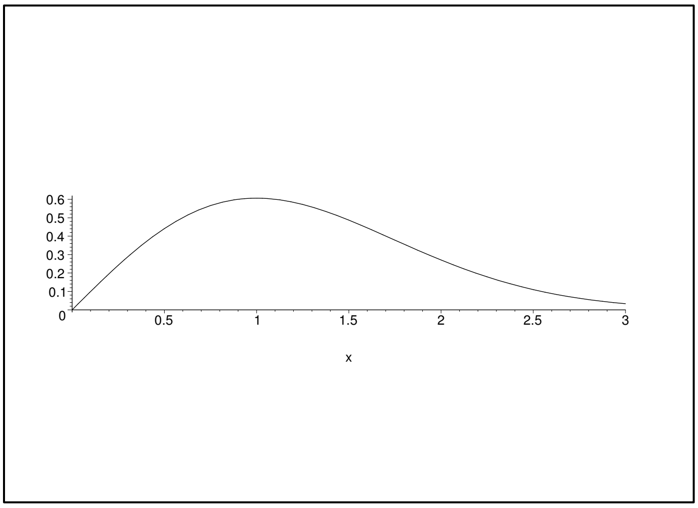


Figure 6: The graph of the frequency f.

1) If $k > 0$, then $f(x) \ge 0$. The requirement for $f(x)$ being a frequency is then reduced to

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = k\int_0^{\infty} x \cdot \exp\left(-\frac{1}{2}\,x^2\right)\,dx = k\int_0^{\infty} e^{-u}\,du = k,$$

where we have used the substitution $u = \frac{1}{2}\,x^2$ with $du = x\,dx$. Consequently, $k = 1$, and

$$f(x) = \begin{cases} x\,\exp\left(-\frac{1}{2}\,x^2\right), & x \geq 0, \\[2ex] 0, & x < 0. \end{cases}$$

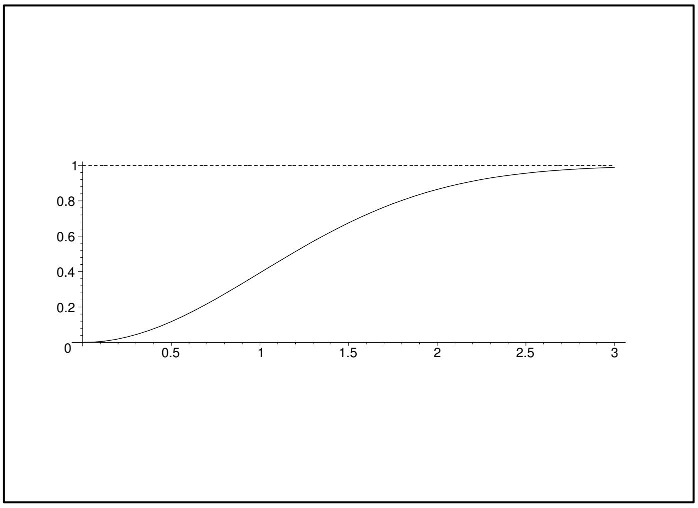


Figure 7: The graph of the distribution functions F.

2) If $x > 0$, we use the substitution $u = \frac{1}{2}\,t^2$ to obtain

$$F(x) = \int_0^x f(t)\,dt = \int_0^x t\,\exp\left(-\frac{1}{2}\,t^2\right)\,dt = \int_0^{\frac{1}{2}x^2} e^{-u}\,du = 1 - \exp\left(-\frac{1}{2}\,x^2\right).$$

Hence, the distribution function of X is

$$F(x) = \begin{cases} 1 - \exp\left(-\frac{1}{2}\,x^2\right), & \text{for } x > 0, \\[2ex] 0, & \text{for } x \leq 0. \end{cases}$$

3) Consider the previous figures. We see that $f(x)$ has a maximum for $x = \frac{1}{\sqrt{e}}$, and a turning point for $F(x)$ at the same point.

4) The median is found from the equation

$$F(x) = 1 - \exp\left(-\frac{1}{2}\,x^2\right) = \frac{1}{2}, \quad \text{i.e.} \quad \exp\left(\frac{1}{2}\,x^2\right) = 2,$$

thus $\frac{1}{2}\,x^2 = \ln 2$, and hence

$$(X) = \sqrt{2\,\ln 2} \approx 1.1774.$$

Example 3.6 *Prove that the function*

$$f(x)=\begin{cases}\dfrac{b}{\theta}\left(\dfrac{x}{\theta}\right)^{b-1}\exp\left(-\left\{\dfrac{x}{\theta}\right\}^{b}\right), & x>0,\\[2mm] 0, & x\le 0,\end{cases}$$

where b and θ are positive constants, is the frequency of a random variable X, and find the distribution function of it.
Prove that $P\{X\le\theta\}$ does not depend on b.
This distribution is called a Weibull distribution.

Clearly, $f(x)\ge 0$.

We get for $x>0$ by the substitution,

$$s=\frac{t}{\theta}\quad\text{followed by}\quad u=s^b=\left(\frac{t}{\theta}\right)^b,$$

that

$$\begin{aligned}F(x) &= \int_0^x \frac{b}{\theta}\left(\frac{t}{\theta}\right)^{b-1}\exp\left(-\left\{\frac{t}{\theta}\right\}^b\right)dt=\int_0^{\frac{x}{\theta}} b\,s^{b-1}\exp\left(-s^b\right)\,ds\\ &= \int_0^{\left(\frac{x}{\theta}\right)^b}\exp(-u)\,du = 1-\exp\left(-\left\{\frac{x}{\theta}\right\}^b\right)\to 1 \qquad \text{for } x\to\infty.\end{aligned}$$

We conclude that

$$F(x)=\begin{cases}1-\exp\left(-\left\{\dfrac{x}{\theta}\right\}^{b}\right), & x>0,\\[2mm] 0, & x\le 0,\end{cases}$$

is the distribution function of a random variable X with $f(x)$ as its frequency.

It follows by insertion that

$$P\{X\le\theta\}=F(\theta)=1-\exp\left(-\left\{\frac{\theta}{\theta}\right\}^b\right)=1-\frac{1}{e}$$

is independent of b.

Example 3.7 *A patient arrives to a doctor's waiting room. The probability is p, where $p \in]0,1[$, that he will be treated immediately; but if he does not, the probability that he must wait longer than the time x is equal to e^{-ax}, where a is some positive constant. Find the distribution function of the random variable X, which indicates the waiting time.*

1) If the patient is treated immediately, then the waiting time is $X = 0$, thus

$$P\{X = 0\} = p.$$

2) The probability that the patient must wait more than x, $x \geq 0$, is

$$P\{X > x\} = P\{\text{pt. must wait}\} \cdot P\{\text{waiting time } > x \mid \text{pt. must wait}\} = (1-p)e^{-ax},$$

hence

$$P\{X \leq x\} = 1 - P\{X > x\} = 1 - (1-p)e^{-ax}.$$

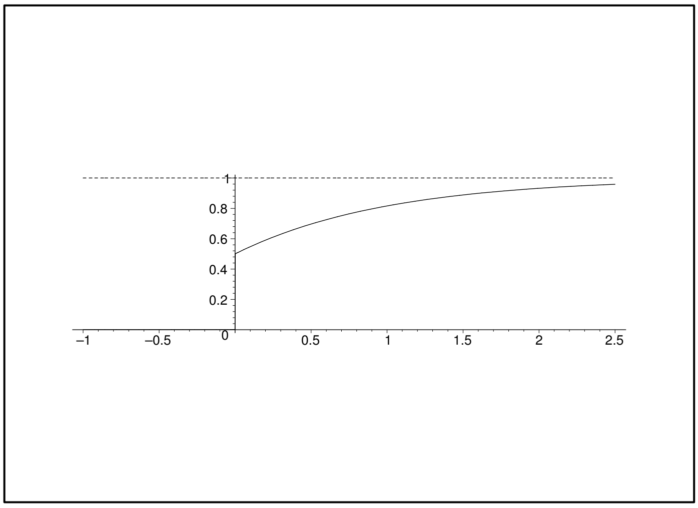


Figure 8: The graph of the distribution function $F(x)$ when $a = 1$ and $p = \frac{1}{2}$.

3) The distribution function $F(x) = P\{X \le x\}$ is here

$$F(x) = \begin{cases} 0, & x < 0, \\ 1-(1-p)e^{-x}, & x \ge 0. \end{cases}$$

The distribution of X is of *mixed type*, i.e. it is neither discrete nor continuous.

4 Frequencies and distributions functions, 2 dimensions

Example 4.1 *Let X and Y be independent random variables with the frequencies*

$$f(x) = x\,e^{-x}, \quad x > 0, \qquad g(y) = e^{-y}, \quad y > 0,$$

(both frequencies are otherwise 0).
Find the frequency of $X + Y$.
Find the mean $E\{X\}$, $E\{Y\}$ og $E\{X + Y\}$.

The frequency of $X + Y$ is given by the convolution integral

$$k(z) = \int_{-\infty}^{\infty} f(x)\,g(z-x)\,dx.$$

This expression is only > 0, when $z > 0$. We have furthermore the constraints $z - x > 0$ and $x > 0$, so the convolution integral is reduced to

$$k(z) = \int_0^z f(x)\,g(z-x)\,dx = \int_0^z x\,e^{-x}\,e^{-(z-x)}\,dx = e^{-z}\int_0^z x\,dx = \frac{1}{2}\,z^2\,e^{-z},$$

and $k(z) = 0$ for $z \le 0$.
The means are

$$\begin{aligned}
E\{X\} \quad &= \quad \textstyle\int_0^\infty x\,f(x)\,dx = \int_0^\infty x^2\,e^{-x} = 2,\\
E\{Y\} \quad &= \quad \textstyle\int_0^\infty y\,g(y)\,dy = \int_0^\infty y\,e^{-y}\,dy = 1,\\
E\{X+Y\} \quad &= \quad \textstyle\int_0^\infty z\,k(z)\,dz = \dfrac{1}{2}\int_0^\infty z^3\,e^{-z}\,dz = 3.
\end{aligned}$$

Remark 4.1 Here we are given that

$$\int_0^\infty x^n e^{-x}\,dx = n! \qquad \text{for } n \in \mathbb{N}_0.$$

This formula is easily proved by induction. When $n = 0$, it is trivial. In general we get by a partial integration and the assumption above of the induction,

$$\int_0^\infty x^{n+1}e^{-x}\,dx = \left[-x^{n2}e^{-x}\right]_0^\infty + (n+1)\int_0^\infty x^n e^{-x}\,dx = (n+1)n! = (n+1)!,$$

and the claim follows. $\Diamond$

Example 4.2 *Check if the function*

$$F(x,y)=\begin{cases} 1-e^{-(x+y)}, & x \ge 0,\, y \ge 0, \\ \\ 0, & \textit{otherwise}, \end{cases}$$

is a distribution function of a 2-dimensional random variable.

Since $F \in C^\infty(\mathbb{R}_+ \times \mathbb{R}_+)$, we have

$$\frac{\partial^2 F}{\partial x \partial y} = -e^{-(x+y)} < 0 \qquad \text{for } (x,y) \in \mathbb{R}_+ \times \mathbb{R}_+,$$

so F cannot be a distribution function. In fact, if so then

$$f(x,y) = \frac{\partial^2 F}{\partial x \partial y} \quad (< 0)$$

should be a frequency, which is not possible, because frequencies are never negative.

ALTERNATIVELY we prove that one of the necessary conditions is not fulfilled. Choose

$$x_1 = y_1 = 1 \qquad \text{og} \qquad x_2 = y_2 = 1+\alpha, \quad \alpha > 0.$$

Then

$$F(1+\alpha,1+\alpha)-F(1,1+\alpha)-F(1+\alpha,1)+F(1,1) = e^{-(1+\alpha)} - e^{-(2+2\alpha)} + e^{-(2+\alpha)} - e^{-2}$$

$$= e^{-2}\left\{2e^{-\alpha} - e^{-2\alpha} - 1\right\} = -e^{-2}\left(1-e^{-\alpha}\right)^2 < 0,$$

and not ≥ 0, as it should be.

Example 4.3 *Prove that the function*

$$f(x,y)=\begin{cases} x\,e^{-x(y+1)}, & x > 0,\, y > 0, \\ \\ 0, & \textit{otherwise}, \end{cases}$$

is a frequency of a 2-dimensional random variable (X,Y).
Find the frequencies and the distribution functions of the random variables X and Y, and find the medians of these two distributions.
Check if the random variables X and Y are independent.

Clearly, $f(x,y) \ge 0$ for every (x,y), and f is continuous, with the exception of the positive part of the x-axis.

1) If $x > 0$ is kept fixed, it follows by a vertical integration,

$$f_X(x) = e^{-x}\int_{y=0}^{\infty} x\,e^{-xy}\,dy = e^{-x},$$

and $f_X(x) = 0$ for $x \le 0$.

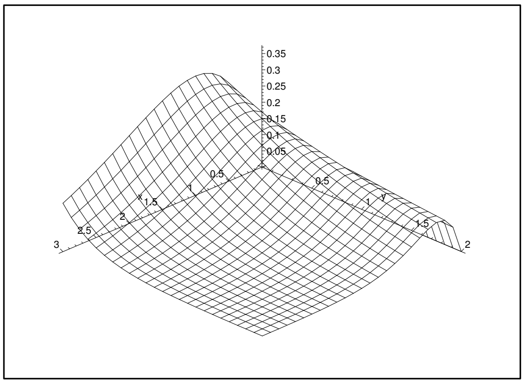


Figure 9: The graph of the frequency $f(x,y)$.

2) If $y>0$ is kept fixed, we get by a horizontal integration, where we use the substitution $z=x(y+1)$, etc.,

$$f_Y(y)=\int_{x=0}^{\infty} x\,e^{-x(y+1)}\,dy=\frac{1}{(y+1)^2},$$

and $f_Y(y)=0$ for $y\le 0$.

3) It follows from

$$\int_0^{\infty} e^{-x}\,dx=1,\qquad \left[\text{possibly } \int_0^{\infty}\frac{1}{(y+1)^2}\,dy=1\right],$$

that $f(x,y)$ is a frequency of a 2-dimensional random variable (X,Y), and that X and Y have the marginal frequencies $f_X(x)$ and $f_Y(y)$, given in (1) and (2), resp..

4) The marginal distribution functions are

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \begin{cases} 1 - e^{-x}, & \text{for } x > 0, \\ 0, & \text{for } x \le 0, \end{cases}$$

and

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)\,dt = \begin{cases} 1 - \dfrac{1}{y+1} = \dfrac{y}{y+1}, & \text{for } y > 0, \\ 0, & \text{for } y \le 0. \end{cases}$$

5) Medians:

a) $F_X(x) = \dfrac{1}{2} = 1 - e^{-x}$ for $e^{-x} = \dfrac{1}{2}$, hence $x = (X) = \ln 2$.

b) $F_y(y) = \dfrac{1}{2} = 1 - \dfrac{1}{y+1}$ for $\dfrac{1}{y+1} = \dfrac{1}{2}$, thus $y = (Y) = 1$.

6) Since

$$f_X(x) \cdot f_Y(y) = \frac{e^{-x}}{(y+1)^2} \neq x\,e^{-x(y+1)} = f(x,y) \qquad \text{for } x,\, y > 0,$$

X and Y are not independent.

Example 4.4 *A 2-dimensional random variable (X,Y) has the frequency*

$$f(x,y) = \begin{cases} c\,x\,y, & 0 < x < y < 1, \\ 0, & \textit{otherwise.} \end{cases}$$

Find the constant c. Find the frequencies and the distribution function of the random variables X and Y. Check if the random variables X and Y are independent. Finally, find the distribution function of the 2-dimensional random variable (X,Y).

1) If $c > 0$, then $f(x,y) \ge 0$. It follows from

$$1 = c \int_0^1 \left\{ \int_0^y xy\,dx \right\} dy = c \int_0^1 \frac{1}{2}\,y^3\,dy = \frac{c}{8},$$

that $c = 8$, hence the frequency is given by

$$f(x,y) = \begin{cases} 8xy, & 0 < x < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

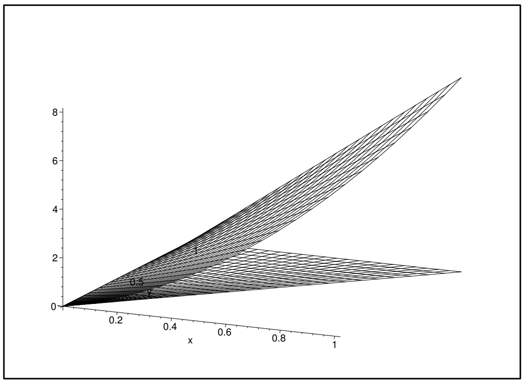


Figure 10: The graph of the frequency $f(x,y)$ over $0 < x < y < 1$.

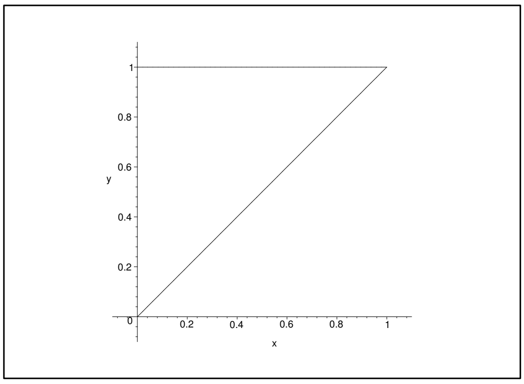


Figure 11: The domain where $f(x,y) > 0$.

2) Clearly, $f_X(x) = 0$ for $x \notin]0,1[$. When $x \in]0,1[$ it follows by a vertical integration that

$$f_X(x) = \int_x^1 8xy\,dy = 8x\left[\frac{y^2}{2}\right]_x^1 = 4x - 4x^3,$$

hence the marginal frequency is

$$f_X(x) = \begin{cases} 4x - 4x^3, & x \in]0,1[, \\ 0, & \text{otherwise.} \end{cases}$$

When $x \in]0,1[$, we get

$$F_X(x) = \int_0^x f_X(t)\,dt = \int_0^x \left\{4t - 4t^3\right\}\,dt = 2x^2 - x^4,$$

thus the marginal distribution function is

$$F_X(x) = \begin{cases} 0, & x \le 0, \\ 2x^2 - x^4, & 0 < x < 1, \\ 1, & x \ge 1. \end{cases}$$

Clearly, $f_Y(y) = 0$ for $y \notin]0,1[$. If $y \in]0,1[$, we get by a horizontal integration that

$$f_Y(y) = \int_0^y 8xy\,dx = 8y\left[\frac{x^2}{2}\right]_0^y = 4y^3,$$

and the marginal frequency is

$$f_Y(y) = \begin{cases} 4y^3, & y \in]0,1[, \\ 0, & \text{otherwise.} \end{cases}$$

Then for $y \in]0,1[$,

$$F_Y(y) = \int_0^y f_Y(t)\,dt = \int_0^y 4t^3\,dt = y^4,$$

hence the marginal distribution function is

$$F_Y(y) = \begin{cases} 0, & y \le 0, \\ y^4, & y \in]0,1[, \\ 1, & y \ge 1. \end{cases}$$

3) Since $f_X(x) \cdot f_Y(y) \neq f(x,y)$, we see that X and Y are not stochastically independent.

Remark 4.2 If in general the domain, in which the frequency $f(x,y) > 0$ is strictly positive, is *not* a rectangle (possibly with infinite sides, so e.g. $\mathbb{R} \times \mathbb{R}$ is in this sense considered as a degenerated rectangle), then the random variables X and Y are *never* stochastic independent. $\Diamond$

4) If x, $y \in]0,1[$, then the distribution function is

$$F(x,y) = \int_0^y \left\{\int_0^x f(t,u)\,dt\right\} du,$$

so $0 \le t \le x \le 1$, and $0 \le u \le y \le 1$. Furthermore, $f(t,u) = 8ty \neq 0$ for $0 < t < u < 1$, and 0 otherwise, so $0 < t < \min\{x,u\}$, and thus

$$F(x,y) = \int_0^y \left\{\int_0^{\min\{x,u\}} 8tu\,dt\right\} du = \int_0^y 4u \cdot (\min\{x,u\})^2\,du = \int_0^y 4u \cdot \min\left\{x^2,u^2\right\}\,du.$$

If $x \ge 0$, we get $0 < u < y \le x \le 1$, hence $\min\left\{x^2,u^2\right\} = u^2$, and thus

$$F(x,y) = \int_0^y 4u \cdot u^2\,du = y^4.$$

If $0 \le x \le y$, then we get instead

$$\begin{aligned} F(x,y) &= \int_0^y 4u \cdot \min\left\{x^2, u^2\right\}\,du = \int_0^x 4u^3\,du + \int_x^y 4u\,x^2\,du \\ &= x^4 + 2x^2\left(y^2 - x^2\right) = 2x^2y^2 - x^4. \end{aligned}$$

Summing up the distribution function of the 2-dimensional random variable (X, Y) is

$$F(x,y) = \begin{cases} 2x^2y^2 - x^4, & 0 \le x \le y \le 1, \\ 2x^2 - x^4, & 0 \le x \le 1 \le y, \\ y^4, & 0 \le y \le x, \\ 1, & 1 \le \min\{x,y\}, \\ 0, & \text{otherwise.} \end{cases}$$

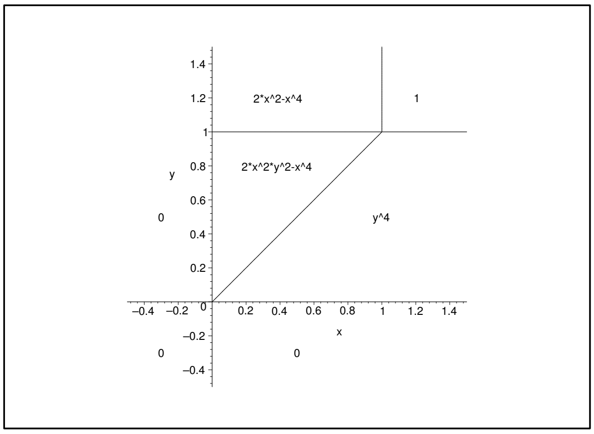


Figure 12: The distribution function $F(x,y)$ of Example 4.4.

Example 4.5 *A 2-dimensional random variable* (X,Y) *has the frequency*

$$f(x,y)=\begin{cases} ct^2, & 0<x<y<1,\\ \\ 0, & \textit{otherwise}.\end{cases}$$

1) Find the constant c.

2) Find the frequencies and the distribution functions of the random variables X *and* Y.

3) Find the simultaneous distribution function of the 2-dimensional random variable (X,Y).

1) If $c>0$, then obviously $f(x,y)\geq 0$. It follows from

$$1=c\int_0^1\left\{\int_0^y y^2\,\mathrm{dx}\right\}dy=c\int_0^1 y^3\,dy=\frac{c}{4},$$

that if $c=4$, then

$$f(x,y)=\begin{cases} 4y^2, & 0<x<y<1,\\ 0, & \text{otherwise},\end{cases}$$

is the frequency of the 2-dimensional random variable (X,Y).

2) By a *vertical integration*, $x\in\,]0,1[$ fixed, we obtain the marginal frequency of X,

$$f_X(x)=\int_x^1 4y^2\,dy=\frac{4}{3}\left(1-x^3\right),$$

thus

$$f_X(x)=\begin{cases} \dfrac{4}{3}\left(1-x^3\right), & \text{for } x\in\,]0,1[,\\ \\ 0, & \text{otherwise}.\end{cases}$$

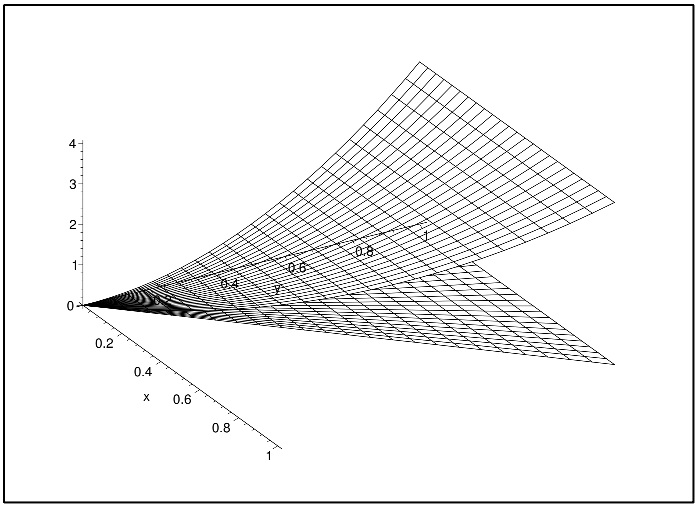


Figure 13: The graph of $f(x, y)$.

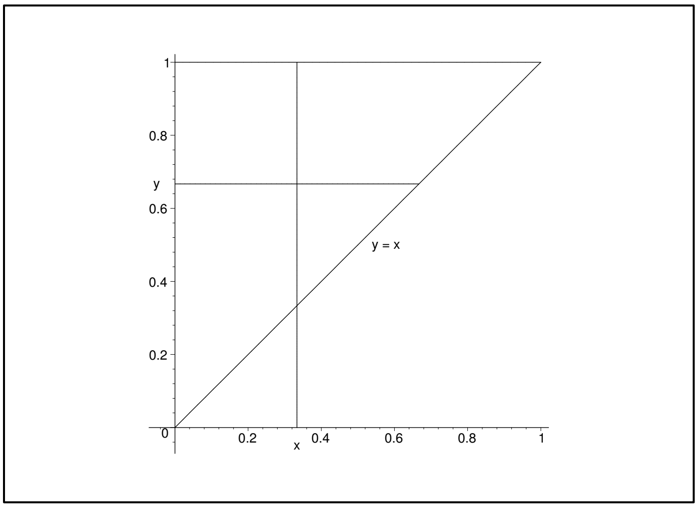


Figure 14: The domain $0 < x < y < 1$.

When $x \in]0, 1[$, the marginal distribution function of X is given by

$$F_X(x) = \int_0^x \frac{4}{3}\left(1 - t^3\right) dt = \frac{4}{3}\,x - \frac{1}{3}\,x^4,$$

hence

$$F_X(x) = \begin{cases} 0, & x \le 0, \\ \frac{1}{3}\left(4x - x^4\right), & 0 < x < 1, \\ 1, & x \ge 1. \end{cases}$$

By a *horizontal integration*, $y \in]0,1[$ fixed, we get the marginal frequency of Y,

$$f_Y(y) = \int_0^y 4y^2 \, \mathrm{dx} = 3y^3,$$

hence

$$f_Y(y) = \begin{cases} 4y^3, & y \in]0,1[, \\ 0, & \text{otherwise.} \end{cases}$$

It follows immediately by an integration that the marginal distribution function is given by

$$F_Y(y) = \begin{cases} 0, & y \le 0, \\ y^4, & y \in]0,1[, \\ 1, & y \ge 1. \end{cases}$$

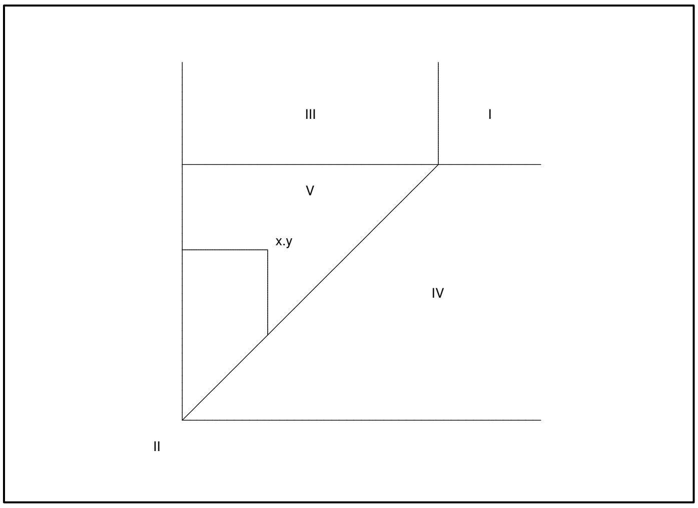


Figure 15: The five sub-domains for the distribution function.

3) When the plane is divided into the five sub-domains I–V, it follows that

I $F(x,y) = 1$ for $x \geq 1$ and $y \geq 1$.

II $F(x,y) = 0$ for $x \leq 0$ or $y \leq 0$.

III $F(x,y) = F_X(x) = \dfrac{1}{3}\left(4x - x^4\right)$ for $0 < x < 1$ and $y \geq 1$.

IV $F(x,y) = F_Y(y) = y^4$ for $0 < y \leq 1$ and $x > y$.

V Only here we need some computations. We keep (x,y) fixed in domain V, thus $0 < x < y < 1$, cf. the figure. First take the *inner* integral [i.e. we first integrate vertically] and then horizontally. Then we get

$$\begin{aligned} F(x,y) &= \int_0^x \left\{ \int_0^y f(t,u)\,du \right\} dt = \int_0^x \left\{ \int_t^y 4u^2\,\mathrm{du} \right\} dt \\ &= \int_0^x \left\{ \frac{4}{3}\,y^3 - \frac{4}{3}\,t^3 \right\} dt = \frac{4}{3}\,xy^3 - \frac{1}{3}\,x^4, \end{aligned}$$

hence

$$F(x,y) = \frac{1}{3}\left(4xy^3 - x^4\right) \qquad \text{for } 0 < x < y < 1.$$

Example 4.6 *A 2-dimensional random variable* (X, Y) *has the frequency*

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

1) *Find the constant* c.
2) *Find the marginal frequencies and the distribution functions of the random variables* X *and* Y.
3) *(A hard question). Find the simultaneous distribution function* $F(x, y)$ *of the 2-dimensional random variable* (X, Y).
(The distribution function $F(x, y)$ *must be computed for* every $(x, y) \in \mathbb{R}^2$*).*

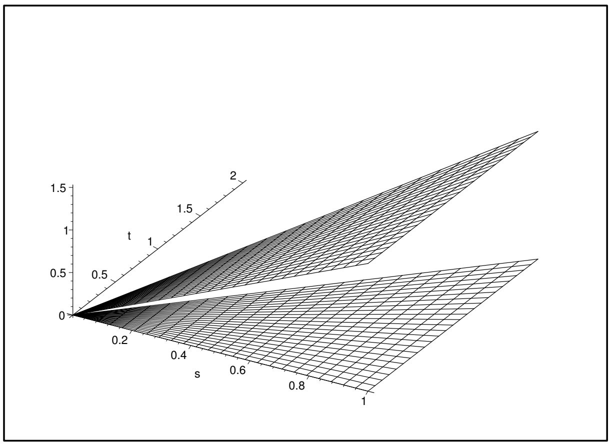


Figure 16: The graph of $f(x, y)$, and its projection A, where $f(x, y) > 0$.

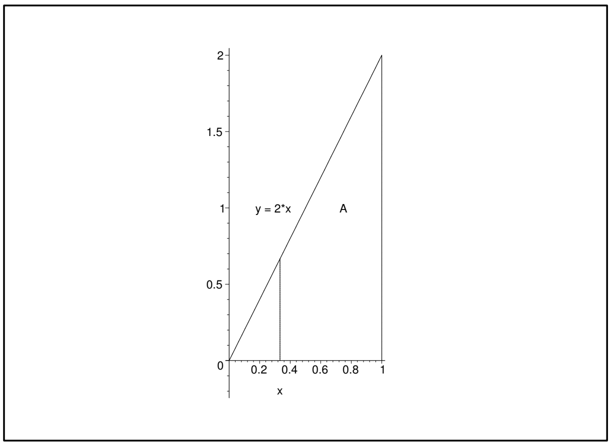


Figure 17: The domain of integration A.

1) By means of a plane integral we get the condition (cf. the figure)

$$1 = \int_A f(x,y)\,dx\,dy = c\int_{x=0}^{1} x\left\{\int_{y=0}^{2x} 1\,dy\right\}dx = c\int_0^1 2x^2\,dx = c\cdot\frac{2}{3}.$$

Therefore, if we choose $c = \frac{3}{2}$, then $f(x,y) \geq 0$ everywhere, and its integral is 1, so the frequency is

$$f(x,y) = \begin{cases} \dfrac{3}{2}\,x, & 0 < y < 2x < 2, \\ 0, & \text{otherwise.} \end{cases}$$

2) The marginal frequencies are found to be

$$f_X(x) = \int_{y=0}^{2} x\frac{3}{2}\,x\,dy = 3x^2 \qquad \text{for } 0 < x < 1,$$

and

$$f_Y(y) = \int_{x=\frac{y}{2}}^{1} \frac{3}{2}\,x\,dx = \left[\frac{3}{4}\,x^2\right]_{x=\frac{y}{2}}^{1} = \frac{3}{4} - \frac{3}{16}\,y^2, \quad \text{for } 0 < y < 2,$$

hence

$$f_X(x) = \begin{cases} 3x^2, & 0 < x < 1, \\ 0, & \text{otherwise,} \end{cases} \qquad f_Y(y) = \begin{cases} \dfrac{3}{4} - \dfrac{3}{16}\,y^2, & 0 < y < 2, \\ 0, & \text{otherwise.} \end{cases}$$

We find the remaining distribution functions by an integration:

$$F_X(x) = \begin{cases} 0, & x < 0, \\ x^3, & 0 \leq x \leq 1, \\ 1, & x > 1, \end{cases} \qquad F_Y(y) = \begin{cases} 0, & y < 0, \\ \dfrac{3}{4}\,y - \dfrac{1}{16}\,y^3, & 0 \leq y \leq 2, \\ 1, & y > 2. \end{cases}$$

3) (The hard question). We divide $\mathbb{R}^2$ into the five domains I–V, cf. the figure.

Clearly,

$$F(x,y) = 0 \qquad \text{in domain I,}$$

$$F(x,y) = 1 \qquad \text{in domain II.}$$

In domain III (i.e. A) we get by first integrating horizontally, $\frac{u}{2} \leq t \leq x$, and then vertically, $0 \leq u \leq y$,

$$\begin{aligned} F(x,y) &= \frac{3}{2}\int_{u=0}^{y}\left\{\int_{t=\frac{u}{2}}^{x} t\,dt\right\}du = \frac{3}{4}\int_{u=0}^{y}\left\{x^2 - \frac{u^2}{4}\right\}du \\ &= \frac{3}{4}\left[x^2u - \frac{1}{12}\,u^3\right]_{u=0}^{y} = \frac{3}{4}\,x^2y - \frac{1}{16}\,y^3. \end{aligned}$$

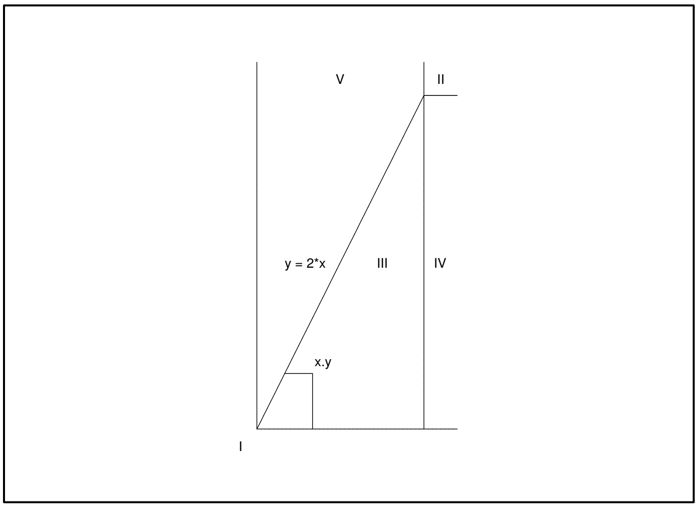


Figure 18: The five domains of the distribution function.

We get in domain IV,

$$F(x,y) = F(1,y) = F_Y(y) = \frac{3}{4}\,y - \frac{1}{16}\,y^3,$$

and in domain V,

$$F(x,y) = F(x,2x) = \frac{3}{4}\,2x^3 - \frac{1}{16}\cdot 8x^4\,(= F_X(x)) = x^3.$$

Summing up we obtain the distribution function

$$F(x,y) = \begin{cases} 0, & \text{for } x \le 0 \text{ or } y \le 0, \\[2mm] 1, & \text{for } x \ge 1 \text{ and } y \ge 2, \\[2mm] \dfrac{3}{4}\,x^2y - \dfrac{1}{16}\,y^3, & \text{for } 0 \le y \le 2x \le 2, \\[2mm] \dfrac{3}{4}\,y - \dfrac{1}{16}\,y^3, & \text{for } x \ge 1 \text{ and } 0 \le y \le 2, \\[2mm] x^3, & \text{for } y \ge 2x \text{ and } 0 \le x \le 1. \end{cases}$$

Example 4.7 *A 2-dimensional random variable* (X, Y) *has the frequency*

$$f(x,y) = \begin{cases} cx, & 0 \le x \le 1,\, 0 \le y \le 1-x, \\ 0, & \textit{otherwise}. \end{cases}$$

1) Find the constant c.

2) Find the frequencies and the distribution functions of the random variables X *and* Y.

3) Find the simultaneous distribution function of (X, Y).

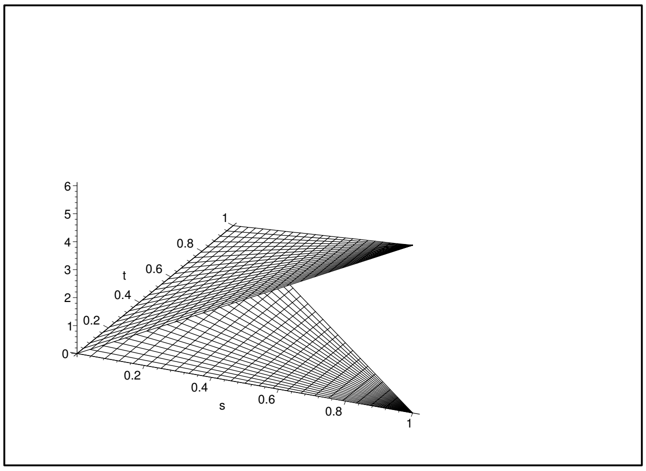


Figure 19: The graph of the frequency $f(x)$.

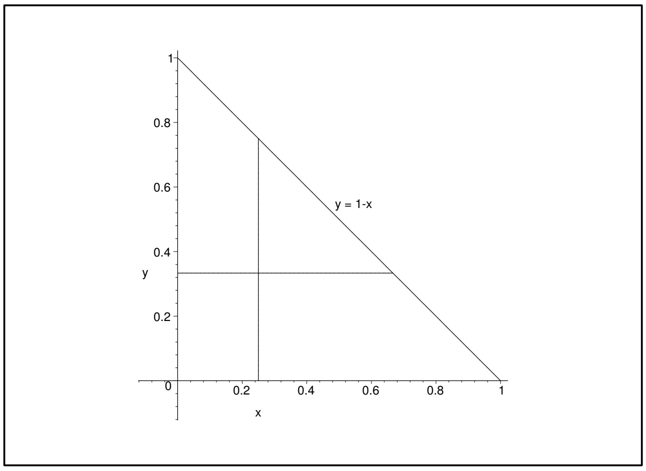


Figure 20: The domain of integration of the frequency $f(x)$.

1) If $c > 0$, then $f(x,y) \geq 0$. It follows from the condition

$$1 = c\int_0^1 \left\{\int_0^{1-x} x\,\mathrm{dy}\right\} dx = c\int_0^1 \left(x - x^2\right)\,dx = c\left\{\frac{1}{2} - \frac{1}{3}\right\} = \frac{c}{6},$$

that if $c = 6$, then the frequency of (X,Y) is given by

$$f(x,y) = \begin{cases} 6x, & 0 \leq x \leq 1,\, 0 \leq y \leq 1-x, \\ \\ 0, & \text{otherwise.} \end{cases}$$

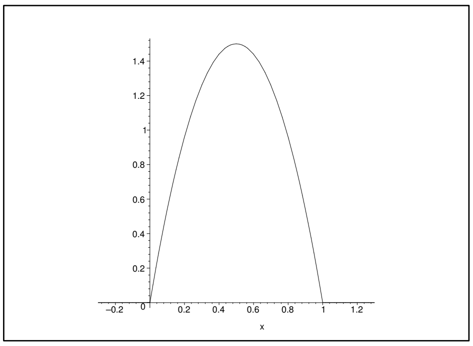


Figure 21: The graph of the frequency $f_X(x)$ of X.

2) It follows by a vertical integration, $x \in]0,1[$ fixed, that

$$f_X(x) = \int_0^{1-x} 6x \, \mathrm{dy} = 6x - 6x_2,$$

so the frequency of X is

$$f_X(x) = \begin{cases} 6x - 6x^2, & x \in]0,1[, \\ 0, & \text{otherwise}. \end{cases}$$

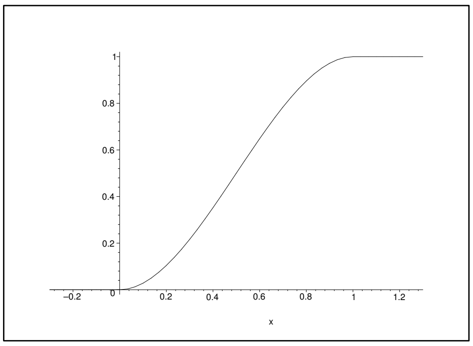


Figure 22: The distribution function $F_X(x)$ of X.

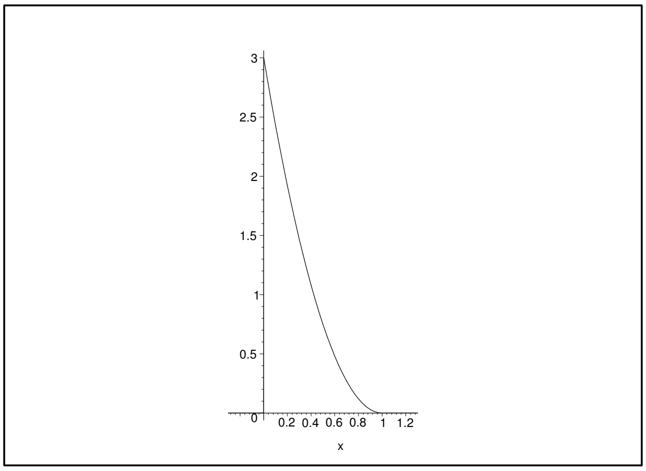


Figure 23: The graph of the frequency $f_Y(y)$.

A horizontal integration, $y \in]0,1[$ fixed, gives

$$f_Y(y) = \int_0^{1-y} 6x \, dx = 3(1-y)^2,$$

hence the frequency of Y is

$$f_Y(y) = \begin{cases} 3(1-y)^2, & 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

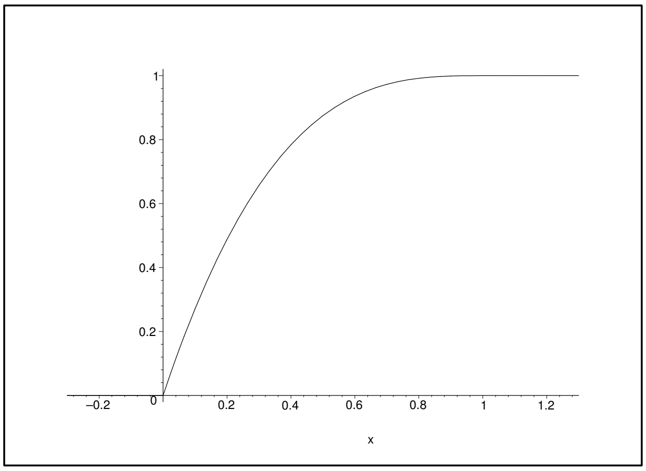


Figure 24: The distribution function $F_Y(y)$.

By an integration, for $y \in]0,1[$ fixed,

$$F_Y(y) = \int_0^y 3(1-t)^2\,dt = \left[-(1-t)^3\right]_0^y = 1-(1-y)^3 = 3y-3y^2+y^3,$$

so the distribution function of Y is

$$F_Y(u) = \begin{cases} 0, & y \le 0, \\ 1-(1-y)^3, & 0 < y < 1, \\ 1, & y \ge 1. \end{cases}$$

3) If we divide the plane into the domains I–VI, it follows that

I $F(x,y) = 1$ for $x \ge 1$ and $y \ge 1$.

II $F(x,y) = 0$ for $x \le 0$ or $y \le 0$.

III $F(x,y) = F_X(x) = 3x^2 - 2x^3$ for $0 < x < 1$ and $y \ge 1$.

IV $F(x,y) = F_Y(y) = 1-(1-y)^3$ for $x \ge 1$ and $0 < y < 1$.

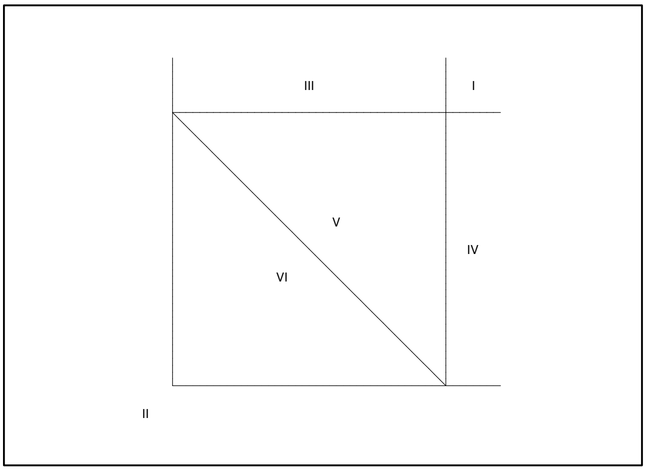


Figure 25: The domains I–VI.

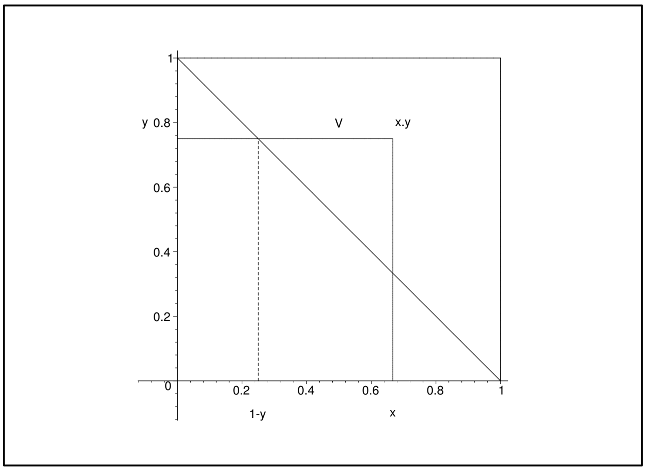


Figure 26: The domain of integration in case V.

V Let $0 \le x \le 1$ and $1-x \le y \le 1$. Then

$$\begin{aligned}
F(x,y) &= \int_0^{1-y} \left\{ \int_0^y 6u\,du \right\} dt + \int_{1-y}^x \left\{ \int_0^{1-t} 6t\,du \right\} dt \\
&= \int_0^{1-y} 3y^2\,dt + \int_{1-y}^x \left(6t - 6t^2\right)\,dt \\
&= 3y^2(1-y) + \left[3t^2 - 2t^3\right]_{1-y}^x \\
&= 3y^2(1-y) + 3x^2 - 2x^3 - 3(1-y)^2 + 2(1-y)^3 \\
&= 3x^2 - 2x^3 + (1-y)\left\{3y - 3y^2 - 3 + 3y + 2 - 4y + 2y^2\right\} \\
&= 3x^2 - 2x^3 + (1-y)\left(-1 + 2y - y^2\right) \\
&= 3x^2 - 2x^3 - (1-y)^3.
\end{aligned}$$

VI Finally, if $0 \le x \le 1$ and $0 \le y \le 1-x$, then

$$F(x,y) = \int_0^x \left\{ \int_0^y 6t\, du \right\} dt = 3x^2 y.$$

Summing up we get

$$F(x,y) = \begin{cases} 1, & \text{for } x \ge 1 \text{ and } y \ge 1, \\[2mm] 0, & \text{for } x \le 0 \text{ or } y \le 0, \\[2mm] 3x^2 - 2x^3, & \text{for } 0 < x < 1 \text{ and } y \ge 1, \\[2mm] 1-(1-y)^3, & \text{for } x \ge 1 \text{ and } 0 < y < 1, \\[2mm] 3x^2 - 2x^3 - (1-y)^3, & \text{for } 0 \le x \le 1 \text{ and } 1-x \le y \le 1, \\[2mm] 3x^2 y, & \text{for } 0 \le x \le 1 \text{ and } 0 \le y \le 1-x. \end{cases}$$

Example 4.8 *Let X_1 and X_2 be independent and identically distributed random variables of the frequencies*

$$f(x) = \begin{cases} \frac{1}{2}\,x, & 0 < x < 2, \\ 0, & \textit{otherwise}, \end{cases}$$

and let the 2-dimensional random variable

$$(Y_1, Y_2) = \tau\,(X_1, X_2)$$

be given by

$$Y_1 = X_1 X_2, \qquad Y_2 = \frac{X_1}{X_2}.$$

1. *Compute the means $E\{X_1\}$ and $E\left\{\frac{1}{X_1}\right\}$.*

2. *Compute the means of Y_1 and for Y_2.*

The vector function τ, given by

$$\tau\,(x_1, x_2) = \left(x_1 x_2, \frac{x_1}{x_2}\right),$$

maps $D =]0,2[\times]0,2[$ bijectively onto

$$D' = \left\{(y_1, y_2) \in \mathbb{R}^2 \mid 0 < y_1 < 4y_2,\, y_1 y_2 < 4\right\}.$$

3. *Sketch D', and find the simultaneous frequency $k\,(y_1, y_2)$ of (Y_1, Y_2).*

4. *Find the marginal frequencies of Y_1 and Y_2.*

5. *Are Y_1 and Y_2 independent?*

1) It follows that

$$E\{X_1\} = \frac{1}{2}\int_0^2 x^2\,dx = \frac{1}{2}\left[\frac{x^3}{3}\right]_0^2 = \frac{4}{3},$$

and

$$E\left\{\frac{1}{X_1}\right\} = \frac{1}{2}\int_0^2 \frac{x}{x}\,dx = 1.$$

2) Since X_1 and X_2 are independent, we conclude from (1) that

$$E\{Y_1\} = E\{X_1 X_2\} = E\{E_1\}\,E\{X_2\} = \left(\frac{4}{3}\right)^3 = \frac{16}{9},$$

and

$$E\{Y_2\} = E\left\{\frac{X_1}{X_2}\right\} = E\{X_1\} \cdot E\left\{\frac{1}{X_2}\right\} = \frac{4}{3}\cdot 1 = \frac{4}{3}.$$

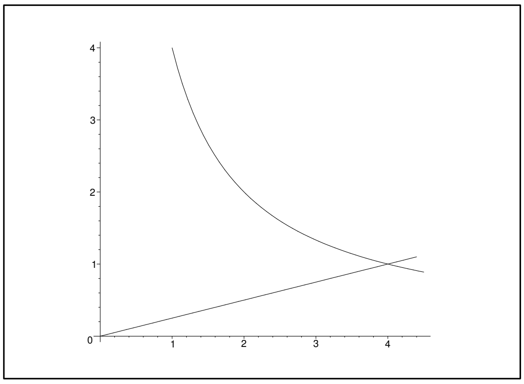


Figure 27: The domain D' lies between the y_2-axis, the hyperbola $y_1 y_2 = 4$ and the straight line $y_1 = 4y_2$.

3) It follows from $y_1 = x_1 x_2$ and $y_2 = \dfrac{x_1}{x_2}$ that

$$x_1 = \sqrt{y_1 y_2} \qquad \text{and} \qquad x_2 = \sqrt{\frac{y_1}{y_2}}.$$

Hence the Jacobian becomes

$$\frac{\partial(x_1,x_2)}{\partial(y_1,y_2)} = \begin{vmatrix} \dfrac{1}{2}\sqrt{\dfrac{y_2}{y_1}} & \dfrac{1}{2}\sqrt{\dfrac{y_1}{y_2}} \\ \dfrac{1}{2}\dfrac{1}{\sqrt{y_1 y_2}} & -\dfrac{1}{2}\dfrac{1}{y_2}\sqrt{\dfrac{y_1}{y_2}} \end{vmatrix} = -\frac{1}{4}\left(\frac{1}{y_2}+\frac{1}{y_2}\right) = -\frac{1}{2y_2}.$$

The simultaneous frequency of (X_1, X_2) is

$$g(x_1,x_2)\, f(x_1)\cdot f(x_2) = \begin{cases} \dfrac{1}{4}\, x_1 x_2, & \text{for } (x_1,x_2) \in]0,2[\times]0,2[, \\ 0 & \text{otherwise,} \end{cases}$$

so the simultaneous frequency of (Y_1, Y_2) is

$$k(y_1,y_2) = \begin{cases} \dfrac{1}{8}\,\dfrac{y_1}{y_2} & \text{for } (y_1,y_2) \in D', \\ 0 & \text{otherwise.} \end{cases}$$

4) The marginal frequency of Y_1 for $0 < y_1 < 4$ is given by

$$k_{Y_1}(y_1) = \frac{y_1}{8}\int_{\frac{y_1}{4}}^{\frac{4}{y_1}} \frac{1}{y_2}\, dy_2 = \frac{y_1}{8}\left[\ln y_2\right]_{\frac{y_1}{4}}^{\frac{4}{y_1}} = \frac{y_1}{4}\ln\left(\frac{4}{y_1}\right) = \frac{y_1}{4}\left(\ln 4 - \ln y_1\right),$$

and $= 0$ otherwise.
The marginal frequency of Y_2 for $y_2 \in]0,1]$ is given by

$$k_{Y_2}(y_2) = \frac{1}{8y_2}\int_0^{4y_2} y_1\, dy_1 = \frac{1}{16y_2}\cdot 16y_2^2 = y_2.$$

If instead $y_2 \in]1, \infty[$, then

$$k_{Y_2}(y_2) = \frac{1}{8y_2} \int_0^{\frac{4}{y_2}} y_1 \, dy_1 = \frac{1}{y_2^3}.$$

Summing up we get

$$k_{Y_1}(y_1) = \begin{cases} \dfrac{y_1}{4} \ln\left(\dfrac{4}{y_1}\right), & 0 < y_1 < 4, \\ \\ 0, & \text{otherwise,} \end{cases}$$

and

$$k_{Y_2}(y_2) = \begin{cases} y_2, & \text{for } y_2 \in]0,1], \\ \\ \dfrac{1}{y_2^3}, & \text{for } y_2 \in]1,\infty[, \\ \\ 0, & \text{for } y_2 \leq 0. \end{cases}$$

5) Since D' is not a rectangular domain, we conclude that Y_1 and Y_2 cannot be independent. There is clearly a trap here, because we get for $(y_1, y_2) \in D'$,

$$k(y_1, y_2) = \frac{1}{8} \cdot y_1 \cdot \frac{1}{y_2},$$

in which y_1 and y_2 apparently are separated.

Example 4.9 *Let (X_1, X_2) be a 2-dimensional random variable of the frequency*

$$h(x_1, x_2) = \begin{cases} \dfrac{1}{2}\,(x_1 + x_2)\exp\left(-(x_1 + x_2)\right), & x_1 > 0,\ x_2, \\ \\ 0, & \textit{otherwise}. \end{cases}$$

1. *Compute the marginal frequencies of X_1 and X_2.*

Then introduce the random variables Y_1 and Y_2 by

$$Y_1 = X_1 + X_2, \qquad Y_2 = X_1 - X_2.$$

2. *Prove that Y_1 are Y_2 non-correlated.*

The vector function τ given by

$$\tau(x_1, x_2) = (x_1 + x_2, x_1 - x_2)$$

maps $\mathbb{R}_+ \times \mathbb{R}_+$ bijectively onto

$$D' = \left\{(y_1, y_2) \in \mathbb{R}^2 \mid |y_2| < y_1\right\}.$$

3. *Compute the simultaneous frequency $k(y_1, y_2)$ of (Y_1, Y_2).*

4. *Compute the marginal frequency of Y_1, and find the mean of Y_1.*

5. *Find the marginal frequency of Y_2.*

1) If $x_1 > 0$, then

$$h_{X_1}(x_1) = \frac{1}{2}\,e^{-x_1}\int_0^{\infty}\left(x_1 e^{-x_2} + x_2 e^{-x_2}\right)dx_2 = \frac{1}{2}\,(x_1 + 1)\,e^{-x_1},$$

and $h_{X_1}(x_1) = 0$ for $x_1 \le 0$.
Then by symmetry, X_1 and X_2 have the same distribution, hence

$$h_{X_2}(x_2) = \begin{cases} \dfrac{1}{2}\,(x_2 + 1)\,e^{-x_2}, & \text{for } x_2 > 0, \\ \\ 0, & \text{for } x_2 \le 0. \end{cases}$$

2) It follows from (1) that $V\{X_1\} = V\{X_2\}$, thus

$$\operatorname{Cov}(Y_1, Y_2) = \operatorname{Cov}(X_1 + X_2, X_1 - X_2) = V\{X_1\} - V\{X_2\} + \operatorname{Cov}(X_1, X_2) = 0,$$

which shows that Y_1 and Y_2 are non-correlated.
NOTICE that they are not independent, cf. (3), because the domain is not a rectangle parallel to the coordinate axes.

3) It follows from $y_1 = x_1 + x_2$ and $y_2 = x_1 - x_2$ that

$$x_1 = \frac{1}{2}\,(y_1 + y_2) \qquad \text{and} \qquad x_2 = \frac{1}{2}\,(y_1 - y_2)\,,$$

so the Jacobian becomes

$$\frac{\partial\,(x_1, x_2)}{\partial\,(y_1, y_2)} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

Hence,

$$k\,(y_1, y_2) = \begin{cases} \frac{1}{4}\,y_1 e^{-y_1}, & \text{for } |y_2| < y_1, \\ 0, & \text{otherwise.} \end{cases}$$

4) The marginal frequency of Y_1 is

$$f_{Y_1}\,(y_1) = \frac{1}{4}\,y_1 e^{-y_1} \int_{-y_1}^{y_1} dy_2 = \frac{1}{2}\,y_1^2 e^{-y_1} \qquad \text{for } y_1 > 0,$$

and $f_{Y_1}\,(y_1) = 0$ otherwise.
Since $Y_1 \in \Gamma(3,1)$ is gamma distributed, we get

$$E\,\{Y_1\} = 3 \cdot 1 = 2,$$

which can also be found directly from

$$E\,\{Y_1\} = \frac{1}{2} \int_0^{\infty} y_1^3 e^{-y_1}\,dy_1 = \frac{3!}{2} = 3.$$

5) If $y_2 \in \mathbb{R}$, then

$$f_{Y_2}\,(y_2) = \frac{1}{4} \int_{|y_2|}^{\infty} y_1 e^{-y_1}\,dy_1 = \frac{1}{4}\,\left[-\,(y_1+1)\,e^{-y_1}\right]_{|y_2|}^{\infty} = \frac{1}{4}\,(|y_2|+1)\,e^{-|y_2|}.$$

Example 4.10 *Let* (X_1, X_2) *be a 2-dimensional random variable of the frequency*

$$h(x_1, x_2) = \begin{cases} \frac{1}{2}\,(x_1+1)\;e^{-(x_1+x_2)}, & x_1 > 0, \quad x_2 > 0, \\ \\ 0, & \textit{otherwise.} \end{cases}$$

1. *Compute the marginal frequencies of* X_1 *and* X_2*.*

2. *Compute the means of* X_1 *and* X_2*.*

We introduce the random variables Y_1 *and* Y_2 *by*

$$Y_1 = X_1 + X_2, \qquad Y_2 = \frac{X_2}{X_1 + X_2}.$$

The vector function τ*, given by*

$$\tau(x_1, x_2) = \left(x_1 + x_2, \frac{x_2}{x_1 + x_2} \right),$$

maps $\mathbb{R}_+ \times \mathbb{R}_+$ *bijectively onto* $D' = \mathbb{R}_+ \times\,]0, 1[$*.*

3. *Compute the simultaneous frequency* $k(y_1, y_2)$ *of* (Y_1, Y_2)*.*

4. *Compute the marginal frequencies of* Y_1 *and* Y_2*, and check if* Y_1 *and* Y_2 *are independent.*

5. *Compute the means of* Y_1 *and* Y_2*.*

1) Since $h(x_1, x_2)$ has a nice factorization,

$$h(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2),$$

where

$$f_{X_1}(x_1) = \begin{cases} \frac{1}{2}\,(x_1+1)\,e^{-x_1}, & x_1 > 0, \\ \\ 0, & x_1 \le 0, \end{cases} \qquad f_{X_2}(x_2) = \begin{cases} e^{-x_2}, & x_2 > 0, \\ \\ 0, & x_2 \le 0, \end{cases}$$

and $f_{X_1}(x_1) \ge 0$ and $f_{X_2}(x_2) \ge 0$, where

$$\int_{-\infty}^{\infty} f_{X_1}(x_1)\; dx_1 = 1, \qquad \int_{-\infty}^{\infty} f_{X_2}(x_2)\; dx_2 = 1,$$

we have

a) found the marginal frequencies,

b) and shown that X_1 and X_2 are stochastically independent.

Alternatively we compute for $x_1 > 0$, resp. $x_2 > 0$,

$$f_{X_1}(x_1) = \int_0^{\infty} h(x_1, x_2)\; dx_2 = \frac{1}{2}\,(x_1+1)\,e^{-x_1} \int_0^{\infty} e^{-x_2}\, dx_2 = \frac{1}{2}\,(x_1+1)\,e^{-x_1},$$

and

$$f_{X_2}(x_2) = \int_0^\infty h(x_1, x_2)\, dx_1 = \frac{1}{2}\, e^{-x_2} \int_0^\infty (x_1 + 1)\, e^{-x_1}\, dx_1 = e^{-x_2}.$$

Here we have applied that

$$\int_0^\infty x^n e^{-x}\, dx = n! \qquad \text{for } n \in \mathbb{N}_0.$$

2) The means are

$$E\{X_1\} = \int_0^\infty x_1 f_{X_1}(x_1)\, dx_1 = \frac{1}{2} \int_0^\infty \left(x_1^2 + x_1\right) e^{-x_1}\, dx_1 = \frac{1}{2}\,(2! + 1!) = \frac{3}{2},$$

and

$$E\{X_2\} = \int_0^\infty x_2 f_{X_2}(x_2)\, dx_2 = \int_0^\infty x_2\, e^{-x_2}\, dx_2 = 1! = 1.$$

3) The transform formula is memorized by

$$h(x_1, x_2)\, dx_1\, dx_2 = k(y_1, y_2) \left| \frac{\partial (x_1, x_2)}{\partial (y_1, y_2)} \right| dy_1\, dy_2.$$

This formula shows that the task is to find x_1 and x_2 expressed by (y_1, y_2). It follows from

$$\begin{cases} y_1 = x_1 + x_2, \\ y_2 = \dfrac{x_2}{x_1 + x_2}, \end{cases} \qquad \text{that} \qquad \begin{cases} y_1 = x_1 + x_2, \\ y_1 y_2 = x_2, \end{cases}$$

hence

$$\begin{cases} x_1 = y_1 \left(1 - y_2\right), \\ x_2 = y_1 y_2. \end{cases}$$

Thus we get the weight function

$$\frac{\partial \left(x_1, x_2\right)}{\partial \left(y_1, y_2\right)} = \begin{vmatrix} 1 - y_2 & -y_1 \\ y_2 & y_1 \end{vmatrix} = y_1 - y_1 y_2 + y_1 y_2 = y_1 > 0,$$

because $D' = \mathbb{R}_+ \times \,]0,1[$ is given.
In this domain we get the simultaneous frequency of (Y_1, Y_2),

$$\begin{aligned} k\left(y_1, y_2\right) &= h\left(y_1\left(1 - y_2\right), y_1 y_2\right) \cdot y_1 = \frac{1}{2}\left(y_1 - y_1 y_2 + 1\right) \cdot e^{-y_1} \cdot y_1 \\ &= \frac{1}{2}\left(y_1^2 - y_1^2 y_2 + y_1\right) e^{-y_1} \qquad\qquad \text{for } \left(y_1, y_2\right) \in D', \end{aligned}$$

and $k\left(y_1, y_2\right) = 0$ otherwise.

4) The marginal frequencies of Y_1 and Y_2 are computed for $y_1 > 0$, resp. $y_2 \in \,]0,1[$. (Otherwise they are 0.)

$$\begin{aligned} k_{Y_1}\left(y_1\right) &= \int_{y_2=0}^{1} k\left(y_1, y_2\right) dy_2 = \frac{1}{2}\, e^{-y_1} \int_0^1 \left(y_1^2 - y_1^2 y_2 + y_1\right) dy_2 \\ &= \frac{1}{2}\, e^{-y_1}\left(y_1^2 - \frac{1}{2}\, y_1^2 + y_1\right) = \left(\frac{1}{4}\, y_1^2 + \frac{1}{2}\, y_1\right) e^{-y_1}, \\ k_{Y_2}\left(y_2\right) &= \int_{y_1=0}^{\infty} k\left(y_1, y_2\right) dy_1 = \frac{1}{2}\int_0^{\infty} \left(y_1^2 - y_1^2 y_2 + y_1\right) e^{-y_1}\, dy_1 \\ &= \frac{1}{2}\left(2 - 2y_2 + 1\right) = \frac{3}{2} - y_2. \end{aligned}$$

Since $k\left(y_1, y_2\right) \neq k_{Y_1}\left(y_1\right) \cdot k_{Y_2}\left(y_2\right)$, we see that Y_1 and Y_2 are not independent.

5) The means are

$$E\left\{Y_1\right\} = E\left\{X_1 + X_2\right\} = E\left\{X_1\right\} + E\left\{X_2\right\} = \frac{3}{2} + 1 = \frac{5}{2}.$$

Alternatively,

$$E\left\{Y_1\right\} = \int_0^{\infty} y_1 k_{Y_1}\left(y_1\right) dy_1 = \int_0^{\infty} \left(\frac{1}{4}\, y_1^3 + \frac{1}{2}\, y_1^2\right) e^{-y_1}\, dy_1 = \frac{6}{4} + \frac{2}{2} = \frac{5}{2}.$$

For Y_2 we get

$$E\left\{Y_2\right\} = \int_0^1 y_2 k_{Y_2}\left(y_2\right) dy_2 = \int_0^1 \left(\frac{3}{2}\, y_2 - y_2^2\right) dy_2 = \frac{3}{4} - \frac{1}{3} = \frac{5}{12}.$$

Example 4.11 *Let (X,Y) be a 2-dimensional random variable of the frequency*

$$h(x,y)=\begin{cases} e^{-y}, & \textit{for } 0\le x\le y,\\ 0, & \textit{otherwise.}\end{cases}$$

1) Find the marginal frequencies of the random variables X and Y.

2) Find the means $E\{X\}$ and $E\{Y\}$ of the random variables X and Y.

3) Find the variances $V\{X\}$ and $V\{Y\}$ of the random variables X and Y.

4) Compute $E\{XY\}$, and then the correlation $\varrho(X,Y)$.

5) Find the frequency of $Z=X+Y$.

1) The marginal frequency of X is

$$f_X(x)=\int_{y=x}^{\infty} e^{-y}\,dy=e^{-x} \qquad \text{for } x\ge 0,$$

and $f_X(x)=0$ for $x<0$.
Analogously the marginal frequency of Y is given by

$$f_Y(y)=\int_{x=0}^{y} e^{-y}\,dx=y\,e^{-y} \qquad \text{for } y\ge 0,$$

and $f_Y(y)=0$ for $y\le 0$.
Summing up we get

$$X\in\Gamma(1,1) \qquad \text{and} \qquad Y\in\Gamma(2,1).$$

2) Then

$$E\{X\}=\int_0^{\infty} x\,e^{-x}\,dx=\left[-(x+1)\,e^{-x}\right]_0^{\infty}=1,$$

and

$$E\{Y\}=\int_0^{\infty} y\cdot y\,e^{-y}\,dy=\int_0^{\infty} y^2e^{-y}\,dy=\left[\left(-y^2-2y-2\right)e^{-y}\right]_0^{\infty}=2.$$

3) We first compute

$$E\left\{X^2\right\}=\int_0^{\infty} x^2e^{-x}\,dx=\left[\left(-x^2-2x-2\right)e^{-x}\right]_0^{\infty}=2$$

and

$$E\left\{Y^2\right\}=\int_0^{\infty} y^2\cdot y\,e^{-y}\,dy=\int_0^{\infty} y^3e^{-y}\,dy=3!=6.$$

This gives us the variances

$$\begin{aligned} V\{X\} &= E\left\{X^2\right\}-(E\{X\})^2=2-1^2=1,\\ V\{Y\} &= E\left\{Y^2\right\}-(E\{Y\})^2=6-2^2=2.\end{aligned}$$

4) By a reduction to a double integral we get

$$E\{XY\} = \int\int_{\mathbb{R}^2} xy\, h(x,y)\, dx\, dy = \int_{y=0}^{\infty} \left\{ \int_{x=0}^{y} x\, dx \right\} y\, e^{-y}\, dy = \int_0^{\infty} \frac{1}{2}\, y^3 e^{-y}\, dy = 3,$$

or ALTERNATIVELY

$$\begin{aligned} E\{XY\} &= \int_{x=0}^{\infty} x \left\{ \int_{y=x}^{\infty} y\, e^{-y}\, dy \right\} dx = \int_{x=0}^{\infty} x\, \left[-(y+1)e^{-y} \right]_x^{\infty}\, dy \\ &= \int_0^{\infty} \left(x^2 e^{-x} + x\, e^{-x} \right)\, dx = 2 + 1 = 3. \end{aligned}$$

Then

$$\operatorname{Cov}(X,Y) = E\{XY\} - E\{X\} \cdot E\{Y\} = 3 - 2 \cdot 1 = 1,$$

hence

$$\varrho(X,Y) = \frac{\operatorname{Cov}(X,Y)}{\sqrt{V\{X\}\, V\{Y\}}} = \frac{1}{\sqrt{1 \cdot 2}} = \frac{\sqrt{2}}{2}.$$

5) The random variable $Z = X + Y$ has its values in $]0, \infty[$. If $z > 0$, then the frequency is given by

$$f_Z(z) = \int_{-\infty}^{\infty} h(x, z-x)\, dx,$$

where the condition $0 \le x \le y = z - x$ is reformulated to

$$0 \le x \le \frac{z}{2}.$$

Hence, when $z > 0$,

$$f_Z(z) = \int_0^{\frac{z}{2}} h(x, z-x)\, dx = \int_0^{\frac{z}{2}} e^{-(z-x)}\, dx = e^{-z} \left\{ \exp\left(\frac{z}{2}\right) - 1 \right\} = \exp\left(-\frac{z}{2}\right) - \exp(-z),$$

so

$$f_Z(z) = \begin{cases} \exp\left(-\dfrac{z}{2}\right) - e(-z), & \text{for } z \ge 0, \\ \\ \qquad 0, & \text{for } z < 0. \end{cases}$$

ALTERNATIVELY we compute the distribution function of Z by the following double integral,

$$\begin{aligned} F_Z(z) &= \int_{x=0}^{\frac{z}{2}} \left\{ \int_{y=x}^{z-x} e^{-y}\, dy \right\} dx = \int_{x=0}^{\frac{z}{2}} \left\{ e^{-x} - e^{x-z} \right\} dx \\ &= 1 - \exp\left(-\frac{z}{2}\right) - e^{-z} \left(\exp\left(\frac{z}{2}\right) - 1 \right) = 1 + e^{-z} - 2\, \exp\left(-\frac{z}{2}\right) \\ &= \left\{ 1 - \exp\left(-\frac{z}{2}\right) \right\}^2. \end{aligned}$$

Hence,

$$f_Z(z) = F_Z'(z) = \exp\left(-\frac{z}{2}\right) - e^{-z} \qquad \text{for } z \ge 0,$$

and $f_Z(z) = 0$ for $z < 0$.

ADDITIONAL REMARK. Since

$$E\{Z\} = \int_0^\infty \left\{ z \, \exp\left(-\frac{z}{2}\right) - z\, e^{-z} \right\} dz = 4 - 1 = 3,$$

and

$$E\left\{Z^2\right\} = \int_0^\infty \left\{ z^2 \exp\left(-\frac{z}{2}\right) - z^2 e^{-z} \right\} dz = 16 - 2 = 14,$$

we get

$$V\{Z\} = 14 - 3^2 = 5 = V\{X+Y\}.$$

This gives

$$\operatorname{Cov}(X,Y) = \frac{1}{2}\left(V\{X+Y\} - V\{X\} - V\{Y\}\right) = \frac{1}{2}\{5-1-2\} = 1. \qquad \Diamond$$

Example 4.12 *A 2-dimensional random variable (X,Y) has the simultaneous frequency*

$$f(x,y)=\frac{1}{2}\,xy, \qquad 0<y<x<2 \qquad (\textit{and 0 otherwise}).$$

1) Compute the marginal frequencies of X and Y.

2) Compute the marginal distribution functions of X and Y.

3) Find the means of X and Y.

4) Find the medians of X and Y.

1) *The marginal frequencies:*

a) For fixed $x \in [0,2]$ we integrate with respect to $y \in [0,x]$, which gives

$$f_X(x)=\int_{y=0}^{x}\frac{1}{2}\,xy\,dy=\frac{1}{4}\,x^3, \qquad 0<x<2,$$

and $f_X(x)=0$ otherwise.

b) For fixed $y \in [0,2]$ we integrate with respect to $x \in [y,2]$, which gives

$$f_Y(y)=\int_{x=y}^{2}\frac{1}{2}\,xy\,dx=y-\frac{1}{4}\,y^3, \qquad 0<y<2,$$

and $f_Y(y)=0$ otherwise.

2) We get the distribution functions by integrating the frequencies,

$$F_X(x)=\begin{cases} 0, & x\le 0,\\[2mm] \dfrac{1}{16}\,x^4, & 0<x<2,\\ 1, \\ & x\ge 2, \end{cases}$$

and

$$F_Y(y)=\begin{cases} 0, & y\le 0,\\[2mm] \dfrac{1}{2}\,y^2-\dfrac{1}{16}\,y^4, & 0<y<2,\\[2mm] 1, & y\ge 2. \end{cases}$$

3) Then we compute the means,

$$\begin{aligned} E\{X\} &= \int_0^2 x\,f_X(x)\,dx=\int_0^2\frac{1}{4}\,x^4\,dx=\left[\frac{1}{20}\,x^5\right]_0^2=\frac{32}{20}=\frac{8}{5},\\ E\{Y\} &= \int_9^2 y\,f_Y(y)\,dy=\int_0^2\left(y^2-\frac{1}{4}\,y^4\right)dy=\frac{8}{3}-\frac{8}{5}=\frac{16}{15}. \end{aligned}$$

4) The median of X is found from the equation

$$F_X(x) = \frac{1}{16}\,x^4 = \frac{1}{2},$$

i.e. $x^4 = 8$, hence $(X) = \sqrt[4]{8}$.
The median of Y is found from the equation

$$F_Y(y) = \frac{1}{2}\,y^2 - \frac{1}{16}\,y^4 = \frac{1}{2},$$

i.e.

$$y^4 - 8y^2 + 8 = 0.$$

Since $y^2 \leq 2^2 = 4$, we get $y^2 = 4 - \sqrt{8}$, so

$$(Y) = \sqrt{4 - \sqrt{8}}.$$

Example 4.13 *A rectangle has the edge lengths X_1 and X_2, where X_1 and X_2 are independent random variables, both of the frequency*

$$f(x) = \begin{cases} 3x^2, & 0 < x < 1, \\ 0, & \textit{otherwise.} \end{cases}$$

1. *Find the mean $E\{X_1\}$.*

2. *Find the mean of the circumference of the rectangle, $E\{2X_1 + 2X_2\}$, and the mean of the area of the rectangle, $E\{X_1 X_2\}$.*

We introduce the random variables Y_1 and Y_2 by

$$Y_1 = X_1 X_2, \qquad Y_2 = \frac{X_1}{X_2}.$$

The vector function τ, given by

$$\tau(x_1, x_2) = \left(x_1 x_2, \frac{x_1}{x_2} \right),$$

maps $D =]0,1[\times]0,1[$ bijectively onto

$$D' = \left\{ (y_1, y_2) \in \mathbb{R}^2 \mid 0 < y_1 < y_2,\ y_1 y_2 < 1 \right\}.$$

3. *Sketch D' and find the simultaneous frequency $k(y_1, y_2)$ for (Y_1, Y_2).*

4. *Compute the marginal frequencies of Y_1 and Y_2.*
(This can be answerede with or without using (3)).

5. *Are Y_1 and Y_2 stochastically independent?*

6. *Find the mean and the median of Y_2, and give an intuitive explanation of that the median is smaller than the mean.*

1) The mean is

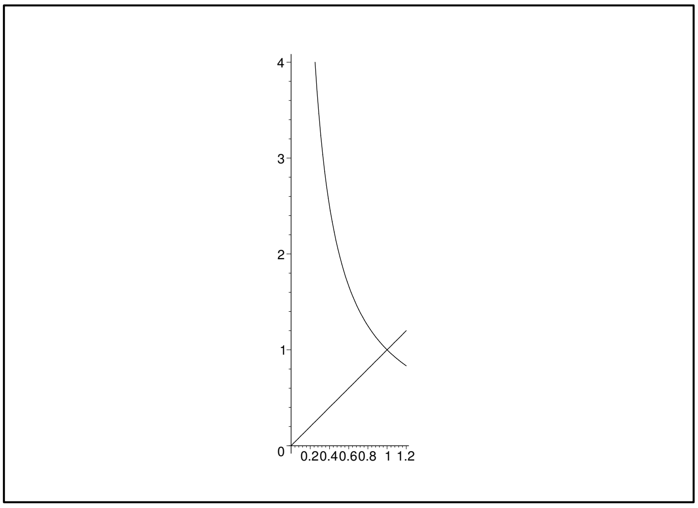
$$E\{X_1\} = E\{X_2\} = \int_0^1 3x^3\,dx = \frac{3}{4}.$$

2) Since X_1 and X_2 are independent, we get

$$E\{2X_1 + 2X_2\} = 4\,E\{X_1\} = 3,$$

and

$$E\{X_1 X_2\} = E\{X_1\} \cdot E\{X_2\} = \left(\frac{3}{4} \right)^2 = \frac{9}{16}.$$



3) From

$$x_1 = \sqrt{y_1 y_2} \qquad \text{og} \qquad x_2 = \sqrt{\frac{y_1}{y_2}},$$

we get the Jacobian

$$\frac{\partial(x_1,x_2)}{\partial(y_1,y_2)} = \begin{vmatrix} \frac{1}{2}\sqrt{\frac{y_2}{y_1}} & \frac{1}{2}\sqrt{\frac{y_1}{y_2}} \\ \frac{1}{2}\,\frac{1}{\sqrt{y_1 y_2}} & -\frac{1}{2}\,\frac{1}{y_2}\sqrt{\frac{y_1}{y_2}} \end{vmatrix} = -\frac{1}{2y_2} < 0,$$

and the simultaneous frequency for $(y_1,y_2) \in D'$, is given by

$$k(y_1,y_2) = 3\left(\sqrt{y_1 y_2}\right)^2 \cdot 3\left(\sqrt{\frac{y_1}{y_2}}\right)^2 \cdot \frac{1}{2y_2} = 9 \cdot y_1 y_2 \cdot \frac{y_1}{y_2} \cdot \frac{1}{2y_2} = \frac{9}{2}\,\frac{y_1^2}{y_2},$$

and $k(y_1,y_2) = 0$ otherwise.

4) The marginal frequency of Y_1 for $y_1 \in \,]0,1[$, is given by

$$k_{Y_1}(y_1) = \frac{9}{2}\,y_1^2 \int_{y_1}^{\frac{1}{y_1}} \frac{dy_2}{y_2} = \frac{9}{2}\,y_1^2\,\left[\ln y_2\right]_{y_1}^{\frac{1}{y_1}} = 9y_1^2 \ln\frac{1}{y_1} = -9y_1^2 \ln y_1,$$

and 0 otherwise.
If $y_2 \in \,]0,1]$, then

$$k_{Y_2}(y_2) = \frac{9}{2y_2}\int_0^{y_2} y_1^2\,dy_1 = \frac{3}{2y_2}\,y_2^3 = \frac{3}{2}\,y_2^2.$$

If $y_2 \in \,]1,\infty[$, then

$$k_{Y_2}(y_2) = \frac{9}{2y_2}\int_0^{\frac{1}{y_2}} y_1^2\,dy_1 = \frac{3}{2y_2}\cdot\frac{1}{y_2^3} = \frac{3}{2}\cdot\frac{1}{y_2^4},$$

hence

$$k_{Y_2}(y_2) = \begin{cases} \dfrac{3}{2}\, y_2^2 & \text{for } y_2 \in\,]0,1], \\[2mm] \dfrac{3}{2}\, \dfrac{1}{y_2^4} & \text{for } y_2 \in\,]1,\infty[, \\[2mm] 0 & \text{otherwise.} \end{cases}$$

5) Since D' is not a rectangular domain, we conclude that Y_1 and Y_2 are not independent. ALTERNATIVELY we see that

$$k(y_1, y_2) \neq k_{Y_1}(y_1) \cdot k_{Y_2}(y_2)\,.$$

6) The mean of Y_2 is

$$E\{Y_2\} = \frac{3}{2}\int_0^1 y_2^3\, dy_2 + \frac{3}{2}\int_1^{\infty} \frac{dy_2}{y_2^3} = \frac{3}{8} + \frac{3}{4} = \frac{9}{8}.$$

It follows from

$$P\{Y_2 < 1\} = \int_0^1 \frac{3}{2}\, y_2^2\, dy_2 = \frac{1}{2},$$

that the median is

$$\langle Y_2 \rangle = 1 < E\{Y_2\}\,.$$

Example 4.14 *Let X_1 and X_2 be independent random variables of the frequencies*

$$f_{X_1}(x_1) = \begin{cases} x_1 e^{-x_1}, & x_1 > 0, \\ 0, & x_1 \le 0, \end{cases}$$

and

$$f_{X_2}(x_2) = \begin{cases} 9x_2 e^{-3x_2}, & x_2 > 0, \\ 0, & x_2 \le 0. \end{cases}$$

1. *Find the means $E\{X_1\}$ and $E\{X_2\}$.*

We introduce the random variables Y_1 and Y_2 by

$$Y_1 = X_1 + 3X_2, \qquad Y_2 = \frac{X_1}{X_1 + 3X_2}.$$

The vector function τ, given by

$$\tau(x_1, x_2) = \left(x_1 + 3x_2, \frac{x_1}{x_1 + 3x_2} \right),$$

maps $\mathbb{R}_+ \times \mathbb{R}_+$ bijectively onto $\mathbb{R}_+ \times \,]0,1[$.

2. *Compute the simultaneous frequency $k(y_1, y_2)$ of (Y_1, Y_2).*

3. *Compute the marginal frequencies of Y_1 and Y_2.*

4. *Check if Y_1 and Y_2 are stochastically independent.*

5. *Compute the mean $E\{Y_2\}$, and prove that the median of Y_2 is equal to $E\{Y_2\}$.*

1) Since $X_1 \in \Gamma(2,1)$, we have $E\{X_1\} = 2 \cdot 1 = 2$.
Since $X_2 \in \Gamma\left(2, \frac{1}{3}\right)$, we have $E\{X_2\} = 2 \cdot \frac{1}{3} = \frac{2}{3}$.

2) It follows from

$$y_1 = x_1 + 3x_2 \qquad \text{and} \qquad y_2 = \frac{x_1}{x_1 + 3x_2}$$

that

$$x_1 = y_1 y_2 \qquad \text{and} \qquad x_2 = \frac{1}{3}(y_1 - x_1) = \frac{1}{3}\, y_1 (1 - y_2).$$

The Jacobian becomes

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} y_2 & y_1 \\ \frac{1}{3}(1 - y_2) & -\frac{1}{3}\, y_1 \end{vmatrix} = -\frac{1}{3}\, y_1 < 0.$$

The simultaneous frequency og $(y_1, y_2) \in \mathbb{R}_+ \times \,]0,1[$ is given by

$$\begin{aligned} k(y_1, y_2) &= y_1 y_2 \cdot e^{-y_1 y_2} \cdot 9 \cdot \frac{1}{3}\, y_1 (1-y_2) \cdot e^{-y_1(1-y_2)} \cdot \frac{1}{3}\, y_1 \\ &= y_1^3 \cdot e^{-y_1} \cdot y_2 (1-y_2) = \left\{ \frac{1}{6}\, y_1^3 e^{-y_1} \right\} \cdot \{6 y_2 (1-y_2)\}, \end{aligned}$$

and $k(y_1, y_2) = 0$ otherwise.

3) It follows immediately of the splitting of (2) that

$$k_{Y_1}(y_1) = \begin{cases} \dfrac{1}{6}\, y_1^3 e^{-y_1} & \text{for } y_1 > 0, \\ \\ 0 & \text{otherwise,} \end{cases}$$

and

$$k_{Y_2}(y_2) = \begin{cases} 6 y_2 (1-y_2) & \text{for } 0 < y_2 < 1, \\ \\ 0 & \text{otherwise,} \end{cases}$$

hence $Y_1 \in \Gamma(4,1)$ is gamma distributed, and $Y_2 \in Be(2,2)$ is beta distributed.

4) It follows from

$$k(y_1, y_2) = k_{Y_1}(y_1) \cdot k_{Y_2}(y_2),$$

that Y_1 and Y_2 are independent.

5) By symmetry, the median is $\langle Y_2 \rangle = \dfrac{1}{2}$, and the mean is

$$E\{Y_2\} = \frac{2}{2+2} = \frac{1}{2} = \langle Y_2 \rangle .$$

Alternatively,

$$E\{Y_2\} = 6 \int_0^1 \left(y_2^2 - y_2^3 \right) dy_2 = 6 \left(\frac{1}{3} - \frac{1}{4} \right) = \frac{6}{12} = \frac{1}{2}.$$

Example 4.15 *Let X_1 and X_2 be independent random variables of the frequencies*

$$f_{X_1}(x_1) = \begin{cases} e^{-x_1}, & x_1 > 0, \\ \\ 0, & x_1 \le 0, \end{cases}$$

and

$$f_{X_2}(x_2) = \begin{cases} x_2 e^{-x_2}, & x_2 > 0, \\ \\ 0, & x_2 \le 0. \end{cases}$$

1. *Find the means $E\{X_1\}$, $E\{X_2\}$ and $E\left\{\frac{1}{X_2}\right\}$.*

We introduce the random variables Y_1 and Y_2 by

$$Y_1 = X_1 + X_2, \qquad Y_2 = \frac{X_1}{X_2}.$$

The vector function τ, given by

$$\tau(x_1, x_2) = \left(x_1 + x_2, \frac{x_1}{x_2} \right),$$

maps $\mathbb{R}_+ \times \mathbb{R}_+$ bijectively onto itself.

2. *Find the simultaneous frequency $k(y_1, y_2)$ of (Y_1, Y_2).*

3. *Find the marginal frequencies of Y_1 and Y_2.*
(The question can be answered with or without use of the result of (2).)

4. *Check if Y_1 and Y_2 are independent.*

5. *Find the mean $E\{Y_2\}$.*

6. *Find the median of Y_2.*

1. Since $X_1 \in \Gamma(1,1)$ and $X_2 \in \Gamma(2,1)$, it follows immediately that

$$E\{X_1\} = 1 \qquad \text{and} \qquad E\{X_2\} = 2.$$

Alternatively,

$$E\{X_1\} = \int_0^\infty x_1 e^{-x_1}\, dx_1 = 1 \quad \text{and} \quad E\{X_2\} = \int_0^\infty x_2^2 e^{-x_2}\, dx_2 = 2! = 2.$$

Finally,

$$E\left\{\frac{1}{X_2}\right\} = \int_0^\infty \frac{1}{x_2} \cdot x_2\, e^{-x_2}\, dx_2 = \int_0^\infty e^{-x_2}\, dx_2 = 1.$$

2. We solve the equations $y_1 = x_1 + x_2$ and $y_2 = \dfrac{x_1}{x_2}$, i.e. $x_1 = x_2 y_2$, with respect to (x_1, x_2). By insertion,

$$y_1 = x_2 y_2 + x_2 = x_2 (1 + y_2), \qquad \text{i.e. } x_2 = \frac{y_1}{1 + y_2},$$

hence

$$x_1 = \frac{y_1 y_2}{1 + y_2} = y_1 - \frac{y_1}{1 + y_2} \qquad \text{and} \qquad x_2 = \frac{y_1}{1 + y_2}.$$

The solution is unique, because $y_2 \neq -1$. From $(y_1, y_2) \in \mathbb{R}_+ \times \mathbb{R}_+$ follows that $(x_1, x_2) \in \mathbb{R}_+ \times \mathbb{R}_+$, and vice versa, so τ maps the domain $\mathbb{R}_+ \times \mathbb{R}_+$ bijectively onto itself.

The Jacobian is

$$\frac{\partial (x_1, x_2)}{\partial (y_1, y_2)} = \begin{vmatrix} \dfrac{y_2}{1 + y_2} & \dfrac{y_2}{(1 + y_2)^2} \\[2ex] \dfrac{1}{1 + y_2} & -\dfrac{y_1}{(1 + y_2)^2} \end{vmatrix} = -\frac{1}{(1 + y_2)^3} (y_2 y_1 + y_1) = -\frac{y_1}{(1 + y_2)^2} < 0.$$

Now, X_1 and X_2 are independent, so the simultaneous frequency of (X_1, X_2) is given by

$$f(x_1, x_2) = \begin{cases} x_2\, e^{-(x_1+x_2)}, & \text{for } x_1 > 0 \text{ and } x_2 > 0, \\[2mm] 0, & \text{otherwise.} \end{cases}$$

Thus the simultaneous frequency of (Y_1, Y_2) is

$$k(y_1, y_2) = \frac{y_1}{1+y_2}\, e^{-y_1} \cdot \frac{y_1}{(1+y_2)^2} = \frac{1}{2}\, y_1^2\, e^{-y_1} \cdot \frac{2}{(1+y_2)^3} \qquad \text{for } y_1 > 0 \text{ and } y_2 > 0,$$

and $k(y_1, y_2) = 0$ otherwise.

3. & 4. It follows from the splitting of (2) that

$$k_{Y_1}(y_1) = \begin{cases} \dfrac{1}{2}\, y_1^2 e^{-y_1} & \text{for } y_1 > 0, \\[2mm] 0, & \text{otherwise,} \end{cases}$$

and

$$k_{Y_2}(y_2) = \begin{cases} \dfrac{2}{(1+y_2)^3} & \text{for } y_2 > 0, \\[2mm] 0 & \text{otherwise.} \end{cases}$$

It follows that Y_1 and Y_2 are independent.
Notice that $Y_1 \in \Gamma(3, 1)$, and also by some elaboration that $2Y_2 \in F(2, 4)$.

3. Alternatively it follows for $y_1 > 0$ that

$$k_{Y_1}(y_1) = \int_0^{y_1} e^{-(y_1-t)} \cdot t \cdot e^{-t}\, dt = e^{-y_1} \int_0^{y_1} t\, dt = \frac{1}{2}\, y_1^2 e^{-y_1},$$

and $k_{Y_1}(y_1) = 0$ for $y_1 \le 0$.
If $y_2 > 0$, we get by a formula that

$$k_{Y_2}(y_2) = \int_0^{\infty} e^{-y_2 t} \cdot t\, e^{-y} \cdot |t|\, dt = \int_0^{\infty} t^2 e^{-(1+y_2)t}\, dt = \frac{2!}{(1+y_2)^3} = \frac{2}{(1+y_2)^3},$$

and $k_{Y_2}(y_2) = 0$ for $y_2 \le 0$.

5. Since X_1 and X_2 are independent, the mean is

$$E\{Y_2\} = E\left\{\frac{X_1}{X_2}\right\} = E\{X_1\} \cdot E\left\{\frac{1}{X_2}\right\} = 1 \cdot 1 = 1.$$

Alternatively,

$$\begin{aligned} E\{Y_2\} &= \int_0^{\infty} \frac{2y_2}{(1+y_2)^3}\, dy_2 = 2\int_0^{\infty} \left\{ \frac{1}{(1+y_2)^2} - \frac{1}{(1+y_2)^3} \right\} dy_2 \\ &= 2\left[-\frac{1}{1+y_2} + \frac{1}{2}\, \frac{1}{(1+y_2)^2} \right]_0^{\infty} = 2\left(1 - \frac{1}{2}\right) = 1. \end{aligned}$$

ALTERNATIVELY, $2Y_2 \in F(2,4)$, so

$$E\{2Y_2\} = \frac{n_2}{n_2-2} = \frac{4}{2} = 2, \qquad \text{hence} \qquad E\{Y_2\} = 1.$$

6. The distribution function of $y_2 > 0$ is given by

$$F_{Y_2}(y_2) = \int_0^{y_2} \frac{2}{(1+t)^3}\,dt = \left[-\frac{1}{(1+t)^2}\right]_0^{y_2} = 1 - \frac{1}{(1+y_2)^2},$$

so the median is determined by

$$\frac{1}{(1+y_2)^2} = \frac{1}{2},$$

hence

$$\langle Y_2 \rangle = \sqrt{2} - 1.$$

5 Functions of random variables, in general

Example 5.1 *Let X_1 and X_2 be random variables, and let $Y_1 = \varphi_1(X_1)$ and $Y_2 = \varphi_2(X_2)$.*

1) Assume that X_1 and X_2 are independent. Is it possible to conclude that Y_1 and Y_2 are independent?

2) Assume that X_1 and X_2 are dependent. Does it follow that Y_1 and Y_2 are dependent?

1) The answer is 'yes'. Let $\varphi^{-1}(C) = \{t \mid \varphi(t) \in C\}$ denote the inverse set map. Then

$$\begin{aligned} &P\{\varphi_1(X_1) \in A \wedge \varphi_2(X_2) \in B\} \\ &\qquad = P\left\{X_1 \in \varphi_1^{-1}(A) \wedge X_2 \in \varphi_2^{-1}(B)\right\} \\ &\qquad = P\left\{X_1 \in \varphi_1^{-1}(A)\right\} \cdot P\left\{X_2 \in \varphi_2^{-1}(B)\right) , \qquad \text{because } X_1,\, X_2 \text{ are independent} \\ &\qquad = P\{\varphi_1(X_1) \in A\} \cdot P\{\varphi_2(X_2) \in B\}, \end{aligned}$$

and we conclude that $\varphi_1(X_1)$ and $\varphi_2(X_2)$ are stochastically independent.

2) The answer is 'no'! It suffices to give an example. Let

$$\varphi_1(X_1) = c_1 \qquad \text{and} \qquad \varphi_2(X_2) = c_2$$

be constant maps. Then

$$P\{\varphi_1(X_1) \in A \wedge \varphi_2(X_2) \in B\} = \left\{ \begin{array}{ll} 1, & \text{if } c_1 \in A \text{ and } c_2 \in B, \\ & \\ 0, & \text{otherwise,} \end{array} \right\}$$

$$= P\{c_1 \in A\} \cdot P\{c_2 \in B\} = P\{\varphi_1(X_1) \in A\} \cdot P\{\varphi_2(X_2) \in B\},$$

proving that $\varphi_1(X_1)$ and $\varphi(X_2)$ are independent, no matter if X_1 and X_2 are independent or not.

Example 5.2 *A discrete random variable* (X, Y) *has its distribution given by the following table*

$Y \setminus X$	1	2	3
1	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
2	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{12}$
3	$\frac{1}{12}$	$\frac{1}{12}$	0

Find the marginal distributions of X *and* Y.
Compute $P\{X \cdot Y$ *is even*$\}$.
Compute $P\{X \geq Y\}$.
Are X *and* Y *independent?*

1) All probabilities are ≥ 0, and their sum is 1, so the table describes a distribution.

$Y \setminus X$	1	2	3	f_Y
1	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{3}$
2	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{2}$
3	$\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{1}{6}$
f_X	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$	1

We conclude from the table that the marginal distributions are

$$P\{X = 1\} = \frac{1}{3}, \qquad P\{X = 2\} = \frac{1}{2}, \qquad P\{X = 3\} = \frac{1}{6},$$

and

$$P\{Y = 1\} = \frac{1}{3}, \qquad P\{Y = 2\} = \frac{1}{2}, \qquad P\{Y = 3\} = \frac{1}{6}.$$

2) By a counting of the table we get

$$\begin{aligned} P\{X \cdot Y \text{ is even}\} &= P\{X = 2\} + P\{X = 1 \wedge Y = 2\} + P\{X = 3 \wedge Y = 2\} \\ &= \frac{1}{2} + \frac{1}{6} + \frac{1}{2} = \frac{3}{4}. \end{aligned}$$

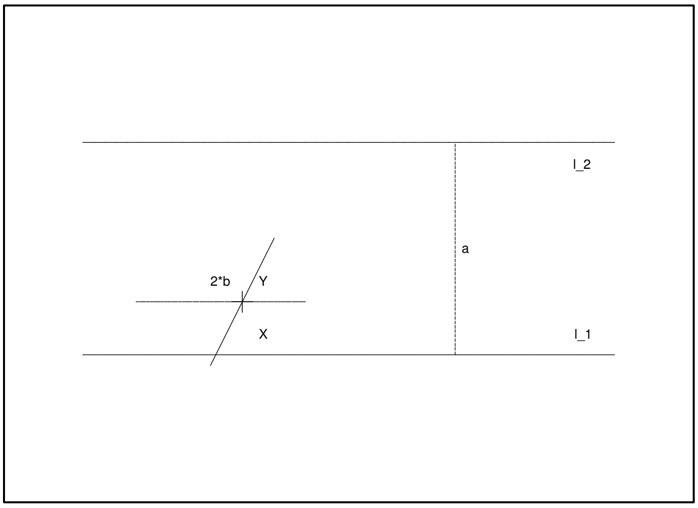
3) We get in the same way

$$\begin{aligned} P\{X \geq Y\} &= P\{Y = 1\} + P\{Y = 2 \wedge X = 2\} \\ &\qquad + P\{Y = 2 \wedge X = 3\} + P\{Y = 3 \wedge X = 3\} \\ &= \frac{1}{3} + \frac{1}{4} + \frac{1}{12} + 0 = \frac{8}{12} = \frac{2}{3}. \end{aligned}$$

4) The random variables X and Y are *not* independent. We have e.g.

$$P\{X = 3 \wedge Y = 3\} = 0 \quad \text{and} \quad P\{X = 3\} \cdot P\{Y = 3\} = \frac{1}{36} \neq 0.$$

Example 5.3 *In a plane we draw two parallel lines ℓ_1 and ℓ_2 of the distance a. A needle of length $2b$, where $b < a$, is thrown such that it falls randomly between the two lines in the following sense:*



The midpoint of the needle has the distance X from ℓ_1, where X is rectangularly distributed over $]0,a[$, and the needle forms an angle Y with the two parallel lines, where Y is rectangularly distributed over $]0,\pi[$. Finally, X and Y are independent.

1) Find a condition – expressed by X, Y, b – which describes that the needle intersects the line ℓ_1.

2) Prove that the probability that the needle intersects ℓ_1 is $\frac{2b}{a} \cdot \frac{1}{\pi}$.

Remark 5.1 This example is called Buffon's needle problem, and it can be traced back to 1777, when Georges-Louis Leclerc, count of Buffon, published a paper on the subject. If a needle is thrown at random many times, then the fraction when the needle intersects ℓ_1, is approximately equal to $\frac{2b}{a} \cdot \frac{1}{\pi}$. If this fraction is denoted by f, then we have the approximation $\pi \approx \frac{2b}{a} \cdot \frac{1}{f}$. Since then many people have tried to find π in this way. In 1850 the astronomer Wolfe threw a needle 5000 times with the values $a = 45$ mm, $b = 36$ mm. He obtained intersection in 2532 of the cases, hence $f = \frac{25332}{5000}$. This gives the approximation 3.160 of π, which is quite fair. In 1901 Lazzarini published a paper with a far better approximation of π. He used $a = 3$ cm, $b = 2,5$ cm, the needle was thrown 3408 times, and he obtained intersection 1808 times. In this case we get the approximation $\pi \approx \frac{5}{3} \cdot \frac{3408}{1808} = 3.1415919\dots$, which is astonishingly in agreement with $\pi = 3.1415926\dots$. However, Lazzarini was extremely lucky in getting his paper published. Some mathematicians have later pointed out the fairly strange number 3408 of throws, and they also noted that Lazzarini's fraction can be reduced to $\frac{355}{113}$, which long has been known as one of the very best rational approximations of π. Therefore, mathematicians of today are convinced that the paper was a swindle.

1) It follows by the geometry that the needle intersects ℓ_1, if $X \le b \cdot \sin Y$.

2) We shall find $P\{X \le b \cdot \sin Y\}$. Since X is rectangularly distributed over $]0, a[$, and Y is rectangularly distributed over $]0, \pi[$, we get

$$f_X(x) = \begin{cases} \dfrac{1}{a}, & \text{for } x \in\,]0, a[, \\[2mm] 0, & \text{otherwise,} \end{cases} \qquad f_Y(y) = \begin{cases} \dfrac{1}{\pi}, & \text{for } y \in\,]0, \pi[, \\[2mm] 0, & \text{otherwise.} \end{cases}$$

Then

$$P\{b\,\sin Y > X\} = \frac{1}{\pi\,a}\int_{y=0}^{\pi}\left\{\int_{x=0}^{b\,\sin y} 1\,dx\right\}dy = \frac{1}{\pi\,a}\int_{y=0}^{\pi} b\,\sin y\,dy = \frac{2b}{a}\cdot\frac{1}{\pi}.$$

Alternatively,

$$P\{X \le b \cdot \sin Y\} = P\{X - b \cdot \sin Y \le 0\},$$

so we can instead find the distribution function of $Z = X - b \cdot \sin Y$.

We shall, however, first find the distribution function $G(y)$ of $-b \cdot \sin Y$. This is given by

$$G(y) = P\{-b \cdot \sin Y \leq y\} = P\left\{\sin Y \geq -\frac{y}{b}\right\}.$$

If $y \geq 0$, then $G(y) = 1$, and if $y \leq -b$, then $G(y) = 0$. Finally, if $y \in]-b, 0[$, then

$$\begin{aligned} G(y) &= P\left\{\operatorname{Arcsin}\left(-\frac{y}{b}\right) \leq Y \leq \pi - \operatorname{Arcsin}\left(-\frac{y}{b}\right)\right\} \\ &= \frac{1}{\pi}\left\{\pi - 2\operatorname{Arcsin}\left(-\frac{y}{b}\right)\right\} = 1 + \frac{2}{\pi}\operatorname{Arcsin}\left(\frac{y}{b}\right), \end{aligned}$$

hence the frequency is

$$g(y) = G'(y) = \frac{2}{\pi} \cdot \frac{1}{\sqrt{1 - \left\{\frac{y}{b}\right\}^2}} \cdot \frac{1}{b} = \frac{2}{\pi b} \frac{1}{\sqrt{1 - \left\{\frac{y}{b}\right\}^2}}, \quad y \in]-b, 0[,$$

and $= 0$ otherwise.

Since X and Y, and hence also X and $-b \sin Y$ are independent, we conclude that $Z = X - b \sin Y$ has the frequency

$$h(s) = \int_{-\infty}^{\infty} f_X(s - x)\, g(x)\, dx, \qquad s \in \mathbb{R}.$$

Thus, if $b < a$,

$$\begin{aligned} P\{X \leq b \cdot \sin Y\} &= P\{Z \leq 0\} = \int_{-\infty}^{0} h(s)\, ds = \int_{-\infty}^{0} \left\{\int_{-\infty}^{\infty} f_X(s - x)\, g(x)\, dx\right\} ds \\ &= \int_{-\infty}^{\infty} g(x) \left\{\int_{-\infty}^{0} f_X(s - x)\, ds\right\} dx = \int_{-\infty}^{\infty} g(x) \left\{\int_{-\infty}^{-x} f_X(s)\, ds\right\} dx \\ &= \frac{2}{\pi b} \int_{-b}^{0} \frac{1}{\sqrt{1 - \left\{\frac{x}{b}\right\}^2}} \left\{\int_{-\infty}^{-x} \frac{1}{a} \chi_{[0,a]}(s)\, ds\right\} dx \\ &= \frac{2}{\pi a b} \int_{0}^{b} \frac{1}{\sqrt{1 - \left\{\frac{x}{b}\right\}^2}} \left\{\int_{0}^{x} \chi_{[0,a]}(s)\, ds\right\} dx, \end{aligned}$$

where $\chi_{[0,a]}(s) = 1$, if $s \in [0, a]$, and $= 0$ otherwise. Now,

$$\int_{0}^{a} \chi_{[0,a]}(s)\, ds = \begin{cases} a & \text{for } x \geq a, \\ x, & \text{for } x \in [0, a], \\ 0, & \text{for } x < 0, \end{cases}$$

so we get for $b < a$,

$$\begin{aligned} P\{X \leq b \cdot \sin Y\} &= \frac{2}{\pi a b} \int_{0}^{b} \frac{x}{\sqrt{1 - \left\{\frac{x}{b}\right\}^2}}\, dx = \frac{2b^2}{\pi a b} \int_{0}^{1} \frac{y}{\sqrt{1 - y^2}}\, dy \\ &= \frac{2b}{\pi a} \left[-\sqrt{1 - y^2}\right]_{0}^{1} = \frac{2b}{a} \cdot \frac{1}{\pi}, \end{aligned}$$

which is the searched result.

Remark 5.2 If the needle is thrown a great number of times, then the relative frequency f that it intersects ℓ_1 will approximately be $\frac{2b}{a} \cdot \frac{1}{\pi}$, so we conclude that

$$\pi \sim \frac{2b}{a} \cdot \frac{1}{f}.$$

This formula has earlier been used in the attempt of experimentally to find π, however, without great success. The results have either been too poor, or one has cheated (like e.g. Lazzarini). $\diamond$

Remark 5.3 One can also go through this example *without* the assumption that $b < a$; but in this case the computations become really tough, because the curve $x = b \cdot \sin y$ then intersects the curve $x = a$. $\diamond$

6 Inequalities between two random variables

Example 6.1 *Two persons A and B have the intension of meeting between 8 AM and 9 AM. Both A and B arrive at the meeting place at a randomly chosen time between 8 AM and 9 AM. Furthermore, they have agreed that none of them will wait in more than 10 minutes. Find the probability that they meet.*
If instead, A and B have agreed that A will wait 15 minutes for B, while B will wait 5 minutes for A, what is then the probability that they meet?
HINT. *The arrival times of A and B are rectangularly distributed.*

Let X be the arrival time of A, and let Y be the arrival time of B. Then X and Y are independent random variables, which are both rectangularly distributed over an interval of length 1 hour, represented by the interval $]0,1[$.

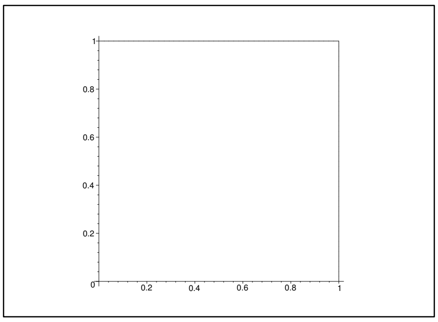


Figure 28: The domain where the simultaneous frequency is 1.

The simultaneous frequency is

$$f(x,y) = \begin{cases} 1, & \text{for } (x,y) \in]0,1[\times]0,1[, \\ 0, & \text{otherwise.} \end{cases}$$

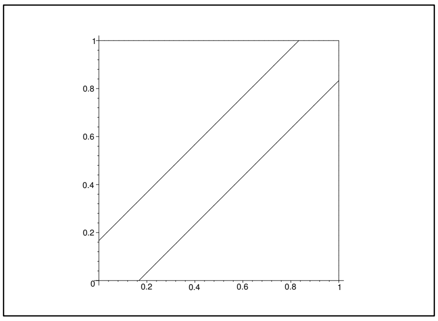


Figure 29: The domain C is the diagonal strip.

1) Since 10 minutes $= \dfrac{1}{6}$ hour, the event corresponds to $|X-Y| < \dfrac{1}{6}$. The probability is equal to the area of C, hence

$$P\left\{|X-Y| < \frac{1}{6}\right\} = 1 - 2 \cdot \frac{1}{2}\left(\frac{5}{6}\right)^2 = \frac{11}{36} = 0.306.$$

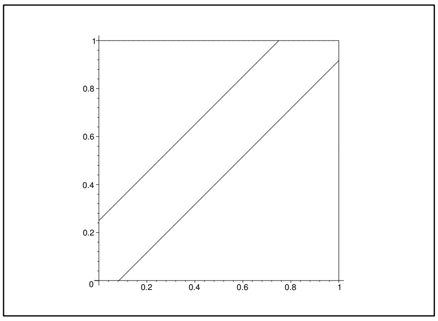


Figure 30: The domain D is the translated diagonal strip.

2) The event corresponds to $-\dfrac{1}{12} < Y - X < \dfrac{1}{4}$. The probability is equal to the area of D on the figure,

$$1 - \frac{1}{2}\left(\frac{3}{4}\right)^2 - \frac{1}{2}\left(\frac{11}{12}\right)^2 = 1 - \frac{101}{144} = \frac{43}{144} = 0.299.$$

Alternatively, (1) is solved in the following way:

$$\begin{aligned}
&P\{\text{meeting}\}\\
&\quad = P\{A \text{ arrives } \textit{first} \text{ between } 8^{00} \text{ and } 8^{50}, \text{ and } B \text{ at most 10 min. later}\}\\
&\qquad + P\{B \text{ arrives } \textit{first} \text{ between } 8^{00} \text{ and } 8^{50}, \text{ and } A \text{ at most 10 min. later}\}\\
&\qquad + P\{A \text{ and } B \text{ both arrive between } 8^{50} \text{ and } 9^{00}\}\\
&\quad = \frac{5}{6}\cdot\frac{1}{6} + \frac{5}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{11}{36}.
\end{aligned}$$

Alternatively, we find the frequency $h(z)$ of $Z = X - Y$. Since $f(x,y) = 1_{[0,1]^2}(x,y)$, we get

$$h(z) = \int_{-\infty}^{\infty} f(x, x-z)\,dx = \int_0^1 1_{[0,1]^2}(x, x-z)\,dx = \int_0^1 1_{[0,1]}(x-z)\,dx,$$

The integrand is only $\neq 0$, if $x \in]0,1[$ and $x - z \in]0,1[$, i.e. $x \in]z, z+1[$, thus for $z \in]-1,1[$, cf. the figure.

(i) For $z \in]-1,0[$ fås $h(z) = \int_0^{z+1} dx = z+1$.

(ii) For $z \in]0,1[$ fås $h(z) = \int_z^1 dx = 1 - z$.

(iii) If $z \notin]-1,0[$, then $h(z) = 0$.

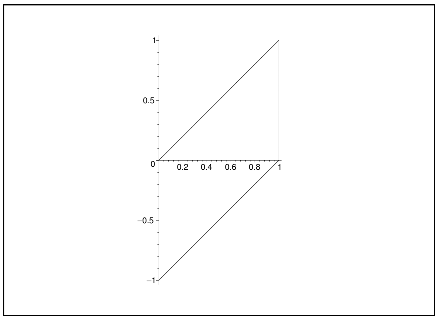


Figure 31: The domain of integration for $h(z)$.

Then the task can be treated in the following way:

1)

$$
\begin{aligned}
P\left\{|X-Y|<\frac{1}{6}\right\} &= P\left\{-\frac{1}{6}<Z<\frac{1}{6}\right\} = \int_{-\frac{1}{6}}^{\frac{1}{6}} h(z)\,dz \\
&= \int_{-\frac{1}{6}}^{0}(z+1)\,dz + \int_0^{\frac{1}{6}}(1-z)\,dz = \left[\frac{1}{2}(z+1)^2\right]_{-\frac{1}{6}}^{0} - \left[\frac{1}{2}(1-z)^2\right]_0^{\frac{1}{6}} \\
&= \frac{1}{2}\left\{1-\left(\frac{5}{6}\right)^2\right\} - \frac{1}{2}\left\{\left(\frac{5}{6}\right)^2-1\right\} = 1-\left(\frac{5}{6}\right)^2 = \frac{11}{36}.
\end{aligned}
$$

2)

$$
\begin{aligned}
P\left\{-\frac{1}{12}<Z<\frac{1}{4}\right\} &= \int_{-\frac{1}{12}}^{0}(z+1)\,dz + \int_0^{\frac{1}{4}}(1-z)\,dz = \left[\frac{1}{2}(z+1)^2\right]_{-\frac{1}{12}}^{0} - \left[\frac{1}{2}(1-z)^2\right]_0^{\frac{1}{4}} \\
&= \frac{1}{2}\left\{1-\left(\frac{11}{12}\right)^2\right\} + \frac{1}{2}\left\{1-\left(\frac{3}{4}\right)^2\right\} = 1-\frac{1}{2}\left\{\left(\frac{11}{12}\right)^2+\left(\frac{3}{4}\right)^2\right\} = \frac{43}{144}.
\end{aligned}
$$

Example 6.2 *Henry and John arrive independently of each other to a meeting point. Both Henry and John arrive at randomly chosen times between 8 AM and 9 AM.*

1) Find the probability that Henry arrives as the first one.

2) Find the probability that John arrives more than 10 minutes after Henry.

3) Find the probability that the difference between their arrival times is at most 5 minutes.

HINT. *The arrival times of Henry and John are rectangularly distributed.*

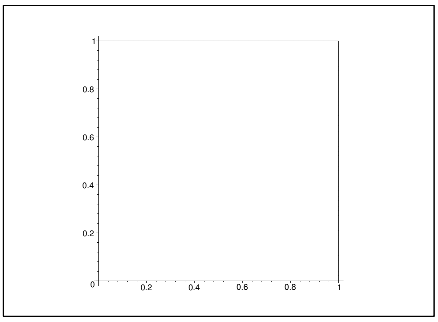


Figure 32: The domain where the simultaneous distribution function is 1.

Let the random variable X denote Henry's arrival time, and let the random variable Y denote John's arrival time. Since X and Y are independent and rectangularly distributed over e.g. $]0,1[$, the simultaneous frequency is

$$f(x,y) = \begin{cases} 1, & (x,y) \in]0,1[\times]0,1[, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\int_A f(x,y) f(x,y)\,dx\,dy = \text{ area}(A), \qquad \text{for } A \subseteqq]0,1[\times]0,1[.$$

1) By an area consideration we get

$$P\{\text{Henry arrives first}\} = P\{X < Y\} = \text{ the area of the upper triangle } = \frac{1}{2}.$$

2) Since 10 minutes $= \dfrac{1}{6}$ hour, we get in the same way

$$P\{\text{John arrives more that 10 min. after Henry}\}$$
$$= P\left\{Y > X + \frac{1}{6}\right\} = \text{ area of the upper triangle } = \frac{1}{2}\left(\frac{5}{6}\right)^2 = \frac{25}{72}.$$

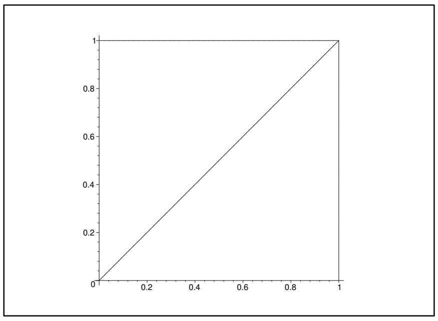


Figure 33: The domain given by $X < Y$ is the upper triangle.

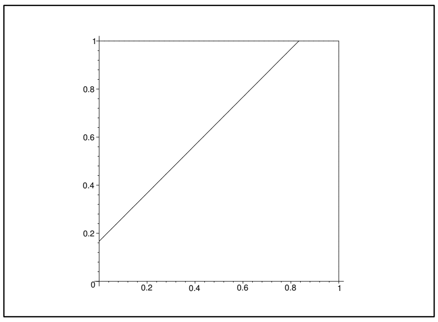


Figure 34: The domain given by $Y > X + \dfrac{1}{6}$ is the upper triangle.

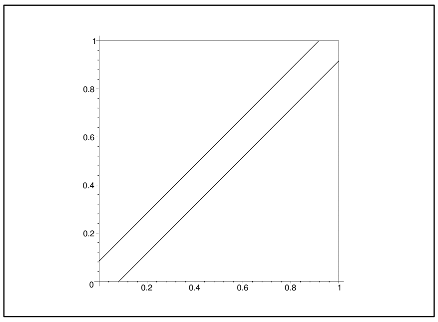


Figure 35: The domain where the difference is at most 5 minutes is represented by the domain around the diagonal.

3) Since 5 minutes $= \dfrac{1}{12}$ hour, we get the condition $-\dfrac{1}{12} < X - Y < \dfrac{1}{12}$, and the probability is again obtained by an area consideration,

$$
\begin{aligned}
&P\{\text{the difference is at most 5 min.}\} = P\left\{|X-Y| < \frac{1}{12}\right\} \\
&\qquad = \text{ area of the diagonal strip } = 1 - 2 \cdot \frac{1}{2}\left(\frac{11}{12}\right)^2 = \frac{23}{144}.
\end{aligned}
$$

Example 6.3 *Two persons A and B arrive at a meeting point between 7 AM and 8 AM. Their arrivals are independent of each other, and they both have a tendency of arriving at the end of the interval which for convenience is put equal to $]0,1[$. (We adjust the time at 7 AM).*
The arrival time of A is denoted by X, and we assume that its frequency is

$$f(x)=\begin{cases} 2x, & 0<x<1,\\ 0, & otherwise,\end{cases}$$

while the arrival time of B is denoted by Y, and is also assumed to have the frequency

$$g(y)=\begin{cases} 2y, & 0<x<1,\\ 0, & otherwise,\end{cases}$$

A will at most wait 20 minutes for B, while B is a very impatient person who does not want to wait at all.
Find the probability that the two persons meet.

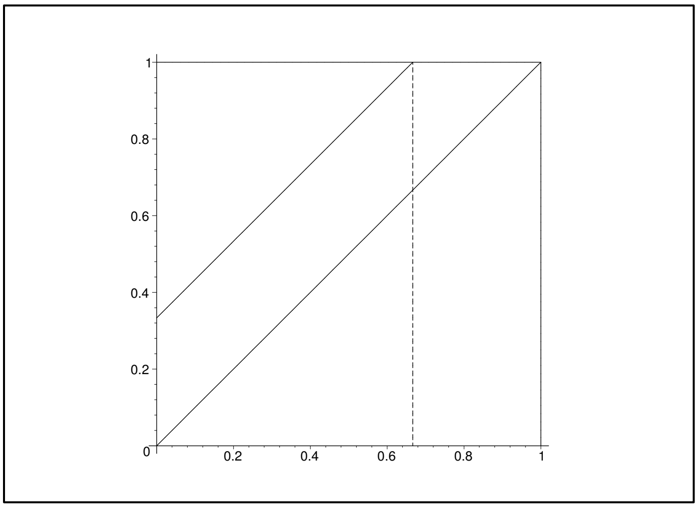


Figure 36: The domain of integration is the diagonal strip.

Since X and Y are independent, the frequency of the 2-dimensional random variable (X,Y) is given by

$$h(x,y)=f(x)g(y)=\begin{cases} 4xy, & 0<x,\, y<1,\\ \\ 0, & \text{otherwise,}\end{cases}$$

Since 20 minutes $=\dfrac{1}{3}$ hour, the task is to find $P\left\{X\le Y\le X+\dfrac{1}{3}\right\}$, i.e. the integral of $h(x,y)$ over the diagonal strip on the figure.

The domain of integration is split at $x=\dfrac{2}{3}$. Then by first integrating vertically (the inner integral, so x is kept fixed),

$$
\begin{aligned}
P\left\{X \le Y \le x+\frac{1}{3}\right\} &= \int_0^{\frac{2}{3}} \left\{\int_x^{x+\frac{1}{3}} 4xy\,dy\right\} dx + \int_{\frac{2}{3}}^1 \left\{\int_x^1 4xy\,dy\right\} dx \\
&= \int_0^{\frac{2}{3}} 2x\left\{\left(x+\frac{1}{3}\right)^2 - x^2\right\} dx + \int_{\frac{2}{3}}^1 \left\{2x-2x^3\right\} dx \\
&= \int_0^{\frac{2}{3}} \left\{\frac{4}{3}\,x^2+\frac{2}{9}\,x\right\} dx + \left[x^2-\frac{1}{2}\,x^4\right]_{\frac{2}{3}}^1 = \left[\frac{4}{9}\,x^3+\frac{1}{9}\,x^2\right]_0^{\frac{2}{3}} + \left(1-\frac{1}{2}\right) - \left(\frac{4}{9}-\frac{8}{81}\right) \\
&= \frac{4}{9}\cdot\frac{8}{27}+\frac{1}{9}\cdot\frac{4}{9}+\frac{1}{2}-\frac{36-8}{81} = \frac{32}{243}+\frac{12}{243}+\frac{1}{2}-\frac{28}{81} \\
&= \frac{44}{243}+\frac{1}{2}-\frac{84}{243} = \frac{1}{2}-\frac{40}{243} = \frac{243-80}{486} = \frac{163}{486} \approx 0,335.
\end{aligned}
$$

Example 6.4 *According to their schedules, 2 trains A and B shall arrive to a station at the same time on each their line. Train A shall stay at the station for 5 minutes, and train B for 4 minutes. However, the trains are very often delayed up to 20 minutes, so we assume that the arrival time of train A is rectangularly distributed over $[0, 20]$ (measured in minutes), and the arrival time of train B is also rectangularly distributed over $[0, 20]$. The delay time is counted from the planned arrival time.*

1) Find the probability that train A arrives before train B.

2) Find the probability that the two trains meet at the station.

3) Find the probability that train A arrives before train B and departs after train B.

Since the arrival time X of train A and the arrival time Y of train B are independent and both rectangularly distributed over $[0, 20]$, the simultaneous frequency of (X, Y) is

$$
f(x,y) = \begin{cases} \dfrac{1}{400}, & \text{for } 0 < x,\ y < 20, \\[2mm] 0, & \text{otherwise.} \end{cases}
$$

1) It follows by an area consideration of weight $\dfrac{1}{400}$ that

$$P\{X < Y\} = \frac{1}{2}.$$

2) If the two trains meet at the station, then $X-4 < Y < X+5$, so (X, Y) lies in the diagonal strip. Then by an area consideration,

$$
\begin{aligned}
P\{X-4<Y<X+5\} &= \frac{1}{400}\left\{20^2-\frac{1}{2}\cdot 15^2-\frac{1}{2}\cdot 16^2\right\} = \frac{1}{400}\left\{400-\frac{225}{2}-\frac{256}{2}\right\} \\
&= 1-\frac{481}{800} = \frac{319}{800}.
\end{aligned}
$$

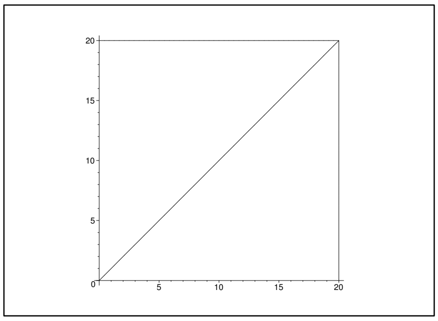


Figure 37: The event $X < Y$ is represented by the upper triangle.

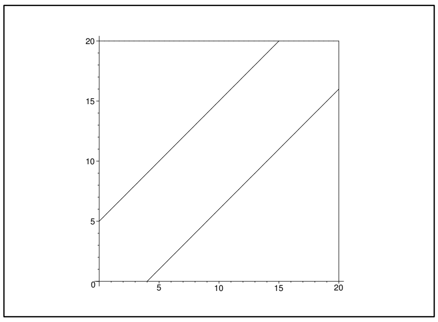


Figure 38: The event $X - 4 < Y < X + 5$ is represented by the diagonal strip.

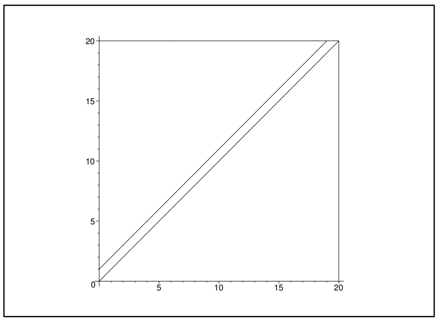


Figure 39: The event $X < Y < X + 1$ is represented by the diagonal strip.

3) If A arrives before B, i.e. $X < Y$, and departs after B, i.e. $X + 5 > Y + 4$, then $X < Y < X + 1$. The probability can again be found by an area consideration,

$$
\begin{aligned}
P\{X < Y < X+1\} &= \frac{1}{2} - \frac{1}{2} \cdot \left(\frac{19}{20}\right)^2 = \frac{1}{2}\left\{1^2 - \left(\frac{19}{20}\right)^2\right\} \\
&= \frac{1}{2}\left(1 + \frac{19}{20}\right)\left(1 - \frac{19}{20}\right) = \frac{1}{2} \cdot \frac{39}{20}\frac{1}{20} = \frac{39}{800}.
\end{aligned}
$$

Example 6.5 *Henry and Peter throw dices. Every minute* ($t = 1, 2, 3, \dots$) *Henry throws a dice (probability* $\frac{1}{6}$ *of getting a six), and Peter also throws another dice every minute. We define the random variables* X *and* Y *by*

$X = k$, *if Henry obtains his first six in throw number* k,

$Y = k$, *if Peter obtains his first six in throw number* k.

1. *Find* $P\{X = k\}$, $k \in \mathbb{N}$, *and find the mean* $E\{X\}$.

2. *Find for every* $k \in \mathbb{N}$ *the probability* $P\{X = k \wedge Y = k\}$, *and then find* $P\{X = Y\}$.

3, *Compute the two probabilities* $P\{X < Y\}$ *and* $P\{Y < X\}$.

4. *We define a random variable* Z *by*

$Z = k$, *if Henry obtains his second six in throw number* k.

Find $P\{Z = k\}$, $k = 2, 3, 4, \dots$.

5. *Find for* $k = 2, 3, 4, \dots$,

$$P\{Z = k \wedge Y > k\},$$

and then find the probability that Henry gets at least two sixes before Peter obtains his first six.

1) Since X (and also Y) is geometric distributed with $p = \dfrac{1}{6}$, we get

$$P\{X = k\} = \frac{1}{6}\left(\frac{5}{6}\right)^{k-1}, \quad k \in \mathbb{N}, \quad \text{and} \quad E\{X\} = 6.$$

2) Since X and Y are independent, we get for $k \in \mathbb{N}$ that

$$P\{X = k \wedge Y = k\} = P\{X = k\} \cdot P\{Y = k\} = \left\{\frac{1}{6}\left(\frac{5}{6}\right)^{k-1}\right\}^2 = \frac{1}{36}\left(\frac{25}{36}\right)^{k-1}.$$

Then by a summation,

$$P\{X = Y\} = \sum_{k=1}^{\infty} P\{X = k \wedge Y = k\} = \frac{1}{36}\sum_{k=1}^{\infty}\left(\frac{25}{36}\right)^{k-1} = \frac{1}{36} \cdot \frac{1}{1 - \dfrac{25}{36}} = \frac{1}{11}.$$

3) Clearly,

$$P\{X < Y\} + P\{Y < X\} + P\{X = Y\} = 1.$$

It follows from the symmetry that $P\{X < Y\} = P\{Y < X\}$, so

$$P\{X < Y\} = P\{Y < X\} = \frac{1}{2}(1 - P\{X = Y\}) = \frac{1}{2}\left(1 - \frac{1}{11}\right) = \frac{5}{11}.$$

ALTERNATIVELY,

$$\begin{aligned} P\{X<Y\} &= \sum_{k=1}^{\infty} P\{X=k \wedge Y>k\} = \sum_{k=1}^{\infty} P\{X=k\}\cdot P\{Y>k\} \\ &= \sum_{k=1}^{\infty} \frac{1}{6}\left(\frac{5}{6}\right)^{k-1}\cdot\left(\frac{5}{6}\right)^{k} = \frac{1}{6}\cdot\frac{5}{6}\sum_{k=1}^{\infty}\left(\frac{25}{36}\right)^{k-1} = \frac{5}{36}\cdot\frac{36}{11} = \frac{5}{11}. \end{aligned}$$

4) The random variable Z can be written $Z = X_1 + X_2$, where X_1 and X_2 are independent of the same distribution as X. Then we get for $k \geq 2$,

$$\begin{aligned} P\{Z=k\} &= \sum_{i=1}^{k-1} P\left\{X_1=i\right\}\cdot P\left\{X_k=k-i\right\} = \sum_{i=1}^{k-1} \frac{1}{6}\left(\frac{5}{6}\right)^{i-1}\cdot\frac{1}{6}\left(\frac{5}{6}\right)^{k-i-1} \\ &= \frac{1}{6}\sum_{i=1}^{k-1}\left(\frac{5}{6}\right)^{k-2} = (k-1)\cdot\frac{1}{36}\left(\frac{5}{6}\right)^{k-2}. \end{aligned}$$

ALTERNATIVELY, $Z = k$ means that in the first $k-1$ throws we have obtained precisely one six, and that we in the k-th throw get another six, hence

$$P\{Z=k\} = (k-1)\cdot\frac{1}{6}\left(\frac{5}{6}\right)^{k-2}\cdot\frac{1}{6} = (k-1)\cdot\frac{1}{36}\left(\frac{5}{6}\right)^{k-2}.$$

5) Here,

$$\begin{aligned} P\{Z<Y\} &= \sum_{k=2}^{\infty} P\{Z=k \wedge Y>k\} = \frac{1}{36}\sum_{k=2}^{\infty}(k-1)\left(\frac{25}{36}\right)^{k-2}\cdot\frac{25}{36} \\ &= \frac{1}{36}\cdot\frac{25}{36}\cdot\frac{1}{\left(\frac{11}{36}\right)^2} = \frac{25}{121}. \end{aligned}$$

7 Functions $Y = f(X)$ of random variables

Example 7.1 *Let X be rectangularly distributed over $]0,a[$, where $a > 0$. Find the distribution function and the frequency of the random variable $Y = X^2 + X$.*

The frequency of X is

$$f(x) = \begin{cases} \dfrac{1}{a}, & 0 < x < a, \\ \\ 0, & \text{otherwise.} \end{cases}$$

The function $y = \tau(x) = x^2 + x$ maps $]0,a[$ increasingly onto $\left]0, a+a^2\right[$. The inverse map is

$$x = \tau^{-1}(y) = -\frac{1}{2} + \sqrt{\frac{1}{4}+y}, \quad \text{where } \left(\tau^{-1}\right)'(y) = \frac{1}{2\sqrt{\dfrac{1}{4}+y}}.$$

Then the frequency of Y is

$$g(y) = \begin{cases} \dfrac{1}{a} \cdot \dfrac{1}{2\sqrt{\dfrac{1}{4}+y}} = \dfrac{1}{a} \cdot \dfrac{1}{\sqrt{1+4y}}, & 0 < y < a^2 + a, \\ \\ 0, & \text{otherwise,} \end{cases}$$

and the distribution function $G(y) = \int_{-\infty}^{y} g(u)\,du$ is

$$G(y) = \begin{cases} 0, & y \le 0, \\ \\ \dfrac{1}{a}\,\tau^{-1}(y) = \dfrac{1}{a}\left\{\sqrt{\dfrac{1}{4}+y} - \dfrac{1}{2}\right\}, & 0 < y < a^2 + a, \\ \\ 1, & y > a^2 + a. \end{cases}$$

Alternatively, we get for $y \in \left]0, a^2 + a\right[$,

$$G(y) = P\{Y \le y\} = P\left\{X \le -\frac{1}{2} + \sqrt{\frac{1}{4}+y}\right\} = \frac{1}{a}\left\{\sqrt{\frac{1}{4}+y} - \frac{1}{2}\right\},$$

hence in the same interval,

$$g(y) = G'(y) = \frac{1}{a} \cdot \frac{1}{2\sqrt{\dfrac{1}{4}+y}} = \frac{1}{a} \cdot \frac{1}{\sqrt{1+4y}}.$$

Example 7.2 *A line segment of length 1 is randomly divided into two pieces of the lengths X and $1-X$, where we assume that X is rectangularly distributed over the interval $]0,1[$. We form a rectangle of edge lengths X and $1-X$. Find the probability that the area of this rectangle is bigger than $\frac{1}{8}$.*

The area of the rectangle is $Y = X(1-X)$. We shall find the probability that this expression is bigger than $\frac{1}{8}$.

Now,

$$x(1-x) = -x^2 + x > \frac{1}{8},$$

if and only if

$$\frac{1}{2} - \frac{\sqrt{2}}{4} < x < \frac{1}{2} + \frac{\sqrt{2}}{4}.$$

Since X is rectangularly distributed, we get

$$P\left\{\frac{1}{2} - \frac{\sqrt{2}}{4} < X < \frac{1}{2} + \frac{\sqrt{2}}{4}\right\} = \int_{\frac{1}{2}-\frac{\sqrt{2}}{4}}^{\frac{1}{2}+\frac{\sqrt{2}}{4}} 1\,dx = \frac{\sqrt{2}}{2}.$$

Remark 7.1 *It is possible in general* to find the distribution function of

$$Y = f(X) = X(1-X).$$

If $0 < y < \frac{1}{4}$, then

$$\begin{aligned} P\{Y \le y\} &= P\left\{\left\{X \le \frac{1}{2} - \sqrt{\frac{1}{4} - y}\right\} \cup \left\{X \ge \frac{1}{2} + \sqrt{\frac{1}{2} - y}\right\}\right\} \\ &= 1 - 2\sqrt{\frac{1}{4} - y} = 1 - \sqrt{1-4y}, \end{aligned}$$

so

$$F_Y(y) = \begin{cases} 0, & y \le 0, \\ 1 - \sqrt{1-4y}, & 0 < y < \frac{1}{4}, \\ 1, & y \ge \frac{1}{4}, \end{cases}$$

and

$$f_Y(y) = \begin{cases} \dfrac{2y}{\sqrt{1-4y}}, & 0 < y < \frac{1}{4}, \\ 0, & \text{otherwise.} \end{cases}$$

Note that the probability of $Y = X(1-X)$ being bigger than y is

$$P\{Y > y\} = 1 - F_Y(y) = \begin{cases} 1, & \text{for } y \le 0, \\[2mm] \sqrt{1-4y}, & \text{for } 0 < y < \dfrac{1}{4}, \\[2mm] 0, & \text{for } y \ge \dfrac{1}{4}. \end{cases}$$

If $y = \dfrac{1}{8}$ we get

$$P\left\{Y > \frac{1}{8}\right\} = \sqrt{1 - \frac{1}{2}} = \frac{\sqrt{2}}{2}. \qquad \Diamond$$

Example 7.3 *Let the random variable X be rectangularly distributed over the interval $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$. Find the distribution functions and the frequencies of the random variables*

$$Y = \sin X, \qquad Z = \cos X, \qquad U = \tan X.$$

Since X is rectangularly distributed over $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, the frequency is given by

$$f_X(x) = \begin{cases} \dfrac{1}{\pi} & \text{for } x \in \left]-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right[, \\[2mm] 0 & \text{otherwise.} \end{cases}$$

1) $Y = \sin X$.
If $|y| < 1$, then

$$P\{Y \le y\} = P\{X \le \text{ Arcsin } y\} = \frac{1}{\pi}\left\{\text{Arcsin } y + \frac{\pi}{2}\right\} = \frac{1}{\pi}\text{ Arcsin } y + \frac{1}{2},$$

hence the distribution function of Y is

$$F_Y(y) = \begin{cases} 0, & y \le -1, \\[2mm] \dfrac{1}{\pi}\text{ Arcsin } y + \dfrac{1}{2}, & -1 < y < 1, \\[2mm] 1, & y \ge 1, \end{cases}$$

and the frequency is then obtained by a differentiation,

$$f_Y(y) = \begin{cases} \dfrac{1}{\pi\sqrt{1-y^2}}, & -1 < y < 1, \\[2mm] 0, & \text{otherwise.} \end{cases}$$

Alternatively, we first find the frequency $f_Y(y)$ of Y. Since $y = \sin x$ maps $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ *increasingly* onto $]-1, 1[$, it has an inverse map,

$$x = \tau^{-1}(y) = \text{ Arcsin } y \quad \text{with} \quad \left(\tau^{-1}\right)'(y) = \frac{1}{\sqrt{1-y^2}}.$$

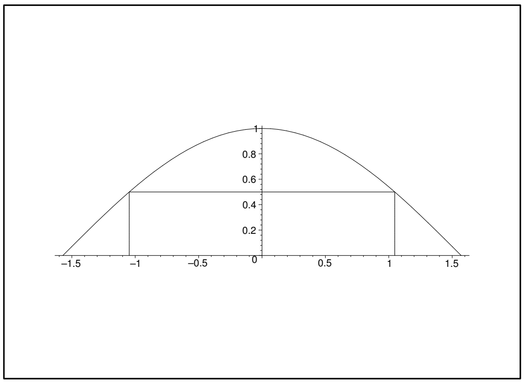


Figure 40: The graph of $z = \cos x$ for $x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$.

Since $f\left(\tau^{-1}(y)\right) = \dfrac{1}{\pi}$, we get that

$$f_Y(y) = \begin{cases} \dfrac{1}{\pi\sqrt{1-y^2}}, & -1 < y < 1, \\ \\ 0, & \text{otherwise.} \end{cases}$$

and if $-1 < y < 1$, we get the distribution function

$$F_Y(y) = P\{Y \le y\} = \int_{-1}^{y} \frac{1}{\pi\sqrt{1-y^2}}\, dy = \left[\frac{1}{\pi}\,\text{Arcsin}\; y\right]_{-1}^{y} = \frac{1}{\pi}\,\text{Arcsin}\; y + \frac{1}{2}.$$

2) $Z = \cos X$.
In this case, only $z \in]0,1[$ is of interest. It follows by the symmetry – cf. the figure – that

$$\begin{aligned} P\{Z \le z\} &= 2\,P\left\{\text{Arccos}\; z \le X < \frac{\pi}{2}\right\} = 2 \cdot \frac{1}{\pi}\left\{\frac{\pi}{2} - \ \text{Arccos z}\right\} \\ &= 1 - \frac{2}{\pi}\,\text{Arccos}\; z = \frac{2}{\pi}\,\text{Arcsin}\; z. \end{aligned}$$

Hence, the distribution function is

$$F_Z(z) = \begin{cases} 0, & z \le 0, \\ \\ \dfrac{2}{\pi}\,\text{Arcsin}\; z, & 0 < z < 1, \\ \\ 1, & z \ge 1, \end{cases}$$

from which we get the frequency by differentiation

$$f_Z(z) = \begin{cases} \dfrac{2}{\pi\sqrt{1-z^2}}, & 0 < z < 1, \\ \\ 0, & \text{otherwise.} \end{cases}$$

Since $z = \cos x$ is *not* monotonous in $\left]-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right[$, we cannot apply the usual argument.

3) $U = \tan X$.
If $u \in \mathbb{R}$, then

$$P\{U \leq u\} = P\{X \leq \text{ Arctan } u\} = \frac{1}{\pi}\left\{\text{Arctan } u + \frac{\pi}{2}\right\} = \frac{1}{\pi}\,\text{Arctan } u + \frac{1}{2},$$

hence

$$F_U(u) = \frac{1}{\pi}\,\text{Arctan } u + \frac{1}{2} \quad \text{og} \quad f_U(u) = F_U'(u) = \frac{1}{\pi\left(1+u^2\right)}, \qquad u \in \mathbb{R}.$$

Alternatively, like in (1) it is possible to find the frequency of U, because $u = \tau(x) = \tan x$ maps $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ *increasingly* onto $\mathbb{R}$ with the inverse map $x = \tau^{-1}(u) = \text{Arctan}(u)$, where

$$\left(\tau^{-1}\right)'(u) = \frac{1}{1+u^2}.$$

Then apply the standard formula.

The distribution of U is a *Cauchy distribution*.

Example 7.4 *Assume that the random variable X is rectangularly distributed over the interval $]0,\pi[$. Find the distribution functions and the frequencies of the random variables*

$$Y = \frac{1}{X}, \qquad Z = \cos X, \qquad U = \sin X.$$

When X is rectangularly distributed over $]0,\pi[$, then

$$f_X(x) = \begin{cases} \dfrac{1}{\pi} & \text{for } x \in]0,\pi[, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad F_X(x) = \begin{cases} 0 & \text{for } x \le 0, \\ \dfrac{x}{\pi} & \text{for } x \in]0,\pi[, \\ 1 & \text{for } x \ge \pi. \end{cases}$$

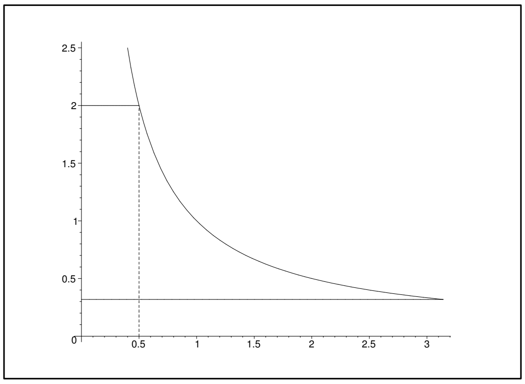


Figure 41: The graph of $y = \dfrac{1}{x}$ in the interval $]0,\pi[$.

1) The image of $]0,\pi[$ by the map $y = \dfrac{1}{x}$ is $\left]\dfrac{1}{\pi},\infty\right[$. If $y > \dfrac{1}{\pi}$, then we get the distribution function

$$F_Y(y) = P\{Y \le y\} = P\left\{\frac{1}{X} \le y\right\} = P\left\{X \ge \frac{1}{y}\right\} = 1 - P\left\{X < \frac{1}{y}\right\} = 1 - \frac{1}{\pi y},$$

hence

$$F_Y(y) = \begin{cases} 1 - \dfrac{1}{\pi y} & \text{for } y > \dfrac{1}{\pi}, \\ 0 & \text{for } y \le \dfrac{1}{\pi}, \end{cases} \qquad \text{and} \qquad f_Y(y) = \begin{cases} \dfrac{1}{\pi y^2} & \text{for } y > \dfrac{1}{\pi}, \\ 0 & \text{for } y \le \dfrac{1}{\pi}. \end{cases}$$

2) The image of $]0,\pi[$ by $z = \cos x$ is $]-1,1[$. If $z \in]-1,1[$, then we get the distribution function

$$\begin{aligned} F_Z(z) &= P\{Z \le z\} = P\{\cos X \le z\} = P\{X \ge \text{ Arccos } z\} \\ &= 1 - P\{X < \text{ Arccos } z\} = 1 - \frac{1}{\pi}\text{Arccos } z, \end{aligned}$$

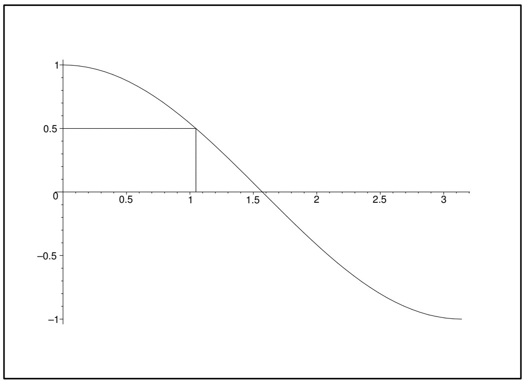


Figure 42: The graph of $z = \cos x$ for $x \in]0, \pi[$.

hence

$$F_Z(z) = \begin{cases} 0, & z \le -1, \\ 1 - \dfrac{1}{\pi} \text{Arccos}\, z, & -1 < z < 1, \\ 1, & z \ge 1, \end{cases}$$

and

$$f_Z(z) = \begin{cases} \dfrac{1}{\pi} \dfrac{1}{\sqrt{1-z^2}}, & -1 < z < 1, \\ 0, & \text{otherwise.} \end{cases}$$

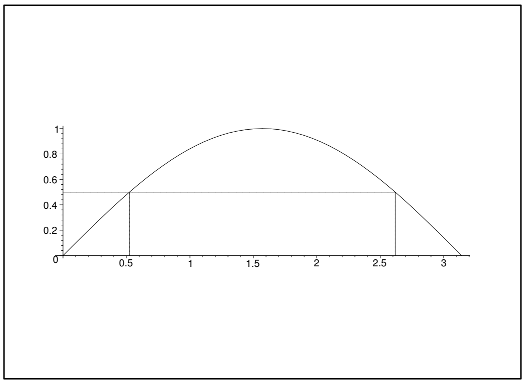


Figure 43: The graph of $u = \sin x$ for $x \in]0, \pi[$.

3) The image of $]0,\pi[$ by $u = \sin x$ is $]0,1[$. If $u \in]0,1[$, then we get the distribution function

$$\begin{aligned}
F_U(u) &= P\{U \le u\} = P\{\sin X \le u\} \\
&= P\{X \le \text{ Arcsin } u\} + P\{X \ge \pi - \text{ Arcsin } u\} \\
&= P\{X \le \text{ Arcsin } u\} + 1 - P\{X < \pi - \text{ Arcsin } u\} \\
&= 1 + \frac{1}{\pi}\text{Arcsin } u - \frac{1}{\pi}\{\pi - \text{ Arcsin } u\} = \frac{2}{\pi}\text{Arcsin } u,
\end{aligned}$$

hence

$$F_U(u) = \begin{cases} 0, & u \le 0, \\ \frac{2}{\pi}\text{Arcsin } u, & 0 < u < 1, \\ 1, & u \ge 1, \end{cases}$$

and

$$f_U(u) = \begin{cases} \frac{2}{\pi}\frac{1}{\sqrt{1-u^2}}, & 0 < u < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Alternatively, we may apply the usual formula in (1) and (2), but not in (3), because sinus is not a bijective onto the given interval.

If $t(x)$ is a bijective transformation, and $x = x(t)$ is the inverse, then we have in the form of differentials,

$$f_X(x)\,dx = f_X(x(t))\left|\frac{dx}{dt}\right| dt = f_T(t)\,dt.$$

Hence, we shall always find the inverse map $x = x(t)$.

1) If $x \in]0,\pi[$, then $y = \frac{1}{x} \in \left]\frac{1}{\pi},\infty\right[$. The inverse map is given by $x = \frac{1}{y}$, thus

$$\frac{dx}{dy} = -\frac{1}{y^2} < 0,$$

and we get for $y > \frac{1}{\pi}$ that

$$f_Y(y) = f_X\left(\frac{1}{y}\right)\cdot\left|-\frac{1}{y^2}\right| = \frac{1}{\pi}\cdot\frac{1}{y^2},$$

hence by an integration for $y > \frac{1}{\pi}$,

$$F_Y(y) = \int_{\frac{1}{\pi}}^{y}\frac{1}{\pi}\cdot\frac{1}{\eta^2}\,d\eta = \frac{1}{\pi}\left[-\frac{1}{\eta}\right]_{\frac{1}{\pi}}^{y} = 1 - \frac{1}{\pi y}.$$

100

2) If $x \in]0, \pi[$, then $z = \cos x \in]-1, 1[$ bijectively. The inverse is given by $x = \text{Arccos } z$, thus

$$\frac{dx}{dz} = -\frac{1}{\sqrt{1-z^2}} < 0.$$

If $z \in]-1, 1[$, then

$$f_Z(z) = f_X(\text{Arccos } z) \cdot \left| -\frac{1}{\sqrt{1-z^2}} \right| = \frac{1}{\pi\sqrt{1-z^2}},$$

so if $z \in]-1, 1[$, then we get by an integration that

$$F_Z(z) = \int_{-1}^{z} \frac{1}{\pi} \cdot \frac{1}{\sqrt{1-\zeta^2}}\, d\zeta = \frac{1}{\pi}\left[-\text{ Arccos } \zeta\right]_{-1}^{z} = \frac{1}{\pi}\left\{\pi - \text{ Arccos } z\right\} = 1 - \frac{1}{\pi}\text{ Arccos } z.$$

Example 7.5 *Assume that the random variable X has the frequency*

$$f(x) = \begin{cases} \dfrac{3}{2\pi}, & 0 < x < \dfrac{\pi}{2}, \\[2mm] \dfrac{1}{2\pi}, & \dfrac{\pi}{2} < x < \pi, \\[2mm] 0, & \text{otherwise.} \end{cases}$$

Find the distribution function of X.
Then find the distribution functions and the frequencies of each of the random variables

$$Y = \frac{1}{X}, \qquad Z = \sin X.$$

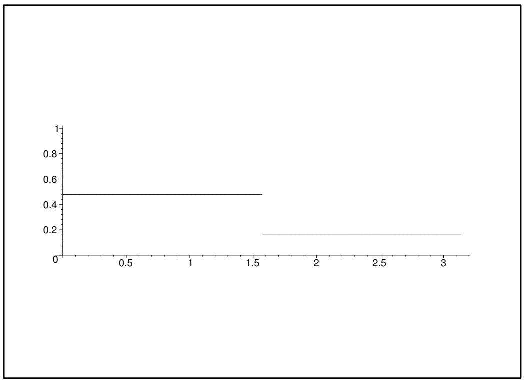
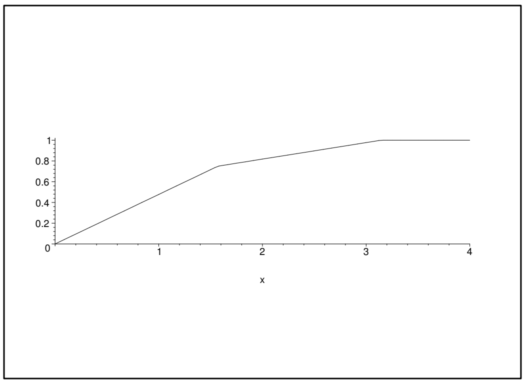


Figure 44: The frequency $f(x)$.



1) If $x \le 0$, then $F(x) = 0$.
If $0 < x \le \dfrac{\pi}{2}$, then $F(x) = \dfrac{3x}{2\pi}$.
In particular, $F\left(\dfrac{\pi}{2}\right) = \dfrac{3}{4}$, hence for $\dfrac{\pi}{2} < x < \pi$,

$$F(x) = F\left(\frac{\pi}{2}\right) + \int_{\frac{\pi}{2}}^{x} \frac{1}{2\pi}\,dx = \frac{3}{4} + \frac{1}{2\pi}\left\{x - \frac{\pi}{2}\right\} = \frac{1}{2} + \frac{x}{2\pi}.$$

If $x \ge \pi$, then $F(x) = 1$.
Summing up we get the distribution function

$$F(x) = \begin{cases} 0, & x \le 0, \\ \dfrac{3x}{2\pi}, & 0 < x \le \dfrac{\pi}{2}, \\ \dfrac{1}{2} + \dfrac{x}{2\pi}, & \dfrac{\pi}{2} < x < \pi, \\ 1, & x \ge \pi. \end{cases}$$

2) The interval $]0,\pi[$ is by $y = \dfrac{1}{x}$ mapped bijectively onto $\left]\dfrac{1}{\pi}, \infty\right[$.

$1^{\textbf{st}}$ **variant.** If $y > \dfrac{1}{\pi}$, then we get the distribution function

$$\begin{aligned} F_Y(y) &= P\{Y \le y\} = P\left\{\frac{1}{x} \le y\right\} = P\left\{X \ge \frac{1}{y}\right\} = 1 - P\left\{X < \frac{1}{y}\right\} \\ &= \begin{cases} 1 - \dfrac{3}{2\pi y} & \text{for } y > \dfrac{2}{\pi}, \\ \dfrac{1}{2} - \dfrac{1}{2\pi y}, & \text{for } \dfrac{1}{\pi} < y \le \dfrac{2}{\pi}, \\ 0, & \text{for } y \le \dfrac{1}{\pi}. \end{cases} \end{aligned}$$

The frequency is obtained by a differentiation,

$$f_Y(y) = \begin{cases} \dfrac{3}{2\pi y^2}, & \text{for } y > \dfrac{2}{\pi}, \\ \dfrac{1}{2\pi y^2}, & \text{for } \dfrac{1}{\pi} < y < \dfrac{2}{\pi}, \\ 0, & \text{otherwise.} \end{cases}$$

2$^{\text{nd}}$ variant. Since $x = \dfrac{1}{y}$ and $\dfrac{dx}{dy} = -\dfrac{1}{y^2}$, it follows that

$$f_Y(y) = F_X\left(\frac{1}{y}\right) \cdot \frac{1}{y^2} = \begin{cases} \dfrac{3}{2\pi y^2} & \text{for } y \geq \dfrac{2}{\pi}, \\[2ex] \dfrac{1}{2\pi y^2} & \text{for } \dfrac{1}{\pi} < y < \dfrac{2}{\pi}, \\[2ex] 0 & \text{otherwise.} \end{cases}$$

If $y \leq \dfrac{1}{\pi}$, then $F_Y(y) = 0$.

We get for $\dfrac{1}{\pi} \leq y \leq \dfrac{2}{\pi}$ that

$$F_Y(y) = \int_{\frac{1}{\pi}}^{y} \frac{d\eta}{2\pi\eta^2} = \left[-\frac{1}{2\pi\eta}\right]_{\frac{1}{\pi}}^{y} = \frac{1}{2} - \frac{1}{2\pi y}.$$

If $y > \dfrac{2}{\pi}$, then

$$\begin{aligned} F_Y(y) &= F_Y\left(\frac{2}{\pi}\right) + \int_{\frac{2}{\pi}}^{y} \frac{3}{2\pi\eta^2}\, d\eta = \frac{1}{2} - \frac{1}{4} + \frac{3}{2\pi}\left[-\frac{1}{\eta}\right]_{\frac{2}{\pi}}^{y} \\ &= \frac{1}{4} + \frac{3}{4} - \frac{3}{2\pi y} = 1 - \frac{3}{2\pi y}. \end{aligned}$$

3) The function $z = \sin x$ is not bijective, so we cannot apply the usual theorem. Since the image of $]0, \pi[$ by $z = \sin x$ is $]0, 1[$, we get for $z \in]0, 1[$ that

$$\begin{aligned} F_Z(z) &= P\{Z \leq z\} = P\{\sin X \leq z\} \\ &= P\{X \leq \text{ Arcsin } z\} + P\{X \geq \pi - \text{ Arcsin } z\} \\ &= 1 + P\{X \leq \text{ Arcsin } z\} - P\{X < \pi - \text{ Arcsin } z\}. \end{aligned}$$

Since Arcsin $z \in \left]0, \dfrac{\pi}{2}\right[$, we must have $\pi -$ Arcsin $z \in \left]\dfrac{\pi}{2}, \pi\right[$. Then it follows from $F_X(x) = P\{X \leq x\}$, found in (1) that

$$F_Z(z) = 1 + \frac{3}{2\pi} \text{ Arcsin } z - \left\{\frac{1}{2} + \frac{1}{2\pi}(\pi - \text{ Arcsin } z)\right\} = \frac{2}{\pi} \text{ Arcsin } z,$$

thus

$$F_Z(z) = \begin{cases} 0, & \text{for } z \leq 0, \\[2ex] \dfrac{2}{\pi} \text{ Arcsin } z, & \text{for } 0 < z < 1, \\[2ex] 1, & \text{for } z \geq 1, \end{cases}$$

and hence

$$f_Z(z) = \begin{cases} \dfrac{1}{\pi} \cdot \dfrac{1}{\sqrt{1 - z^2}}, & \text{for } 0 < z < 1, \\[2ex] 0, & \text{otherwise.} \end{cases}$$

Example 7.6 *The random variable X has the frequency*

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

Find the frequencies of the random variables

$$Y = \sinh X, \qquad Z = \cosh X.$$

We first note that $\sinh : \mathbb{R}_+ \to \mathbb{R}_+$ and $\cosh : \mathbb{R}_+ \to]1, +\infty[$ are bijective and monotonous.

1) $Y = \sinh X$. In this case, $y = \tau(x) = \sinh x$, $x \in \mathbb{R}_+$, hence

$$x = \tau^{-1}(y) = \text{ Arsinh } y = \ln\left(y + \sqrt{1+y^2}\right), \qquad y > 0,$$

and

$$\frac{dx}{dy} = \left(\tau^{-1}(y)\right)' = \frac{1}{\sqrt{1+y^2}}, \qquad y > 0.$$

It follows from the usual theorem that the frequency is given for $y > 0$ by

$$g(y) = f\left(\tau^{-1}(y)\right) \cdot \left|\left(\tau^{-1}\right)'(y)\right| = e^{\ln\left(y+\sqrt{1+y^2}\right)} \cdot \frac{1}{\sqrt{1+y^2}} = \frac{1}{y+\sqrt{1+y^2}} \cdot \frac{1}{\sqrt{1+y^2}},$$

thus

$$g(y) = \begin{cases} \dfrac{1}{y+\sqrt{1+y^2}} \cdot \dfrac{1}{\sqrt{1+y^2}}, & \text{for } y > 0, \\[2mm] 0, & \text{otherwise.} \end{cases}$$

2) $Z = \cosh X$. In this case, $z = \tau(x) = \cosh x$, $x \in \mathbb{R}_+$, thus

$$x = \tau^{-1}(z) = \text{ Arcosh } z = \ln\left(z+\sqrt{z^2-1}\right), \qquad z > 1,$$

and

$$\frac{dx}{dz} = \left(\tau^{-1}(z)\right)' = \frac{1}{\sqrt{z^2-1}}. \qquad z > 1.$$

Applying the theorem we get for $z > 1$ the frequency

$$h(z) = f\left(\tau^{-1}(z)\right) \cdot \left|\left(\tau^{-1}\right)'(z)\right| = e^{-\ln\left(z+\sqrt{z^2-1}\right)} \cdot \frac{1}{\sqrt{z^2-1}} = \frac{1}{z+\sqrt{z^2-1}} \cdot \frac{1}{\sqrt{z^2-1}},$$

hence

$$h(z) = \begin{cases} \dfrac{1}{z+\sqrt{z^2-1}} \cdot \dfrac{1}{\sqrt{z^2-1}}, & \text{for } z > 1, \\[2mm] 0, & \text{otherwise.} \end{cases}$$

8 Functions of two random variables, $f(X,Y)$

Example 8.1 *1) Let X and Y be independent random variables with their frequencies*

$$f_X(x) = \frac{k}{\pi\left(k^2+x^2\right)}, \quad x \in \mathbb{R}, \quad f_Y(y) = \frac{1}{\pi\left(1+y^2\right)}, \quad t \in \mathbb{R},$$

where k denotes some positive constant. Prove that $X+Y$ has the frequency

$$g(x) = \frac{k+1}{\pi\left\{(k+1)^2+x^2\right\}}, \qquad x \in \mathbb{R}.$$

2) Let X_1 and X_2 be independent random variables of frequencies

$$f_1(x_1) = \frac{a_1}{\pi\left(a_1^2+x_1^2\right)}, \quad x_1 \in \mathbb{R}, \quad f_2(x_2) = \frac{a_2}{\pi\left(a_2^2+x_2^2\right)}, \quad x_2 \in \mathbb{R},$$

where a_1 and a_2 denote positive constants. Find by using the result of (1) the frequency of X_1+X_2.

3) Let Y_1 and Y_2 be independent random variables of the frequencies

$$g_1(y_1) = \frac{a_1}{\pi\left\{a_1^2+\left(y_1-b_1\right)^2\right\}}, \quad y_1 \in \mathbb{R},$$

$$g_2(y_2) = \frac{a_2}{\pi\left\{a_2^2+\left(y_2-b_2\right)^2\right\}}, \quad y_2 \in \mathbb{R},$$

where a_1, $a_2 \in \mathbb{R}_+$ and b_1, $b_2 \in \mathbb{R}$. Find by using the result of (2) the frequency of Y_1+Y_2.

1) (The hard question). The frequency $g(x)$ of $X+Y$ is given by the convolution

$$g(x) = \frac{1}{\pi^2}\int_{-\infty}^{\infty} \frac{k}{k^2+t^2}\cdot\frac{1}{1+(t-x)^2}\,dt.$$

A decomposition gives us the structure

$$(1) \quad \frac{k}{k^2+t^2}\cdot\frac{1}{1+(t-x)^2} = \frac{a+bt}{k^2+t^2}+\frac{c+d(t-x)}{1+(t-x)^2},$$

which has the integral

$$\frac{a}{k}\,\text{Arctan}\left(\frac{t}{k}\right)+c\,\text{Arctan}(t-x)+\frac{b}{2}\,\ln\left(k^2+t^2\right)+\frac{d}{2}\,\ln\left\{1+(t-x)^2\right\}.$$

The integral is clearly convergent, so $d=-b$, and the logarithmic terms disappear by taking the limit.

We conclude that

$$g(x) = \frac{1}{\pi}\left\{\frac{a}{k}+c\right\},$$

where $a = a(k,x)$ and $c = c(k,x)$ depend on both k and x.

If we put $d=-b$ into (1), then

$$\frac{k}{k^2+t^2}\cdot\frac{1}{1+x^2-2xt+t^2}=\frac{a+bt}{k^2+t^2}+\frac{bx+c-bt}{1+x^2-2xt+t^2},$$

hence

(2) $k=(a+bt)\left(1+x^2-2x\cdot t+t^2\right)+\left(k^2+t^2\right)(bx+c-bt).$

The constant term of this equation is

$$k=\left(1+x^2\right)a+k^2x\cdot b+k^2c,$$

thus

$$\left(1+x^2\right)\,\frac{a}{k}+kx\cdot b+k\cdot c=1.$$

Since we want to find $\dfrac{a}{k}+c$, we rewrite this as

(3) $\left(1+x^2\right)\left\{\dfrac{a}{k}+c\right\}+kx\cdot b+\left(k-1-x^2\right)c=1.$

The coefficient of t i (2) gives the equation

$$-2xa+\left(1+x^2\right)b-k^2b=0,$$

which is rewritten as

(4) $-2kx\left\{\dfrac{a}{k}+c\right\}+\left(1+x^2-k^2\right)b+2kx\cdot c=0.$

The coefficient of t^2 in (2) implies the equation

$$a-xb+c=0,$$

which is rewritten in the following way (cf. the above)

(5) $k\left\{\dfrac{a}{k}+c\right\}-x\cdot b+(1-k)\cdot c=0.$

Summing up we obtain the following linear system of the three unknowns $\dfrac{a}{k}+c$, b and c,

$$\left\{\begin{array}{ccccccc}
\left(1+x^2\right)\left\{\dfrac{a}{k}+c\right\} & + & kx\cdot b & = & \left(k-1-x^2\right)c & = & 1,\\
-2kx\left\{\dfrac{a}{k}+c\right\} & + & \left(1+x^2-k^2\right)b & + & 2kx\cdot c & = & 0,\\
k\left\{\dfrac{a}{k}+c\right\} & - & x\cdot b & + & (1-k)c & = & 0,
\end{array}\right.$$

Then by Cramer's formula,

$$\frac{a}{k}+c=\frac{\left|\begin{array}{ccc}1 & kx & k-1-x^2\\ 0 & 1+x^2-k^2 & 2kx\\ 0 & -x & 1-k\end{array}\right|}{\left|\begin{array}{ccc}1+x^2 & kx & k-1-x^2\\ -2kx & 1+x^2-k^2 & 2kx\\ k & -x & 1-k\end{array}\right|}.$$

First compute the numerator

$$\begin{aligned}\left|\begin{array}{cc} 1+x^2-k^2 & 2kx \\ -x & 1-k \end{array}\right| &= (1-k)\left(1+x^2\right)-(1-k)k^2+2kx^2 \\ &= 1+x^2-k-kx^2+2kx^2-k^2+k^3 \\ &= x^2(1+k)+(1-k)\left(1-k^2\right) \\ &= (k+1)\left\{x^2+(k-1)^2\right\}.\end{aligned}$$

When the third column is replaced by the sum of the first and the third column we see that the denominator is reduced to

$$\begin{aligned}&\left|\begin{array}{ccc} 1+x^2 & kx & k \\ -2kx & 1+x^2-k^2 & 0 \\ k & -x & 1 \end{array}\right| = \left|\begin{array}{ccc} 1+x^2-k^2 & 2kx & 0 \\ -2kx & 1+x^2-k^2 & 0 \\ k & -x & 1 \end{array}\right| \\ &= \left|\begin{array}{cc} 1+x^2-k^2 & 2kx \\ -2kx & 1+x^2-k^2 \end{array}\right| = \left(1+x^2-k^2\right)^2+4k^2x^2 \\ &= x^4+x^2\left(2-2k^2+4k^2\right)+k^4-2k^2+1 \\ &= x^2\left\{x^2+(k-1)^2\right\}+x^2\left(2k^2+2-k^2+2k-1\right)+\left(k^2-1\right)^2 \\ &= x^2\left\{x^2+(k-1)^2\right\}+x^2(k+1)^2+(k+1)^2(k-1)^2 \\ &= x^2\left\{x^2+(k-1)^2\right\}+(k+1)^2\left\{x^2+(k-1)^2\right\} \\ &= \left\{x^2+(k-1)^2\right\}\left\{x^2+(k+1)^2\right\}.\end{aligned}$$

Hence, if $(x,y) \neq (0,1)$,

$$\frac{a}{k}+c=\frac{(k+1)\left\{x^2+(k-1)^2\right\}}{\{x^2+(k+1)^2\}\left\{x^2+(k-1)^2\right\}}=\frac{k+1}{x^2+(k+1)^2},$$

which is extended by continuity to $(x,k)=(0,1)$.

Thus, the frequency is given by

$$g(x)=\frac{1}{\pi}\left\{\frac{a}{k}+c\right\}=\frac{1}{\pi}\cdot\frac{k+1}{x^2+(k+1)^2}$$

as required.

2) The frequency of X_1+X_2 is

$$\begin{aligned}
f(x) &= \frac{a_1a_2}{\pi^2}\int_{-\infty}^{\infty}\frac{1}{a_1^2+t^2}\cdot\frac{1}{a_2^2+(t-x)^2}\,dt, && u=\frac{t}{a_1},\\
&= \frac{a_2}{\pi^2}\int_{-\infty}^{\infty}\frac{1}{1+\left(\dfrac{t}{a_1}\right)^2}\cdot\frac{1}{a_2^2+a_1^2\left(\dfrac{t}{a_1}-\dfrac{x}{a_1}\right)^2}\,d\left(\frac{t}{a_1}\right)\\
&= \frac{1}{\pi^2a_1}\int_{-\infty}^{\infty}\frac{1}{1+u^2}\cdot\frac{\dfrac{a_2}{a_1}}{\left(\dfrac{a_2}{a_1}\right)^2+\left(u-\dfrac{x}{a_1}\right)^2}\,du && k=\frac{a_2}{a_1}\text{ and (1)},\\
&= \frac{1}{\pi a_1}\cdot\frac{\dfrac{a_2}{a_1}+1}{\left(\dfrac{a_2}{a_1}+1\right)^2+\left(\dfrac{x}{a_1}\right)^2}=\frac{1}{\pi}\cdot\frac{a_1+a_2}{(a_1+a_2)^2+x^2}.
\end{aligned}$$

3) In this case we get the frequency

$$\begin{aligned}
g(y) &= \frac{a_1a_2}{\pi^2}\int_{-\infty}^{\infty}\frac{1}{a_1^2+(t-b_1)^2}\cdot\frac{1}{a_2^2+(y-t-b_2)^2}\,dt && u=t-b_1,\\
&= \frac{a_1a_2}{\pi^2}\int_{-\infty}^{\infty}\frac{1}{a_1^2+u^2}\cdot\frac{1}{a_2^2+(y-u-b_1-b_2)^2}\,du\\
&= \frac{a_1a_2}{\pi^2}\int_{-\infty}^{\infty}\frac{1}{a_1^2+u^2}\cdot\frac{1}{a_2^2+(u-\{y-b_1-b_2\})^2}\,du\\
&= \frac{1}{\pi}\cdot\frac{a_1+a_2}{(a_1+a_2)^2+(y-\{b_1+b_2\})^2},
\end{aligned}$$

where we have applied (2).

Example 8.2 *Let X and Y be independent Cauchy distributed random variable of the frequencies*

$$f_X(x) = \frac{1}{\pi\left(1+x^2\right)}, \quad x \in \mathbb{R}, \quad f_Y(y) = \frac{1}{\pi\left(1+y^2\right)}, \quad y \in \mathbb{R}.$$

Prove that the random variable $Z = XY$ has the frequency

$$f_Z(z) = \frac{2}{\pi^2} \cdot \frac{\ln|z|}{z^2-1}, \qquad z \in \mathbb{R},$$

(suitably modified for $z = -1,\ 0,\ 1$).
HINT*: One may apply that*

$$\left(z^2-1\right) \cdot \frac{1}{1+z^2} \cdot \frac{1}{z^2+x^2} = \frac{1}{1+x^2} - \frac{1}{z^2+x^2}.$$

If $z \neq -1,\ 0,\ 1$, then the frequency of $Z = XY$ is given by

$$\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} g(x)\, g\left(\frac{z}{x}\right) \frac{1}{|x|}\, dx = \frac{1}{\pi^2} \int_{-\infty}^{\infty} \frac{1}{1+x^2} \cdot \frac{1}{1+\left(\dfrac{z}{x}\right)^2} \cdot \frac{1}{|x|}\, dx \\
&= \frac{2}{\pi} \int_0^{\infty} \frac{1}{1+x^2} \cdot \frac{1}{z^2+x^2}\, x\, dx \qquad \text{symmetry; then } u = x^2, \\
&= \frac{1}{\pi^2} \int_0^{\infty} \frac{1}{1+u} \cdot \frac{1}{z^2+u}\, du \\
&= \frac{1}{\pi^2} \int_0^{\infty} \left\{ \frac{1}{z^2-1} \cdot \frac{1}{1+u} - \frac{1}{z^2-1} \cdot \frac{1}{z^2+u} \right\} du \\
&= \frac{1}{\pi^2} \cdot \frac{1}{z^2-1} \lim_{A \to \infty} \int_0 \left\{ \frac{1}{1+u} - \frac{1}{z^2+u} \right\} du \\
&= \frac{1}{\pi^2} \cdot \frac{1}{z^2-1} \lim_{A \to \infty} \left\{ \ln\left(\frac{u+1}{u+z^2}\right) - \ln\left(\frac{1}{z^2}\right) \right\} = \frac{2}{\pi^2} \cdot \frac{\ln|z|}{z^2-1}.
\end{aligned}$$

Since the exceptional set $\{-1, 0, 1\}$ is a null set, we can choose $f_Z(z) = 0$, which is sufficient at these points.

Note that

$$\lim_{z \to 0} f_Z(z) = \infty,$$

and that it follows by l'Hospital's rule that

$$\lim_{z \to \pm 1} \frac{2}{\pi^2} \cdot \frac{\ln|z|}{z^2-1} = \frac{2}{\pi^2} \lim_{z \to \pm 1} \frac{\frac{1}{z}}{2z} = \frac{2}{\pi^2} \lim_{z \to \pm 1} \frac{1}{2z^2} = \frac{1}{\pi^2}.$$

Example 8.3 *Let X and Y be independent random variables, both rectangularly distributed over the interval* $]0,1[$.

1) Find the frequency of the random variable XY.

2) Find the frequency of the random variable $\dfrac{X}{Y}$.

3) Find $P\{Y > 2X\}$.

The two independent random variables X and Y have the same frequency,

$$f(x)=\begin{cases}1, & x\in\,]0,1[,\\[2mm] 0, & \text{otherwise.}\end{cases}$$

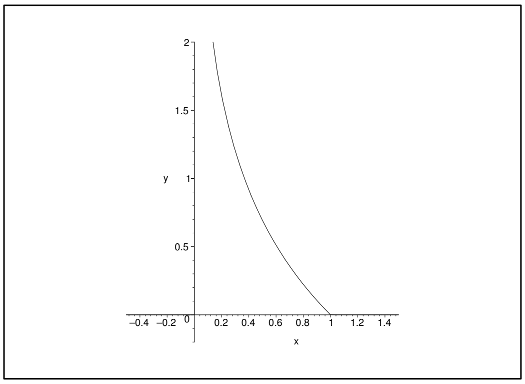


Figure 45: The graph of the frequency $g(s)$ of XY.

1) Since the values of XY lie in $]0,1[$, the frequency is for $s\in\,]0,1[$,

$$g(s)=\int_{x=s}^{1} f(x)\,f\left(\frac{s}{x}\right)\,\frac{1}{x}\,dx=\int_{x=s}^{1}\frac{1}{x}\,dx=-\ln s,$$

thus

$$g(s)=\begin{cases}-\ln s, & 0<s<1,\\[2mm] 0, & \text{otherwise.}\end{cases}$$

2) Since the values of $\dfrac{X}{Y}$ lie in $]0,\infty[$, it follows by an application of a formula for $s\in\,]0,\infty[$ that the frequency is given by

$$h(s)=\int_0^{\infty} f(s\,x)\,f(x)\,x\,dx.$$

This expression is only $\neq 0$, if $sx\in\,]0,1[$ and $x\in\,]0,1[$, hence if $0<x<1$ and $0<x<\dfrac{1}{s}$. Then we must split the investigation:

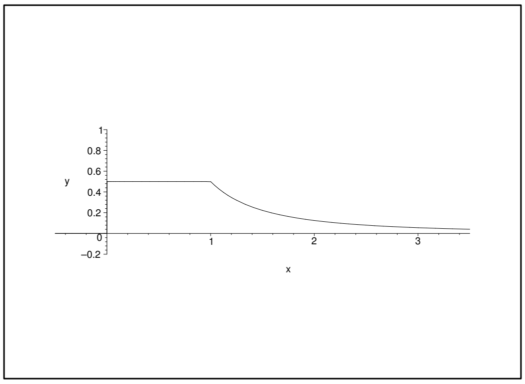


Figure 46: The graph of the frequency $h(s)$ of $\dfrac{X}{Y}$.

a) If $0 < s \leq 1$, then $1 \leq \dfrac{1}{s}$, hence

$$h(s) = \int_0^1 1 \cdot 1 \cdot x \, dx = \frac{1}{2}.$$

b) If $1 < s < \infty$, then instead

$$h(s) = \int_0^{\frac{1}{s}} x \, dx = \frac{1}{2s^2}.$$

Summing up we get

$$h(s)=\begin{cases}\dfrac{1}{2}, & 0<s\le 1,\\[2mm] \dfrac{1}{2s^2}, & 1<s<\infty,\\[2mm] 0, & s\le 0.\end{cases}$$

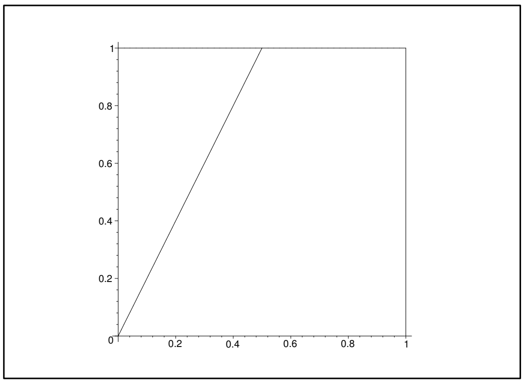


Figure 47: The line $y=2x$ defines the upper triangle A, where $y>2x$.

3) 1^{st} **variant.** It follows from the geometry that

$$P\{Y>2X\}=\ \text{area}(A)=\frac{1}{2}\cdot\frac{1}{2}\cdot 1=\frac{1}{4}.$$

2^{nd} **variant.** It follows from (2) that

$$P\{Y>2X\}=P\left\{\frac{X}{Y}<\frac{1}{2}\right\}=\frac{1}{2}.$$

An alternative solution is the following:

1) Since XY has its values lying in $]0,1[$, it follows from the figure that if $s\in]0,1[$. then

$$P\{XY\le s\}=\ \text{areal}(A)=\int_0^1 1\,dx+\int_s^1\frac{s}{x}\,dx=s-s\,\ln s,$$

and the frequency is obtained by a differentiation,

$$g(s)=\begin{cases}-\ln s, & 0<s<1,\\[2mm] 0, & \text{otherwise.}\end{cases}$$

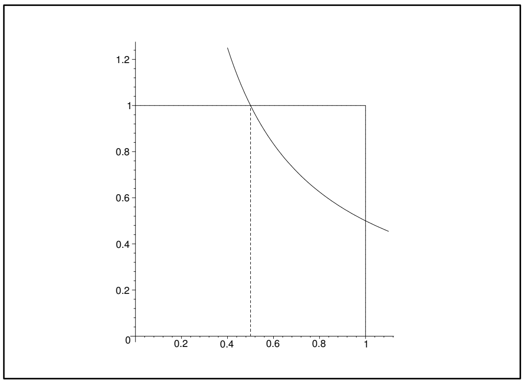


Figure 48: The curve $xy = s$ defines the domain A.

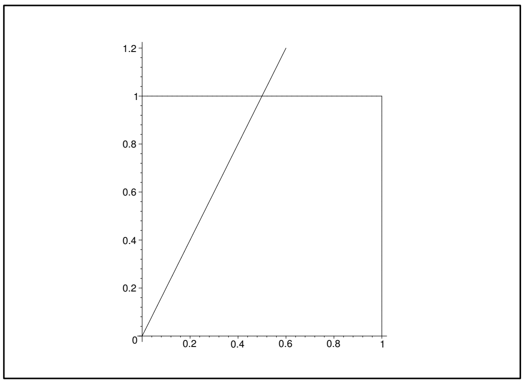


Figure 49: The domain A lies above the line $\dfrac{x}{y} = s$, $0 < s < 1$.

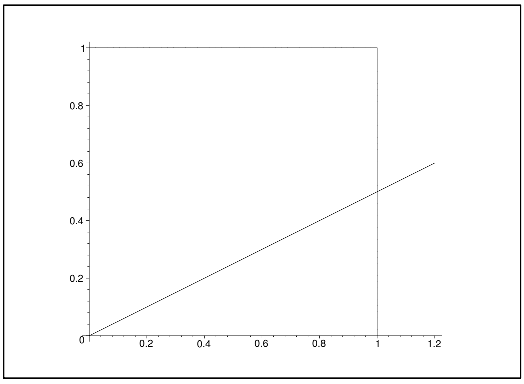


Figure 50: The domain B lies above the line $\dfrac{x}{y} = s$, $s > 1$.

2) It follows that the values of $\dfrac{X}{Y}$ lie in $]0,\infty[$. If $s \in]0,1[$, then it follows from the first figure that

$$P\left\{\frac{X}{Y} \le s\right\} = \text{ area}(A) = \frac{1}{2}\,s, \qquad 0 < s < 1.$$

If $s > 1$, then it follows from the second figure that

$$P\left\{\frac{X}{Y} \le s\right\} = \text{ area}(B) = 1 - \frac{1}{2s}, \qquad s \ge 1.$$

Finally, the frequency is obtained by a differentiation,

$$h(s) = \begin{cases} \dfrac{1}{2}, & 0 < s < 1, \\[2mm] \dfrac{1}{2s^2}, & 1 \le s, \\[2mm] 0, & s \le 0. \end{cases}$$

Example 8.4 *Assume that X and Y are independent random variable, both attaining the values 1, 2, 3, 4, 5, 6. Assume that*

$$f_i = P\{X = i\}, \qquad g_i = P\{Y = i\}, \qquad i = 1,\, 2,\, 3,\, 4,\, 5,\, 6.$$

It is well-known that if $f_i = g_i = \dfrac{1}{6}$, $i = 1$, 2, 3, 4, 5, 6, then the probabilities $P\{X + Y = k\}$, $k = 2$, 3, ... , 12, are not all the same.
Prove that no matter how f_i and g_i are chosen, we can never *obtain $P\{X + Y = k\} = \dfrac{1}{11}$ for all $k = 2$, 3, ... , 12.*

Assume that

$$P\{X + Y = k\} = \frac{1}{11} \qquad \text{for all } k = 2,\, 3,\, \dots ,12.$$

If $k = 2$, then

$$\frac{1}{11} = f_1 g_1, \qquad \text{hence } f_1 > 0 \text{ and } g_1 > 0.$$

If $k = 12$, then

$$\frac{1}{11} = f_6 g_6, \qquad \text{hencer } f_6 > 0 \text{ and } g_6 > 0.$$

If $k = 7$, then

$$\frac{1}{11} = f_6 g_1 + \{f_5 g_2 + f_4 g_3 + f_3 g_4 + f_2 g_5\} + f_1 g_6.$$

By subtracting the equation for $k = 2$ from the equation for $k = 7$, it follows by a rearrangement that

$$(f_1 - f_6)\, g_1 = \{f_5 g_2 + f_4 g_3 + f_3 g_4 + f_2 g_5\} + f_1 g_6 > 0,$$

because $\{\cdots\} \geq 0$ and $f_1 g_6 > 0$. Since $g_1 > 0$, we must have $f_1 > f_6$.

If we subtract the equation for $k = 12$ from the equation for $k = 7$, then by a rearrangement.

$$(f_6 - f_1)\, g_6 = \{f_5 g_2 + f_4 g_3 + f_3 g_4 + f_2 g_5\} + f_6 g_1 > 0$$

for similar reasons. We conclude that $f_6 > f_1$.

These two claims cannot be simultaneously fulfilled, so the assumption must be wrong.

It even follows from the proof above that $P\{X + Y = 2\}$, $P\{X + Y = 7\}$ and $P\{X + Y = 12\}$ can never have the same value, thus we can get a stronger result.

ALTERNATIVELY we assume that we can choose the f_i and the g_j in such a way that the probabilities are equal, i.e.

$$P\{X + Y = k\} = \frac{1}{11}, \qquad k = 2, \ldots, 12.$$

Then in particular,

$$P\{X + Y = 2\} = f_1 g_1 = \frac{1}{11} \quad \text{and} \quad P\{X + Y = 12\} = f_6 g_6 = \frac{1}{11},$$

hence $f_1 g_1 = f_6 g_6$. This is reformulated in the following way

$$\frac{f_1}{f_6} = \frac{g_6}{g_1} = x.$$

Considering the case $k = 7$ we get

$$\frac{1}{11} = P\{X+Y=7\} \geq f_1 g_6 + f_6 g_1 = f_1 g_1 \left\{ \frac{g_6}{g_1} + \frac{f_6}{f_1} \right\} = \frac{1}{11} \left\{ x + \frac{1}{x} \right\}.$$

Since $x + \dfrac{1}{x} > 1$ (actually $x + \dfrac{1}{x} \geq 2$, when $x > 0$), this is not possible, and we have obtained a contradiction, and the claim follows.

Example 8.5 *Let the 2-dimensional random variable* (X_1, X_2) *have its frequency* $h(x_1, x_2)$ *given by*

$$h(x_1, x_2) = \begin{cases} \dfrac{1}{\pi r^2}, & x_1^2 + x_2^2 < r^2, \\[2ex] 0, & \textit{otherwise}, \end{cases}$$

(a uniform distribution over the disc $x_1^2 + x_2^2 < r^2$*).*
Let the random variables Y_1 *and* Y_2 *be given by*

$$X_1 = Y_1 \cos Y_2, \quad X_2 = Y_1 \sin Y_2, \qquad 0 \leq Y_1 < r, \quad 0 \leq Y_2 < 2\pi.$$

Find the frequency of the 2-dimensional random variable (Y_1, Y_2)*, and find the marginal frequencies. Are* Y_1 *and* Y_2 *independent?*

Remark 8.1 This clearly corresponds to the transformation between rectangular and polar coordinates over a fixed disc. $\Diamond$

It follows that

$$\mathbf{x} = (x_1, x_2) = \varphi(\mathbf{y}) = (y_1 \cos y_2, y_1 \sin y_2), \qquad y_1 \in [0, r[, \quad y_2 \in [0, 2\pi[.$$

The corresponding Jacobian is

$$\begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\[2ex] \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \cos y_2 & -y_1 \sin y_2 \\[2ex] \sin y_2 & y_1 \cos y_2 \end{vmatrix} = y_1 \geq 0.$$

Hence,

$$k(y_1, y_2) = \begin{cases} \dfrac{y_1}{\pi r^2}, & (y_1, y_2) \in [0, r] \times [0, 2\pi[, \\[2ex] 0, & \text{otherwise}. \end{cases}$$

The frequencies of the marginal distributions are then

$$f_{Y_1}(y_1) = \begin{cases} \int_0^{2\pi} \dfrac{y_1}{\pi r^2}\, dy_2 = 2\pi \cdot \dfrac{y_1}{\pi r^2} = \dfrac{2y_1}{r^2}, & y_1 \in [0, r[, \\[2ex] 0, & \text{otherwise}, \end{cases}$$

and

$$f_{Y_2}(y_2) = \begin{cases} \int_0^r \dfrac{y_1}{\pi r^2}\, dy_1 = \dfrac{1}{2\pi}, & y_2 \in [0, 2\pi[, \\ \\ 0, & \text{otherwise.} \end{cases}$$

It follows immediately that

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

hence Y_1 and Y_2 are stochastically independent.

Example 8.6 *Let the 2-dimensional random variable* (X_1, X_2) *have the frequency*

$$h(x_1, x_2) = \begin{cases} x_1 + x_2, & \textit{for } 0 < x_1 < 1 \textit{ og } 0 < x_2 < 1, \\ \\ \quad 0, & \textit{otherwise,} \end{cases}$$

and let $(Y_1, Y_2) = \tau(X_1, X_2)$ *be given by*

$$Y_1 = X_1 + X_2, \qquad Y_2 = X_2.$$

1) Prove that τ *maps* $]0,1[\times]0,1[$ *bijectively onto the domain*

$$D' = \left\{(y_1, y_2) \in \mathbb{R}^2 \mid 0 < y_2 < 1,\ y_2 < y_1 < y_2 + 1\right\}.$$

2) Find the frequency $k(y_1, y_2)$ *of* (Y_1, Y_2).

3) Find the marginal frequencies of Y_1 *and* Y_2.

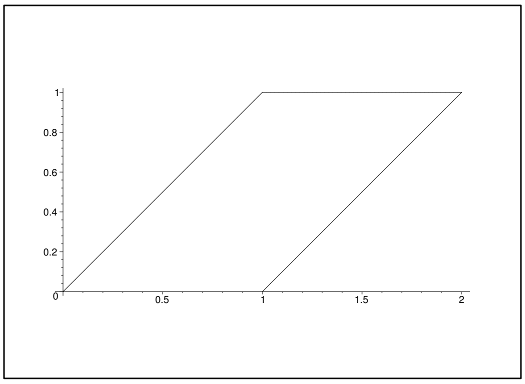


Figure 51: The domain D'.

Since $\tau(x_1, x_2) = (x_1 + x_2, x_2)$, i.e.

$$y_1 = x_1 + x_2, \qquad y_2 = x_2,$$

it follows that the inverse map τ^{-1} exists,

$$\tau^{-1}(y_1, y_2) = (y_1 - y_2, y_2),$$

thus

$$x_1 = y_1 - y_2, \qquad x_2 = y_2.$$

The conditions $0 < x_1 < 1$ and $0 < x_2 < 1$ can now be replaced by

$$0 < y_1 - y_2 < 1, \qquad 0 < y_2 < 1,$$

hence

$$y_2 < y_1 < y_2 + 1 \qquad \text{and} \qquad 0 < y_2 < 1,$$

and we have proved that τ maps $]0,1[\times]0,1[$ bijectively onto

$$D' = \left\{(y_1, y_2) \in \mathbb{R}^2 \mid 0 < y_2 < 1,\, y_2 < y_1 < y_2 + 1\right\}.$$

The *Jacobian* is then given by

$$J_\tau(y_1, y_2) = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$$

Then (Y_1, Y_2) i D' has the frequency

$$k(y_1, y_2) = h(y_1 - y_2, y_2) \cdot 1 = y_1,$$

thus

$$k(y_1, y_2) = \begin{cases} y_1, & \text{for } 0 < y_2 < 1 \text{ and } y_2 < y_1 < y_2 + 1, \\ 0, & \text{otherwise.} \end{cases}$$

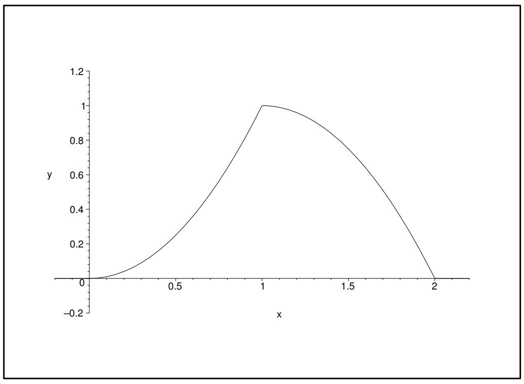


Figure 52: The graph of $F_{Y_1}(y_1)$.

Marginal frequencies.

1) We get for Y_1 by a vertical integration

a) If $0 < y_1 \leq 1$, then

$$f_{Y_1}(y_1) = \int_{y_2=0}^{y_1} k(y_1, y_2)\, dy_2 = \int_{y_2=0}^{y_1} y_1\, dy_2 = y_1^2.$$

b) If $1 < y_1 < 2$, then

$$f_{Y_1}(y_1) = \int_{y_2=y_1-1}^{1} k(y_1, y_2)\, dy_2 = \int_{y_2=y_1-1}^{1} y_1\, dy_2 = y_1(2-y_1),$$

hence summing up,

$$f_{Y_1}(y_1) = \begin{cases} y_1^2, & 0 < y_1 \le 1, \\ y_1(2-y_1) = 1-(y_1-1)^2, & 1 < y_1 < 2, \\ 0, & \text{otherwise.} \end{cases}$$

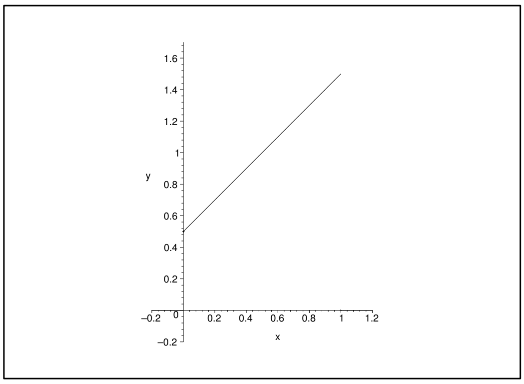


Figure 53: The graph of $f_{Y_2}(y_2)$.

2) For Y_2 it follows by a horizontal integration for $0 < y_2 < 1$ that

$$\begin{aligned} f_{Y_2}(y_2) &= \int_{y_1=y_2}^{y_2+1} k(y_1,y_2)\, dy_1 = \int_{y_1=y_2}^{y_2+1} y_1\, dy_1 = \left[\frac{1}{2}\, y_1^2\right]_{y_1=y_2}^{y_2+1} \\ &= \frac{1}{2}\left\{(y_2+1)^2 - y_2^2\right\} = \frac{1}{2}\,(2y_2+1)\cdot 1 = y_2 + \frac{1}{2}, \end{aligned}$$

thus summing up

$$f_{Y_2}(y_2) = \begin{cases} y_2 + \dfrac{1}{2}, & 0 < y_2 < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Example 8.7 *Assume that the 2-dimensional random variable* (X_1, X_2) *has the frequency*

$$h(x_1, x_2) = \begin{cases} 2\,\exp\left(-\left(x_1+x_2\right)\right), & (x_1, x_2) \in D, \\ 0, & \textit{otherwise.} \end{cases}$$

where

$$D = \left\{(x_1, x_2) \in \mathbb{R}^2 \mid 0 < x_2 < x_1 < \infty\right\},$$

and let $(Y_1, Y_2) = \tau(X_1, X_2)$ *be given by*

$$Y_1 = X_1 + X_2, \qquad Y_2 = (X_1 - X_2)^2 .$$

1) Prove that τ *maps* D *bijectively onto the domain*

$$D' = \left\{(y_1, y_2) \in \mathbb{R}^2 \mid 0 < y_1 < \infty,\, 0 < y_2 < y_1^2\right\}.$$

2) Find the frequency $k(y_1, y_2)$ *of* (Y_1, Y_2).

3) Find the marginal frequencies of Y_1 *and* Y_2.

4) Are Y_1 *and* Y_2 *independent random variables?*

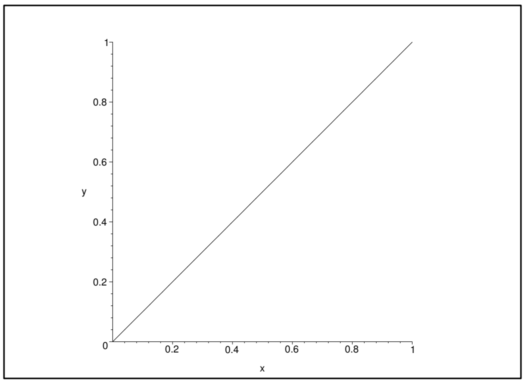


Figure 54: The domain D lies between the X_1 axis and the line $x_2 = x_1$.

1) It follows from

$$y_1 = x_1 + x_2 \quad \text{and} \quad y_2 = (x_1 - x_2)^2 \quad \text{og} \quad x_1 - x_2 > 0$$

that

$$x_1 + x_2 = y_1 \qquad \text{and} \qquad x_1 - x_2 = +\sqrt{y_2},$$

hence

$$x_1 = \frac{1}{2}\left\{y_1 + \sqrt{y_2}\right\} \quad \text{and} \quad x_2 = \frac{1}{2}\left\{y_1 - \sqrt{y_2}\right\}.$$

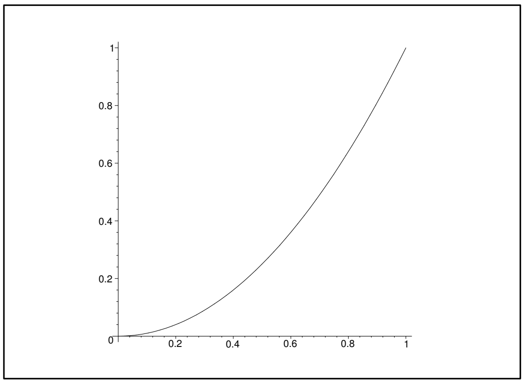


Figure 55: The domain D' lies between the Y_1 axis and the parabola $y_2 = y_1^2$.

Since (x_1, x_2) is uniquely determined by (y_1, y_2), we conclude that τ is bijective. The boundary curve $x_2 = x_1$ is mapped into $y_2 = 0$. The boundary curve $x_2 = 0$ is mapped into $y_1 = \sqrt{y_2}$, i.e. $y_2 = y_1^2$, $y_1 \geq 0$. Since $x_2 > 0$ in D, we must have $y_1 > \sqrt{y_2}$, thus $0 < y_2 < y_1^2$, and we have proved that τ maps D bijectively onto D'.

2) We next compute the Jacobian,

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\[2ex] \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{2} & \dfrac{1}{4}\,\dfrac{1}{\sqrt{y_2}} \\[2ex] \dfrac{1}{2} & -\dfrac{1}{4}\,\dfrac{1}{\sqrt{y_2}} \end{vmatrix} = -\frac{1}{4}\,\frac{1}{\sqrt{y_2}} < 0.$$

It follows from $x_1 + x_2 = y_1$ that the frequency of (Y_1, Y_2) is given by

$$k(y_1, y_2) = \begin{cases} 2\,e^{-y_1} \cdot \dfrac{1}{4}\,\dfrac{1}{\sqrt{y_2}} = \dfrac{1}{2}\,\dfrac{e^{-y_1}}{\sqrt{y_2}} & \text{for } (y_1, y_2) \in D', \\[2ex] 0, & \text{for } (y_1, y_2) \notin D'. \end{cases}$$

3) The marginal frequency of Y_1 is obtained by a vertical integration,

$$f_{Y_1}(y_1) = \frac{1}{2}\int_0^{y_1^2} \frac{e^{-y_1}}{\sqrt{y_2}}\,dy_2 = e^{-y_1}\,\left[\sqrt{y_2}\right]_0^{y_1^2} = y_1\,e^{-y_1} \quad \text{for } y_1 > 0,$$

and $f_{Y_1}(y_1) = 0$ for $y \leq 0$.

The frequency of Y_2 is obtained by a horizontal integration,

$$f_{Y_2}(y_2) = \frac{1}{2}\int_{\sqrt{y_2}}^{\infty} \frac{e^{-y_1}}{\sqrt{y_2}}\,dy_1 = \frac{1}{2}\,\frac{e^{-\sqrt{y_2}}}{\sqrt{y_2}} \qquad \text{for } y_2 > 0,$$

and $f_{Y_2}(y_2) = 0$ for $y_2 \leq 0$.

4) Since D' is not a rectangle, it follows immediately that Y_1 and Y_2 are not independent. It also follows from

$$f_{Y_1}(y_1) \cdot f_{Y_2}(y_2) \neq k(y_1, y_2).$$

Example 8.8 *Assume that X_1 and X_2 are independent identically distributed random variables of the frequency*

$$f(x) = \begin{cases} x\,e^{-x}, & x > 0, \\ \\ 0, & \textit{otherwise}. \end{cases}$$

1. *Compute the means $E\{X_1\}$ and $E\left\{\dfrac{1}{X_1}\right\}$.*

2. *Compute the probability $P\{X_2 > X_1\}$.*

Define the random variables Y_1 and Y_2 by

$$Y_1 = X_1 + X_2, \qquad Y_2 = \frac{X_1}{X_2}.$$

It will without proof be given that the vector function τ given by

$$\tau(x_1, x_2) = \left(x_1 + x_2, \frac{x_1}{x_2}\right)$$

maps $\mathbb{R}_+ \times \mathbb{R}_+$ bijectively onto itself.

3. *Find the simultaneous frequency $k(y_1, y_2)$ of (Y_1, Y_2).*

4. *Find the marginal frequencies of Y_1 and Y_2.*
(This question can be answered both with and without an application of the answer of question **3.***).*

5. *Check if Y_1 and Y_2 are independent.*

6. *Compute the mean $E\{Y_2\}$.*

7. *Find, e.g. by an application of question* **2.***, the median of Y_2.*

8. *Give an intuitive explanation of why the median of Y_2 is smaller than the mean of Y_2.*

1) The means are

$$E\{X_1\} = \int_0^\infty x^2 e^{-x}\,dx = 2 \qquad \text{and} \qquad E\left\{\frac{1}{X_1}\right\} = \int_0^\infty \frac{x}{x}\,e^{-x}\,dx = 1.$$

2) By the symmetry,

$$P\{X_2 > X_1\} = P\{X_1 > X_2\} = \frac{1}{2}\left(P\{X_2 > X_1\} + P\{X_1 > X_2\}\right) = \frac{1}{2}.$$

Alternatively, the frequency of $Y_2 = \dfrac{X_1}{X_2}$ is zero for $y_2 \le 0$, and when $y_2 > 0$, then

$$\begin{aligned} f_{Y_2}(y_2) &= \int_{-\infty}^\infty f(y_2 x)\,f(x)\cdot|x|\,dx = \int_0^\infty y_2 x\,e^{-y_2 x}\cdot x\,e^{-x}\cdot x\,dx \\ &= y_2 \int_0^\infty x^3 e^{-(1+y_2)x}\,dx = \frac{y_2}{(1+y_2)^4}\int_0^\infty t^3 e^{-t}\,dt = \frac{6y_2}{(1+y_2)^4}, \end{aligned}$$

hence

$$
\begin{aligned}
P\left\{X_2>X_1\right\} &= P\left\{\frac{X_1}{X_2}<1\right\}=\int_0^1 \frac{6y_2}{\left(1+y_2\right)^4}\,dy_2=6\int_0^1\left\{\frac{1}{\left(1+y_2\right)^3}-\frac{1}{\left(1+y_2\right)^4}\right\}dy_2\\
&= 6\left[-\frac{1}{2}\,\frac{1}{\left(1+y_2\right)^2}+\frac{1}{3}\,\frac{1}{\left(1+y_2\right)^3}\right]_0^1=-\frac{3}{2^2}+\frac{2}{2^3}+\frac{3}{1^2}-\frac{2}{1^3}\\
&= 1-\frac{3}{4}+\frac{1}{4}=\frac{1}{2}.
\end{aligned}
$$

3) It follows from $y_1=x_1+x_2$ and $y_2=\dfrac{x_1}{x_2}$ that

$$x_1=y_2x_2 \qquad \text{and} \qquad y_1=x_1+x_2=\left(y_2+1\right)x_2,$$

hence

$$x_1=\frac{y_1y_2}{y_2+1}=y_1-\frac{y_1}{y_2+1} \qquad \text{and} \qquad x_2=\frac{y_1}{y_2+1}.$$

The Jacobian is

$$\frac{\partial\left(x_1,x_2\right)}{\partial\left(y_1,y_2\right)}=\left|\begin{array}{cc}\dfrac{y_2}{y_2+1} & \dfrac{y_1}{\left(y_2+1\right)^2}\\[2ex] \dfrac{1}{y_2+1} & -\dfrac{y_1}{\left(y_2+1\right)^2}\end{array}\right|=-\frac{y_1}{\left(y_2+1\right)^3}\left(y_2+1\right)=-\frac{y_1}{\left(y_2+1\right)^2}<0.$$

The simultaneous frequency of $\left(X_1,X_2\right)$ is

$$g\left(x_1,x_2\right)=\begin{cases} x_1x_2\,e^{-\left(x_1+x_2\right)} & \text{for } x_1>0 \text{ og } x_2>0,\\[1ex] 0 & \text{otherwise,}\end{cases}$$

hence the simultaneous frequency of $\left(Y_1,Y_2\right)$ is 0 for $y_1\le 0$ or $y_2\le 0$, and

$$
\begin{aligned}
k\left(y_1,y_2\right) &= \frac{y_1y_2}{y_2+1}\cdot\frac{y_1}{y_2+1}\cdot e^{-y_1}\cdot\frac{y_1}{\left(y_2+1\right)^4}\\
&= \frac{y_1^3y_2}{\left(y_2+1\right)^4}\,e^{-y_1}, \qquad \text{for } y_1>0 \text{ and } y_2>0,
\end{aligned}
$$

which also can be written

$$k\left(y_1,y_2\right)=\begin{cases}\dfrac{1}{6}\,y_1^3e^{-y_1}\cdot\dfrac{6y_2}{\left(y_2+1\right)^4} & \text{for } y_1>0 \text{ and } y_2>0,\\[2ex] 0 & \text{otherwise.}\end{cases}$$

4) It follows from the rewriting of **3.** (possibly by the second variant of **2.**) that the marginal frequencies are

$$k_{Y_1}\left(y_1\right)=\begin{cases}\dfrac{1}{6}\,y_1^3e^{-y_1} & \text{for } y_1>0,\\[2ex] 0 & \text{for } y_1\le 0,\end{cases}$$

$$k_{Y_2}(y_2) = \begin{cases} \dfrac{6y_2}{(y_2+1)^4} & \text{for } y_2 > 0, \\ \\ 0 & \text{for } y_2 \le 0. \end{cases}$$

5) It is obvious that Y_1 and Y_2 are independent because

$$k(y_1, y_2) = k_{Y_1}(y_1) \cdot k_{Y_2}(y_2).$$

6) Since X_1 and X_2 are independent, the mean is

$$E\{Y_2\} = E\left\{\frac{X_1}{X_2}\right\} = E\{X_1\} \cdot E\left\{\frac{1}{X_2}\right\} = 2 \cdot 1 = 2.$$

7) By question **2**,

$$\frac{1}{2} = P\{X_2 > X_1\} = P\left\{\frac{X_1}{X_2} < 1\right\} = P\{Y_2 < 1\},$$

hence the median is $\langle Y_2 \rangle = 1$.

ALTERNATIVELY, Y_2 has the distribution function

$$\begin{aligned} K_{Y_2}(y_2) &= \int_0^{y_2} \frac{6t}{(t+1)^4}\,dt = \int_1^{y_2+1} \frac{6(u-1)}{u^4}\,du = 6\int_1^{y_2+1} \left\{\frac{1}{u^3} - \frac{1}{u^4}\right\} du \\ &= 6\left[-\frac{1}{2}\,\frac{1}{u^2} + \frac{1}{3}\,\frac{1}{u^3}\right]_1^{y_2+1} = 1 - \frac{3}{(y_2+1)^2} + \frac{2}{(y_2+1)^3} = 1 - \frac{3y_2+1}{(y_2+1)^3}. \end{aligned}$$

If we put $K_{Y_2}(y_2) = \dfrac{1}{2}$, then

$$\frac{3y_2+1}{(y_2+1)^3} = \frac{1}{2}, \qquad \text{dvs.} \qquad 6y_2 + 2 = (y_2+1)^3,$$

or

$$y_2^3 + 3y_2^2 - 3y_2 - 1 = (y_2 - 1)\left(y_2^2 + 4y_2 + 1\right) = 0.$$

The only positive solution is $y_2 = 1$, which is the median.

8) The mass of probability is divided into two equal parts by the median 1. However, the mass of probability is scattered more to the right of 1 then to the left of 1. Thus, the mean must lie to the right of 1.

9 Means and moments of higher order

Example 9.1 *Let X be a random variable of values in $\mathbb{N}_0$. Prove that $E\{X\}$ exists, if and only if*

$$\sum_{k=0}^{\infty} P\{X > k\} < \infty,$$

and that in the affirmative case,

$$E\{X\} = \sum_{k=0}^{\infty} P\{X > k\}.$$

It is often easier to apply this formula by computation of means.

First note that a *necessary and sufficient condition* of the existence of $E\{X\}$ and is equal to

$$E\{X\} = \sum_{n=0}^{\infty} n\, P\{X = n\} = \sum_{n=1}^{\infty} n\, P\{X = n\},$$

is that

$$\sum_{n=1}^{\infty} n\, P\{X = n\} \text{ is absolutely convergent.}$$

Since the terms are non-negative, this is equivalent with

$$\sum_{n=1}^{\infty} n\, P\{X=n\} \text{ is (just) convergent.}$$

If $\sum_{k=0}^{\infty} P\{X>k\}$ is convergent, then by a (correct) interchanging of the order of summation (because all terms are ≥ 0),

$$\sum_{k=0}^{\infty} P\{X>k\} = \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} P\{X=n\} = \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} P\{X=n\} = \sum_{n=1}^{\infty} n\, P\{X=n\} = E\{X\}.$$

If conversely $E\{X\}$ exists, we just repeat the computations above in the reverse order.

Example 9.2 *Two persons A and B play the following game:*
They each throw two coins. The winner is he who gets most heads in a throw. The game is a draw, if they obtain an equal number of heads.

1. *What is the probability q that the game is a draw?*

2. *What is the probability p_A that A wins?*

If the game is a draw, the game is continued in the same way. One stops first time one of the two players wins.

3. *What is the probability that A wins in game number k?*

4. *Find the mean of the number of games.*

$B \setminus A$	TT	TH	HT	HH
TT	0	1	1	1
TH	-1	0	0	1
HT	-1	0	0	1
HH	-1	-1	-1	0

Table 1: If A wins, we write 1. If B wins, we write -1. In case of a draw we write 0.

1) We write 1 if A wins, and -1 if B wins. Finally, 0 means a draw. Tail is denoted by T, and head by H. Since the 16 possibilities all have the same probability, we get by simply counting

$$q = P\{\text{a draw}\} = \frac{6}{16} = \frac{3}{8}.$$

2) It follows from (1) and the symmetry of A and B that

$$P\{A \text{ wins}\} = p_A = p_B = P\{B \text{ wins}\} = \frac{5}{16}.$$

3) If A wins in game number k, then the first $k-1$ games must all have been draws, hence

$$P\{A \text{ wins in the } k^{\text{th}} \text{ game}\} = q^{k-1} \cdot p_A = \left(\frac{3}{8}\right)^{k-1} \cdot \frac{5}{16}.$$

4) Let X denote the number of games. Then

$$P\{X=k\} = P\{A \text{ wins in game number } k\} + P\{B \text{ wins in game number } k\} = \frac{5}{8} \cdot \left(\frac{3}{8}\right)^{k-1}.$$

When $|x|<1$, then

$$\sum_{k=1}^{\infty} k\, x^{k-1} = \frac{d}{dx}\left\{\sum_{k=0}^{\infty} x^k\right\} = \frac{d}{dx}\left(\frac{1}{1-x}\right) = \frac{1}{(1-x)^2}.$$

Using this result we get the mean

$$E\{X\} = \frac{5}{8}\sum_{k=1}^{\infty} k\left(\frac{3}{8}\right)^{k-1} = \frac{5}{8} \cdot \frac{1}{\left(1-\frac{3}{8}\right)^2} = \frac{8}{5}.$$

Example 9.3 *A box contains N balls of the numbers from 1 to N. Select at random n balls with replacement. Let X_n denote the random variable which indicates the largest selected number.*
Find the distribution of X_n.
Find the mean $E\{X_n\}$, and prove for large N that this mean is approximately equal to $\dfrac{n}{n+1}\,N$.

Let X denote one selection at random of n numbers. Since all numbers have the same probability, the distribution function is given by

$$F_X(k) = P\{X \le k\} = \frac{k}{N}, \qquad k = 1,\, 2,\, \dots,\, N.$$

Thus we derive the distribution function of X_n,

$$F_{X_n}(k) = P\left\{\max_{j=1,\dots,n} X_j \le k\right\} = (P\{X \le k\})^n = \left(\frac{k}{N}\right)^n, \qquad k = 1,\, 2,\, \dots,\, N,$$

hence

$$p_k = P\{X_n = k\} = P\{X_n \le k\} - P\{X_n \le k-1\} = \frac{k^n-(k-1)^n}{N^n}.$$

The mean is

$$\begin{aligned} E\{X_n\} &= \sum_{k=1}^{N} k\, p_k = \frac{1}{N^n}\left\{\sum_{k=1}^{N} k^{n+1} - \sum_{k=1\,(=2)}^{N} k(k-1)^n\right\} = \frac{1}{N^n}\left\{\sum_{k=1}^{N} k^{n+1} - \sum_{k=1}^{N-1}(k+1)k^n\right\} \\ &= \frac{1}{N^n}\left\{N^{n+1} - \sum_{k=1} N-1k^n\right\} = N\left\{1 - \frac{1}{N}\sum_{k=1}^{N-1}\left(\frac{k}{N}\right)^n\right\}. \end{aligned}$$

Then notice that $\frac{1}{N}\sum_{k=1}^{N-1}\left(\frac{k}{N}\right)^n$ can be interpreted as an approximating sum of the integral $\int_0^1 x^n\,dx = \frac{1}{n+1}$, hence

$$\frac{1}{N}\sum_{k=1}^{N-1}\left(\frac{k}{N}\right)^n \to \int_0^1 x^n\,dx = \frac{1}{n+1} \qquad \text{for } N \to \infty.$$

Then

$$E\{X_n\} = N\left\{1 - \frac{1}{n+1} + \frac{1}{N}\,\varepsilon\left(\frac{1}{N}\right)\right\} \approx \frac{n}{n+1}\,N \qquad \text{for store } N.$$

Example 9.4 *Let X be a random variable of the distribution function $F(x)$, the frequency $f(x)$ and the mean μ. Prove that*

$$\mu = \int_0^\infty \{1 - F(x)\}dx - \int_{-\infty}^0 F(x)\,dx.$$

It is given that $\int_{-\infty}^\infty |x|\,f(x)\,dx < \infty$, and that

$$(6)\;\; \mu = \int_{-\infty}^\infty x\,f(x)\,dx = \int_{-\infty}^0 x\,f(x)\,dx + \int_0^\infty x\,f(x)\,dx.$$

Let $A > 0$. Then

$$\int_{-A}^0 x\,f(x)\,dx = [x\,F(x)]_{-A}^0 - \int_{-A}^0 F(x)\,dx = A\,F(-A) - \int_{-A}^0 F(x)\,dx,$$

and

$$\int_0^A x\,f(x)\,dx = [x\{F(x)-1\}]_0^A + \int_0^A \{1 - F(x)\}\,dx = -A\{1 - F(A)\} + \int_0^A (1 - F(x))\,dx.$$

Since

$$0 \le A\,F(-A) = A\int_{-\infty}^{-A} f(x)\,dx \le \int_{-\infty}^{-A} |x|\,f(x)\,dx \to 0 \quad \text{for } A \to \infty,$$

we conclude that $A\,F(-A) \to 0$ for $A \to \infty$. Since $\int_{-\infty}^0 x\,f(x)\,dx$ is absolutely convergent, it follows by taking the limit $A \to \infty$ that

$$\int_{-\infty}^0 x\,f(x)\,dx = \lim_{A\to\infty}\left\{A\,F(-A) - \int_{-A}^0 F(x)\,dx\right\} = -\int_{-\infty}^0 F(x)\,dx.$$

Analogously,

$$0 \le A\{1 - F(A)\} = A\int_A^\infty f(x)\,dx \le \int_A^\infty |x|\,f(x)\,dx \to 0 \quad \text{for } A \to \infty,$$

so when $A \to \infty$, we conclude in the same way that

$$\int_0^{\infty} x\, f(x)\, dx = \lim_{A \to \infty} \left\{ -A\{1 - F(A)\} + \int_0^A \{1 - F(x)\}\, dx \right\} = \int_0^{\infty} \{1 - F(x)\}\, dx,$$

where the integrals are even absolutely convergent.

Finally, by insertion into (6) we get

$$\mu = \int_{-\infty}^0 x\, f(x)\, dx + \int_0^{\infty} x\, f(x)\, dx = \int_0^{\infty} \{1 - F(x)\}\, dx - \int_{-\infty}^0 F(x)\, dx$$

as required.

Alternatively, a more streamlined, though also more sophisticated method is the following. We see that

$$\int_0^{\infty} \{1 - F(x)\}\, dx = \int_{x=0}^{\infty} \left\{ \int_{y=x}^{\infty} f(y)\, dy \right\} dx = \int_{y=0}^{\infty} f(y) \left\{ \int_{x=0}^{y} 1\, dx \right\} dy = \int_0^{\infty} y\, f(y)\, dy,$$

and

$$\begin{aligned} -\int_{-\infty}^0 F(x)\, dx &= -\int_{x=-\infty}^{0} \left\{ \int_{y=-\infty}^{x} f(y)\, dy \right\} dx = \int_{y=-\infty}^{0} f(y) \left\{ \int_{x=y}^{0} (-1)\, dx \right\} dy \\ &= \int_{-\infty}^0 y\, f(y)\, dy. \end{aligned}$$

By adding these two expressions, we formally obtain that

$$\int_0^{\infty} \{1 - F(x)\}\, dx - \int_{-\infty}^{0} F(x)\, dx = \int_{-\infty}^{\infty} y\, f(y)\, dy = \mu.$$

However, we have assumed that the mean exists, which implies that all the integrals above are absolutely convergent, so the formal calculation is also real.

Example 9.5 *Let X be an non-negative random variable of the distribution function $F(x)$ and frequency $f(x)$. Prove that*

$$(7)\quad E\left\{X^k\right\} = k \int_0^{\infty} x^{k-1}\{1 - F(x)\}\, dx, \qquad k \in \mathbb{N}.$$

(If the k^{th} moment does not exist, then both the right hand side and the left hand side of (7) are equal to ∞.)
Find a similar formula, if the random variable X is non-positive.
Remark. *One can prove that formula (7) holds for every non-negative random variable X.*

1) Assume that X is non-negative and that $E\left\{X^k\right\}$ exists, i.e.

$$0 \le E\left\{X^k\right\} = \int_0^{\infty} x^k f(x)\, dx < \infty.$$

Let $A > 0$. Then by a partial integration,

$$\begin{aligned}\int_0^A x^k f(x)\, dx &= \left[x^k\{F(x) - 1\}\right]_0^A + k \int_0^A x^{k-1}\{1 - F(x)\}\, dx \\ &= -A^k\{1 - F(A)\} + k \int_0^A x^{k-1}\{1 - F(x)\}\, dx.\end{aligned}$$

Now $\int_0^{\infty} x^k f(x)\, dx < \infty$, so we get the estimate

$$0 \le A^k\{1 - F(A)\} = A^k \int_A^{\infty} f(x)\, dx \le \int_A^{\infty} x^k\, f(x)\, dx \to 0 \quad \text{for } A \to \infty.$$

Then by taking the limit $A \to \infty$,

$$E\left\{X^k\right\} = \int_0^{\infty} x^k f(x)\, dx = k \int_0^{\infty} x^{k-1}\{1 - F(x)\}\, dx.$$

It is trivial that

$$\int_0^A x^k f(x)\, dx \le k \int_0^A x^{k-1}\{1 - F(x)\}\, dx,$$

so we infer that if $E\left\{X^k\right\} = \infty$, then $k \int_0^{\infty} x^{k-1}\{1 - F(x)\}\, dx = \infty$.

2) Then assume that X is non-positive and that $E\left\{X^k\right\}$ exists, i.e.

$$0 \le \left|E\left\{X^k\right\}\right| = \int_{-\infty}^{0} |x|^k f(x)\,dx < \infty.$$

If $A > 0$, then

$$\int_{-A}^{0} x^k f(x)\,dx = \left[x^k F(x)\right]_{-A}^{0} - k\int_{-A}^{0} x^{k-1}F(x)\,dx = -(-A)^k F(-A) - k\int_{-A}^{0} x^{k-1}F(x)\,dx.$$

Since

$$0 \le A^k F(-A) = A^k \int_{-\infty}^{-A} f(x)\,dx \le \int_{-\infty}^{-A} |x|^k f(x)\,dx \to 0 \quad \text{for } A \to \infty,$$

it follows by taking the limit that

$$(8)\ \ E\left\{X^k\right\} = \int_{-\infty}^{0} x^k f(x)\,dx = -k\int_{-\infty}^{0} x^{k-1}F(x)\,dx.$$

Clearly,

$$\int_{-A}^{0} x^k f(x)\,dx, \quad (-A)^k F(-A) \quad \text{og} \quad -k\int_{-A}^{0} x^{k-1}F(x)\,dx$$

have all the same sign. Hence,

$$\left|\int_{-A}^{0} x^k f(x)\,dx\right| \le \left|k\int_{-A}^{0} x^{k-1}F(x)\,dx\right|.$$

Therefore, if $E\left\{X^k\right\}$ does not exist, then both integrals of (8) are divergent.

Alternatively and more streamlined (and also more sophisticated), because one at first does not care so much for the convergence of the integrals (this should of course be done at last), we have the following proof:
When $k \in \mathbb{N}$, and X is non-negative of the distribution function $F(x)$ and the frequency $f(x)$, then

$$\begin{aligned} k\int_{0}^{\infty} x^{k-1}\{1-F(x)\}\,dx &= k\int_{x=0}^{\infty} x^{k-1}\left\{\int_{y=x}^{\infty} f(y)\,dy\right\}dx \\ &= \int_{y=0}^{\infty} f(y)\left\{\int_{x=0}^{y} k\,x^{k-1}\,dx\right\}dy = \int_{0}^{\infty} y^k f(y)\,dy = E\left\{X^k\right\}. \end{aligned}$$

Then notice that if the k^{th} moment does not exist, then all the integrals involved are divergent of the value ∞. Since the integrand is non-negative, we can interchange the order of integration. Conversely, if the k^{th} moment exists, then all the involved integrals are convergent with a non-negative integrand, hence uniformly convergent, and all the computations are legal.

Then let $X \le 0$ have the distribution function $F(x)$ and the frequency $f(x)$. Then for $k \in \mathbb{N}$,

$$\begin{aligned} -k\int_{-\infty}^{0} x^{k-1}F(x)\,dx &= -k\int_{x=-\infty}^{0} x^{k-1}\left\{\int_{y=-\infty}^{x} f(y)\,dy\right\}dx \\ &= \int_{y=-\infty}^{0} f(y)\left\{\int_{x=y}^{0}\left(-k\,x^{k-1}\right)dx\right\}dy = \int_{-\infty}^{0} y^k f(y)\,dy = E\left\{X^k\right\}. \end{aligned}$$

Thus we get in this case the formula

$$E\left\{X^k\right\} = -k \int_{-\infty}^{0} x^{k-1} F(x)\,dx.$$

We still have to remark that in the case of divergence the left hand side and the right hand side are either both $-\infty$ or both $+\infty$. It the k^{th} moment exists, then all integrals are absolutely convergent.

Example 9.6 *Let X and Y be non-negative random variables of distribution functions F_X and F_Y, means $E\{X\}$ and $E\{Y\}$ variances $V\{X\}$ and $V\{Y\}$.*
The random variable X is said to be stochastically larger than Y, if

$$F_X(x) \le F_Y(x) \qquad \textit{for all } x \in \mathbb{R}.$$

1) If so, prove that $E\{Y\} \le E\{X\}$.

2) Can one also conclude that $V\{Y\} \le V\{X\}$?

1) Since X and Y are non-negative, it follows from **Example 9.4** or **Example 9.5** that

$$E\{Y\} = \int_0^\infty \{1 - F_Y(x)\}\,dx \le \int_0^\infty \{1 - F_X(x)\}\,dx = E\{X\}.$$

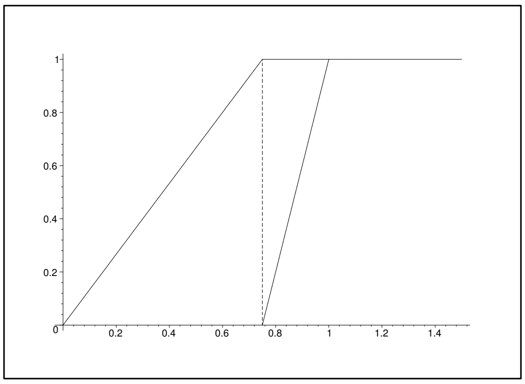


Figure 56: Illustration of $F_X(x) \le F_Y(x)$ in (2).

2) The answer is "no"! We construct a counterexample. Let $a \in \left]\frac{1}{2}, 1\right[$, and let X be rectangularly distributed over $]a, 1[$, and let Y be rectangularly distributed over $]0, a[$. Then clearly X and Y are non-negative, and $F_X(x) \le F_Y(y)$, cf. the figure, hence X is stochastically larger than Y.

This is in agreement with

$$E\{Y\} = \frac{a}{2} < \frac{1+a}{2} = E\{X\}.$$

For the variances, however, we get

$$V\{Y\} = \frac{a^2}{12} > \frac{(1-a)^2}{12} = V\{X\},$$

because $a > 1 - a > 0$ for $\frac{1}{2} < a < 1$.

Example 9.7 *Let X be a random variable satisfying*

$$E\{X\} = E\left\{X^2\right\} = 1.$$

Find the distribution function of X.

Since

$$V\{X\} = E\left\{X^2\right\} - (E\{X\})^2 = 1 - 1 = 0,$$

it follows that X is a constant, and since $E\{X\} = 1$, we get $X \equiv 1$. The distribution function is then

$$F_X(x) = \begin{cases} 1 & \text{for } x \geq 1, \\ 0 & \text{for } x < 1. \end{cases}$$

Example 9.8 *Let X be a random variable, for which the k^{th} moment exists for some $k \in \mathbb{N} \setminus \{1\}$. Prove the following generalization of Chebyshev's inequality:*
For every $a \in \mathbb{R}_+$,

$$P\{|X - E\{X\}| \geq a\} \leq \frac{1}{a^k} E\left\{|X - E\{X\}|^k\right\}.$$

When $E\left\{X^k\right\}$ exists for some $k \in \mathbb{N} \setminus \{1\}$, then $E\left\{X^j\right\}$ exists for every $j = 1, \dots, k$. In fact, $|x|^j \leq 1$ for $|x| \leq 1$, and $|x|^j \leq |x|^k$ for $|x| \geq 1$, so

$$\int_{-\infty}^{\infty} |x|^j f(x)\, dx \leq \int_{\{|x|\leq 1\}} 1 \cdot f(x)\, dx + \int_{\{|x|\geq 1\}} |x|^k f(x)\, dx < \infty.$$

In particular, $E\{X\}$ exists.

Then

$$E\left\{|X - E\{X\}|^k\right\} = \int_{-\infty}^{\infty} |x - E\{X\}|^k f(x)\, dx \geq \int_{\{|x-E\{X\}\geq a\}} a^k f(x)\, dx = a^k P\{|X - E\{X\}| \geq a\},$$

and the inequality follows by a division by a^k.

Example 9.9 *Let X be a random variable.*

1) Let $g : \mathbb{R} \to \mathbb{R}$ be an even, non-negative function, which is increasing on $[0, \infty[$. Prove that

$$P\{|X| \geq a\} \leq \frac{E\{g(X)\}}{g(a)} \qquad \textit{for every } a \in \mathbb{R}_+.$$

2) Let $g : \mathbb{R} \to \mathbb{R}$ be a non-negative, increasing function. Prove that

$$P\{X \geq a\} \leq \frac{E\{g(X)\}}{g(a)} \qquad \textit{for every } a \in \mathbb{R}.$$

We always assume that $E\{g(X)\}$ exists.

In the main proof we assume that X is of continuous type, i.e.

$$E\{g(X)\} = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx.$$

The similar proofs when X is either of discrete type or of mixed type are obtained by simple modifications of the main proof.

1) First by a splitting

$$P\{|X| \geq a\} = P\{X \geq a\} + P\{X \leq -a\} = \int_a^{\infty} f(x)\, dx + \int_{-\infty}^{-a} f(x)\, dx.$$

According to the assumptions, $g(x) \geq 0$ and

$$g(a) \leq g(x) \text{ for } x \geq a \quad \text{and} \quad g(-a) = g(a) \leq g(x) \text{ for } x \leq -a,$$

hence

$$g(a)P\{|X| \geq a\} = g(a) \int_a^{\infty} f(x)\,dx + g(-a) \int_{-\infty}^{-a} f(x)\,dx$$
$$\leq \int_a^{\infty} g(x)\,f(x)\,dx + \int_{-\infty}^{-a} g(x)\,f(x)\,dx \leq \int_{-\infty}^{\infty} g(x)\,f(x)\,dx = E\{g(X)\}.$$

If $g(a) \neq 0$, the result is obtained by a division by $g(a) > 0$.

2) Similarly,

$$g(a)P\{X \geq a\} = g(a) \int_a^{\infty} f(x)\,dx \leq \int_a^{\infty} g(x)\,f(x)\,dx \leq \int_{-\infty}^{\infty} g(x)\,f(x)\,dx = E\{g(X)\}.$$

If $g(a) \neq 0$, the result follows by a division by $g(a) > 0$.

Example 9.10 *A random variable X is assumed to have the mean μ and the variance σ^2. Prove that if a is a median of X, then*

$$|a - \mu| \leq \sqrt{2} \cdot \sigma.$$

HINT. *Apply Chebyshev's inequality.*

When a is a median, then

$$F(a) \geq \frac{1}{2} \qquad \text{and} \qquad F(a-) \leq \frac{1}{2}.$$

One of the sets $\{x \leq a\}$ and $\{x \geq a\}$ must necessarily be contained in the set $\{|x - \mu| \geq |a - \mu|\}$, and since $P\{X \leq a\} \geq \frac{1}{2}$ and $P\{X \geq a\} \geq \frac{1}{2}$, we get

$$\frac{1}{2} \leq P\{|X - \mu| \geq |a - \mu|\} \leq \frac{\sigma^2}{|a - \mu|^2}, \text{for } a \neq \mu.$$

Then by a rearrangement,

$$|a - \mu| \leq \sqrt{2} \cdot \sigma.$$

If $a = \mu$, there is of course nothing to prove.

Example 9.11 *Let X be a random variable, for which all moments $m_k = E\left\{X^k\right\}$ exist. We define for $k \in \mathbb{N}$ the k^{th} decreasing moment*

$$m_{(k)} = E\{X(X-1)(X-2)\cdots(X-k+1)\}.$$

1) Find for $n = 1,\ 2,\ 3$, the n^{th} decreasing as a linear combination of the k^{th} moments $k \leq n$.

2) Find for $n = 1,\ 2,\ 3$, the n^{th} moment as a linear combination of the k^{th} decreasing moments for $k \leq n$.

A. If $n = 1$, then

$$m_{(1)} = E\{X\} = m_1.$$

B. If $n = 2$, then

$$m_{(2)} = E\{X(X-1)\} = E\left\{X^2\right\} - E\{X\} = m_2 - m_1,$$

and hence

$$m_2 = m_{(2)} + m_{(1)}.$$

C. If $n = 3$, then

$$\begin{aligned} m_{(3)} &= E\{X(X-1)(X-2)\} = E\left\{X^3\right\} - 3E\left\{X^2\right\} + 2E\{X\} \\ &= m_3 - 3m_2 + 2m_1, \end{aligned}$$

and conversely,

$$m_3 = m_{(3)} + 3m_2 - 2m_{(1)} = m_{(3)} + 3m_{(2)} + m_{(1)}.$$

Example 9.12 *Let X be a random variable, for which all moments $m_k = E\left\{X^k\right\}$ exist. It is then well-known that all the central moments*

$$v_k = E\left\{(X-m_1)^k\right\}.$$

also exist.

1) Express for $n = 2,\ 3,\ 4$, the n^{th} central moment by the k^{th} moments for $k \leq n$.

2) Express for $n = 2,\ 3,\ 4$, the n^{th} moment by the k^{th} central moments for $k \leq n$ and by m_1.

A. If $n = 2$, then

$$v_2 = E\left\{(X-m_1)^2\right\} = E\left\{X^2 - 2m_1X + m_1^2\right\} = m_2 - m_1^2,$$

and thus

$$m_2 = v_2 + m_1^2.$$

B. If $n = 3$, then

$$\begin{aligned} v_3 &= E\left\{(X-m_1)^3\right\} = E\left\{X^3 - 3m_1X^2 + 3m_1X^2 + 3m_1^2X - m_1^3\right\} \\ &= m_3 - 3m_1m_2 + 2m_1^3, \end{aligned}$$

and thus

$$m_3 = v_3 + 3m_1\left(v_2 + m_1^2\right) - 2m_1^3 = v_3 + 3m_1v_2 + m_1^3.$$

C. If $n = 4$, then

$$\begin{aligned} v_4 &= E\left\{(X-n_1)^4\right\} = E\left\{X^4 - 4m_1X^3 + 6m_1^2X^2 - 4m_1^3X + m_1^4\right\} \\ &= m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4, \end{aligned}$$

and thus

$$\begin{aligned} m_4 &= v_4 + 4m_1\left\{v_3 + 3m_1v_2 + m_1^3\right\} - 6m_1^2\left\{v_2 + m_1^2\right\} + 3m_1^4 \\ &= v_4 + 4m_1v_3 + 6m_1^2v_2 + m_1^4. \end{aligned}$$

Example 9.13 *A real function φ defined on an open interval I is said to be convex, if for all $a \in I$ there exists a real number c_a, such that*

$$\varphi(x) - \varphi(a) \geq c_a(x-a) \qquad \textit{for every } x \in I$$

(if φ is differentiable at the point a, then $c_a = \varphi'(a)$).
Let X be a random variable, which only has values in the open interval I, and let $\varphi : I \to \mathbb{R}$ be convex. Assuming that both $E\{X\}$ and $E\{\varphi(X)\}$ exist, prove Jensen's inequality

$$E\{\varphi(X)\} \geq \varphi(E\{X\}).$$

1) Let us first check where the assumption of convexity can be applied. Hence, we insert as a test any $a \in I$,

$$E\{\varphi(X)\} - \varphi(a) = E\{\varphi(X) - \varphi(a)\} \geq c_a\, E\{X-a\} = c_a(E\{X\} - a).$$

2) Since $E\{X\} \in I$, we get by choosing $a = E\{X\}$,

$$E\{\varphi(X)\} - \varphi(E\{X\}) \geq 0, \quad \text{i.e.} \quad E\{\varphi(X)\} \geq \varphi(E\{X\}).$$

10 Mean and variance in special cases

Example 10.1 *A random variable X has the frequency*

$$f(x) = \begin{cases} a\,e^{-ax}, & x > 0, \\ \\ 0, & x \le 0, \end{cases}$$

where a is a positive constant.
Find the frequency of $Y = X^2$, and find the mean and variance of Y.

The distribution function $G(y)$ of $Y = X^2$ is 0 for $y \le 0$. If $y > 0$, then

$$\begin{aligned} G(y) &= P\{Y \le y\} = P\left\{X^2 \le y\right\} = P\left\{-\sqrt{y} \le X \le \sqrt{y}\right\} = P\left\{X \le \sqrt{y}\right\} \\ &= \int_0^{\sqrt{y}} a\,e^{-ax}\,dx = 1 - e^{-a\sqrt{y}}, \qquad \text{for } y > 0. \end{aligned}$$

Thus, the corresponding frequency is

$$g(y) = \begin{cases} G'(y) = \dfrac{a}{2\sqrt{y}}\,\exp\left(-a\sqrt{y}\right), & \text{for } y > 0, \\ \\ 0, & \text{for } y \le 0. \end{cases}$$

Then the mean of Y is

$$E\{Y\} = E\left\{X^2\right\} = \int_0^\infty x^2 \cdot a\,e^{-ax}\,dx = \frac{1}{a^2}\int_0^\infty t^2\,e^{-t}\,dt = \frac{2}{a^2}.$$

The variance of Y is

$$\begin{aligned} V\{Y\} &= E\left\{Y^2\right\} - (E\{Y\})^2 = E\left\{X^4\right\} - \left(\frac{2}{a^2}\right)^2 = \int_0^\infty x^4\,a\,e^{-ax}\,dx - \frac{4}{a^4} \\ &= \frac{1}{a^4}\int_0^\infty t^4\,e^{-t}\,dt - \frac{4}{a^4} = \frac{4!}{a^4} - \frac{4}{a^4} = \frac{24-4}{a^4} = \frac{20}{a^4}. \end{aligned}$$

Example 10.2 *Let X be rectangularly distributed over $]-h, h[$.*
Compute for $k \in \mathbb{N}$ the moments

$$E\left\{X^k\right\} \qquad \textit{and} \qquad E\left\{|X|^k\right\}.$$

We first notice that

$$E\left\{|X|^k\right\} = \frac{1}{2h}\int_{-h}^{h} |x|^k\,dx = \frac{1}{h}\int_0^h x^k\,dx = \frac{h^k}{k+1}.$$

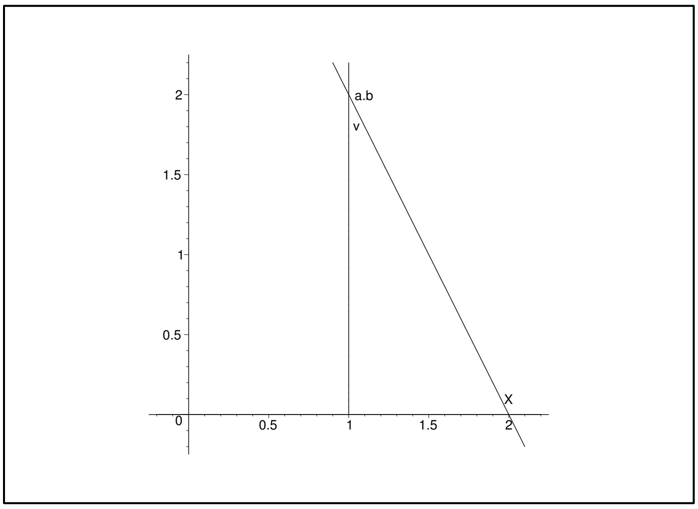
Then by the symmetry,

$$E\left\{X^{2k+1}\right\} = 0, \qquad k \in \mathbb{N}_0, \quad 2k+1 \text{ odd},$$

and

$$E\left\{X^{2k}\right\} = E\left\{|X|^{2k}\right\} = \frac{h^{2k}}{2k+1}.$$

Example 10.3 *Draw from the point (a,b), where $b>0$, a line of the angle θ with the line $x=a$, where $\theta \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$. The line intersects the x axis at a point of the abscissa X.*



Assuming that $\Theta(=v)$ is rectangularly distributed over $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, find the frequency of X.
Check if X has a mean.
The distribution of X is called a Cauchy distribution.

The relationship between the random variables X and Θ is

$$\tan\Theta = \frac{X-a}{b}, \quad \text{i.e.} \quad \Theta = \tau^{-1}(X) = \text{ Arctan } \frac{X-a}{b},$$

[and $X = \tau(\Theta) = a + b\,\tan\Theta$]. Since Θ has the frequency

$$f(\theta) = \begin{cases} \dfrac{1}{\pi}, & -\dfrac{\pi}{2} < \theta < \dfrac{\pi}{2}, \\ 0, & \text{otherwise,} \end{cases}$$

and $\tau^{-1}(x) \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ for every $x \in \mathbb{R}$, we derive that the frequency of X is

$$g(x) = f\left(\tau^{-1}(x)\right) \cdot \left|\frac{d\theta}{dx}\right| = \frac{1}{\pi} \cdot \frac{1}{1+\left(\frac{x-a}{b}\right)^2} \cdot \frac{1}{b} = \frac{b}{\pi\left\{b^2+(x-a)^2\right\}}, \qquad x \in \mathbb{R}.$$

It follows from

$$\int x\,g(x)\,dx = \frac{b}{\pi}\int \frac{x-a+a}{b^2+(x-a)^2}\,dx = \frac{b}{2\pi}\,\ln\left\{b^2+(x-a)^2\right\} + \frac{a}{\pi}\,\text{Arctan}\left(\frac{x-a}{b}\right),$$

that

$$\int_0^A x\cdot g(x)\,dx \to \infty \quad \text{for } A\to\infty, \qquad \text{(improper integral)},$$

so X has *no mean.*

ALTERNATIVELY,

$$|x|\cdot g(x) \sim k\cdot\frac{1}{|x|} \qquad \text{for large } x,$$

hence

$$\int_{-\infty}^{\infty} |x|\,g(x)\,dx = \infty.$$

Example 10.4 *A line segment of length 1 is divided randomly into two parts of lengths X and $1-X$, where we assume that X is rectangularly distributed over $]0,1[$.*
Let Y denote the length of the smallest of the two line segments, and let Z denote the length of the largest line segment.

1) Find the distribution of Y and the distribution of Z.

2) Find the mean and variance of Z.

1) The distribution function of X is

$$F_X(x) = \begin{cases} 0, & \text{for } x\le 0,\\[2mm] x, & \text{for } 0<x<1,\\[2mm] 1, & \text{for } x\ge 1.\end{cases}$$

Obviously, $Y=\min\{X,1-X\}$ takes its values in $\left]0,\frac{1}{2}\right]$. If $y\in\left]0,\frac{1}{2}\right]$, then

$$P\{X\le y \text{ or } 1-X\le y\} = P\{0<X\le y\}+P\{1-y\le X<1\} = 2y,$$

hence

$$F_Y(y) = \begin{cases} 1 & \text{for } y\ge\frac{1}{2},\\[2mm] 2y & \text{for } 0<y<\frac{1}{2},\\[2mm] 0 & \text{for } y\le 0,\end{cases} \qquad \text{and} \quad f_Y(y) = \begin{cases} 2 & \text{for } 0<y<\frac{1}{2},\\[2mm] 0 & \text{otherwise.}\end{cases}$$

Analogously, $Z = \max\{X, 1-X\}$ takes its values in $\left[\frac{1}{2}, 1\right[$. If $z \in \left[\frac{1}{2}, 1\right[$, then

$$F_Z(z) = P\{X \le z \text{ and } 1 - X \le z\} = P\{1 - z \le X \le z\} = 2z - 1,$$

hence

$$F_Z(z) = \begin{cases} 1 & \text{for } z \ge 1, \\ 2z-1 & \text{for } \frac{1}{2} < z < 1, \\ 0 & \text{for } z \le \frac{1}{2}, \end{cases} \qquad \text{and} \quad f_Z(z) = \begin{cases} 2 & \text{for } \frac{1}{2} < z < 1, \\ 0 & \text{otherwise.} \end{cases}$$

2) Since Y is rectangularly distributed over $\left]0, \frac{1}{2}\right[$, we get

$$E\{Y\} = \frac{1}{4} \qquad \text{and} \qquad V\{Y\} = \frac{1}{12} \cdot \left(\frac{1}{2}\right)^2 = \frac{1}{48}.$$

Now, Z is rectangularly distributed over $\left]\frac{1}{2}, 1\right[$, so

$$E\{Z\} = \frac{3}{4} \qquad \text{and} \qquad V\{Z\} = \frac{1}{12} \cdot \left(\frac{1}{2}\right)^2 = \frac{1}{48}.$$

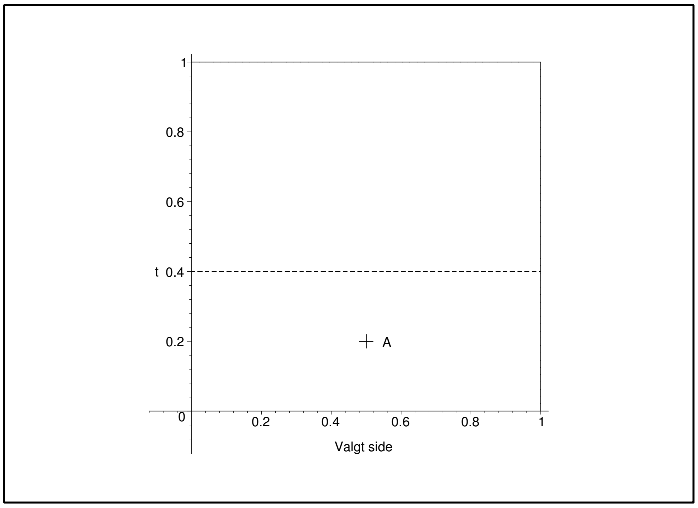
Example 10.5 *A point A is chosen randomly in the unit square $]0,1[\times]0,1[$. This means that if X and Y denote the abscissa and the ordinate, resp., of the point A, then (X,Y) has the simultaneous frequency*

$$f(x,y)=\begin{cases}1, & \textit{in }]0,1[\times]0,1[,\\[2mm] 0, & \textit{otherwise.}\end{cases}$$

1) *Find the probability of the event that the distance from A to a given edge of the square is $\le t$.*

2) *Let U denote the distance from A to the* closest *edge of the square. Find the distribution function and the frequency of U.*

3) *Find the mean and the variance of U.*



1) Obviously, the probability that the distance from A to e.g. $]0,1[$ on the x axis, is $<t$ for $0<t<1$, hence, the corresponding random variable is rectangularly distributed over $]0,1[$.

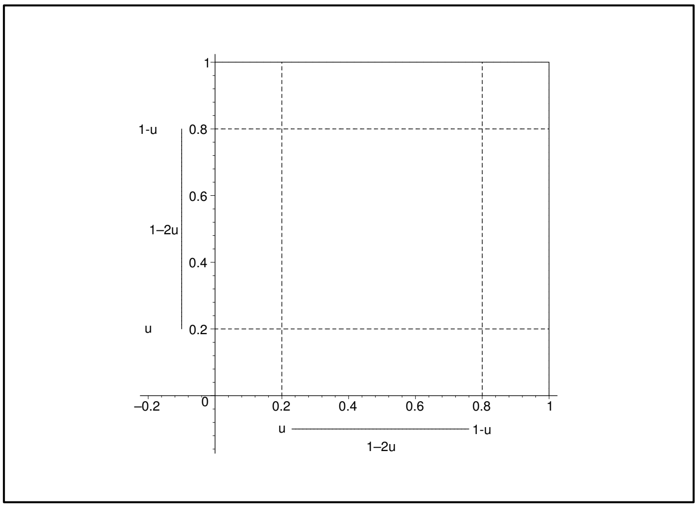
2) If U denotes the distance from A to the closest edge, then U has its values in $]0,\frac{1}{2}[$. If $u\in]0,\frac{1}{2}[$, then $U\le u$, if and only if A lies in the unions of the domains between a dotted line and the closest parallel edge, so by considering an area,

$$F_U(u)=1-(1-2u)^2=4\left(u-u^2\right)=4y-4u^2,\qquad 0<u<\frac{1}{2}.$$

We conclude that the frequency is

$$f_U(u)=F_U'(u)=4-8u,\qquad 0<u<\frac{1}{2},$$

and $f_U(u)=0$ otherwise.



3) The mean is

$$E\{U\} = \int_0^{\frac{1}{2}} u(4-8u)\,du = \int_0^{\frac{1}{2}} \left(4u - 8u^2\right) du = \left[2u^2 - \frac{8}{3}\,u^3\right]_0^{\frac{1}{2}} = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}.$$

Then compute

$$E\left\{U^2\right\} = \int_0^{\frac{1}{2}} u^2(4-8u)\,du = \int_0^{\frac{1}{2}} \left(4u^2 - 8u^3\right) du = \left[\frac{4}{3}\,u^3 - 2u^4\right]_0^{\frac{1}{2}} = \frac{1}{6} - \frac{1}{8} = \frac{1}{24},$$

so the variance is

$$V\{U\} = E\left\{U^2\right\} - (E\{U\})^2 = \frac{1}{24} - \frac{1}{36} = \frac{1}{72}.$$

Example 10.6 *The function f is for $0 < x < 1$ given by*

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}},$$

while the function is equal to 0 for any other value of x.

1) Prove that $f(x)$ is the frequency of a random variable X.

2) Find the mean and the variance of the random variable X.

3) Find the frequency of the random variable $Y = \sqrt{X}$.

4) Find the mean of the random variable Y.

The distribution of X is called the Arcussinus distribution.

1) Obviously, $f(x) \geq 0$ for every $x \in \mathbb{R}$. Then

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\,dx &= \frac{1}{\pi}\int_0^1 \frac{dx}{\sqrt{x(1-x)}} = \frac{1}{\pi}\int_{-\frac{1}{2}} \frac{1}{2}\frac{dy}{\sqrt{\left(y+\frac{1}{2}\right)\left(\frac{1}{2}-y\right)}} = \frac{1}{\pi}\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{dy}{\sqrt{\frac{1}{4}-y^2}} \\ &= \frac{1}{\pi}\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{2\,dy}{\sqrt{1-(2y)^2}} = \frac{1}{\pi}\int_{-1}^{1} \frac{dt}{\sqrt{1-t^2}} = \frac{1}{\pi}\left[\text{Arcsin } t\right]_{-1}^{1} = 1,\end{aligned}$$

thus $f(x)$ is the frequency of a random variable.

2) Since $f(x) = 0$ outside a bounded interval, all moments exist. In particular,

$$\begin{aligned}E\{X\} &= \int_0^1 x\,f(x)\,dx = \frac{1}{\pi}\int_0^1 \frac{x-\frac{1}{2}+\frac{1}{2}}{\sqrt{x-x^2}}\,dx \\ &= -\frac{1}{2\pi}\int_{x=0}^1 \frac{d\left(x^2-x\right)}{\sqrt{x-x^2}} + \frac{1}{2}\cdot\int_0^1 f(x)\,dx = 0 + \frac{1}{2} = \frac{1}{2},\end{aligned}$$

which can also be seen graphically, because the graph of $f(x)$ is clearly symmetric with respect to the line $x = \frac{1}{2}$.
Furthermore,

$$E\{X(X-1)\} = \frac{1}{\pi}\int_0^1 \frac{x(x-1)}{\sqrt{x(1-x)}}\,dx = -\frac{1}{\pi}\int_0^1 \sqrt{x(1-x)}\,dx.$$

Since the graph of the integrand $\sqrt{x(1-x)}$ is a half circle of centre $\left(\frac{1}{2},0\right)$ and radius $r = \frac{1}{2}$, we have

$$E\{X(X-1)\} = -\frac{1}{\pi}\,\text{area}\left\{\text{halvcirkel, radius } \frac{1}{2}\right\} = -\frac{1}{\pi}\cdot\frac{1}{2}\cdot\pi\left(\frac{1}{2}\right)^2 = -\frac{1}{8},$$

hence

$$E\left\{X^2\right\} = E\{X(X-1)\} + E\{X\} = -\frac{1}{8}+\frac{1}{2} = \frac{3}{8},$$

and

$$V\{X\} = E\left\{X^2\right\} - (E\{X\})^2 = \frac{3}{8} - \left(\frac{1}{2}\right)^2 = \frac{3}{8}-\frac{1}{4} = \frac{1}{8}.$$

3) Since $y = \psi(x) = \sqrt{x}$ is a bijective map $\psi :]0,1[\to]0,1[$, with the inverse $x = \varphi(x) = y^2$, where $\dfrac{dy}{dx} = 2y > 0$, we conclude that the frequency of $Y = \sqrt{X}$ is

$$g(y) = \begin{cases} f(\varphi(y)) \cdot \varphi'(y) = \dfrac{2y}{\pi\sqrt{y^2\,(1-y^2)}} = \dfrac{2}{\pi} \cdot \dfrac{1}{\sqrt{1-y^2}}, & \text{for } y \in]0,1[, \\ \\ 0, & \text{otherwise,} \end{cases}$$

corresponding to the distribution function

$$G(y) = \begin{cases} 1, & \text{for } y \geq 1, \\ \\ \dfrac{2}{\pi}\,\text{Arcsin}\, y, & \text{for } y \in]0,1[, \\ \\ 0, & \text{for } y \leq 0. \end{cases}$$

4) The mean of Y is

$$E\{Y\} = \frac{2}{\pi} \int_0^1 \frac{y}{\sqrt{1-y^2}}\,dy = \frac{2}{\pi}\left[-\sqrt{1-y^2}\right]_0^1 = \frac{2}{\pi}.$$

Example 10.7 *1) Prove that the function*

$$f(x)=\begin{cases}\dfrac{ab}{a-b}\left(e^{-bx}-e^{-ax}\right), & \text{for } x\geq 0,\\[2mm] 0, & \text{for } x<0,\end{cases}$$

where a and b denote positive constants, $a \neq b$, can be considered as the frequency of a random variable X.

2) Find the distribution function of the random variable X.

3) Find $E\{X\}$ and $V\{X\}$, expressed by a and b .

4) Let a be a fixed number. Prove for every fixed x that

$$\lim_{b\to a} f(x)=g(x),$$

where

$$g(x)=\begin{cases}a^2x\,e^{-ax}, & \text{for } x\geq 0,\\[2mm] 0, & \text{for } x<0.\end{cases}$$

5) Prove that the function $g(x)$ can be considered as a frequency of a random variable Y.

6) Finally, prove that $E\{X\}\to E\{Y\}$ for $b\to a$.

1) We may assume that $a>b$. Then $f(x)\geq 0$ for $x\in\mathbb{R}$, and

$$\int_{-\infty}^{\infty} f(x)\,dx=\frac{ab}{a-b}\int_0^{\infty}\left\{e^{-bx}-e^{-ax}\right\}dx=\frac{ab}{a-b}\left\{\frac{1}{b}-\frac{1}{a}\right\}=\frac{ab}{a-b}\cdot\frac{a-b}{ab}=1,$$

thus $f(x)$ can be considered as a frequency.

2) The distribution function $F(x)$ of X is 0 for $x\leq 0$. If $x>0$, then

$$\begin{aligned}F(x) &= \frac{ab}{a-b}\int_0^x\left\{e^{-bt}-e^{-at}\right\}dt=\frac{ab}{a-b}\left[-\frac{1}{b}\,e^{-bt}+\frac{1}{a}\,e^{-at}\right]_0^x=\frac{1}{a-b}\left[-a\,e^{-bt}+b\,e^{-at}\right]_0^x\\ &= \frac{1}{a-b}\left\{-a\,e^{-bx}+b\,e^{-ax}+a-b\right\}=1+\frac{1}{a-b}\left(b\,e^{-ax}-a\,e^{-bx}\right).\end{aligned}$$

3) We get by a partial integration, or by means of the Γ integral,

$$\begin{aligned}E\{X\} &= \int_0^{\infty}x\,f(x)\,dx=\frac{ab}{a-b}\int_0^{\infty}x\left\{e^{-bx}-e^{-ax}\right\}dx\\ &= \frac{ab}{a-b}\left\{\frac{1}{b^2}\int_0^{\infty}t\,e^{-t}\,dt-\frac{1}{a^2}\int_0^{\infty}t\,e^{-t}\,dt\right\}=\frac{ab}{a-b}\left(\frac{1}{b^2}-\frac{1}{a^2}\right)\\ &= \frac{ab}{a-b}\cdot\frac{a^2-b^2}{a^2b^2}=\frac{a+b}{ab}=\frac{1}{a}+\frac{1}{b},\end{aligned}$$

and

$$\begin{aligned} E\left\{X^2\right\} &= \int_0^\infty x^2 f(x)\,dx = \frac{ab}{a-b}\int_0^\infty x^2\left\{e^{-bx}-e^{-ax}\right\}dx \\ &= \frac{ab}{a-b}\left\{\frac{1}{b^3}\int_0^\infty t^2 e^{-t}\,dt - \frac{1}{a^3}\int_0^\infty t^2 e^{-t}\,dt\right\} \\ &= \frac{ab}{a-b}\left(\frac{2}{b^3}-\frac{2}{a^3}\right) = \frac{2ab}{a-b}\cdot\frac{a^3-b^3}{(ab)^3} = 2\cdot\frac{a^2+ab+b^2}{(ab)^2} \\ &= 2\left\{\frac{1}{a^2}+\frac{1}{ab}+\frac{1}{b^2}\right\}, \end{aligned}$$

hence

$$V\{X\} = E\left\{X^2\right\} - (E\{X\})^2 = 2\left\{\frac{1}{a^2}+\frac{1}{ab}+\frac{1}{b^2}\right\} - \left\{\frac{1}{a^2}+\frac{2}{ab}+\frac{1}{b^2}\right\} = \frac{1}{a^2}+\frac{1}{b^2}.$$

4) Let $x > 0$ and $a > 0$ be fixed. Then by e.g. l'Hospital's rule,

$$\begin{aligned} &\lim_{b\to a}\frac{ab}{a-b}\left(e^{-bx}-e^{-ax}\right) \\ &\quad = a^2\lim_{b\to a}\frac{e^{-bx}-e^{-ax}}{a-b} = a^2\lim_{b\to a}\frac{-x\,e^{-bx}}{-1} = a^2 x\,e^{-ax}. \end{aligned}$$

For $x \le 0$ we get of course 0, thus

$$\lim_{b\to a} f(x) = g(x).$$

5) Obviously, $g(x) \ge 0$. Since

$$\int_{-\infty}^\infty g(x)\,dx = a^2\int_0^\infty x\,e^{-ax}\,dx = \int_0^\infty t\,e^{-t}\,dt = 1,$$

it follows that $g(x)$ is the frequency of a random variable Y.

6) the mean of Y is

$$E\{Y\} = \int_{-\infty}^\infty x\,g(x)\,dx = a^2\int_0^\infty x\cdot x\,e^{-ax}\,dx = \frac{1}{a}\int_0^\infty t^2 e^{-t}\,dt = \frac{2}{a}.$$

It follows from (3) that

$$\lim_{b\to a} E\{X\} = \lim_{b\to a}\left(\frac{1}{a}+\frac{1}{b}\right) = \frac{2}{a} = E\{Y\}.$$

Remark 10.1 It is possible to give a simper solution. In fact, if X_1 and X_2 are independent random variables of the frequencies

$$f_{X_1}(x) = \begin{cases} a\,e^{-ax}, & x \ge 0, \\ 0, & x < 0, \end{cases} \qquad \text{and} \qquad f_{X_2}(x) = \begin{cases} b\,e^{-bx}, & x \ge 0, \\ 0, & x < 0, \end{cases}$$

then $X = X_1 + X_2$ has the frequency $f(x)$, and

$$E\{X\} = E\{X_1\} + E\{X_2\} = \frac{1}{a} + \frac{1}{b},$$

$$V\{X\} = V\{X_1\} + V\{X_2\} = \frac{1}{a^2} + \frac{1}{b^2}. \quad \Diamond$$

Example 10.8 *Let X and Y be independent random variables with the distributions given by*

$$P\{X = k\} = P\{Y = k\} = pq^k, \qquad k \in \mathbb{N}_0,$$

where $p > 0$, $q > 0$ and $p + q = 1$.

1) Find the means $E\{X\}$ and $E\{Y\}$.

2) Find the variances $V\{X\}$ and $V\{Y\}$.

3) Find $P\{X + Y = k\}$, $k \in \mathbb{N}_0$.

1) The means are

$$E\{X\} = E\{Y\} = \sum_{k=0}^{\infty} k \cdot P\{X = k\} = \sum_{k=1}^{\infty} kpq^k = pq \sum_{k=1} k \cdot q^{k-1} = pq \cdot \frac{1}{(1-q)^2} = \frac{pq}{p^2} = \frac{q}{p}.$$

Here we have used that by a partial differentiation with respect to $q \in]0,1[$ we obtain the important expressions

$$\frac{1}{1-q} = \sum_{k=0}^{\infty} q^k \quad \text{and} \quad \frac{d}{dq}\left(\frac{1}{1-q}\right) = \frac{1}{(1-q)^2} = \sum_{k=1}^{\infty} k \cdot q^{k-1}.$$

2) The variance is found by a smart rearrangement,

$$V\{X\} = E\left\{X^2\right\} - (E\{X\})^2 = E\{X(X-1)\} + E\{X\} - (E\{X\})^2,$$

where

$$\begin{aligned} E\{X(X-1)\} &= \sum_{k=2}^{\infty} k(k-1)\, P\{X = k\} = \sum_{k=2}^{\infty} k(k-1)pq^k = pq^2 \sum_{k=2}^{\infty} k(k-1)q^{k-2} \\ &= pq^2 \frac{d^2}{dq^2}\left\{\sum_{k=0}^{\infty} q^k\right\} = pq^2\, \frac{d^2}{dq^2}\left\{\frac{1}{1-q}\right\} = pq^2 \cdot \frac{2}{(1-q)^3} = \frac{2pq^2}{p^3} = 2 \cdot \frac{q^2}{p^2}, \end{aligned}$$

i.e.

$$V\{X\} = V\{Y\} = 2 \cdot \frac{q^2}{p^2} + \frac{q}{p} - \frac{q^2}{p^2} = \frac{q^2}{p^2} + \frac{q \cdot p}{p^2} = \frac{q(p+q)}{p^2} = \frac{q}{p^2}.$$

Notice that it is easier to compute $E\{X(X-1)\}$ than

$$E\left\{X^2\right\} = \sum_{k=1}^{\infty} k^2 P\{X = k\} = \sum_{k=1}^{\infty} k^2 pq^k.$$

3) Since X and Y are independent, we get for $k \in \mathbb{N}_0$,

$$\begin{aligned} P\{X+Y=k\} &= \sum_{i=0}^{k} P\{X=i \wedge Y=k-i\} = \sum_{i=0}^{k} P\{X=i\} \cdot P\{Y=k-i\} \\ &= \sum_{i=0}^{k} pq \cdot pq^{k-i} = p^2 q^k \sum_{i=0}^{k} 1 = (k+1)p^2 q^k, \qquad\qquad k \in \mathbb{N}_0. \end{aligned}$$

Remark 10.2 The distribution of X is a reduced waiting time distribution. $\Diamond$

Example 10.9 *There are given two components in an instrument. The life times of the components, T_1 and T_2, are independent random variables, both of the frequency*

$$f(t)=\begin{cases} a\,e^{-at}, & t>0,\\ 0, & t\le 0,\end{cases}$$

where a is a positive constant.
We introduce the random variables X_1, X_2 and Y_2 by

$$X_1=\min\{T_1,T_2\}, \qquad X_2=\max\{T_1,T_2\}, \qquad Y_2=X_2-X_1.$$

Here, X_1 denotes the time until the first one of the components fails to function, and X_2 indicates the time until the second component also fails, while Y_2 is the time from the failure of the first component until the failure of the second component.

1. *Find the frequency and the mean of X_1.*

2. *Find the mean of Y_2.*

Let it be given without proof that (X_1,X_2) has the simultaneous frequency

$$h(x_1,x_2)=\begin{cases} 2a^2e^{-a(x_1+x_2)}, & 0<x_1<x_2,\\ 0, & \textit{otherwise.}\end{cases}$$

4. *Find the simultaneous frequency of the 2-dimensional random variable (X_1,Y_2).*

5. *Find the frequency of Y_2.*

6. *Check if the random variables X_1 and Y_2 are independent.*

1) We get for X_1,

$$P\{X_1>x_1\}=P\{T_1>x_1\wedge T_2>x_1\}=P\{T_1>x_1\}\cdot P\{T_2>x_2\}=e^{-2ax_1},$$

thus

$$P\{X_1\le x_1\}=1-e^{-2ax_1}, \qquad x_1>0,$$

and we see that X_1 is exponentially distributed with the frequency

$$f_{X_1}=\begin{cases} 2a\,e^{-2ax_1}, & x_1>0,\\ 0, & x_1\le 0,\end{cases} \qquad \text{and mean } \frac{1}{2a}.$$

2) For X_2 we get

$$\begin{aligned} P\{X_2\le x_2\} &= P\{T_1\le x_2\wedge T_2\le x_2\}=P\{T_1\le x_2\}\cdot P\{T_2\le x_2\}\\ &= \left(1-e^{-ax_2}\right)^2, \qquad x_2>0,\end{aligned}$$

so X_2 has the frequency

$$f_{X_2}(x_2)=2a\,e^{-ax_2}\left(1-e^{-ax_2}\right)=2a\,e^{-ax_2}-2a\,e^{-2ax_2} \qquad \text{for } x_2>0,$$

and

$$f_{X_2}(x_2) = 0 \qquad \text{for } x_2 \le 0.$$

The mean is

$$\begin{aligned} E\{X_2\} &= \int_0^\infty x_2 f_{X_2}(x_2)\, dx_2 = \int_0^\infty \left\{2a\, x_2 e^{-ax_2} - 2a\, x_2 e^{-2ax_2}\right\} dx_2 \\ &= \frac{2}{a} - \frac{1}{2a} = \frac{3}{2a}. \end{aligned}$$

Additional. It is easy to find the mean of X_2 from $X_1 + X_2 = T_1 + T_2$, i.e.

$$E\{X_2\} = E\{T_1\} + E\{T_2\} - E\{X_1\} = \frac{1}{a} + \frac{1}{a} - \frac{1}{2a} = \frac{3}{2a}.$$

3) This is trivial, because

$$E\{Y_2\} = E\{X_2\} - E\{X_1\} = \frac{3}{2a} - \frac{1}{2a} = \frac{1}{a}.$$

4) The simultaneous frequency $k(y_1, y_2)$ of

$$(Y_1, Y_2) = (X_1, X_2 - X_1)$$

can e.g. be found directly from a convenient formula with $a = 1$, $b = 0$, $c = -1$ and $d = -1$,

$$\begin{aligned} k(y_1, y_2) &= h\left(\frac{dy_1 - by_2}{ad - bc}, \frac{-cy_1 + ay_2}{ad - bc}\right) \cdot \frac{1}{|ad - bc|} \\ &= h(y_1, y_1 + y_2) = 2a^2 e^{-a(2y_1 + y_2)} \qquad \text{for } y_1 > 0 \text{ og } y_2 > 0, \end{aligned}$$

and

$$k(y_1, y_2) = 0 \qquad \text{otherwise.}$$

This is also written

$$k(y_1, y_2) = \begin{cases} 2a\, e^{-2ay_1} \cdot a\, e^{-ay_2}, & \text{for } y_1 > 0 \text{ and } y_2 > 0, \\ 0, & \text{otherwise.} \end{cases}$$

5) (and 6.) It follows immediately from **4.** that $Y_1 \;\; (= X_1)$ and Y_2 are independent, and that Y_2 has the frequency

$$k_{Y_2}(y_2) = \begin{cases} a\, e^{-ay_2}, & y_2 > 0, \\ 0, & y_2 \le 0. \end{cases}$$

Index

2-dimensional random variable, 5

almost everywhere, 7
Arcussinus distribution, 149

binomial distribution, 18
Buffon's needle problem, 77

Cauchy distribution, 97, 111, 144
causal distribution, 4
Chebyshev's inequality, 13, 138, 139
conditional distribution, 11
conditional distribution function, 11
conditional probability, 11
continuous distribution, 5, 6
continuous random variable, 5, 6
convergence in distribution, 16
convergence in probability, 16
convex function, 142
correlation, 15
covariance, 15
Cramer's formula, 108

discrete distribution, 4, 6
discrete random variable, 4, 6
distribution function, 4, 20, 33

expectation, 11

fraud in Probability, 77
frequency, 5, 6, 20, 33
function of random variable, 75
function of random variables, 93
function of two random variables, 107

gamma distribution, 61
geometric distribution, 90

Helly-Bray's lemma, 16

independent random variables, 7
inequality between random variables, 81

Jacobian, 10, 54
Jensen's inequality, 142

law of total probability, 11

MAPLE, 22
marginal distribution, 5
marginal frequency, 6, 37, 54, 121
mean, 11, 129
median, 4, 21
moment, 12, 140, 143
moment of higher order, 129

null-set, 7

polar coordinates, 118
probability field, 4

quantile, 4

random variable, 4
rectangular coordinates, 118
rectangular distribution, 93, 95, 98, 112, 143, 144

simultaneous distribution, 5
simultaneous distribution function, 6
simultaneous frequency, 54

transformation theorem, 8

weak law of large numbers, 16