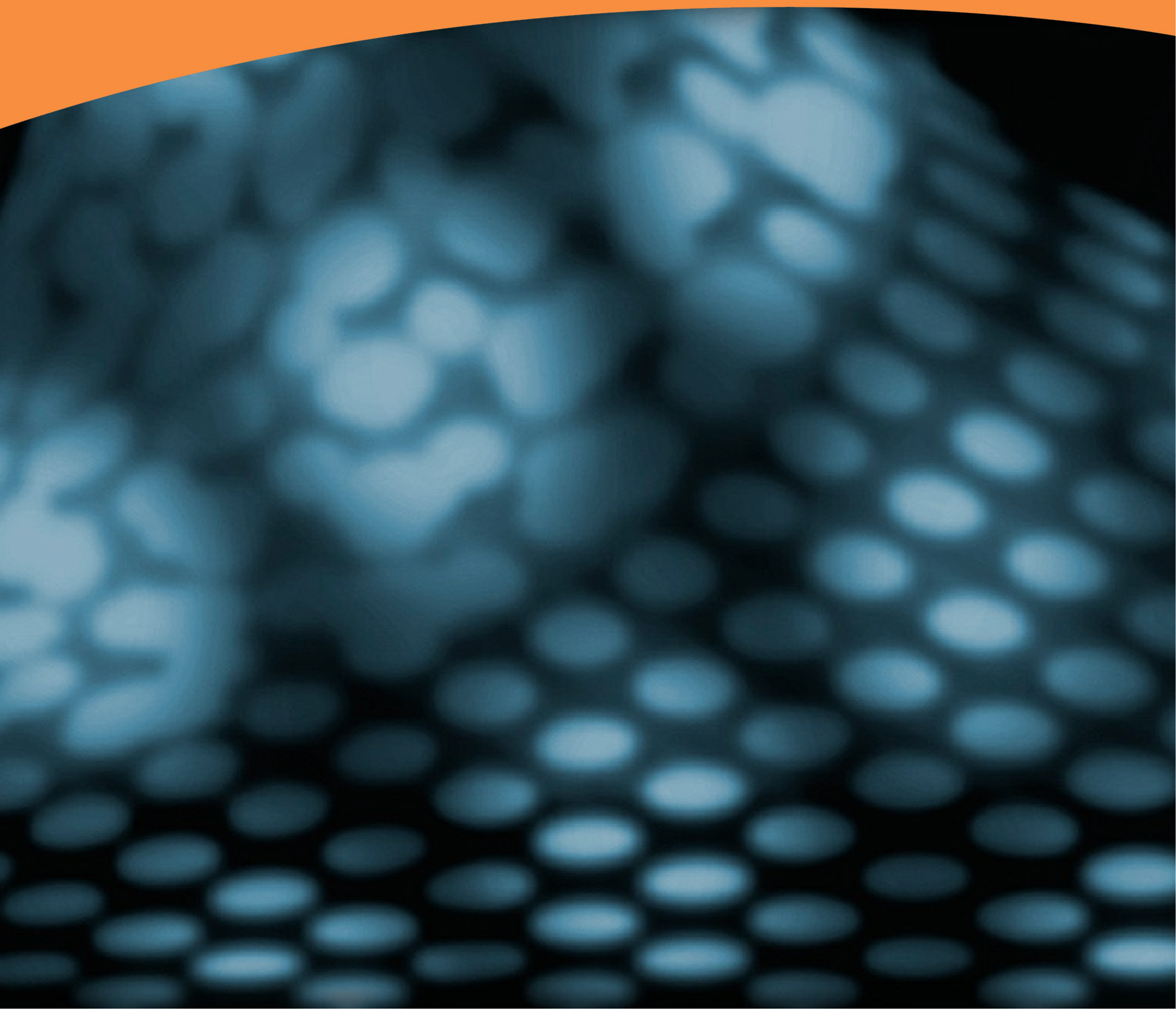
Continuous Distributions

Probability Examples c-6

Leif Mejlbro



bookboon
The eBook company

Leif Mejlbro

Probability Examples c-6
Continuous Distributions

Probability Examples c-6 – Continuous Distributions

ISBN 978-87-7681-522-6

Contents

Introduction

This is the sixth 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 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
27th October 2014

1 Some theoretical background

1.1 The exponential distribution

A random variable X follows an *exponential distribution* with parameter $a > 0$, if its distribution function $F(x)$ is given by

$$F(x) = \begin{cases} 1 - e^{-ax}, & \text{for } x \geq 0, \\ \\ 0, & \text{for } x < 0. \end{cases}$$

The corresponding *frequency* $f(x)$ is given by

$$f(x) = \begin{cases} a\, e^{-ax}, & \text{for } x \geq 0, \\ \\ 0, & \text{for } x < 0. \end{cases}$$

We have for an exponentially distributed random variable X with parameter $a > 0$,

$$E\{X\} = \frac{1}{a} \qquad \text{and} \qquad V\{X\} = \frac{1}{a^2}.$$

In general, if X is exponentially distributed, then

$$P\{X > s + t \mid X > s\} = P\{X > t\}, \qquad \text{for } s,\, t > 0,$$

which is equivalent with the formula

$$P\{X > s + t\} = P\{X > s\} \cdot P\{X > t\}, \qquad \text{for } s,\, t > 0.$$

We say that the exponential distribution is *forgetful*.

In practice, the exponential distribution often occurs as a distribution of *lifetimes*, which is in particular the case in queuing theory. In this case the forgetfulness is of paramount importance.

An exponentially distributed random variable X with parameter $a > 0$ is a special *gamma distribution* (cf. the following), so one also writes,

$$X \in \Gamma\left(1\,,\,\frac{1}{a}\right)$$

for the exponential distribution.

Another type of generalized exponential distributions is the *Weibull distribution* with parameters a, $b > 0$. This is given by the distribution function

$$F(x) = \begin{cases} 1 - \exp\left(-a\, x^b\right), & \text{for } x \geq 0, \\ \\ 0, & \text{for } x < 0. \end{cases}$$

We note that we get the exponential distribution for $b = 1$. The *Weibull distribution* is used in connection with the *theory of reliability*.

1.2 The normal distribution

A random variable X is following a *normal distribution with mean 0 and variance 1*, and we write $X \in \mathrm{N}(0,1)$, if its frequency $\varphi(x)$ is given by

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \qquad \text{for } x \in \mathbb{R}.$$

Its distribution function is traditionally denoted by $\Phi(x)$. It is given by

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)\,dt, \qquad \text{for } x \in \mathbb{R},$$

which cannot be expressed simpler by elementary functions.

We notice that if $X \in \mathrm{N}(0,1)$, then

$$E\{X\} = 0 \qquad \text{and} \qquad V\{X\} = 1,$$

and $\Phi(-x) = 1 - \Phi(x)$, and

$$1 - \Phi(x) \sim \frac{1}{x}\,\varphi(x) \qquad \text{for } x \to +\infty,$$

or more precisely,

$$\left\{\frac{1}{x} - \frac{1}{x^3}\right\} \varphi(x) < 1 - \Phi(x) < \frac{1}{x}\,\varphi(x).$$

A random variable X is following a *normal distribution with mean* μ *and variance* σ^2, and we write $X \in \mathrm{N}\left(\mu\,,\,\sigma^2\right)$, if its frequency $f(x)$ is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\,\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \qquad \text{for } x \in \mathbb{R}.$$

The normal distribution is also called the *Gaußian distribution*. It plays a central role in the theory of probability.

Let X_1 and X_2 be two independent $\mathrm{N}(0,1)$ distributed random variables, and let a, b be two real constants. Then the linear combination is normally distributed with mean 0 and variance $c^2 = a^2+b^2$, i.e.

$$a\,X_1 + b\,X_2 \in \mathrm{N}\left(0, a^2+b^2\right).$$

1.3 2-dimensional normal distributions

A 2-dimensional random variable (Z_1, Z_2) is *normally distributed with the parameters* μ_1, $\mu_2 \in \mathbb{R}$, and σ_1^2, $\sigma_2^2 > 0$ and $\varrho \in [0,1[$, if its frequency is given by

$$f\left(z_1, z_2\right) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}}\,\exp\left(-\frac{1}{2\left(1-\varrho^2\right)}\left\{\left(\frac{z_1-\mu_1}{\sigma_1}\right)^2 - 2\varrho\,\frac{z_1-\mu_1}{\sigma_1}\cdot\frac{z_2-\mu_2}{\sigma_2} + \left(\frac{\sigma_2-\mu_2}{\sigma_2}\right)^2\right\}\right).$$

In this case we write $(Z_1, Z_2) \in \mathrm{N}\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \varrho\right)$.

If $(Z_1, Z_2) \in \mathrm{N}\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \varrho\right)$, then we have for the marginal random variables that

$$Z_1 \in \mathrm{N}\left(\mu_1, \sigma_1^2\right) \qquad \text{and} \qquad Z_2 \in \mathrm{N}\left(\mu_2, \sigma_2^2\right),$$

and

$$E\left\{Z_1\right\} = \mu_1, \quad V\left\{Z_1\right\} = \sigma_1^2, \qquad \text{og} \qquad E\left\{Z_2\right\} = \mu_2, \quad V\left\{Z_2\right\} = \sigma_2^2,$$

and concerning the correlation coefficient between them,

$$\varrho\left(Z_1, Z_2\right) = \varrho.$$

In general, if two random variables are independent, then they are also non-correlated, while the converse does not need to be true. However, if (Z_1, Z_2) is following a 2-dimensional normal distribution, where Z_1 and Z_2 are non-correlated, then Z_1 and Z_2 are in fact also independent, so we obtain a stronger result in this case.

Let (Z_1, Z_2) follow a 2-dimensional normal distribution, and let a, b, c and d be real constants, which satisfy the condition (of the determinant) $ad - bc \neq 0$. Then the 2-dimensional random variable

$$(U_1, U_2) := (aZ_1 + bZ_2, cZ_1 + dZ_2)$$

is again following a 2-dimensional normal distribution.

The theory above has a ring of geometry, and one may get a better description by using matrices for the more general *n-dimensional normal distribution.* We introduce for a general n-dimensional random variable

$$\mathbf{X} = (X_1, X_2, \dots, X_n)$$

the corresponding *mean vector*,

$$\mu := (E\{X_1\}, E\{X_2\}, \dots, E\{X_n\}),$$

and its *covariance matrix*,

$$\mathbf{C} = \left\{ \begin{array}{cccc} V\{X_1\} & \operatorname{Cov}(X_1, X_2) & \cdots & \operatorname{Cov}(X_1, X_n) \\ \operatorname{Cov}(X_2, X_1) & V\{X_2\} & \cdots & \operatorname{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{Cov}(X_n, X_1) & \operatorname{Cov}(X_n, X_2) & \cdots & V\{X_n\} \end{array} \right\},$$

provided that they exist.

Then the frequency of $\mathbf{X} \in \mathrm{N}\big(\mu, \sigma_1^2, \sigma_2^2, \varrho\big)$ is written

$$f(\mathbf{x}) = \frac{1}{\left(\sqrt{2\pi}^2\right)^2 \sqrt{\det \mathbf{C}}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)\,\mathbf{C}^{-1}(\mathbf{x}-\mu)^{\top}\right), \qquad \mathbf{x} \in \mathbb{R}^2,$$

where

$$\mathbf{C} = \begin{pmatrix} \sigma_1^2 & \varrho\sigma_1\sigma_2 \\ \varrho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \qquad \text{and} \qquad \mathbf{C}^{-1} = \frac{1}{1-\varrho^2} \begin{pmatrix} \dfrac{1}{\sigma_1^2} & -\dfrac{\varrho}{\sigma_1\sigma_2} \\ -\dfrac{\varrho}{\sigma_1\sigma_2} & \dfrac{1}{\sigma_2^2} \end{pmatrix}.$$

This leads to the following generalization: An n-dimensional random variable $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is a *non-singular normal distribution with mean vector* μ and a *positive definite covariance matrix* $\mathbf{C}$, if the corresponding frequency $f\mathbf{x})$ is given in the form

$$f(\mathbf{x}) = \frac{1}{\left(\sqrt{2\pi}^2\right)^n \sqrt{\det \mathbf{C}}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)\,\mathbf{C}^{-1}(\mathbf{x}-\mu)^{\top}\right), \qquad \mathbf{x} \in \mathbb{R}^n.$$

1.4 Conditional normal distribution

Let $(X, Y) \in \mathrm{N}\big(\mu, \sigma_1^2, \sigma_2^2, \varrho\big)$ denote a 2-dimensional normal distribution. Then the conditional frequency $f(x \mid y)$ for X, for given $Y = y$, is defined as

$$f(x \mid y) = \frac{1}{\sqrt{2\pi}\,\sigma_1\sqrt{1-\varrho^2}} \exp\left(-\frac{1}{2\sigma_1^2(1-\varrho^2)}\left\{x - \mu_1 - \varrho\,\frac{\sigma_1}{\sigma_2}\,(y-\mu_2)\right\}^2\right).$$

It follows that the conditional distribution of X, given $Y=y$, is a *normal distribution* with

$$\textit{mean: } \mu_1+\varrho\,\frac{\sigma_1}{\sigma_2}\,(y-\mu_2) \qquad \text{and} \qquad \textit{variance: } \sigma_1^2\left(1-\varrho^2\right).$$

This is exploited, whenever one wants to estimate Y as a *linear function* $Y=aX+b$ of X, where we shall find the constants a and b, such that $V\left\{[Y-(aX+b)]^2\right\}$ becomes as small as possible. This is the case, when we choose

$$a=\varrho\,\frac{\sigma_2}{\sigma_1} \qquad \text{and} \qquad b=\mu_2-\varrho\,\frac{\sigma_2}{\sigma_1}\,\mu_1.$$

The line

$$y=ax+b=\mu_2+\varrho\,\frac{\sigma_2}{\sigma_1}\,(x-\mu_1)$$

is called the *regression line of Y with respect to X*. Analogously, the line

$$x=\mu_1+\varrho\,\frac{\sigma_1}{\sigma_2}\,(y-\mu_2)$$

is the *regression line of X with respect to Y*. Summing up, they are of course written

$$\frac{x-\mu_1}{\sigma_1}=\frac{y-\mu_2}{\sigma_2},$$

which is easier to remember in practice.

1.5 Sums of independent normal distributed random variables

The main result is that if the normal distributed random variables X_1, X_2, ..., X_n are mutually independent and normally distributed with $X_i\in \mathrm{N}\left(\mu_i,\sigma_i^2\right)$, then the sum $\sum X_i$ also normally distributed with

$$\sum_{i=1}^n X_i\in \mathrm{N}\left(\sum_{i=1}^n \mu_i,\sum_{i=1}^n\sigma_i^2\right).$$

We express this result by saying that the *normal distribution is reproductive with respect to the parameters* μ and σ^2.

1.6 The Central Limit Theorem

The normal distribution is of paramount importance for the Theory of Probability, in particular because we have the following result, which shows that we can approximate a distribution of sums of independent identically distributed random variables by $\Phi(x)$.

We assume that a sequence of random variables $\{X_i\}$ are all mutually independent and identically distributed of mean μ and variance $\sigma^2>0$. If we put

$$Y_n=X_1+X_2+\cdots+X_n=\sum_{i=1}^n X_i,$$

then

$$\frac{Y_n - n\mu}{\sigma\sqrt{n}}, \qquad n \in \mathbb{N},$$

converges *in distribution* towards $\Phi(x)$ for $n \to +\infty$, i.e.

$$\lim_{n\to+\infty} P\left\{\frac{Y_n - n\mu}{\sigma\sqrt{n}} \le x\right\} = \Phi(x) \qquad \text{for every } x \in \mathbb{R}.$$

This means, roughly speaking, that

Y_n is almost normally distributed $\mathrm{N}\left(n\mu, n\sigma^2\right)$ for large values of n.

1.7 The Maxwell distribution

A random variable X is *Maxwell distributed*, if it has a frequency of the form

$$f(x) = \begin{cases} \sqrt{\frac{1}{\pi}}\,\frac{1}{\sigma^3}\,x^2\,\exp\left(-\frac{x^2}{2\sigma^2}\right), & \text{for } x > 0, \\ 0, & \text{for } x \le 0, \end{cases}$$

where the parameter $\sigma > 0$. For such a distribution,

$$E\{X\} = 2\sqrt{\frac{2}{\pi}}\,\sigma \qquad \text{and} \qquad V\{X\} = \left\{3 - \frac{8}{\pi}\right\}\sigma^2.$$

The Maxwell distribution occurs typically, when three given independent random variables X_1, X_2, X_3, are all $\mathrm{N}\left(0,\sigma^2\right)$ distributed. Then the random variable

$$X = \sqrt{X_1^2 + X_2^2 + X_3}$$

is *Maxwell distributed* of parameter $\sigma > 0$.

Conversely, *if* X_1, X_2 and X_2 are mutually independent identically distributed random variables of a *ball symmetrically* distribution of the 3-dimensional random variable (X_1, X_2, X_3), i.e. the distribution depends only on the radius $r := \sqrt{x_1^2 + x_2^2 + x_2^2}$, then X_1, X_2 and X_3 are all normally distributed random variables of the same type, $X_i \in \mathrm{N}\left(0,\sigma^2\right)$.

1.8 The Gamma distribution

It is well-known that the *Gamma function* is given by

$$\Gamma(\mu) := \int_0^{+\infty} t^{\mu-1}\,e^{-t}\,dt, \qquad \text{for } \mu > 0,$$

with the properties

$$\Gamma(\mu+1) = \mu\,\Gamma(\mu) \quad \text{for } \mu > 0, \qquad \text{in particular, } \Gamma(n+1) = n! \quad \text{for } n \in \mathbb{N}_0.$$

Note also that

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

In connection with the Gamma function we also mention the *Beta function*, which is defined by

$$B\left(\mu_1,\mu_2\right) = \int_0^1 t^{\mu_1-1}(1-t)^{\mu_2-1}\,dt = \frac{\Gamma\left(\mu_1\right)\Gamma\left(\mu_2\right)}{\Gamma\left(\mu_1+\mu_2\right)}, \qquad \text{for } \mu_1,\, \mu_2 > 0.$$

A random variable X is *Gamma distributed* with *form parameter* $\mu > 0$ and *scale parameter* $\alpha > 0$, if its frequency is given by

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\alpha^{\mu}}\,x^{\mu-1}\,\exp\left(-\frac{x}{\alpha}\right), & \text{for } x > 0, \\ 0, & \text{for } x \le 0. \end{cases}.$$

We write $X \in \Gamma(\mu, \alpha)$. Note that

$$E\{X\} = \alpha\mu \qquad \text{and} \qquad V\{X\} = \alpha^2\mu.$$

It was mentioned previously that the *exponential distribution* is a special Gamma distribution, $\Gamma\left(1\,,\,\frac{1}{a}\right)$, of the frequency

$$f(x) = \begin{cases} a\,e^{-ax}, & \text{for } x > 0, \\ \\ 0, & \text{for } x \le 0. \end{cases}$$

If $X_1 \in \Gamma(\mu_1, \alpha)$ and $X_2 \in \Gamma(\mu_2, \alpha)$ are independent Gamma distributed random variables of the same scale parameter, then the sum $X_1 + X_2$ is also Gamma distributed, $X_1 + X_2 \in \Gamma(\mu_1 + \mu_2, \alpha)$. We say that the *Gamma distribution is reproductive in the form parameter for fixed scale parameter.*

If more generally, $X_1, X_2, \ldots, X_n$ are independent Gamma distributed random variables with the same scale parameter, $X_i \in \Gamma(\mu_i, \alpha)$, then their sum is also Gamma distributed,

$$\sum_{i=1}^{n} X_i \in \Gamma\left(\sum_{i=1}^{n} \mu_i\,,\,\alpha\right).$$

1.9 The χ^2 distribution

Let $X_1, X_2, \ldots, X_n$ be identically distributed independent normal distributed random variables, thus $X_i \in \mathrm{N}(\mu, \sigma^2)$. Then

$$Y_n = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 \in \Gamma\left(\frac{n}{2}\,,\,2\right) := \chi^2(n),$$

where we call a distribution from $\Gamma\left(\frac{n}{2}\,,\,2\right) = \chi^2(n)$ a χ^2 *distribution* with n *degrees of freedom.* It follows immediately that if $X_1, X_2, \ldots, X_n$ are independent random variables, where $X_i \in \chi^2(n_i)$, then for the sum

$$\sum_{i=1}^{n} X_i \in \chi^2\left(\sum_{i=1}^{n} n_i\right).$$

If $X \in \chi^2(n)$, then

$$E\{X\} = n \qquad \text{and} \qquad V\{X\} = 2n.$$

Given a sequence $\{X_n\}$ of random variables, where

$$X_n \in \chi^2(n), \qquad \text{for every } n \in \mathbb{N}.$$

Then it follows from the *Central Limit Theorem* that

$$\lim_{n\to+\infty} P\left\{\frac{X_n - n}{\sqrt{2n}} \le x\right\} = \Phi(x) \qquad \text{for every } x \in \mathbb{R}.$$

In particular, $X_n \in \chi^2(n)$ is approximatively normal distributed $\mathrm{N}(n, 2n)$ for large $n \in \mathbb{N}$.

Notice that if the random variable X is $\chi^2(3)$ distributed then the random variable $Y = \sqrt{X}$ is a Maxwell distribution.

1.10 The t distribution

A random variable X is *t-distributed*, or following a *Student distribution*, with n degrees of freedom, if its frequency is given by

$$f(x) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\dfrac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{1}{2}(n+1)}, \qquad x \in \mathbb{R}.$$

We write $X \in t(n)$.

We obtain a *Student distribution* in the following way:

Let $Y \in \mathrm{N}(0,1)$ and $Z \in \chi^2(n)$ be independent random variables. If

$$X := \frac{Y}{\sqrt{\dfrac{Z}{n}}}, \qquad n \in \mathbb{N},$$

then $X \in t(n)$ is *Student distributed* with n degrees of freedom.

If $n = 1$ we get the *Cauchy distribution* of the frequency

$$f(x) = \frac{1}{\pi\left(1+x^2\right)}, \qquad x \in \mathbb{R}.$$

We notice that this distribution does not have a mean.

In general, the t distribution is symmetric with respect to 0, and if $n > 1$, then its mean is $E\{X\} = 0$.

When n is large, then the $\mathrm{N}(0,1)$ distribution is a fair approximation of the t distribution. (In practice usually for $n > 25$).

We mention that the t distribution is important because we often consider n independent identically normal distributed random variables $X_1, X_2, \dots, X_n$, where we know neither the mean nor the variance. Using $X_1, X_2, \dots, X_n$ we want to estimate the mean μ and the variance σ^2. We assume of course that $n > 1$.

As an *estimator* of the mean μ we use the normal distributed random variable

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \in \mathrm{N}\left(\mu\,,\,\frac{\sigma^2}{n}\right).$$

Since the variance decreases, when n increases, it follows immediately that the estimator becomes better, the larger n.

As an estimator of the variance σ^2 we use the random variable

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2,$$

where it is only for technical reasons that we use the denominator $n-1$. This trick assures that $E\left\{S^2\right\} = \sigma^2$.

It can be proved that

$$(n-1)\,\frac{S^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \in \chi^2(n-1).$$

The "loss" of one degree of freedom is due to the fact that we have used one degree of freedom to estimate μ.

It follows after some computations that the random variable

$$\frac{\overline{X}-\mu}{\sqrt{\dfrac{S^2}{n}}} \in \mathrm{t}(n-1)$$

is *Student distributed* with $n-1$ degrees of freedom.

1.11 The F distribution.

A random variable Z is *F distributed*, or *Fisher distributed* with the *degrees of freedom* n_1 and n_2, if its frequency is given in the following way, using the Beta function,

$$f(z) = \frac{\left(\dfrac{n_1}{n_2}\right)^{\frac{1}{2}n_1}}{B\left(\dfrac{n_1}{2}\,,\,\dfrac{n_2}{2}\right)} \cdot \frac{z^{\frac{1}{2}n_1-1}}{\left(1+\dfrac{n_1}{n_2}\,z\right)^{\frac{1}{2}(n_1+n_2)}}, \qquad z>0.$$

We write $Z \in F\left(n_1, n_2\right)$. When $n_2 > 2$, the mean is given by

$$E\{Z\} = \frac{n_2}{n_2-2},$$

while it does not exist for $n_2 \leq 2$.

The *Fisher distribution* occurs in the following casw: If $X_1 \in \chi^2\left(n_1\right)$ and $X_2 \in \chi^2\left(n_2\right)$ are two χ^2 distributed independent random variables, then

$$Z = \frac{\dfrac{X_1}{n_1}}{\dfrac{X_2}{n_2}} \in F\left(n_1\,,\,n_2\right)$$

is F distributed with the degrees of freedom n_1 and n_2. We notice that the random variable Z is used to compare the *normed* random variables $\dfrac{1}{n_1}\,X_1$ and $\dfrac{1}{n_2}\,X_2$.

1.12 Estimation of parameters

We shall shortly describe *Fisher's maximum method*.

Let X be a random variable with distribution function $F\left(x;\alpha_1,\alpha_2,\dots,\alpha_k\right)$ and with k independent parameters α_i, $i=1,\dots,k$. The task is from n independent observations $X_1, \dots, X_n$ of the random variable X with the results of the observations $x_1, \dots, x_n$ to *estimate* these unknown parameters.

1. The discrete case. We consider the function

$$h\left(\alpha_1,\dots,\alpha_k\right) := P\left\{X_1=x_1 \,\wedge\, \cdots \,\wedge\, X_n=x_n\right\} = P\left\{X_1=x_1\right\}\cdots P\left\{X_n=x_n\right\} = \prod_{i=1}^{n} P\left\{X_i=x_i\right\}$$

in the parameters $\alpha_1, \dots, \alpha_k$.

Using methods known from e.g. *Ventus: Calculus 2* series we find the values of $\alpha_1^{\star}, \dots, \alpha_k^{\star}$, for which the function $h(\alpha)$ has its *maximum*, i.e. the values of the parameters, for which

$$P\left\{X_1=x_1 \,\wedge\, \cdots \,\wedge\, X_n=x_n\right\}$$

is largest. These values $\alpha_1^{\star}, \dots, \alpha_k^{\star}$ are called *estimates* of the parameters.

We notice that $\alpha_i^{\star}$ is dependent on the observations $x_1, \dots, x_n$,

$$\alpha_i^{\star} = g_i\left(x_1,\dots,x_n\right), \qquad i=1, \dots, n.$$

The corresponding *random variables*

$$g_i\left(X_1,\ldots,X_n\right)$$

are called *estimators* of the parameters $\alpha_1, \ldots, \alpha_k$. We emphasize that we distinguish between an *estimate*, which is a number, and an *estimator*, which is a random variable.

2. The continuous case. In this case the random variable X is given by the frequency

$$f(x)=f\left(x;\alpha_1,\ldots,\alpha_k\right),$$

which depends on the k unknown parameters $\alpha_1 \ldots, \alpha_k$. Let $X_1, \ldots, X_n$ denote n independent observations of X with the results of the observations $x_1, \ldots, x_n$. Then like in the discrete case we introduce the function

$$h\left(\alpha_1,\ldots,\alpha_k\right)=f\left(x_1\right)\cdots f\left(x_n\right)=\prod_{i=1}^n f\left(x_i\right),$$

where the results of the observations $x_1, \ldots, x_n$ are given, while the parameters $\alpha_1, \ldots, \alpha_k$ are the unknowns. We shall find the values $\alpha_1^\star, \ldots, \alpha_k^\star$ of $\alpha_1, \ldots, \alpha_k$, for which the function $h(\alpha)$ has its maximum. This is again a task, which has been described elsewhere, e.g. in the *Ventus: Calculus 2* series.

Using methods of this type we usually obtain *estimates* (i.e. numbers)

$$\alpha_i^\star=g_i\left(x_1,\ldots,x_n\right),\qquad i=1,\ldots,k,$$

of the parameters $\alpha_1, \ldots, \alpha_k$, with corresponding *estimators* in the same parameters,

$$g_i\left(X_1,\ldots,X_n\right),\qquad i=1,\ldots,k,$$

which are *random variables.*

3. Estimators in general. Let $Y=g\left(X_1,\ldots,X_n\right)$ be an estimator for some parameter α. We say that the estimator is *central*, if

$$E\{Y\}=\alpha,$$

or roughly speaking that Y has the "right mean".

That an estimator is central means that if we by some samples get some estimates (i.e. numbers) of α, then these estimates will close in on α.

If in particular X is normal distributed, then

$$\overline{X}:=\frac{1}{n}\sum_{i=1}^n X_i,$$

is a *central estimator* of the mean, and

$$S^2:=\frac{1}{n-1}\sum_{i=1}^n\left(X_i-\overline{X}\right)^2$$

is a *central estimator* of the variance σ^2. We note that

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \qquad \left(=\frac{n-1}{n}\,S^2\right)$$

is *not* a central estimator of the variance σ^2.

In general,

$$\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$$

is *always* a central estimator of the mean μ, *provided that the distribution has a mean!* Note, however, that one does not always get $\overline{X}$ as an estimator by Fisher's maximum method.

A sequence of estimators $Y_n = g_n\left(X_1,\dots,X_n\right)$ for a parameter α is called *consistent*, if for every $\varepsilon > 0$,

$$P\left\{|Y_n-\alpha| \geq \varepsilon\right\} \to 0 \qquad \text{for } n \to +\infty.$$

It follows from the *Weak Law of Large Numbers* that if the distribution has a mean μ, then the sequence of estimators

$$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$$

is a consistent sequence of estimators for the mean μ.

Let

$$Y = g_1\left(X_1,\dots,X_n\right) \qquad \text{and} \qquad Z = g_2\left(X_1,\dots,X_n\right)$$

be two *central* estimators for the same parameter α. Then Y is said to be more *efficient* than Z, if

$$V\{Y\} < V\{Z\} \qquad \text{for every value of } \alpha.$$

If Y is more *efficient* than Z, then Y has a smaller variance, so roughly speaking, Y will have "more estimates" lying close to α than Z. Therefore, one prefers Y as a (central) estimator in such a case.

2 The Exponential Distribution

Example 2.1 *Let Y, X_1, X_2 be independent positive random variables, all with a continuous distribution, and let Y be exponentially distributed. It is well-known that for all s, $t \in \mathbb{R}_+$,*

(1) $P\{Y > s+t\} = P\{Y > s\} \cdot P\{Y > t\}$,

(the property of forgetfulness).

1) Prove by using (1) that we have the following generalization

(2) $P\{Y > X_1 + X_2\} = P\{Y > X_1\} \cdot P\{Y > X_2\}$.

2) Is (2) also true, if X_1 and X_2 are mutually dependent, while they are both independent of Y?

3) Let Y, X_1, X_2,k $\dots$, X_n be independent positive variables, all with a continuous distribution, and let Y be exponentially distributed.
Prove that

$$P\{Y > X_1 + X_2 + \cdots + X_n\} = \prod_{i=1}^{n} P\{Y > X_1\}.$$

1) Let $f_1(x_1)$ and $f_2(x_2)$ denote the frequencies of X_1 and X_2, resp.. Since X_1 and X_2 are independent and positive, we conclude that the frequency of $Z = X_1 + X_2$ for $z > 0$ is given by

(3) $g(z) = \displaystyle\int_0^z f_1(x) f_2(z-x)\, dx \qquad$ (and $g(z) = 0$ for $z \le 0$).

Let a be the parameter for Y. The frequency $h(t)$ of $Y - Z$ is for $t \ge 0$ (note the the frequency is of no interest for $t < 0$) given by

$$\begin{aligned}
h(t) &= a \int_0^\infty e^{-ay} g(y-t)\, dy = a \cdot e^{-at} \int_0^\infty e^{-az} g(z)\, dz \\
&= a\, e^{-at} \int_0^\infty e^{-az} \left\{ \int_0^z f_1(x) f_2(z-x)\, dx \right\} dz \\
&= a\, e^{-at} \int_0^\infty e^{-ax} f_1(x) \left\{ \int_x^\infty e^{-a(z-x)} f_2(z-x)\, dz \right\} dx \\
&= a\, e^{-at} \cdot \int_0^\infty e^{-ax} f_1(x)\, dx \cdot \int_0^\infty e^{-ay}\, f_2(y)\, dy,
\end{aligned}$$

so

$$P\{Y > X_1 + X_2\} = P\{Y - Z > 0\} = \int_0^\infty h(t)\, dt = \int_0^\infty e^{-ax} f_1(x)\, dx \cdot \int_0^\infty e^{-ay} f_2(y)\, dy.$$

Analogously, the frequency of $Y - X_i$ is given by

$$h_i(t) = a \int_0^\infty e^{-ay} f_i(y-t)\, dy = a \int_t^\infty e^{-ay} f_i(y-t)\, dy = a\, e^{-at} \int_0^\infty e^{-ax}\, f_i(x)\, dx,$$

hence

$$P\{Y > X_i\} = P\{Y - X_i > 0\} = \int_0^{\infty} h_i(t)\,dt = \int_0^{\infty} e^{-ax} f_i(x)\,dx, \qquad i = 1,\, 2,$$

and it follows that

$$P\{Y > X_1 + X_2\} = P\{Y > X_1\} \cdot P\{Y > X_2\}\,.$$

2) Formula (2) has been proved under the assumption of (3), which again presupposes that X_1 and X_2 are independent. We may therefore expect that the answer is "no".

An explicit example, which *shows* that the answer in general is "no" is given by $X_2 = 1 - X_1$, where X_1 is uniformly distributed over $[0,1]$. In this case,

$$P\{Y > X_1 + X_2\} = P\{Y > 1\} = \int_1^{\infty} a\,e^{-ay}\,dy = e^{-a}.$$

It follows that the frequency $h_1(t)$ of $Y - X_1$ is given by

$$h_1(t) = a\,e^{-at} \int_0^1 e^{-ax}\,dx = e^{-at}\left(1 - e^{-a}\right) \qquad \text{for } t > -1,$$

hence

$$P\{Y > X_1\} = \int_0^{\infty} f_1(t)\,dt = \frac{1}{a}\,\left(1 - e^{-a}\right).$$

Since

$$\{Y > X_2 = 1 - X_1\} = \{Y + X_1 > 1\},$$

we first find the frequency of $Y + X_1$, i.e. for $t > 0$,

$$k(t) = \int_0^\infty a\,e^{-ay}\,f_1(t-y)\,dy = \int_{\max\{0,t-1\}}^t a\,e^{-ay}\,dy = \exp(-a\,\max\{0,t-1\}) - e^{-at},$$

thus

$$\begin{aligned} P\{T > X_2\} &= P\{Y + X_1 > 1\} = \int_1^\infty k(t)\,dt = \int_1^\infty \left\{e^{-at}e\;a - e^{-at}\right\}\,dt \\ &= \frac{1}{a}\,(e^a - 1)\cdot e^{-a} = \frac{1}{a}\cdot\left(1 - 1^{-a}\right) = P\{Y > X_2\}, \end{aligned}$$

which of course also can be seen directly by an argument of symmetry.
We finally get

$$P\{Y > X_1\}\cdot P\{Y > X_2\} = \left(\frac{1 - e^{-a}}{a}\right)^2 \neq e^{-a} = P\{Y > X_1 + X_2\},$$

and it follows from this example that the claim does not hold in general.

3) If we put $Z_j = X_1 + \cdots + X_j$, then we get by recursion and (2),

$$\begin{aligned} &P\{Y > X_1 + X_2 + \cdots + X_n\} = P\{Y > Z_n\} = P\{Y > Z_{n-1} + X_n\} \\ &\qquad = P\{Y > Z_{n-1}\}\cdot P\{Y > X_n\} = \cdots = \prod_{i=1}^n P\{Y > X_i\}. \end{aligned}$$

Example 2.2 *Assume that X_1 and X_2 are independent random variables of the frequencies*

$$f_{X_1}(x_1) = \begin{cases} 4\,e^{-4x_1}, & x_1 > 0, \\ 0, & x_1 0, \end{cases} \qquad f_{X_2}(x_2) = \begin{cases} e^{x_2}, & x_2 < 0, \\ 0, & x_2 \ge 0. \end{cases}$$

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

2. *Find the variances $V\{X_1\}$ and $V\{X_2\}$.*

Let the two-dimensional random variable $(Y_1, Y_2) = \tau(X_1, X_2)$ be given by

$$Y_1 = 4X_1 - X_2, \qquad Y_2 = 4X_1 + X_2.$$

3. *Prove that τ maps $\mathbb{R}_+ \times \mathbb{R}_-$ bijectively onto*

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

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

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

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

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

1. and 2. Since X_1 is exponentially distributed, $a = 4$, and $-X_2$ is exponentially distributed with $a = 1$, we get

$$E\{X_1\} = \frac{1}{4}, \qquad V\{X_1\} = \frac{1}{16},$$

and

$$E\{X_2\} = -1, \qquad V\{X_2\} = 1.$$

3. It follows from $y_1 = 4x_1 - x_2$ and $y_2 = 4x_1 + x_2$ that

$$x_1 = \frac{1}{8}(y_1 + y_2) \quad \text{and} \quad x_2 = \frac{1}{2}(y_2 - y_2).$$

Since $y_1 = 4x_1 - x_2 > 0$ and $8x_1 = y_1 + y_2 > 0$, $2x_2 = y_2 - y_1 < 0$, we get

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

4. The Jacobian is

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{1}{8} & \frac{1}{8} \\ -\frac{1}{2} & \frac{1}{2} \end{vmatrix} = \frac{1}{16}\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} = -\frac{1}{8}.$$

Since the simultaneous frequency of (X_1, X_2) is

$$g(x_1, x_2) = \begin{cases} 4\,\exp(-4x_1 + x_2) & \text{for } x_1 > 0 \text{ og } x_2 < 0, \\ 0 & \text{ellers}, \end{cases}$$

the simultaneous frequency $k\left(y_1, y_2\right)$ of $\left(Y_1, Y_2\right)$ is given by

$$k\left(y_1, y_2\right)=\begin{cases} \frac{1}{2} \exp\left(-y_1\right) & \text{i } D', \\ \\ 0 & \text{otherwise.} \end{cases}$$

5. The marginal frequency of Y_1 for $y_1>0$ is given by

$$\int_{-y_1}^{y_1} k\left(y_1, y_2\right)\, dy_2 = y_1 \exp\left(-y_1\right),$$

thus

$$h_{Y_1}\left(y_1\right)=\begin{cases} y_1 \exp\left(-y_1\right) & \text{for } y_1>0, \\ \\ 0 & \text{otherwise.} \end{cases}$$

The marginal frequency of Y_2 is given by

$$h_{Y_2}\left(y_2\right)=\int_{\left|y_2\right|}^{\infty} \frac{1}{2} \exp\left(-y_1\right)\, dy_1 = \frac{1}{2} \exp\left(-\left|y_2\right|\right), \qquad y_2 \in \mathbb{R}.$$

6. Since X_1 and X_2 are independent, and $\text{Cov}(\cdot,\cdot)$ is bilinear, we get

$$\text{Cov}\left(Y_1, Y_2\right) = \text{Cov}\left(4X_1 - X_2, 4X_1 + X_2\right) = 16\,V\left\{X_1\right\} - V\left\{X_2\right\} = 16 \cdot \frac{1}{16} - 1 = 0,$$

hence Y_1 and Y_2 are non-correlated.

7. Since D' is not an axes parallel domain, Y_1 and Y_2 cannot be independent.

ALTERNATIVELY it follows immediately that

$$h\left(y_1\right) \cdot h_{Y_2}\left(y_2\right) \neq k\left(y_1, y_2\right).$$

Example 2.3 *Assume that X_1, X_2, X_3, … are independent random variables, such that X_k for every $k \in \mathbb{N}$ has the frequency*

$$f_k(x) = \begin{cases} k\,e^{-kx}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Furthermore, let

$$Y_n = \sum_{k=1}^{n} \qquad \text{and} \qquad Z_n = Y_n - \ln n, \qquad n \in \mathbb{N}.$$

1) Find mean and variance of X_k, and mean and variance of Y_n.

2) Find the frequency of Y_2.

3) Prove for every $n \in \mathbb{N}$ that the frequency $g_n(y)$ of Y_n is given by

$$g_n(y) = \begin{cases} n\,e^{-y}\left(1 - e^{-y}\right)^{n-1} = n\,e^{-ny}\left(e^{y} - 1\right)^{n-1}, & y \geq 0, \\ 0, & y < 0. \end{cases}$$

HINT: *Use e.g. induction; the essential step is to prove that if the formula is true for some given value of $n_0 \in \mathbb{N}$, then it also holds for the following value $n_0 + 1$.*

4) Find the distribution function $G_n(y)$ for Y_n.

5) Find the distribution function $H_n(z)$ for Z_n.

6) Prove that the sequence $\left(Z_n\right)_{n=1}^{\infty}$ converges in distribution towards a random variable Z, and find the distribution function and the frequency of Z.

1) Every X_k is exponentially distributed with $a = k$, so

$$E\left\{X_k\right\} = \frac{1}{k} \qquad \text{og} \qquad V\left\{X_k\right\} = \frac{1}{k^2}.$$

Since the X_k are mutually independent, we get

$$E\{Y_n\} = \sum_{k=1}^{n} E\{X_k\} = \sum_{k=1}^{n} \frac{1}{k},$$

and

$$V\{Y_n\} = \sum_{k=1}^{n} V\{X_k\} = \sum_{k=1}^{n} \frac{1}{k^2}.$$

2) Since $Y_2 = X_1 + X_2$ only has positive values, we have $g_2(y) = 0$ for $y \le 0$. If $y > 0$, then

$$\begin{aligned} g_2(y) &= \int_0^{\infty} f_1(x)\, f_2(y-x)\, dx = \int_0^y e^{-x} \cdot 2e^{-2(y-x)}\, dx = 2\, e^{-2y} \int_0^y e^x\, dx = 2\, e^{-2y}\left(e^y - 1\right) \\ &= 2\, e^{-y}\left(1 - e^{-y}\right). \end{aligned}$$

3) It is obvious that the formula is true for $n = 1$ and $n = 2$, cf. **2.** Then assume that the formula holds for some $n \in \mathbb{N}$. Since Y_{n+1} only has positive values, we have $g_{n+1}(y) = 0$ for $y \le 0$. If $y > 0$, then

$$\begin{aligned} g_{n+1}(y) &= \int_0^y g_n(x) \cdot f_{n+1}(y-x)\, dx = \int_0^y n\, e^{-nx}\left(e^x - 1\right)^{n-1} \cdot (n+1)\, e^{-(n+1)(y-x)}\, dx \\ &= n(n+1)\, e^{-(n+1)y} \int_0^y \left(e^x - 1\right)^{n-1} \cdot e^x\, dx, \qquad\qquad t = e^x \\ &= (n+1)\, e^{-(n+1)y} \int_1^{e^y} n\,(t-1)^{n-1}\, dt = (n+1)\, e^{-(n+1)y}\left(e^y - 1\right)^n, \end{aligned}$$

which is the formula under consideration with n replaced by $n+1$, and the formula follows by induction.

4) The distribution function of Y_n is $G_n(y) = 0$ for $y \le 0$. If $y > 0$, then

$$G_n(y) = \int_0 g_n(t)\, dt = \int_0^y n\, e^{-t}\left(1 - e^{-t}\right)^{n-1}\, dt = \left[\left(1 - e^{-t}\right)^n\right]_0^y = \left(1 - e^{-y}\right)^n,$$

hence

$$G_n(y) = \begin{cases} \left(1 - e^{-y}\right)^n & \text{for } y > 0, \\ 1 & \text{for } y \le 0. \end{cases}$$

5) By a rearrangement,

$$H_n(z) = P\{Z_n \le z\} = P\{Y_n - \ln n \le z\} = P\{Y_n \le z + \ln n\}.$$

If $z \le -\ln n$, we get $H_n(z) = 0$. If $z > -\ln n$, then instead

$$H_n(z) = G_n(z + \ln n) = \{1 - \exp(-z - \ln n)\}^n = \left(1 - \frac{1}{n}\, e^{-z}\right)^n,$$

so

$$H_n(z) = \begin{cases} \left(1 - \dfrac{1}{n}\, e^{-z}\right)^n & \text{for } z > -\ln n, \\[2mm] 0 & \text{for } z \le -\ln n. \end{cases}$$

6) To any $z \in \mathbb{R}$ there exists an $n_0 \in \mathbb{N}$, such that $z > -\ln n$ for every $n \ge n_0$. Then for such $n \ge n_0$,

$$H_n(z) = \left\{1 - \frac{e^{-z}}{n}\right\}^n \to \exp\left(-e^{-z}\right), \qquad \text{for } n \to \infty,$$

hence (Z_n) converges in distribution towards a random variable Z, the distribution function of which is

$$H(z) = \exp\left(-e^{-z}\right), \qquad z \in \mathbb{R}.$$

The frequency is obtained by a differentiation,

$$h(z) = e^{-z} \cdot \exp\left(-e^{-z}\right) = \exp\left(-z - e^{-z}\right), \qquad z \in \mathbb{R}.$$

Example 2.4 *An instrument A contains two components of lifetimes X_1 and X_2, which are assumed to be independent random variables, both of the frequency*

$$f(x) = \begin{cases} \alpha\, e^{-\alpha x}, & x > 0, \\ 0, & x \le 0, \end{cases}$$

where α is a positive constant. The instrument A can only operate if both components work, so the lifetime X for A is $X = \min\{X_1, X_2\}$.
Another instrument B also contains two components, the lifetimes of which, Y_1 and Y_2, are independent random variables, both of the frequency

$$g(y) = \begin{cases} 4\,\alpha\, e^{-4\alpha y}, & y > 0, \\ 0, & y \le 0. \end{cases}$$

Instrument B first uses one component, and when it breaks down, one immediately changes to the other component, so the lifetime Y of B is $Y = Y_1 + Y_2$.

1) Find frequency, mean and variance of X.

2) Find frequency, mean and variance of Y.

3) Find the frequency and the distribution function of the random variable $U = \dfrac{X}{Y}$.

4) Check if U has a mean, and if this is the case find it.

5) Find $P\{Y > X\}$.

It is immediately seen that X_1 and X_2 are exponentially distributed with parameter $a = \alpha$ and the distribution function

$$F_{X_i} = \begin{cases} 1 - e^{-\alpha\, x} & \text{for } x > 0, \\ 0 & \text{for } x \le 0. \end{cases}$$

Analogously, Y_1 and Y_2 are exponentially distributed with $a = 4\alpha$; however, in this case we do not need to indicate the distribution function.

1) The distribution function of $X = \min\{X_1, X_2\}$ is

$$F_X(x) = 1 - (1 - F_{X_1}(x)) \cdot (1 - F_{X_2}(x)) = \begin{cases} 1 - e^{-2\alpha x} & \text{for } x > 0, \\ 0 & \text{for } x \le 0, \end{cases}$$

hence X is exponentially distributed with $a = 2\alpha$, so

$$f_X(x) = \begin{cases} 2\alpha\, e^{-2\alpha x} & \text{for } x > 0, \\ 0 & \text{for } x \le 0, \end{cases}$$

and

$$E\{X\} = \frac{1}{2\alpha} \qquad \text{and} \qquad V\{X\} = \frac{1}{4\alpha^2}.$$

2) Since Y only has positive values, $g_Y(y) = 0$ for $y \le 0$. If $y > 0$, then

$$g_Y(y) = \int_0^y g(t)\, g(y-t)\, dt = 16\alpha^2 \int_0^y e^{-4\alpha t} \cdot e^{-4\alpha(y-t)}\, dt = 16\alpha^2 y\, e^{-4\alpha y},$$

so $Y \in \Gamma\left(2, \dfrac{1}{4\alpha}\right)$ is Gamma distributed, and

$$E\{Y\} = 2 \cdot \frac{1}{4\alpha} = \frac{1}{2\alpha} \quad \text{og} \quad V\{Y\} = \frac{1}{16\alpha^2} \cdot 2 = \frac{1}{8\alpha^2}.$$

At this stage we cannot assume that the Gamma distribution is known, so we have ALTERNATIVELY by direct computations,

$$E\{Y\} = 16\alpha^2 \int_0^y y^2 e^{-4\alpha y}\, dy = \frac{1}{4\alpha} \int_0^\infty t^2 e^{-t}\, dt = \frac{2}{4\alpha} = \frac{1}{2\alpha},$$

and

$$E\left\{Y^2\right\} = 16\alpha^2 \int_0^\infty y^3 e^{-4\alpha y}\, dy = \frac{1}{16\alpha^2} \int_0^\infty t^3 e^{-t}\, dt = \frac{6}{16\alpha^2} = \frac{3}{8\alpha^2},$$

thus

$$V\{Y\} = E\left\{Y^2\right\} - (E\{Y\})^2 = \frac{3}{8\alpha^2} - \frac{1}{4\alpha^2} = \frac{1}{8\alpha^2}.$$

3) Since $U = \dfrac{X}{Y}$ only has positive values, the frequency is $h(u) = 0$ for $u \le 0$.
Since X and Y are stochastically independent, we get $u > 0$,

$$\begin{aligned} h(u) &= \int_0^\infty f_X(uy) \cdot g_Y(y) \cdot |y|\, dy = \int_0^\infty 2\alpha\, e^{-2\alpha u y} \cdot 16\alpha^2 y\, e^{-4\alpha y} \cdot y\, dy \\ &= 32\alpha^3 \int_0^\infty y^2 \exp(-2\alpha(u+2)y)\, dy = \frac{32\alpha^3}{\{2\alpha(u+2)\}^3} \int_0^\infty t^2\, e^{-t}\, dt \\ &= \frac{32\alpha^3}{8\alpha^3(u+2)^3} \cdot 2! = \frac{8}{(u+2)^3}, \end{aligned}$$

hence

$$h(u) = \begin{cases} \dfrac{8}{(u+2)^3} & \text{for } u > 0, \\ \\ 0 & \text{for } u \le 0. \end{cases}$$

The distribution function is then obtained by an integration,

$$H(u) = \begin{cases} 1 - \dfrac{4}{(u+2)^2} & \text{for } u > 0, \\ \\ 0 & \text{for } u \le 0. \end{cases}$$

4) It follows by an explicit computation that

$$\int_{-\infty}^{\infty} |u|\, h(u)\, du = \int_{-\infty}^{\infty} u\, h(u)\, du = \int_{0}^{\infty} u\, h(u)\, du = 8 \int_0^{\infty} \frac{u+2-2}{(u+2)^3}\, du$$
$$= 8 \int_0^{\infty} \left\{ \frac{1}{(u+2)^2} - \frac{1}{(u+2)^3} \right\} du = 8 \left[-\frac{1}{u+2} + \frac{1}{(u+2)^2} \right]_0^{\infty} = 8 \left(\frac{1}{2} - \frac{1}{4} \right) = 2 = E\{U\},$$

hence the mean *exists* and it is equal to $E\{U\} = 2$.

Alternatively, it follows that

$$E\left\{\frac{1}{Y}\right\} = \int_0^{\infty} \frac{1}{y}\, f_Y(y)\, dy = \int_0^{\infty} 16\alpha^2 e^{-4\alpha y} dy = \frac{16\alpha^2}{4\alpha} = 4\alpha,$$

from which

$$E\{U\} = E\{X\} \cdot E\left\{\frac{1}{Y}\right\} = 2.$$

5) From $U = \dfrac{X}{Y}$ follows that

$$P\{Y > X\} = P\{U < 1\} = H(1) = 1 - \frac{4}{(1+2)^2} = \frac{5}{9}.$$

3 The Normal Distribution

Example 3.1 *Given a random variable X, which is normally distributed of mean 2.12. Given that*

$$P\{X \geq 3\} = 0.324,$$

find the variance of X.

If follows from

$$0.324 = P\{X \geq 3\} = P\{X - \mu \geq 0.88\} = 1 - P\left\{\frac{X-\mu}{\sigma} < \frac{0.88}{\sigma}\right\} = 1 - \Phi\left(\frac{0.88}{\sigma}\right),$$

that

$$\Phi\left(\frac{0.88}{\sigma}\right) = 1 - 0.324 = 0.676 \approx \Phi(0.4567),$$

thus

$$\frac{0.88}{\sigma} = 0.4567,$$

and hence

$$V\{X\} = \sigma^2 = \left(\frac{0.88}{0.4567}\right)^2 = 3.713.$$

Example 3.2 *Given a normally distributed random variable X, for which*

$$P\{X \leq 3\} = 0.9087 \qquad \textit{and} \qquad P\{X \leq 2\} = 0.6030.$$

Find mean and variance of X.

First rearrange the given data in the following way

$$0.9087 = P\{X \leq 3\} = P\left\{\frac{X-\mu}{\sigma} \leq \frac{3-\mu}{\sigma}\right\} = \Phi\left(\frac{3-\mu}{\sigma}\right)$$

and

$$0.6030 = P\{X \leq 2\} = P\left\{\frac{X-\mu}{\sigma} \leq \frac{2-\mu}{\sigma}\right\} = \Phi\left(\frac{2-\mu}{\sigma}\right).$$

Using a table we get the inverse of Φ,

$$\frac{3-\mu}{\sigma} = \Phi^{-1}(0.9087) = 1.333, \qquad \text{thus } 3 - \mu = 1.333\,\sigma,$$

and

$$\frac{2-\mu}{\sigma} = \Phi^{-1}(0-6030) = 0.261, \qquad \text{thus } 2 - \mu = 0.261\,\sigma.$$

Hence, $1.071\,\sigma = 1$, so $\sigma = 0.934$, and

$$V\{X\} = \sigma^2 = 0.872,$$

and

$$E\{X\} = \mu = 3 - 1.333\,\sigma = 1.7555.$$

Example 3.3 *As usual, $\varphi(x)$ and $\Phi(x)$ denote the frequency and the distribution function, resp. in an $N(0,1)$-distribution. Obviously, we have the inequality for every $x > 0$,*

$$\int_x^\infty \varphi(y)\left(1 - \frac{3}{y^4}\right) dy < \int_x^\infty \varphi(y)\,dy < \int_x^\infty \varphi(y)\left(1 + \frac{1}{y^2}\right) dy.$$

1) Apply this inequality to prove that for $x > 0$ we have the inequality

$$\frac{1}{x}\left(1 - \frac{1}{x^2}\right)\varphi(x) < 1 - \Phi(x) < \frac{1}{x}\,\varphi(x),$$

and prove that

$$\lim_{x\to\infty} \frac{1 - \Phi(x)}{\varphi(x)} = 0.$$

2) Prove that

$$\frac{d}{dx}\left(\frac{1 - \Phi(x)}{\varphi(x)}\right) = -1 + x \cdot \frac{1 - \Phi(x)}{\varphi(x)}.$$

3) Prove that

$$\lim_{x\to\infty} \frac{d}{dx}\left(\frac{1 - \Phi(x)}{\varphi(x)}\right) = 0.$$

1) Clearly,

$$\int_x^\infty \varphi(y)\,dy = 1 - \Phi(x).$$

From $x > 0$ we immediately get the right hand side of the inequality, because $1 \le \dfrac{y}{x}$ for $y \ge x$,

$$\int_x^\infty \varphi(y)\,dy \le \frac{1}{x}\int_x^\infty y\,\varphi(y)\,dy = \frac{1}{x}\left[-\varphi(y)\right]_x^\infty = \frac{\varphi(x)}{x},$$

where $\varphi'(y) = -y\,\varphi(y)$.
Furthermore, by a partial integration,

$$\begin{aligned}
\int_x^\infty \varphi(y)\left(1 - \frac{3}{y^4}\right) dy &= \left[\varphi(y)\left(y + \frac{1}{y * 3}\right)\right]_x^\infty - \int_x^\infty \varphi'(y)\left(y + \frac{1}{y^3}\right) dy \\
&= \int_x^\infty \varphi(y)\left(y^2 + \frac{1}{y^2}\right) dy - \varphi(x)\cdot x - \varphi(x)\cdot\frac{1}{x^3} \\
&= \frac{1}{x}\left(1 - \frac{1}{x^2}\right)\varphi(x) + \int_x^\infty \varphi(y)\left(y^2 + \frac{1}{y^2}\right) dy - \varphi(x)\left(x + \frac{1}{x}\right).
\end{aligned}$$

Now,

$$\begin{aligned}\frac{d}{dx}\left\{\varphi(x)\left(x+\frac{1}{x}\right)\right\} &= \varphi'(x)\left(x+\frac{1}{x}\right)+\varphi(x)-\frac{1}{x^2}\,\varphi(x)\\ &= -x^2\,\varphi(x)-\varphi(x)+\varphi(x)-\frac{1}{x^2}\,\varphi(x)=-\varphi(x)\left(x^2+\frac{1}{x^2}\right),\end{aligned}$$

so

$$\begin{aligned}&\int_x^{\infty}\varphi(y)\left(y^2+\frac{1}{y^2}\right)\,dy-\varphi(x)\cdot\left(x+\frac{1}{x}\right)\\ &\qquad=\int_x^{\infty}\varphi(y)\left(y^2+\frac{1}{y^2}\right)\,dy-\int_x^{\infty}\varphi(y)\cdot\left(y^2+\frac{1}{y^2}\right)\,dy=0,\end{aligned}$$

and thus

$$1-\Phi(x)>\int_x^{\infty}\varphi(y)\left(1-\frac{3}{y^4}\right)\,dy=\frac{1}{x}\left(1-\frac{1}{x^2}\right)\varphi(x)+0=\frac{1}{x}\left(1-\frac{1}{x^2}\right)\varphi(x),$$

and we have proved the inequalities.
Now, $\varphi(x)>0$, so it follows from the former result that

$$\frac{1}{x}\left(1-\frac{1}{x^2}\right)<\frac{1-\Phi(x)}{\varphi(x)}<\frac{1}{x}.$$

Both limits tend towards 0 for $x\to\infty$, hence

$$\lim_{x\to\infty}\frac{1-\Phi(x)}{\varphi(x)}=0.$$

2) It follows from $\Phi'(x)=\varphi(x)$, and $\varphi'(x)=-x\,\varphi(x)$ that

$$\begin{aligned}\frac{d}{dx}\left(\frac{1-\Phi(x)}{\varphi(x)}\right) &= \frac{-\Phi'(x)}{\varphi)}-\frac{1-\Phi(x)}{\varphi(x)^2}\cdot\varphi'(x)=-\frac{\varphi(x)}{\varphi(x)}+\frac{1-\Phi(x)}{\varphi(x)^2}\cdot x\,\varphi(x)\\ &= -1+x\cdot\frac{1-\Phi(x)}{\varphi(x)}.\end{aligned}$$

3) If we again use the inequalities proved in (1), we get

$$-1+\frac{x}{\varphi(x)}\cdot\frac{1}{x}\left(1-\frac{1}{x^2}\right)\varphi(x)<\frac{d}{dx}\left(\frac{1-\Phi(x)}{\varphi(x)}\right)<-1+\frac{x}{\varphi(x)}\cdot\frac{1}{x}\,\varphi(x),$$

hence by a reduction,

$$-\frac{1}{x^2}<\frac{d}{dx}\left(\frac{1-\Phi(x)}{\varphi(x)}\right)=0.$$

Both limits tend towards 0 for $x\to\infty$, so

$$\lim_{x\to\infty}\frac{d}{dx}\left(\frac{1-\Phi(x)}{\varphi(x)}\right)=0.$$

Example 3.4 *Let* $X \in N(0,1)$*. Compute for every* $n \in \mathbb{N}$ *the moments*

$$E\{X^n\} \qquad \textit{and} \qquad E\{|X|^n\}.$$

If $n = 2m+1$ is an odd number, then

$$\begin{aligned} E\left\{|X|^{2m+1}\right\} &= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^{2m+1} \exp\left(-\frac{1}{2}\,x^2\right) dx = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^{2m+1} \exp\left(-\frac{1}{2}x^2\right) dx \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \left(2 \cdot \frac{x^2}{2}\right)^m \exp\left(-\frac{1}{2}\,x^2\right) d\left(\frac{x^2}{2}\right) = \frac{2^{m+1}}{\sqrt{2\pi}} \int_0^{\infty} y^m e^{-y}\, dy = \frac{2^{m+1} m!}{\sqrt{2\pi}}. \end{aligned}$$

It follows in particular that all moments exist.

If $n = 2m+1$ is odd, then the integrand is odd, and it follows by the symmetry that

$$E\left\{X^{2m+1}\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{2m+1} \exp\left(-\frac{1}{2}\,x^2\right) dx = 0.$$

If $n = 2m$ is even, then

$$\begin{aligned}
E\left\{X^{2m}\right\} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{2m} \exp\left(-\frac{1}{2}\,x^2\right) dx \\
&= \frac{1}{\sqrt{2\pi}} \left[\frac{x^{2m+1}}{2m+1}\, \exp\left(-\frac{1}{2}\,x^2\right)\right]_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{x^{2m+1}}{2m+1}\, x\, \exp\left(-\frac{1}{2}\,x^2\right) dx \\
&= \frac{1}{2m+1} \cdot \frac{1}{\sqrt{2\pi}} \int_0^{\infty} x^{2m+2} \exp\left(-\frac{1}{2}\,x^2\right) dx = \frac{1}{2m+1}\, E\left\{X^{2m+2}\right\},
\end{aligned}$$

thus

$$E\left\{X^{2m+2}\right\} = (2m+1)\, E\left\{X^{2m}\right\},$$

and hence by a change of variable,

$$E\left\{X^{2m}\right\} = (2m-1)\, E\left\{X^{2m-2}\right\}.$$

Then by recursion,

$$\begin{aligned}
E\left\{X^{2m}\right\} &= E\left\{|X|^{2m}\right\} = (2m-1)\, E\left\{X^{2m-2}\right\} = (2m-1)(2m-3)\cdots 3\cdot 1 \\
&= \frac{2m(2m-1)(2m-2)\cdots 3\cdot 2\cdot 1}{(2\cdot m)\cdot(2(m-1))\cdots(2\cdot 1)} = \frac{(2m)!}{2^m \cdot m!}.
\end{aligned}$$

Example 3.5 *Let X_1 and X_2 be independent random variables, $X_i \in N\left(0, \sigma^2\right)$, $i = 1, 2$. Find the distribution function and the frequency of the random variable*

$$Y = \sqrt{X_1^2 + X_2^2}.$$

Find also the mean and the variance of Y.
The distribution of Y is called the Rayleigh distribution*.*

If $y \le 0$, then $P\{Y \le y\} = 0$. If instead $y > 0$, then we apply that X_1 and X_2 are independent and identically distributed, and use polar coordinates,

$$\begin{aligned}
P\{Y \le y\} &= P\left\{X_1^2 + X_2^2 \le y^2\right\} = \frac{1}{2\pi\sigma^2} \int_{\left\{x_1^2+x_2^2 \le y^2\right\}} \exp\left(-\frac{1}{2\sigma^2}\,\left(x_1^2+x_2^2\right)\right) dx_1\, dx_2 \\
&= \frac{1}{2\pi\sigma^2} \cdot 2\pi \int_0^y r\, \exp\left(-\frac{r^2}{2\sigma^2}\right) dr = 1 - \exp\left(-\frac{y^2}{2\sigma^2}\right).
\end{aligned}$$

Here we have used that the simultaneous frequency of (X_1, X_2) is $\varphi\left(x_1\right) \cdot \varphi\left(x_2\right)$.

Summing up,

$$F_Y(y) = \begin{cases} 0 & \text{for } y \le 0, \\[2ex] 1 - \exp\left(-\dfrac{y^2}{2\sigma^2}\right) & \text{for } y > 0, \end{cases}$$

and

$$f_Y(y) = \begin{cases} 0 & \text{for } y < 0, \\ \dfrac{y}{\sigma^2} \exp\left(-\dfrac{y^2}{2\sigma^2}\right) & \text{for } y \geq 0. \end{cases}$$

The mean is

$$E\{Y\} = \int_0^\infty \frac{y^2}{\sigma^2} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy = \sigma \int_0^\infty z^2 \exp\left(-\frac{z^2}{2}\right) dz = \sigma\sqrt{\frac{\pi}{2}}.$$

Since

$$E\left\{Y^2\right\} = E\left\{X_1^2\right\} + E\left\{X_2^2\right\} = 2\sigma^2,$$

we get the variance

$$V\{Y\} = E\left\{Y^2\right\} - (E\{Y\})^2 = 2\sigma^2 - \sigma^2 \cdot \frac{\pi}{2} = \left(2 - \frac{\pi}{2}\right)\sigma^2.$$

Example 3.6 *Let X_1 and X_2 be independent and identically distributed random variables of the frequency*

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi x}} \exp\left(-\dfrac{x}{2}\right), & x > 0, \\ 0, & x \leq 0. \end{cases}$$

1) Find the frequency of $Y = X_1 + X_2$.

2) Prove that if Z is normally distributed of mean 0 and variance 1, then Z^2 has the frequency $f(x)$.

3) Find mean and variance of Y.

1) If $s > 0$, then the frequency of $Y = X_1 + X_2$ is

$$\begin{aligned} h(s) &= \int_0^s \frac{1}{\sqrt{2\pi x}} \exp\left(-\frac{x}{2}\right) \cdot \frac{1}{\sqrt{2\pi(s-x)}} \exp\left(-\frac{1}{2}(s-x)\right) dx \\ &= \frac{1}{2\pi} \exp\left(-\frac{s}{2}\right) \int_0^s \frac{dx}{\sqrt{x(s-x)}} = \frac{1}{2\pi} \exp\left(-\frac{s}{2}\right) \int_0^s \frac{\frac{1}{s}\, dx}{\sqrt{\frac{x}{s}\left(1-\frac{x}{s}\right)}} \\ &= \frac{1}{2} \exp\left(-\frac{s}{2}\right) \cdot \frac{1}{\pi} \int_0^1 \frac{dt}{\sqrt{t(1-t)}} = C \cdot \frac{1}{2} \exp\left(-\frac{1}{2}\, s\right) = \frac{1}{2} \exp\left(-\frac{1}{2}\, s\right), \end{aligned}$$

because $\int_0^\infty h(s)\, ds = 1$, so $C = 1$. Hence

$$h(s) = \begin{cases} \dfrac{1}{2} \exp\left(-\dfrac{s}{2}\right) & \text{for } s > 0, \\ 0 & \text{for } s \leq 0, \end{cases}$$

and it follows that $Y = X_1 + X_2$ is exponentially distributed.

2) Clearly, $F_{Z^2}(x) = 0$ for $x \leq 0$. If $x > 0$, then

$$F_{Z^2}(x) = P\left\{Z^2 \leq x\right\} = P\{-\sqrt{x} \leq Z \leq \sqrt{x}\} = \Phi(\sqrt{x}) - \Phi(-\sqrt{x}).$$

When $z > 0$, the frequency is obtained by a differentiation,

$$f_{Z^2} = \varphi(\sqrt{x}) \cdot \frac{1}{2\sqrt{x}} - \varphi(-\sqrt{x}) \cdot \left(-\frac{1}{2\sqrt{x}}\right) = \frac{1}{\sqrt{x}}\,\varphi(\sqrt{x}) = \frac{1}{\sqrt{2\pi x}}\,\exp\left(-\frac{1}{2}\,x\right),$$

and it follows from $f_{Z^2}(x) = 0$ for $x \leq 0$, that $f_{Z^2}(x) = f(x)$.

3) Since Y is exponentially distributed with $a = \frac{1}{2}$, we get

$$E\{Y\} = \frac{1}{a} = 2 \qquad \text{and} \qquad V\{Y\} = \frac{1}{a^2} = 4.$$

Remark 3.1 Since X_1 and X_2 are independent and identically distributed, we have

$$E\{Y\} = E\left\{X_1\right\} + E\left\{X_2\right\} \quad \text{og} \quad V\{Y\} = V\left\{X_1\right\} + V\left\{X_2\right\},$$

hence

$$E\left\{X_1\right\} = E\left\{X_2\right\} = 1 \quad \text{og} \quad V\left\{X_1\right\} = V\left\{X_2\right\} = 2. \qquad \Diamond$$

Example 3.7 *Let (X_1, X_2) be a two-dimensional random variable of the frequency*

$$h\left(x_1, x_2\right) = \frac{1}{2\pi}\,\exp\left(-\frac{1}{2}\,\left(x_1^2 + x_2^2\right)\right).$$

Let the two-dimensional random variable (Y_1, Y_2) be given by

$$X_1 = \sqrt{-2\ln Y_1} \cdot \cos\left(2\pi Y_2\right), \qquad X_2 = \sqrt{-2\ln Y_1} \cdot \sin\left(2\pi Y_2\right),$$

where $0 < Y_1 < 1$, $0 < Y_2 < 1$.
Find the frequency of (Y_1, Y_2).
Are Y_1 and Y_2 independent?

The Jacobian is

$$\frac{\partial\left(x_1, x_2\right)}{\partial\left(y_1, y_2\right)} = \begin{vmatrix} \dfrac{1}{2}\,\dfrac{1}{\sqrt{-2\ln y_1}}\left\{-\dfrac{2}{y_1}\right\}\cos\left(2\pi y_2\right) & -2\pi\sqrt{-2\ln y_1} \cdot \sin\left(2\pi y_2\right) \\ \dfrac{1}{2}\,\dfrac{1}{\sqrt{-2\ln y_1}}\left\{-\dfrac{2}{y_1}\right\}\sin\left(2\pi y_2\right) & 2\pi\sqrt{-2\ln y_1} \cdot \cos\left(2\pi y_2\right) \end{vmatrix} = -\frac{2\pi}{y_1}.$$

We get from $x_1^2 + x_2^2 = -2\ln y_1$ that the frequency of (Y_1, Y_2) is

$$k\left(y_1, y_2\right) = \frac{1}{2\pi}\,\exp\left(\ln y_1\right) \cdot \left|-\frac{2\pi}{y_1}\right| = 1 \qquad \text{for}\ \ (y_1, y_2) \in \,]0, 1[\, \times \,]0, 1[.$$

It follows immediately that the marginal frequencies are

$$K_{Y_i}(y_i) = 1 \qquad \text{for } y_i \in]0,1[, \qquad i = 1,\, 2,$$

and $k_{Y_i}(y_i) = 0$ otherwise. Hence

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

so Y_1 and Y_2 are independent.

Example 3.8 *A random variable $U \in N(0,1)$ is normally distributed of mean 0 and variance 1.*

1. *Prove that $E\left\{U^4\right\} = 3$.*

A two-dimensional random variable (X,Y) is following a two-dimensional normal distribution with

$$E\{X\} = E\{Y\} = 0, \qquad V\{X\} = V\{Y\} = \sigma^2 > 0, \qquad \varrho(X,Y) = \varrho, \quad (|\varrho| < 1),$$

i.e. $(X,Y) \in N\left(0, 0, \sigma^2, \sigma^2, \varrho\right)$.

2. *Compute the real constant a, such that the random variables $X - aY$ and Y are non-correlated.*

3. *Explain why the random variables $X - aY$ and Y are stochastically independent for the value of a found in* **2.**

4. *Compute $E\left\{X^2Y^2\right\}$ by using that $X = (X - aY) + aY$.*

1) Since $\varphi'(u) = -u\,\varphi(u)$, we get by partial integration,

$$\begin{aligned}
E\left\{U^4\right\} &= \int_{-\infty}^{\infty} u^4 \varphi(u)\,du = -\int_{-\infty}^{\infty} u^3 \varphi'(u)\,du \\
&= \left[-u^3\varphi(u)\right]_{-\infty}^{\infty} + 3\int_{-\infty}^{\infty} u^2\varphi(u)\,du = 0 - 3\int_{-\infty}^{\infty} u \cdot \varphi'(u)\,du \\
&= \left[-3u\,\varphi(u)\right]_{-\infty}^{\infty} + 3\int_{-\infty}^{\infty} 1 \cdot \varphi(u)\,du = 0 + 3 = 3.
\end{aligned}$$

2) Then by a direct computation,

$$\operatorname{Cov}(X - aY, Y) = \operatorname{Cov}(X,Y) - a\,V\{Y\} = \varrho\,\sqrt{V\{X\}\,V\{Y\}} - a\,V\{Y\} = (\varrho - a)\sigma^2.$$

This expression is equal to 0 for $a = \varrho$.

3) From $E\{X\} = E\{Y\}$ follows that

$$\begin{aligned}
V\{X - \varrho Y\} &= E\left\{(X - \varrho Y)^2\right\} = E\left\{X^2\right\} - 2\varrho\,E\{XY\} + E\left\{Y^2\right\} \\
&= V\{X\} - 2\varrho\ \operatorname{Cov}(X,Y) + V\{Y\} = \sigma^2 + \sigma^2 - 2\varrho^2\sigma^2 \\
&= 2\left(1 - \varrho^2\right)\sigma^2,
\end{aligned}$$

so the covariance matrix of $(X - \varrho Y, Y)$ is given by

$$\mathbf{C} = \begin{pmatrix} 2\left(1 - \varrho^2\right)\sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix},$$

and the frequency is

$$f(x,y)=\frac{1}{2\pi\sqrt{2\left(1-\varrho^2\right)\sigma^2}}\,\exp\left(-\frac{1}{2}\left\{\left(\frac{x}{\sqrt{2\left(1-\varrho^2\right)\sigma}}\right)^2+\left(\frac{y}{\varrho}\right)^2\right\}\right).$$

The structure of the frequency shows that $X-\varrho Y$ and Y are independent.

4) By using the given trick,

$$\begin{aligned}
E\left\{X^2Y^2\right\} &= E\left\{[(X-\varrho Y)+\varrho Y]^2Y^2\right\}\\
&= E\left\{(X-\varrho Y)^2Y^2+2\varrho(X-\varrho Y)Y^3+\varrho^2Y^4\right\}\\
&= E\left\{(X-\varrho Y)^2\right\}E\left\{Y^2\right\}+2\varrho\,E\{X-\varrho Y\}\,E\left\{Y^3\right\}+\varrho^2E\left\{Y^4\right\}\\
&= \left\{\sigma^2+\varrho^2\sigma^2-2\varrho\cdot\varrho\sigma^2\right\}\sigma^2+0+\varrho^2\cdot 3\sigma^4\\
&= \sigma^4\left(1-\varrho^2\right)+3\varrho^2\sigma^4=\sigma^4\left(1+2\varrho^2\right).
\end{aligned}$$

Here we have applied that $X-\varrho Y$ and Y are independent ant that

$$E\left\{Y^3\right\}=0 \qquad\text{and}\qquad E\left\{Y^4\right\}=3\sigma^4.$$

Example 3.9 *Let $X_1,\ X_2,\ \dots,\ X_n$ be mutually independent $N(0,1)$-distributed random variables, and let $S_k = \sum_{i=1}^{k} X_i$, $k < n$.*

1) Find the distribution of S_n.

2) Find the simultaneous distribution of (S_m, S_n), when $m < n$.

1) From $S_n \in N(0,n)$ follows that the distribution function is

$$F_n(s) = \Phi\left(\frac{s}{\sqrt{n}}\right).$$

2) Analogously we get $S_m \in N(0,m)$. The covariance matrix is

$$\mathbf{C} = \begin{pmatrix} V\{S_m\} & \operatorname{Cov}(S_m, S_n) \\ \operatorname{Cov}(S_m, S_n) & V\{S_n\} \end{pmatrix},$$

where $V\{S_m\} = m$ and $V\{S_n\} = n$, and

$$\operatorname{Cov}(S_m, S_n) = \operatorname{Cov}\left(S_m, S_m + \sum_{i=m+1}^{n} X_i\right) = V\{S_m\} = m,$$

thus

$$\mathbf{C} = \begin{pmatrix} m & m \\ M & n \end{pmatrix}, \qquad \det \mathbf{C} = m(n-m),$$

and hence

$$\mathbf{C}^{-1} = \frac{1}{m(n-m)} \begin{pmatrix} n & -m \\ -m & m \end{pmatrix},$$

so

$$\begin{aligned} (x,\, y)\, \mathbf{C}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{m(n-m)} \left(nx^2 - 2mxy + my^2\right) \\ &= \frac{1}{1 - \frac{m}{n}} \left\{ \left(\frac{x}{\sqrt{m}}\right)^2 - 2 \cdot \sqrt{\frac{m}{n}} \cdot \frac{x}{\sqrt{m}} \cdot \frac{y}{\sqrt{n}} + \left(\frac{y}{\sqrt{n}}\right)^2 \right\}. \end{aligned}$$

This result corresponds to

$$\sigma_1 = \sqrt{m}, \qquad \sigma_2 = \sqrt{n} \quad \text{og} \quad \varrho = \sqrt{\frac{m}{n}},$$

so the distribution is

$$(S_m, S_n) \in N\left(0, 0, m, n, \sqrt{\frac{m}{n}}\right),$$

and the frequency is

$$\begin{aligned} f(x,y) &= \frac{1}{2\pi\sqrt{m(n-m)}} \exp\left(-\frac{nx^2 - 2mxy + my^2}{2m(n-m)}\right) \\ &= \frac{1}{2\pi\sqrt{m}\cdot\sqrt{n}\cdot\sqrt{1-\frac{m}{n}}} \cdot \exp\left(-\frac{1}{2}\left\{\left(\frac{x}{\sqrt{m}}\right)^2 - 2\sqrt{\frac{m}{n}} \cdot \frac{x}{\sqrt{m}} \cdot \frac{y}{\sqrt{n}} + \left(\frac{y}{\sqrt{n}}\right)^2\right\}\right). \end{aligned}$$

Example 3.10 *Let (X_1, X_2) follow the two-dimensional normal distribution*

$$(X_1, X_2) \in N\left(0, 0, \sigma_1^2, \sigma_2^2, \varrho\right),$$

and let the random variable Y be given by $Y = X_1/X_2$.
Prove that Y is Cauchy distributed of median $\varrho \, \dfrac{\sigma_1}{\sigma_2}$.
Find in particular the frequency in the case $\sigma_1 = \sigma_2$, $\varrho = 0$.

Since $(X_1, X_2) \in N\left(0, 0, \sigma_1^2, \sigma_2^2, \varrho\right)$, we immediately get the frequency,

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}} \cdot \exp\left[-\frac{1}{2(1-\varrho^2)}\left\{\left(\frac{x_1}{\sigma_1}\right)^2 - 2\varrho\,\frac{x_1}{\sigma_1}\,\frac{x_2}{\sigma_2} + \left(\frac{x_2}{\sigma_2}\right)^2\right\}\right].$$

Then by a known formula, the frequency of $Y = X_1/X_2$ is given by

$$\begin{aligned}
g(y) &= \int_{-\infty}^{\infty} f(yx, x) \cdot |x|\, dx \\
&= \frac{1}{\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}} \cdot \int_0^{\infty} \exp\left[-\frac{1}{2\,(1-\varrho^2)}\left\{\frac{y^2}{\sigma_1^2}\,x^2 - 2\,\frac{\varrho y}{\sigma_1\sigma_2}\,x^2 + \frac{1}{\sigma_2^2}\,x^2\right\}\right] |x|\, dx \\
&= \frac{1}{\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}} \cdot \int_0^{\infty} \exp\left[-\frac{1}{1-\varrho^2}\left\{\left(\frac{y}{\sigma_1} - \frac{\varrho}{\sigma_2}\right)^2 + \left(\frac{\sqrt{1-\varrho^2}}{\sigma_2}\right)^2\right\}\frac{x^2}{2}\right] x\, dx \\
&= \frac{1}{\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}} \cdot \frac{1-\varrho^2}{\left(\dfrac{y}{\sigma_1} - \dfrac{\varrho}{\sigma_2}\right)^2 + \left(\dfrac{\sqrt{1-\varrho^2}}{\sigma_2}\right)^2} \\
&= \frac{1}{\pi} \cdot \frac{\sigma_1}{\sigma_2}\sqrt{1-\varrho^2} \cdot \frac{1}{\left(y - \dfrac{\varrho\sigma_1}{\sigma_2}\right)^2 + \left(\dfrac{\sigma_1}{\sigma_2}\sqrt{1-\varrho^2}\right)^2}, \qquad y \in \mathbb{R},
\end{aligned}$$

which is the frequency of a Cauchy distribution of median $\varrho \, \dfrac{\sigma_1}{\sigma_2}$.

If $\varrho = 0$ and $\sigma_1 = \sigma_2 = \sigma$, then

$$g(y) = \frac{1}{\pi} \cdot \frac{1}{1+y^2}, \qquad y \in \mathbb{R}.$$

Example 3.11 *Let X_1 and X_2 be independent $N(0,1)$-distributed random variables, i.e. (X_1, X_2) has the two-dimensional frequency*

$$h(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}\left(x_1^2 + x_2^2\right)\right), \qquad (x_1, x_2) \in \mathbb{R}^2.$$

Let the random variables Y_1 and Y_2 be given by

$$Y_1 = \frac{X_1}{X_2}, \qquad Y_2 = \sqrt{X_1^2 + X_2^2}.$$

1) Find frequency and distribution function of Y_1.

2) Find frequency and distribution function of Y_2.

3) Compute for $y_1 \in \mathbb{R}$ and $y_2 > 0$ the probability

$$P\{Y_1 \le y_1,\, Y_2 \le y_2\}.$$

HINT: *Express the probability as a plane integral of h over a suitable set, and then compute the integral by using polar coordinates.*

4) Find the simultaneous frequency of (Y_1, Y_2).

1) By a standard formula the frequency of X/Y is given by

$$\begin{aligned} f(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left(z^2+1\right)x^2\right)\cdot |x|\,dx = \frac{1}{\pi}\int_0^{\infty}\exp\left(-\left(z^2+1\right)\frac{x^2}{2}\right)\cdot x\,dx \\ &= \frac{1}{\pi}\cdot\frac{1}{z^2+1}, \end{aligned}$$

thus we have a Cauchy-distribution, and the distribution function is

$$F(z) = \frac{1}{\pi}\left[\operatorname{Arctan} t\right]_{-\infty}^{z} = \frac{1}{\pi}\operatorname{Arctan} z + \frac{1}{2}.$$

2) The distribution function of Y_2 is 0 for $r \le 0$. If $r > 0$, then

$$\begin{aligned} G(r) &= P\{Y_2 \le r\} = \int_{\left\{\sqrt{x_1^2+x_2^2}\le r\right\}} \frac{1}{2\pi}\exp\left(-\frac{1}{2}\left(x_1^2+x_2^2\right)\right)dx_1\,dx_2 \\ &= \frac{1}{2\pi}\cdot 2\pi\int_0^r \exp\left(-\frac{1}{2}\varrho^2\right)\varrho\,d\varrho = \left[-\exp\left(-\frac{1}{2}\varrho^2\right)\right]_0^r = 1-\exp\left(-\frac{1}{2}r^2\right), \end{aligned}$$

so the distribution function is

$$G(r) = \begin{cases} 0 & \text{for } r \le 0, \\[2mm] 1-\exp\left(-\dfrac{1}{2}r^2\right) & \text{for } r > 0. \end{cases}$$

The corresponding frequency is

$$g(r) = \begin{cases} 0 & \text{for } r \le 0, \\[2mm] r\exp\left(-\dfrac{1}{2}r^2\right) & \text{for } r > 0. \end{cases}$$

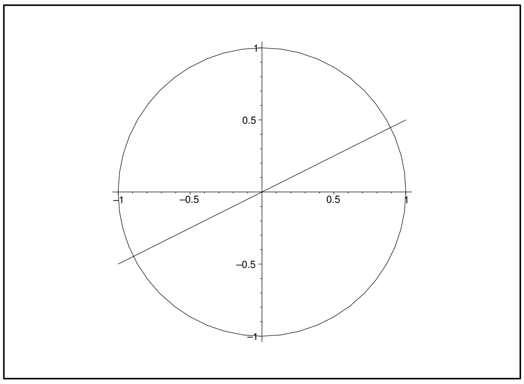


Figure 1: When $y_1 = 2$, the domain of integration is the union of two circular sections of the complementary angle $\varphi_0 = \text{Arccot}\, y_1$.

3) It follows by a geometrical analysis that $P\{Y_1 \le y_1,\, Y_2 \le y_2\}$ is the integral of $h(x_1, x_2)$ over the union of two circular sections of the same angle. The upper circular section has a part of the negative x_1-axis as part of its boundary, and the complementary angle is

$$\varphi_0 = \text{Arccot}\, y_1.$$

We get by using polar coordinates

$$P\{Y_1 \le y_1,\, Y_2 \le y_2\} \;=\; 2\int_{\varphi_0}^{\pi}\left\{\int_0^{y_2}\frac{1}{2\pi}\,\exp\left(-\frac{1}{2}\,\varrho^2\right)\varrho\,d\varrho\right\}d\varphi$$
$$= \left\{1-\exp\left(-\frac{1}{2}\,y_2^2\right)\right\}\cdot\left\{1-\frac{1}{\pi}\,\text{Arccot}\;y_1\right\}.$$

4) The simultaneous frequency of (Y_1, Y_2) is

$$k\,(y_1,y_2) = \frac{\partial^2}{\partial y_1 \partial y_2}\,P\{Y_1 \le y_1,\, Y_2 \le y_2\} = \frac{1}{\pi}\cdot\frac{1}{1+y_1^2}\cdot y_2\,\exp\left(-\frac{1}{2}\,y_2^2\right).$$

It follows that Y_1 and Y_2 are stochastically independent.

Example 3.12 *Give an example of a two-dimensional random variable (X, y) which has the following three properties:*

1) X is normally distributed, $X \in N(0,1)$,

2) Y is normally distributed, $Y \in N(0,1)$,

3) (X,Y) does not follow a two-dimensional normal distribution.

HINT: *Try to find a frequency $f(x,y)$, which is 0, if $xy < 0$.*

If we put

$$f(x,y) = \begin{cases} \dfrac{1}{\pi}\,\exp\left(-\dfrac{1}{2}\left(x^2+y^2\right)\right) & \text{for } xy \ge 0, \\[2ex] 0 & \text{for } xy < 0, \end{cases}$$

then all three conditions are satisfied.

Example 3.13 *Let X and Y be independent random variables of the same frequency f, which is positive and continuously differentiable, and assume that*

(4) $f(x)\,f(y) = g\left(x^2+y^2\right)$ *for every (x,y).*

(The meaning of (4) is that the distribution of (X,Y) is rotational invariant).
Prove that

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\exp\left(-\frac{x^2}{2\sigma^2}\right),$$

hence, X and Y are normally distributed.
HINT: *Differentiate (4) with respect to x and derive that*

$$\frac{f'(x)}{f(x)} = \textit{constant}\cdot x.$$

When (4) is differentiated with respect to x and y, resp., we get

$$f'(x)f(y) = 2x\, g'\left(x^2+y^2\right) \quad \text{and} \quad f(x)f'(y) = 2y\, g'\left(x^2+y^2\right).$$

If $x \neq 0$ and $y \neq 0$, we get

$$g'\left(x^2+y^2\right) = \frac{1}{2x}\, f'(x)f(y) = \frac{1}{2y}\, f(x)f'(x).$$

Since $f(x)\cdot f(y) \neq 0$, it follows by separating the variables that

$$\frac{1}{x}\,\frac{f'(x)}{f(x)} = \frac{1}{y}\,\frac{f'(y)}{f(y)}.$$

The left hand side only depends on x, and the right hand side only depends on y, and since they are equal, they must be equal to a constant c, thus

$$\frac{f'(x)}{f(x)} = \frac{d}{dx}\,\ln f(x) = c\cdot x,$$

and whence by integration,

$$f(x) = k\cdot\exp\left(c\,\frac{1}{2}\,x^2\right) = k\cdot\exp\left(-\frac{1}{2}\left\{\frac{x}{\sigma}\right\}^2\right) = \frac{1}{\sqrt{2\pi}\,\sigma}\,\exp\left(-\frac{1}{2}\left\{\frac{x}{\sigma}\right\}^2\right).$$

In fact,

1) since $\int_{-\infty}^{\infty} f(x)\,dx < \infty$, we see that $c = -\dfrac{1}{\sigma^2}$ must be negative, and

2) since $\int_{-\infty}^{\infty} f(x)\,dx = 1$, we have $k = \dfrac{1}{\sqrt{2\pi}\,\sigma}$, and

3) $f(0) = \dfrac{1}{\sqrt{2\pi}\,\sigma}$ by an continuous extension.

4 The Central Limit Theorem

Example 4.1 *Prove that*

$$e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}\to\frac{1}{2}\qquad \textit{for } n\to\infty.$$

HINT: *Apply the Central Limit Theorem on a sequence of independent Poisson distributed random variables.*

Let X_n, $n\in\mathbb{N}$ be independent Poisson distributed random variables with $\lambda=1$, i.e.

$$P\left\{X_n=k\right\}=\frac{\lambda^k}{k!}\,e^{-\lambda}=\frac{1}{k!}\cdot e^{-1},\qquad k\in\mathbb{N}_0.$$

The Poisson distribution is reproductive, so $Y_n=X_1+\cdots+X_n$ is also Poisson distributed, and $Y_n\in P(1+\cdots+1)=P(n)$. This means that the distribution function of Y_n is

$$P\left\{Y_n=k\right\}=\frac{n^k}{k!}\,e^{-n},\qquad k\in\mathbb{N}_0,$$

hence

$$E\left\{Y_n\right\}=\lambda_n=n\quad\text{and}\quad s_n^2=V\left\{Y_n\right\}=n.$$

Then by the Central Limit Theorem,

$$\lim_{n\to\infty}P\left\{\frac{Y_n-E\left\{Y_n\right\}}{s_n}\le x\right\}=\Phi(x)\qquad\text{for every } x\in\mathbb{R}.$$

This means that

$$\begin{aligned}\lim_{n\to\infty}P\left\{\frac{Y_n-n}{\sqrt{n}}\le x\right\}&=\lim_{n\to\infty}P\left\{Y_n\le n+x\sqrt{n}\right\}=\lim_{n\to\infty}\sum_{k=0}^{n+[x\sqrt{n}]}P\left\{Y_n=k\right\}\\&=\lim_{n\to\infty}\sum_{k=0}^{n+[x\sqrt{n}]}P\left\{Y_n=k\right\}=\lim_{n\to\infty}\sum_{k=0}^{n+[x\sqrt{n}]}\frac{n^k}{k!}\,e^{-n}=\Phi(x).\end{aligned}$$

If we choose $x=0$, then

$$\lim_{n\to\infty}e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}=\Phi(0)=\frac{1}{2}.$$

Example 4.2 *A coin is thrown 10,000 gange, where we get the result head in 4979 of the throws. Using this information one wants to judge if the coin may be considered as "honest" (i.e. the probability is $\frac{1}{2}$ for heads). One may follow the following procedure:*

1) *Assuming that the coin is honest, apply the Central Limit Theorem to find the probability that the number of heads, X, lies between 4979 and 5021 (both numbers included).*

2) *Accepting the hypothesis that the coin is honest if the probability above is ≤ 0.95, check if the coin can be considered as honest.*

3) *Repeat the test with another coin, in which case we get $5000 - N$ heads among the 10,000 throws. Find the largest number N, for which the hypothesis that this coin is honest will be accepted by the method described above.*

1) When we assume that the coin is "honest", then the number X of heads is binomially distributed, $X \in B\left(10,000, \frac{1}{2}\right)$ with

$$E\{X\} = 5000 \quad \text{and} \quad V\{X\} = 10,000 \cdot \frac{1}{2} \cdot \frac{1}{2} = 2500 = 50^2.$$

By the *Central Limit Theorem* X is approximated by

$$Y \in N\left(5000, 50^2\right),$$

thus

$$\begin{aligned} P\{4979 \leq X \leq 5021\} &\approx P\{4978.5 \leq Y \leq 5021.5\} = P\{-21,5 \leq Y - 5000 \leq 21.5\} \\ &= P\left\{-0.43 \leq \frac{Y-5000}{50} \leq 0.43\right\} = 2\,\Phi(0.43) - 1 = 0.3328. \end{aligned}$$

2) Since $0.3328 < 0.95$, we shall *accept the hypothesis that the coin is honest.*

3) We find in the same way as in **(1)** that

$$P\{5000 - N \le X \le 5000 + N\} \approx P\left\{5000 - N - \frac{1}{2} \le Y \le 5000 + N + \frac{1}{2}\right\}$$
$$= P\left\{\left|\frac{Y - 5000}{50}\right| \le \frac{N + \frac{1}{2}}{50}\right\} = 2\,\Phi\left(\frac{N + \frac{1}{2}}{50}\right) - 1.$$

We shall find the largest integer N, for which

$$2\,\Phi\left(\frac{N + \frac{1}{2}}{50}\right) - 1 \le 0.95,$$

thus

$$\Phi\left(\frac{N + \frac{1}{2}}{50}\right) \le 0.975.$$

By using tables we get

$$\frac{N + \frac{1}{2}}{50} \le 1.96, \quad \text{i.e.} \quad N + \frac{1}{2} \le 98.$$

the largest integer N, for which the hypothesis is accepted, is 97.

Example 4.3 *Given the random variables* $X_1, X_2, \dots, X_{900}$, *which are mutually independent of the distribution*

$$P\{X_i = 1\} = \frac{1}{10}, \quad P\{X_i = 0\} = \frac{9}{10}, \qquad i = 1, 2, \dots, 900.$$

Put

$$X = \sum_{i=1}^{900} X_i.$$

1) Find the distribution of X.

2) Find mean and variance of X.

3) Apply the Central Limit Theorem to find the largest integer M, *for which*

$$P\{90 - M \le X \le 90 + M\} \le 0.95.$$

4) In a given table of "random numbers" there is a page of 900 numbers, of which 70 are zeros, which apparently gives a too small number of zeros, if the numbers should have been chosen "randomly" (because we would have expected a number close to 90).
Expain why it is reasonable to reject the hypothesis of randomness if the number 70 does not belong to the interval $[90 - M, 90 + M]$, *where* M *is given by* **(3)**.
Shall one in the present case reject the hypothesis of randomness?

1) Clearly,

$$X = \sum_{i=1}^{900} X_i \in B\left(900, \frac{1}{10}\right).$$

2) Since the distribution is known, it follows immediately that

$$E\{X\} = 90, \qquad V\{X\} = 81 = 9^2.$$

3) Choosing $Y \in N\left(90, 9^2\right)$, cf. **(2)**, it follows from the Central Limit Theorem that

$$P\{90 - M \le X \le 90 + M\}$$

is approximately equal to

$$P\left\{90 - M - \frac{1}{2} \le Y \le 90 + M + \frac{1}{2}\right\} = P\left\{\left|\frac{Y - 90}{9}\right| \le \frac{M + \frac{1}{2}}{9}\right\} = 2\,\Phi\left(\frac{M + \frac{1}{2}}{9}\right) - 1.$$

This is again smaller than or equal to 0.95, if

$$\Phi\left(\frac{M + \frac{1}{2}}{9}\right) \le 0.975, \quad \text{i.e.} \quad \frac{M + \frac{1}{2}}{9} \le 1.96,$$

hence $M \le 17$.
Thus the largest integer is $M = 17$.

4) If we have randomness, then the probability of being *outside* the interval $[90 - M, 90 + M]$ smaller than 0.05.

   Since one does not believe in an event of so small probability, the hypothesis is rejected.

   We have in the case under consideration that

$$[90 - M, 90 + M] = [73, 107].$$

   Since 70 does not lie in this interval, we *reject* the hypothesis of randomness.

Example 4.4 *A schoolteacher wants to demonstrate to his pupils that one by throwing a coin many times will obtain heads in approximately half of the throws. They agree to throw the coin 100 times and consider the result as acceptable if the number of heads, X, satisfies the inequalities $48 \leq X \leq 52$. Considering results obtained by an application of the Central Limit Theorem one shall*

1) Compute $P\{48 \leq X \leq 52\}$.

2) Find the smallest integer M, such that

$$P\{50 - M \leq X \leq 50 + M\} \geq 0.95.$$

3) They then decide to throw the coin $50N$ times and consider the result as acceptable if the number of heads, Y, satisfies

$$24N \leq Y \leq 26N.$$

Find by using the Central Limit Theorem the smallest integer for which the probability of the event above is ≥ 0.95.

1) Since $X \in B\left(100, \frac{1}{2}\right)$, we see that X can be approximated by $N(50, 25)$, so

$$\begin{aligned} P\{48 \leq X \leq 52\} &= F\left(52\,\frac{1}{2}\right) - F\left(47\,\frac{1}{2}\right) \approx \Phi\left(\frac{\frac{5}{2}}{5}\right) - \Phi\left(-\frac{\frac{5}{2}}{5}\right) = 2\,\Phi\left(\frac{1}{2}\right) - 1 \\ &= 2 \cdot 0.6915 - 1 = 0.3830. \end{aligned}$$

Remark 4.1 We get by using a pocket calculator

$$\begin{aligned} P\{48 \leq X \leq 52\} &= \left\{\binom{100}{48} + \binom{100}{49} + \binom{100}{50} + \binom{100}{51} + \binom{100}{52}\right\} \frac{1}{2^{100}} \\ &= \left\{2\binom{100}{48} + 2\binom{100}{49} + \binom{100}{50}\right\} \cdot \frac{1}{2^{100}} = 0,38270. \end{aligned}$$

so we see that the approximation above is indeed a good one. $\Diamond$

2) Since $\Phi(1.96) = 0.975$, and

$$P\{50 - M \le X \le 50 + M\} \approx \Phi\left(\frac{M + \frac{1}{2}}{5}\right) - \Phi\left(-\frac{M + \frac{1}{2}}{5}\right) = 2\,\Phi\left(\frac{2M+1}{10}\right) - 1 \ge 0.95$$

for

$$\frac{2M+1}{10} \ge 1.96, \qquad \text{i.e.} \qquad M \ge 9.3,$$

we conclude that $M = 10$ is the smallest integer for which

$$P\{50 - M \le X \le 50 + M\} \ge 0.95.$$

3) Since $Y \in B\left(50N, \frac{1}{2}\right)$ can be approximated by an $N\left(25N, 25 \cdot \frac{N}{2}\right)$-distribution with $s_N = 5\sqrt{\frac{N}{2}}$, it follows that

$$P\{24N \le Y \le 26N\} = F\left(26N + \frac{1}{2}\right) - F\left(24N - \frac{1}{2}\right) \approx 2\,\Phi\left(\frac{2N+1}{5\sqrt{2N}}\right) - 1.$$

We get in the same way as in **(2)** the condition

$$\frac{2N+1}{5\sqrt{2N}} \ge 1.96,$$

which can also be written as an inequality containing a polynomial of second order in $\sqrt{2N}$,

$$\left(\sqrt{2N}\right)^2 - 9.8\sqrt{2N} + 1 \ge 0.$$

The roots of the equation $z^2 - 9.8\,z + 1 = 0$ are $z = 4.9 \pm \sqrt{4 - 9^2 - 1}$. Now,

$$4.9 - \sqrt{4.9^2 - 1} = 0.10313 < 1,$$

so only $+$ can be used. Hence,

$$2N \ge z^2 = 4.9^2 + 4.9^2 - 1 + 2 \cdot 4.9\sqrt{4 - 9^2 - 1} = 2 \cdot 4.9^2 - 1 + 9.8\sqrt{4.9^2 - 1},$$

from which

$$N \ge 4.9^2 - 0.5 + 4.9\sqrt{4.9^2 - 1} = 47.015.$$

The smallest integer, which fulfils this condition is $N = 48$.

Example 4.5 *Let X_1, X_2, ... be independent identically distributed random variables of mean μ and positive variance σ^2, and let*

$$Y_n = \sum_{i=1}^{n} X_i, \qquad n \in \mathbb{N}.$$

Find for every $x > 0$ by using the Central Limit Theorem

$$\lim_{n\to\infty} P\left\{\left|\frac{Y_n - n\mu}{\sqrt{n}}\right| \le x\right\} \quad \textit{and} \quad \lim_{n\to\infty} P\left\{\left(\frac{Y_n - n\mu}{\sqrt{n}}\right)^2 \le x\right\}.$$

The results shall be expressed in the distribution function $\Phi(x)$.

It follows from

$$E\{Y_n\} = \sum_{i=1}^{n} E\{X_i\} = n\mu$$

and

$$V\{Y_n\} = \sum_{i=1}^{n} V\{X_i\} = n\sigma^n, \qquad \text{i.e. } s_n = \sigma\sqrt{n},$$

by the Central Limit Theorem that

$$\lim_{n\to\infty} P\left\{\left|\frac{Y_n - n\mu}{\sqrt{n}}\right| \le x\right\} = \lim_{n\to\infty} P\left\{-\frac{x}{\sigma} \le \frac{Y_n - n\mu}{\sqrt{n}} \le \frac{x}{\sigma}\right\} = \begin{cases} 2\,\Phi\left(\dfrac{x}{\sigma}\right) - 1, & x > 0, \\[2mm] 0, & x \le 0. \end{cases}$$

Now,

$$\lim_{n\to\infty} P\left\{\left(\frac{Y_n - n\mu}{\sqrt{n}}\right)^2 \le x\right\} = \lim_{n\to\infty} P\left\{\left|\frac{Y_n - n\mu}{\sqrt{n}}\right| \le \sqrt{x}\right\},$$

so it follows from the above that

$$\lim_{n\to\infty} P\left\{\left(\frac{Y_n - n\mu}{\sqrt{n}}\right)^2 \le x\right\} = \begin{cases} 2\,\Phi\left(\dfrac{\sqrt{x}}{\sigma}\right) - 1, & \text{for } x > 0, \\ 0, & \text{for } x \le 0. \end{cases}$$

Example 4.6 *Given a roulette, where the possible event of each game is either red, black or green of the probabilities r, b and g, where*

$$r > 0, \quad b > 0, \quad g > 0 \quad \textit{and} \quad r + b + g = 1.$$

When we continually play on this roulette we assume that the games are independent.
Let $n \in \mathbb{N}$ be a fixed number, and let X_n, Y_n and Z_n denote the number of games among the n games which results in either red, black or green.

1) Find the distribution functions of X_n, Y_n and Z_n.

2) Find the variances $V\{X_n\}$, $V\{Y_n\}$ and $V\{Z_n\}$.

3) Find, e.g. by using that $X_n + Y_n + Z_n = n$, the variance $V\{X_n + Y_n\}$.

4) Compute the correlation $\varrho(X_n, Y_n)$.

5) Allowing n to vary, we shall prove that the sequence $\left(\frac{1}{n}X_n\right)$ for $n \to \infty$ converges in probability towards the constant r, and that the sequence $\left(\frac{1}{n}Y_n\right)$ converges towards the constant b.

6) Prove for any $\varepsilon \in \mathbb{R}_+$ and any $a > \frac{1}{2}$,

$$P\{|X_n - nr| < \varepsilon\, n^a \ \textit{and}\ |Y_n - nb| < \varepsilon\, n^a\} \to 1 \qquad \textit{for } n \to \infty.$$

1) Clearly, $X_n \in B(n, r)$, $Y_n \in B(n, b)$ and $Z_n \in B(n, g)$ are all binomially distributed with

$$P\{X_n = k\} = \binom{n}{k} r^k (1-r)^{n-k}, \qquad k = 0,\,1,\,2,\,\dots,\,n,$$

$$P\{Y_n = k\} = \binom{n}{k} b^k (1-b)^{n-k}, \qquad k = 0,\,1,\,2,\,\dots,\,n,$$

$$P\{Z_n = k\} = \binom{n}{k} g^k (1-g)^{n-k}, \qquad k = 0,\,1,\,2,\,\dots,\,n.$$

2) The variance of $U \in B(n,p)$ is $V\{U\} = np(1-p)$, hence

$$V\{X_n\} = nr(1-r), \qquad V\{Y_n\} = nb(1-b), \qquad V\{Z_n\} = ng(1-g).$$

3) It follows from $X_n + Y_n + Z_n = n$ that

$$\begin{aligned} V\{X_n + Y_n\} &= V\{n - Z_n\} = V\{n\} + V\{Z_n\} = 0 + ng(1-g) \\ &= ng(1-g) = n(1-r-b)(r+b). \end{aligned}$$

4) From

$$\begin{aligned} X_n Y_n &= \frac{1}{2}\left\{(X_n+Y_n)^2 - X_n^2 - Y_n^2\right\} = \frac{1}{2}\left\{(n-Z_n)^2 - X_n^2 - Y_n^2\right\} \\ &= \frac{1}{2}\left\{Z_n^2 - 2n\, Z_n + n^2 - X_n^2 - Y_n^2\right\}, \end{aligned}$$

follows that

$$\begin{aligned} \mathrm{Cov}\,(X_n, Y_n) &= E\{X_n Y_n\} - E\{X_n\}\, E\{Y_n\} \\ &= \frac{1}{2}\, E\left\{Z_n^2\right\} - n\, E\{X_n\} + \frac{n^2}{2} - \frac{1}{2}\, E\left\{X_n^2\right\} - \frac{1}{2}\, E\left\{Y_n^2\right\} - E\{X_n\}\, E\{Y_n\} \\ &= \frac{1}{2}\left[E\left\{Z_n^2\right\} - (E\{Z_n\})^2\right] + \frac{1}{2}\left[(E\{Z_n\})^2 - 2n\, E\{Z_n\} + n^2\right] \\ &\qquad -\frac{1}{2}\left[E\left\{X_n^2\right\} - (E\{X_n\})^2\right] - \frac{1}{2}\,(E\{X_n\})^2 \\ &\qquad -\frac{1}{2}\left[E\left\{Y_n^2\right\} - (E\{Y_n\})^2\right] - \frac{1}{2}\,(E\{Y_n\})^2 - E\{X_n\}\; E\{Y_n\} \\ &= \frac{1}{2}\, V\{Z_n\} + \frac{1}{2}\,(n - E\{Z_n\})^2 - \frac{1}{2}\, V\{X_n\} - \frac{1}{2}\,(E\{X_n\})^2 \\ &\qquad -\frac{1}{2}\, V\{Y_n\} - \frac{1}{2}\,(E\{Y_n\})^2 - E\{X_n\}\; E\{Y_n\} \\ &= \frac{1}{2}\, V\{Z_n\} - \frac{1}{2}\, V\{X_n\} - \frac{1}{2}\, V\{Y_n\} \\ &\qquad +\frac{1}{2}\,(n - E\{Z_n\})^2 - \frac{1}{2}\,(E\{X_n\} + E\{Y_n\})^2 \\ &= \frac{1}{2}\, V\{ZS_n\} - \frac{1}{2}\, V\{X_n\} - \frac{1}{2}\, V\{Y_n\} \\ &= \frac{n}{2}\,\{g(1-g) - r(1-r) - b(1-b)\} \\ &= \frac{n}{2}\,\left\{r + b - r^2 - b^2 - 2rb - r + r^2 - b + b^2\right\} \\ &= -nrb. \end{aligned}$$

Since

$$\sqrt{V\{X_n\}\; V\{Y_n\}} = n\sqrt{rb(1-r)(1-b)},$$

the correlation is

$$\varrho = \frac{-nrb}{n\sqrt{rb(1-r)(1-b)}} = -\sqrt{\frac{r \cdot b}{(1-r)(1-b)}}.$$

ALTERNATIVELY it is immediately seen that

$$\begin{aligned}\operatorname{Cov}(X_n, Y_n) &= \frac{1}{2}\left(V\{X_n+Y_n\} - V\{X_n\} - V\{Y_n\}\right)\\ &= \frac{n}{2}\{gr+gb-rb-rg-br-bg\} = -nrb.\end{aligned}$$

5) It suffices to prove that $\left(\frac{1}{n}X_n\right) \xrightarrow{D} r$ (in distribution), because the proof for $\left(\frac{1}{n}Y_n\right) \xrightarrow{D} b$ is analogous.
Since

$$E\left\{\frac{1}{n}X_n\right\} = r \quad \text{and} \quad V\left\{\frac{1}{n}X_n\right\} = \frac{r(1-r)}{n} = s_n,$$

it follows from the Central Limit Theorem that

$$\lim_{n\to\infty} P\left\{\frac{\frac{1}{n}X_n - r}{\sqrt{\frac{r(1-r)}{n}}} \le y\right\} = \lim_{n\to\infty} P\left\{X_n - nr \le \sqrt{n}\cdot\sqrt{r(1-r)}\cdot y\right\}.$$

Now,

$$\begin{aligned}F_n(x) &= P\left\{\frac{1}{n}X_n \le x\right\} = P\{X_n - nr \le nx - nr\}\\ &= P\left\{X_n - nr \le \sqrt{n}\cdot\sqrt{r(1-r)}\cdot\frac{\sqrt{n}\cdot(x-r)}{\sqrt{r(1-r)}}\right\}.\end{aligned}$$

hence

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty}\Phi\left(\frac{\sqrt{n}}{\sqrt{r(1-r)}}(x-r)\right) = \begin{cases} 0 & \text{for } x<r,\\ 1 & \text{for } x>r,\end{cases}$$

and the claim follows.

ALTERNATIVELY one may use Chebyshev's inequality. We have

$$\begin{aligned}P\left\{\left|\frac{1}{n}X_n - r\right| > \varepsilon\right\} &= P\{|X_n - nr| > n\varepsilon\} \le \frac{nr(b+g)}{n^2\varepsilon^2}\\ &= \frac{1}{n}r(b+g)\cdot\frac{1}{\varepsilon^2} \to 0 \qquad \text{for } n\to\infty,\end{aligned}$$

proving that $\left(\frac{1}{n}X_n\right)$ converges in probability towards r.
It follows similarly that $\left(\frac{1}{n}Y_n\right)$ converges in probability towards b.

6) If $a > \dfrac{1}{2}$ and $\varepsilon > 0$, then

$$P\left\{|X_n - nr| \le \varepsilon\, n^a\right\} = P\left\{|X_n - nr| \le \sqrt{n}\cdot\sqrt{r(1-r)}\cdot\frac{\varepsilon}{\sqrt{r(1-r)}}\cdot n^{a-\frac{1}{2}}\right\}$$
$$\approx \Phi\left(\frac{\varepsilon}{\sqrt{r(1-r)}}\cdot n^{a-\frac{1}{2}}\right) \to 1 \qquad \text{for } n \to \infty.$$

Analogously,

$$P\left\{|Y_n - nb| < \varepsilon\, n^a\right\} \approx \Phi\left(\frac{\varepsilon}{\sqrt{b(1-b)}}\, n^{a-\frac{1}{2}}\right) \to 1 \qquad \text{for } n \to \infty.$$

Hence, for every $\delta > 0$ there is an n_0, such that

$$P\left\{|X_n - nr| < \varepsilon\, n^a\right\} \ge 1 - \frac{\delta}{2}$$

and

$$P\{|Y_n - nb| < \varepsilon\, n^a\} \geq 1 - \frac{\delta}{2}$$

for $n \geq n_0$. Then

$$\begin{aligned} &P\{|X_n - nr| < \varepsilon\, n^a \text{ and } |Y_n - nb| < \varepsilon\, n^a\} \\ &\qquad \geq P\{|X_n - n\,r| < \varepsilon\, n^a\} - (1 - P\{|Y_n - n\,b| < \varepsilon\, n^a\}) \\ &\qquad \geq 1 - \frac{\delta}{2} - \frac{\delta}{2} = 1 - \delta, \qquad \text{for } n \geq n_0. \end{aligned}$$

This holds for every $\delta > 0$, so by taking the limit,

$$\lim_{n\to\infty} P\{|X_n - nr| < \varepsilon\, n^a \text{ and } |Y_n - nb| < \varepsilon\, n^a\} = 1.$$

Alternatively we may also here apply Chebyshev's inequality. The complementary event has the probability

$$\begin{aligned} &P\{\{|X_n - nr| \geq \varepsilon n^a\} \cup \{|Y_n - nb| \geq \varepsilon n^a\}\} \leq P\{|X_n - nr| \geq \varepsilon n^a\} + P\{|Y_n - nb| \geq \varepsilon n^a\} \\ &\qquad \leq \frac{V\{X_n\}}{\varepsilon^2 n^{2a}} + \frac{V\{Y_n\}}{\varepsilon^2 n^{2a}} = \frac{nr(b+g)}{\varepsilon^2 n^{2a}} + \frac{nb(r+g)}{\varepsilon^2 n^{2a}} \leq \frac{1}{\varepsilon^2}\, n^{1-2a} \to 0 \qquad \text{for } n \to \infty, \end{aligned}$$

(because $a > \frac{1}{2}$). Hence, we find the probability of the event,

$$P\{|X_n - nr| < \varepsilon n^a \,\wedge\, |Y_n - nb| < \varepsilon n^a\} \to 1 \quad \text{for } n \to \infty.$$

Example 4.7 *In an experiment an event S (success) occurs with the probability p, where $0 < p < 1$, and the complementary event F (failure) occurs with the probability $q = 1 - p$. The experiment is repeated under identical conditions and such that the events are mutually independent. We define for every $r \in \mathbb{N}$ the random variable Y_r by*

$Y_r = k$, *if we have had precisely k failures before the r-th success.*

1) Prove that Y_1 has its distribution given by

$$P\{Y_1 = k\} = pq^k, \qquad k \in \mathbb{N}_0.$$

2) Prove that Y_r has its distribution given by

$$P\{Y_r = k\} = \binom{k+r-1}{k} p^r q^k, \qquad k \in \mathbb{N}_0.$$

3) Explain why Y_r can be written in the form

$$Y_r = \sum_{i=1}^{r} X_i,$$

where the X_i are independent, identically distributed random variables.

4) Prove by applying Chebyshev's inequality that for every $\varepsilon > 0$ and every $a > \dfrac{1}{2}$,

$$\lim_{r \to \infty} P\left\{\left|Y_r - r\,\frac{p}{q}\right| \le \varepsilon\, r^a\right\} = 1.$$

5) Prove by applying the Central Limit Theorem that for every $\varepsilon > 0$,

$$\lim_{r \to \infty} P\left\{\left|Y_r - r\,\frac{q}{p}\right| \le \varepsilon\,\sqrt{r}\right\} = f(\varepsilon),$$

where $0 < f(\varepsilon) < 1$.

1) The event $\{Y_1 = k\}$ means that we have had precisely k failures before the first success, i.e. in totally $k+1$ experiments, so

$$P\{Y_1 = k\} = \binom{k+1-1}{k} p^1 q^k = pq^k, \qquad k \in \mathbb{N}_0.$$

2) The event $\{Y_r = k\}$ means that we have had k failures and $r-1$ successes in the first $k+r-1$ experiments, at that the $(k+r)$-th experiment is a success. The occurrences of the k failures can be chosen in $\binom{k+r-1}{k}$ ways, so

$$P\{Y_r = k\} = \binom{k+r-1}{k} p^r q^k, \qquad k \in \mathbb{N}_0.$$

This describes a negative binomial distribution, $Y_r \in NB(r,p)$.

3) Since $Y_1 = X_1$ has the distribution

$$P\{X_1 = k\} = P\{Y_1 = k\} = pq^k, \qquad k \in \mathbb{N}_0,$$

we see that

$$P\{X_1 = k-1\} = P\{X_1 + 1 = k\} = pq^{k-1}, \qquad k \in \mathbb{N},$$

is geometrically distributed. Then $Y_r + r$ is Pascal distributed,

$$Y_r + r = \sum_{i=1}^{r} (X_i + 1) \in \text{Pas}(r, p),$$

provided that all $X_i + 1$ are independent and identically distributed, thus

$$Y_r = \sum_{i=1}^{r} X_i, \qquad X_i \in NB(1,p), \text{ i.e. } P\{X_i = k\} = pq^k, \quad k \in \mathbb{N}_0,$$

and the negative binomial distribution is reproductive in the form parameter.

4) By Chebyshev's inequality,

$$P\{|Y - \mu| \geq b\} \leq \frac{\sigma^2}{b^2}, \qquad \text{where } \mu = E\{Y\} \text{ and } \sigma^2 = V\{Y\}.$$

It follows from (3) and a formula that

$$\mu_r = E\{Y_r\} = E\{Y_r + r\} - r = \frac{r}{p} - r = r\,\frac{1-p}{p} = r \cdot \frac{q}{p},$$

and

$$\sigma_r^2 = V\{Y_r\} = V\{Y_1 + r\} = \frac{rq}{p^2}.$$

Then by insertion into Chebyshev's inequality,

$$(5)\quad P\left\{\left|Y_r - r\,\frac{q}{p}\right| \geq b\right\} \leq \frac{rq}{b^2p^2}.$$

Choosing $b = \varepsilon\, r^a$, where $a > \frac{1}{2}$, we get by (5),

$$0 \leq P\left\{\left|Y_r - r\,\frac{q}{p}\right| > \varepsilon\, r^a\right\} \leq \frac{q}{p^2\varepsilon} \cdot \frac{1}{r^{2a-1}} \to 0 \qquad \text{for } r \to \infty,$$

thus

$$\lim_{r\to\infty} P\left\{\left|Y_r - r\,\frac{q}{p}\right| \leq \varepsilon\, r^a\right\} = 1 - \lim_{r\to\infty} P\left\{\left|Y_r - r\,\frac{q}{p}\right| > \varepsilon\, r^a\right\} = 1.$$

5) By the Central Limit Theorem,

$$\lim_{r\to\infty} P\left\{\left|Y_r - r\,\frac{q}{p}\right| \leq \varepsilon\sqrt{r}\right\} = \lim_{r\to\infty} P\left\{\left\|\frac{Y_r - r\,\frac{q}{p}}{\sigma_1\sqrt{r}}\right| \leq \frac{\varepsilon\sqrt{r}}{\sigma_1\sqrt{r}}\right\} = \lim_{r\to\infty} P\left\{\left|\frac{Y_r - r\,\frac{q}{p}}{\sigma_1\sqrt{r}}\right| \leq \frac{\varepsilon p}{\sqrt{q}}\right\}$$

$$= 2\,\Phi\left(\varepsilon\,\frac{p}{\sqrt{q}}\right) - 1 = f(\varepsilon).$$

Here we have used that

$$Y_r = \sum_{i=1}^{r} X_i \quad \text{and} \quad E\{X_i\} = \frac{q}{p} \quad \text{and} \quad \sigma_1^2 = dfracqp^2.$$

If Z is $N(0,1)$-distributed, then clearly,

$$f(\varepsilon) = 2\,\Phi\left(\varepsilon\,\frac{p}{\sqrt{q}}\right) - 1 = P\left\{|Z| \le \varepsilon\,\frac{p}{\sqrt{q}}\right\} \in]0,1[.$$

Example 4.8 *Let $X_1, X_2, \dots$ be independent random variables, all rectangularly distributed over the interval $]0,1[$.*
Introduce for every $n \in \mathbb{N}$ the random variables

$$Y_n = \sum_{i=1}^{n} X_i, \qquad Z_n = \sum_{i=1}^{n} X_i^2.$$

1) *Compute the means $E\{Y_n\}$ and $E\{Z_n\}$.*

2) *Compute the variances $V\{Y_n\}$ and $V\{Z_n\}$.*

3) *Compute for $j = 1, 2, \dots, n$, the covariance $\operatorname{Cov}\left(Y_n, X_j^2\right)$.*

4) *Compute the correlation $\varrho(Y_n, Z_n)$, and prove that it does not depend on n.*

5) *Let n vary. Prove for every $\varepsilon \in \mathbb{R}_+$ and every $a > \frac{1}{2}$ that*

$$P\left\{\left|Y_n - \frac{n}{2}\right| < \varepsilon\, n^a\right\} \to 1 \qquad \textit{for } n \to \infty,$$

and

$$P\left\{\left|Z_n - \frac{n}{3}\right| < \varepsilon\, n^a\right\} \to 1 \qquad n \to \infty.$$

6) *Find*

$$\lim_{n\to\infty} P\left\{\left|Z_n - \frac{n}{3}\right| < \frac{\sqrt{n}}{4}\right\},$$

(3 decimals).

1) Obviously, $E\{Y_n\} = \frac{n}{2}$. Furthermore,

$$E\{Z_n\} = \sum_{i=1}^{n} E\left\{X_i^2\right\} = \sum_{i=1}^{n} \int_0^1 x_i^2\, dx_i = \frac{n}{3}.$$

2) Clearly, $V\{Y_n\} = \dfrac{n}{12}$. Since

$$E\left\{\left(X_i^2\right)^2\right\} = \int_0^1 x_i^4\,dx_i = \frac{1}{5} \qquad \text{and} \qquad E\left\{X_i^2\right\} = \frac{1}{3},$$

we get

$$V\left\{X_i^2\right\} = \frac{1}{5} - \frac{1}{9} = \frac{4}{45},$$

so

$$V\{Z_n\} = \frac{4n}{45}.$$

3) A direct computation gives

$$\text{Cov}\left(X_n, X_j^2\right) = \text{ Cov}\left(X_j, X_j^2\right) = E\left\{X_j^3\right\} - E\left\{X_j\right\}\, E\left\{X_j^2\right\} = \frac{1}{4} - \frac{1}{2}\cdot\frac{1}{3} = \frac{1}{12}.$$

4) By (3),

$$\text{Cov}\left(Y_n, Z_n\right) = \sum_{j=1}^{n} \text{ Cov}\left(Y_n, X_j^2\right) = \frac{n}{12},$$

thus

$$\varrho\left(Y_n, Z_n\right) = \frac{n}{12} \cdot \frac{1}{\sqrt{\frac{n}{12}} \cdot \sqrt{\frac{4n}{45}}} = \sqrt{\frac{45}{48}} = \sqrt{\frac{15}{16}} = \frac{1}{4}\sqrt{15},$$

independently of n.

5) By Chebyshev's inequality,

$$P\left\{\left|Y_n - \frac{n}{2}\right| \geq \varepsilon\, n^a\right\} \leq \frac{\frac{n}{12}}{\varepsilon^2 n^{2a}} = \frac{1}{12\varepsilon^2 n^{2a-1}} \to 0 \qquad \text{for } n \to \infty,$$

so

$$\lim_{n\to\infty} P\left\{\left|Y_n - \frac{n}{2}\right| < \varepsilon\, n^a\right\} = 1 - 0 = 1.$$

It follows in the same way that

$$P\left\{\left|Z_n - \frac{n}{3}\right| \geq \varepsilon\, n^a\right\} \leq \frac{\frac{4n}{45}}{\varepsilon^2 n^{2a}} = \frac{4}{45\varepsilon^2} \cdot \frac{1}{n^{2a-1}} \to 0 \qquad \text{for } n \to \infty,$$

so

$$\lim_{n\to\infty} P\left\{\left|Z_n - \frac{n}{3}\right| < \varepsilon\, n^a\right\} = 1 - 0 = 1.$$

6) By the Central Limit Theorem,

$$\begin{aligned}
P\left\{\left|Z_n - \frac{n}{3}\right| < \frac{\sqrt{n}}{4}\right\} &= P\left\{\frac{\left|Z_n - \frac{n}{3}\right|}{\sqrt{\frac{4n}{45}}} \le \frac{\frac{\sqrt{n}}{4}}{\sqrt{\frac{4n}{45}}}\right\} \to 2\,\Phi\left(\frac{3}{8}\sqrt{5}\right) - 1 \\
&= 2\,\Phi(0.83853) - 1 \approx 2 \cdot 0.799 - 1 = 0 - 598.
\end{aligned}$$

Example 4.9 *Let U_1 and U_2 be two mutually independent random variables of the means μ_1 and μ_2, resp., and the variances σ_1^2 and σ_2^2, resp., and let $U = U_1 \cdot U_2$.*

1. *Prove that the variance of U is $\sigma_1^2\sigma_2^2 + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2$.*

A rectangle has the edges X_1 and X_2, where X_1 and X_2 are independent, identically distributed random variables of frequency

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

Let $Y = X_1 \cdot X_2$ denote the area of the rectangle.

2. *Find, by e.g. using the result of* **1.**, *the mean and variance of Y.*

Let Y_1, Y_2, … be a sequence of mutually independent random variables, all following the same distribution as Y, and let

$$Z_n = \sum_{i=1}^{n} Y_i, \qquad n \in \mathbb{N}.$$

3. *Find, by means of Chebyshev's inequality a positive constant a, such that*

$$P\left\{\left|Z_n - \frac{4n}{9}\right| < \frac{\sqrt{n}}{4}\right\} \ge a \qquad \textit{for every } n \in \mathbb{N}.$$

4. *Find by means of the Central Limit Theorem,*

$$\lim_{n\to\infty} P\left\{\left|Z_n - \frac{4n}{9}\right| < \frac{\sqrt{n}}{4}\right\}, \qquad \textit{(3 decimals).}$$

1) It follows from

$$E\left\{U^2\right\} = E\left\{U_1^2\right\} \cdot E\left\{U_2^2\right\} = \left(\sigma_1^2 + \mu_1^2\right) \cdot \left(\sigma_2^2 + \mu_2^2\right),$$

that

$$V\{U\} = E\left\{U^2\right\} - (E\{U\})^2 = \sigma_1^2\sigma_2^2 + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \mu_1^2\mu_2^2 - (\mu_1\mu_2)^2 = \sigma_1^2\sigma_2^2 + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2.$$

2) In this case we have (at least) two variants.

a) First compute

$$E\{X_1\} = \int_0^1 2x^2\,dx = \frac{2}{3} \quad \text{and} \quad E\{X_1^2\} = \int_0^1 2x^3\,dx = \frac{1}{2},$$

thus

$$V\{X_1\} = E\{X_1^2\} - (E\{X_1\})^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}.$$

We obtain the same results for X_2.
Then

$$E\{Y\} = E\{X_1\} \cdot E\{X_2\} = \frac{4}{9},$$

and by **1.**,

$$V\{Y\} = \frac{1}{18} \cdot \frac{1}{18} + 2 \cdot \frac{4}{9} \cdot \frac{1}{18} = \frac{17}{324}.$$

b) Alternative solution. First find the frequency of $Y = X_1 \cdot X_2$. If $0 < y < 1$, then

$$f_Y(y) = \int_{x=y}^1 f_{X_1}(x)\, f_{X_2}\left(\frac{y}{x}\right) \frac{1}{x}\,dx = \int_{x=y}^1 2x \cdot 2\,\frac{y}{x} \cdot \frac{1}{x}\,dx = 4y \int_{x=y}^1 \frac{1}{x}\,dx = -4y\,\ln y.$$

Hence

$$E\{Y\} = -4\int_0^1 y^2 \ln y\,dy = -4\left[\frac{y^3}{3}\,\ln y - \int \frac{y^2}{3}\,dy\right]_0^1 = \frac{4}{9},$$

$$E\{Y^2\} = -4\int_0^1 y^3 \ln y\,dy = -4\left[\frac{y^4}{4}\,\ln y - \int \frac{y_3}{4}\,dy\right]_0^1 = \frac{1}{4},$$

so

$$V\{Y\} = \frac{1}{4} - \left(\frac{4}{9}\right)^2 = \frac{1}{4} - \frac{16}{81} = \frac{17}{214}.$$

3) Put

$$Z_n = \sum_{i=1}^n Y_i.$$

Then

$$E\{Z_n\} = n \cdot \frac{4}{9} \qquad \text{and} \qquad V\{Z_n\} = n \cdot \frac{17}{324}.$$

It follows from Chebyshev's inequality that

$$P\left\{\left|Z_n - \frac{4n}{9}\right| \geq \frac{\sqrt{n}}{4}\right\} \leq n \cdot \frac{17}{324} \cdot \frac{16}{n} = \frac{68}{81},$$

hence

$$P\left\{\left|Z_n - \frac{4n}{9}\right| < \frac{\sqrt{n}}{4}\right\} \geq 1 - \frac{68}{81} = \frac{13}{81} \qquad (\approx 0,160).$$

4) Obviously, the assumptions of the Central Limit Theorem are satisfied, so

$$P\left\{\left|Z_n - \frac{4n}{9}\right| < \frac{\sqrt{n}}{4}\right\} = P\left\{-\frac{18}{4\sqrt{17}} < \frac{Z_n - 4\cdot\frac{n}{9}}{\sqrt{n}\cdot\frac{\sqrt{17}}{18}} < \frac{18}{4\sqrt{17}}\right\} \to 2\,\Phi\left(\frac{18}{4\sqrt{17}}\right) - 1$$
$$= 2\,\Phi(1.0914) - 1 = 2\cdot 0.8624 - 1 = 0.7248.$$

Hence, with 3 decimals,

$$\lim_{n\to\infty} P\left\{\left|Z_n - \frac{4n}{9}\right| < \frac{\sqrt{n}}{4}\right\} = 0.725.$$

Example 4.10 *Let X_1 and X_2 be independent random variable of the frequencies*

$$f_{X_1}(x_1) = e^{-2|x_1|}, \quad x_1 \in \mathbb{R}, \qquad f_{X_2}(x_2) = \begin{cases} e^{-x_2}, & x_2 > 0, \\ 0, & x_2 \le 0. \end{cases}$$

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

2. *Find the variances $V\{X_1\}$ and $V\{X_2\}$.*

3. *Prove that the random variables $X_1 + X_2$ and $2X_1 - X_2$ are non-correlated.*

4. *Compute the frequency of the random variable $X_1 + X_2$.*

Let $U_1, U_2, \ldots, U_{200}$ be mutually independent random variables, all following the same distribution as X_1, and let

$$V = \sum_{i=1}^{200} U_i.$$

5. *Find by using the Central Limit Theorem an approximate expression of*

$$P\{-20 \le V \le 20\}.$$

1) The integral $\int_{-\infty}^{\infty} |x_1|\, e^{-2|x_1|}\, dx_1 < \infty$ is convergent, and the integrand $x_1 e^{-2|x_1|}$ is odd. Thus

$$E\{X_1\} = \int_{-\infty}^{\infty} x_1 e^{-2|x_1|}\, dx_1 = 0.$$

Furthermore, by using a Γ-integral,

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

2) From $E\{X_1\} = 0$ follows that

$$V\{X_1\} = E\left\{X_1^2\right\} = \int_{\infty}^{\infty} x_1^2 e^{-2|x_1|}\, dx_1 = 2\int_0^{\infty} x_1^2 e^{-2x_1}\, dx_1 = \frac{1}{4}\cdot 2! = \frac{1}{2}.$$

Since

$$E\left\{X_2^2\right\} = \int_0^{\infty} x_2^2 e^{-x_2}\, dx_2 = 2,$$

we get

$$V\{X_2\} = 2 - 1 = 1.$$

3) Now, X_1 and X_2 are independent, and $\mathrm{Cov}(\cdot,\cdot)$ is bilinear, hence

$$\mathrm{Cov}\left(X_1 + X_2, 2X_1 - X_2\right) = 2\,\mathrm{Cov}\left(X_1, X_2\right) - \ \mathrm{Cov}\left(X_2, X_2\right) = 2\,V\{X_1\} - V\{X_2\} = 0.$$

This shows that $X_1 + X_2$ and $2X_1 - X_2$ are non-correlated (though they are *not* independent.)

4) Since X_1 and X_2 are independent, the frequency of $Y = X_1 + X_2$ is given by the convolution integral

$$g(y) = \int_{-\infty}^{\infty} f_{X_1}(x)\, f_{X_2}(y - x)\, dx.$$

The integrand is positive, if $y > x$. We split the investigation according to whether $y \le 0$ or $y > 0$.

a) If $y \le 0$, then

$$g(y) = \int_{-\infty}^{y} e^{2x} e^{-(y-x)}\,dx = e^{-y}\int_{-\infty}^{y} e^{3x}\,dx = \frac{1}{3}\,e^{2y}.$$

b) If instead $y > 0$, then

$$\begin{aligned} g(y) &= \int_{-\infty}^{0} e^{2x} e^{-(y-x)}\,dx + \int_{0}^{y} e^{-2x} e^{-(y-x)}\,dx = e^{-y}\int_{-\infty}^{0} e^{3x}\,dx + e^{-y}\int_{0}^{y} e^{-x}\,dx \\ &= \frac{1}{3}\,e^{-y} + \left(1-e^{-y}\right)e^{-y} = \frac{4}{3}\,e^{-y} - e^{-2y}. \end{aligned}$$

Summing up we get

$$g(y) = \begin{cases} \dfrac{1}{3}\,e^{2y}, & \text{for } y \le 0, \\[2mm] \dfrac{4}{3}\,e^{-y} - e^{-2y}, & \text{for } y > 0. \end{cases}$$

5) It follows from the Central Limit Theorem that V is almost normally distributed with mean 0 and variance $100 = 10^2$. When we put $W \in N\left(0, 10^2\right)$, the wanted probability is almost equal to

$$P\{-20 \le W \le 20\} = P\left\{-2 \le \frac{W}{10} \le 2\right\} = 2\,\Phi(2) - 1 = 2 \cdot 0.9772 - 1 = 0.9544.$$

Example 4.11 *An instrument A contains two components, the lifetimes of which, X_1 and X_2, are independent random variables, both of frequency*

$$f(x)=\begin{cases}\alpha\, e^{-\alpha x}, & x>0,\\[2mm] 0, & x\le 0,\end{cases}$$

where α is a positive constant. The instrument A does only work, when at least one of the two components operates, i.e. the lifetime X of A is $X=\max\{X_1,X_2\}$.
Another instrument B also contains two components, the lifetimes of which, Y_1 and Y_2, are independent random variables, both of frequency

$$g(y)=\begin{cases}\dfrac{\alpha}{3}\exp\left(-\dfrac{\alpha}{3}\,y\right), & y>0,\\[2mm] 0, & y\le 0.\end{cases}$$

The instrument B is only working, if both components are working, i.e. the lifetime Y of B is $Y=\min\{Y_1,Y_2\}$.

1) Find the frequency of X.

2) Find mean and variance of X.

3) Find frequency, mean and variance of Y.

4) Find the simultaneous frequency of (X,Y), and then find $P\{Y>X\}$.

5) Consider 180 instruments of the same type as A. The random variable Z denotes the sum of the lifetimes of the 180 instrument. Find by using the Central Limit Theorem an approximate expression of

$$P\left\{\frac{240}{\alpha}<Z<\frac{300}{\alpha}\right\}.$$

1) According to a known formula, $X=\max\{X_1,X_2\}$ has the distribution function

$$F_X(x)=\begin{cases}F_{X_1}(x)\cdot F_{X_2}(x)=\left(1-e^{-\alpha\,x}\right)^2 & \text{for } x>0,\\[2mm] 0 & \text{for } x\le 0,\end{cases}$$

so the frequency is

$$f(x)=\begin{cases}2\alpha\, e^{-\alpha x}\left(1-e^{-\alpha x}\right)=2\alpha e^{-\alpha x}-2\alpha\, e^{-2\alpha x} & \text{for } x>0,\\[2mm] 0 & \text{for } x\le 0.\end{cases}$$

2) The mean is

$$E\{X\}=\int_0^{\infty}2\alpha\, x\, e^{-\alpha\, x}\,dx-\int_0^{\infty}2\alpha\, x\, e^{-2\alpha\, x}\,dx=\frac{2}{\alpha}\cdot 1!-\frac{1}{2\alpha}\cdot 1!=\frac{3}{2\alpha}.$$

Since

$$\begin{aligned} E\left\{X^2\right\} &= \int_0^{\infty} 2\alpha\, x^2 e^{-\alpha\, x}\, dx - \int_0^{\infty} 2\alpha\, x^2 e^{-2\alpha x}\, dx \\ &= \frac{2}{\alpha^2}\cdot 2! - \frac{2}{(2\alpha)^2}\cdot 2! = \frac{1}{\alpha^2}\left(2-\frac{1}{4}\right) = \frac{14}{4\alpha^2}, \end{aligned}$$

we get the variance

$$V\{X\} = \frac{14}{4\alpha^2} - \frac{14}{4\alpha^2} - \frac{9}{4\alpha^2} = \frac{5}{4\alpha^2}.$$

3) It follows by a formula that $Y = \min\{Y_1, Y_2\}$ has the distribution function

$$\begin{aligned} F_Y(y) &= 1-(1-F_{Y_1}(y))\,(1-F_{Y_2}(y)) = 1-(1-F_{Y_1}(y))^2 \\ &= \begin{cases} 1-\exp\left(-\dfrac{2}{3}\,\alpha\, y\right), & \text{for } y>0, \\ 0, & \text{for } y \le 0. \end{cases} \end{aligned}$$

The frequency is obtained by a differentiation,

$$f_Y(y) = \begin{cases} \dfrac{2\alpha}{3}\,\exp\left(-\dfrac{2\alpha}{3}\,y\right) & \text{for } y>0, \\ 0 & \text{for } y \le 0, \end{cases}$$

proving that Y is exponentially distributed with

$$E\{Y\} = \frac{3}{2\alpha} \qquad \text{and} \qquad V\{Y\} = \frac{9}{4\alpha^2}.$$

4) Since X and Y are stochastically independent, the frequency of (X,Y) is given by

$$h(x,y) = f_X(x)\cdot f_Y(y) = \begin{cases} \dfrac{4\alpha^2}{3}\,e^{-\alpha x}\left(1-e^{-\alpha x}\right)\cdot exp\left(-\dfrac{2\alpha}{3}\,y\right) & \text{for } x>0 \text{ and } y>0, \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$\begin{aligned} &P\{Y>X\} \\ &= \int_0^{\infty}\left\{\int_0^{y} f_X(x)\cdot f_Y(y)\,dx\right\}dy = \int_0^{\infty}\left[F_X(x)\right]_0^y\, f_Y(y)\,dy \\ &= \int_0^{\infty}\left(1-e^{-\alpha y}\right)^2\cdot\frac{2\alpha}{3}\cdot\exp\left(-\frac{2\alpha}{3}\,y\right)\,dy \\ &= \frac{2\alpha}{3}\int_0^{\infty}\left\{\exp\left(-\frac{2\alpha}{3}\,y\right) - 2\,\exp\left(-\frac{5\alpha}{3}\,y\right) + \exp\left(-\frac{8\alpha}{3}\,y\right)\right\}dy \\ &= \frac{2\alpha}{3}\left\{\frac{3}{2\alpha} - 2\cdot\frac{3}{5\alpha} + \frac{3}{8\alpha}\right\} = \frac{2}{3}\left\{\frac{3}{2} - \frac{6}{5} + \frac{3}{8}\right\} = 1 - \frac{4}{5} + \frac{1}{4} \\ &= \frac{1}{5} + \frac{1}{4} = \frac{9}{20}. \end{aligned}$$

5) Since

$$\mu = E\{Z\} = 180\, E\{X\} = \frac{270}{\alpha},$$

and

$$\sigma^2 = V\{Z\} = 180\, V\{X\} = \frac{225}{\alpha^2} = \left(\frac{15}{\alpha}\right)^2,$$

it follows by the Central Limit theorem that

$$P\left\{\frac{240}{\alpha} < Z < \frac{300}{\alpha}\right\} = P\left\{-\frac{30}{\alpha} < Z - \mu < \frac{30}{\alpha}\right\} = P\left\{-2 < \frac{Z-\mu}{\sigma} < 2\right\}$$
$$\approx 2\,\Phi(2) - 1 = 2 \cdot 0.9772 - 1 = 0.9544.$$

Example 4.12 *Let X_1, X_2, X_3, ... be mutually independent random variables with their distribution given by*

$$P\{X_i = 1\} = P\{X_i = -1\} = \frac{1}{2}, \qquad i \in \mathbb{N},$$

and let

$$Y_n = \sum_{i=1}^{n} X_i, \qquad n \in \mathbb{N}.$$

1) Prove that if U_i is Bernoulli distributed with the probability parameter $\frac{1}{2}$, $U_i \in B\left(1, \frac{1}{2}\right)$, then $2U_i - 1$ has the same distribution as X_i.

2) Find by e.g. applying the result of **1.**, *the distribution of the random variable Y_5.*

3) Find the characteristic function of X_i, and then find the characteristic function of Y_n.

4) Prove that the sequence $\left(\frac{1}{\sqrt{n}}\, Y_n\right)$ converges in distribution towards a random variable Y, and find the distribution of Y.
Hint: *One may either use the Central Limit Theorem, or the result of* **3.**).

5) Find an approximate expression of $P\{-30 \le Y_{400} \le 30\}$.

1) When $U_i \in B\left(1, \frac{1}{2}\right)$, then

$$P\{U_i = 0\} = P\{2U_i - 1 = 1\} = \frac{1}{2} = P\{U_i = 1\} = P\{2U_i - 1 = 1\},$$

hence $2U - i - 1$ has the same distribution as X_i.

2) It follows that

$$Y_5 = \sum_{i=1}^{5} X_i = \sum_{i=1}^{5} (2U_i - 1) = 2\sum_{i=1}^{5} U_i - 5 = 2\,U - 5,$$

where

$$U = \sum_{i=1}^{5} U_i \in B\left(5, \frac{1}{2}\right)$$

is binomially distributed, hence

$$P\{U = k\} = \binom{5}{k} \frac{1}{2^5}, \qquad k = 0,\, 1,\, 2,\, 3,\, 4,\, 5.$$

Then

$$\begin{array}{lcccl}
P\left\{Y_5 = -5\right\} & = & P\{U = 0\} & = & \frac{1}{32}, \\
P\left\{Y_5 = -3\right\} & = & P\{U = 1\} & = & \frac{5}{32}, \\
P\left\{Y_5 = -1\right\} & = & P\{U = 2\} & = & \frac{5}{16}, \\
P\left\{Y_5 = 1\right\} & = & P\{U = 3\} & = & \frac{5}{16}, \\
P\left\{Y_5 = 3\right\} & = & P\{U = 4\} & = & \frac{5}{32}, \\
P\left\{Y_5 = 5\right\} & = & P\{U = 5\} & = & \frac{1}{32}.
\end{array}$$

3) The characteristic function of X_i is

$$k_{X_i}(\omega) = E\left\{e^{i\omega X_i}\right\} = \frac{1}{2}\,e^{i\omega} + \frac{1}{2}\,e^{-i\omega} = \cos\omega.$$

The characteristic function of $Y_n = \sum_{i=1}^n X_i$ is then

$$k_{Y_n} = \prod_{i=1}^n k_{X_i}(\omega) = \cos^n \omega.$$

4) First variant. The characteristic function of $\frac{1}{\sqrt{n}}\,Y_n = Z_n$ is

$$k_{Z_n}(\omega) = \cos^n\left(\frac{\omega}{\sqrt{n}}\right) = \left\{1 - \frac{\omega^2}{2}\cdot\frac{1}{n} + \frac{1}{n}\,\varepsilon\left(\frac{1}{n}\right)\right\}^n \to \exp\left(-\frac{\omega^2}{2}\right) \qquad \text{for } n \to \infty.$$

The limit function exists and is continuous at 0, hence the sequence $\left(\frac{1}{\sqrt{n}}\,Y_n\right)$ of random variables converges in distribution towards a random variable Y with the characteristic function

$$k_Y(\omega) = \exp\left(-\frac{\omega^2}{2}\right),$$

and we conclude that $Y \in N(0,1)$ is normally distributed.
Second variant. Now

$$E\{Z_n\} = \frac{1}{\sqrt{n}}\sum_{i=1}^n E\{X_i\} = 0,$$

and

$$V\{Z_n\} = \frac{1}{n}\cdot n\cdot V\{X_1\} = V\{2U_1 - 1\} = 4\,V\{U_1\} = 4\cdot\frac{1}{2}\cdot\frac{1}{2} = 1,$$

it follows by the Central Limit Theorem that

$$F_{Z_n}(z) = P\{Z_n \le z\} = P\left\{\frac{Z_n - E\{Z_n\}}{\sqrt{V\{Z_n\}}} \le z\right\} \to \Phi(z) \qquad \text{for } n \to \infty,$$

proving that $Z_n \to Y$ in distribution, where $Y \in N(0,1)$ is normally distributed.

5) Since $Y_{400} = \frac{1}{20}\,Z_{400}$, it follows from the Second variant in **4.** above that

$$P\{-30 < Y_{400} < 30\} = P\left\{-\frac{3}{2} < Z_{400} < \frac{3}{2}\right\} \approx 2\Phi\left(\frac{3}{2}\right) - 1 \approx 2\cdot 0.9332 - 1 = 0.8664.$$

Here, it is however easier to apply the continuity correction, so we get the interval $]-30.5\,,\,30.5[$.
In this case we have the approximation

$$\begin{aligned} P\{-30.5 < Y_{400} < 30.5\} &= P\left\{-\frac{305}{200} < Z_{400} < \frac{305}{200}\right\} \\ &\approx 2\Phi(1.525) - 1 = 2\cdot 0.9364 - 1 = 0.8728. \end{aligned}$$

Example 4.13 *A random variable X has the frequency*

$$f(x)=\begin{cases}1-\dfrac{x}{2}, & 0<x<2,\\[2mm] 0, & \textit{otherwise.}\end{cases}$$

1) Prove that $E\{X\}=E\left\{X^2\right\}$.

2) Find the variance $V\{X\}$.

3) Let $X_1, X_2, \dots, X_{450}$ be mutually independent random variables, all distributed like X above, and let $Y=\sum_{i=1}^{450}X_i$.
Find by using the Central Limit Theorem a number x, such that

$$P\{300-x<Y<300+x\}$$

is (almost) 0.95.

1) By some simple computations,

$$E\{X\}=\int_0^2 x\,f(x)\,dx=\int_0^2\left(x-\frac{1}{2}\,x\right)\,dx=\left[\frac{1}{2}\,x^2-\frac{1}{6}\,x^3\right]_0^2=2-\frac{4}{3}=\frac{2}{3},$$

and

$$E\left\{X^2\right\}=\int_0^2 x^2 f(x)\,dx=\int_0^2\left(x^2-\frac{1}{2}\,x^3\right)\,dx=\left[\frac{1}{3}\,x^3-\frac{1}{8}\,x^4\right]_0^2=\frac{8}{3}-2=\frac{2}{3},$$

hence

$$E\{X\}=E\left\{X^2\right\}=\frac{2}{3}.$$

2) Another simple computation gives

$$V\{X\}=E\left\{X^2\right\}-(E\{X\})^2=\frac{2}{3}-\frac{4}{9}=\frac{2}{9}.$$

3) If $Y=\sum_{i=1}^{450}X_i$, then

$$E\{Y\}=450\cdot\frac{2}{3}=300\qquad\text{and}\qquad V\{Y\}=450\cdot\frac{2}{9}=100=10^2.$$

If $Z\in N\left(300,10^2\right)$, it follows by the Central Limit Theorem that

$$\begin{aligned}P\{300-x<Y<300+x\}&\approx P\{300-x<Z<300+x\}\\&=P\left\{-\frac{x}{10}<\frac{Z-300}{10}<\frac{x}{10}\right\}=2\,\Phi\left(\frac{x}{10}\right)-1.\end{aligned}$$

From

$$2\,\Phi\left(\frac{x}{10}\right)-1=0,95,\qquad\text{dvs.}\qquad\Phi\left(\frac{x}{10}\right)=0,975,$$

follows that

$$\frac{x}{10}=1.96,\qquad\text{thus}\qquad x=19.6.$$

Example 4.14 *Let X_1, X_2, X_3, ... be mutually independent random variables, all of the distribution given by*

$$P\{X_i = 2\} = \frac{1}{3}, \qquad P\{X_i = -1\} = \frac{2}{3}, \qquad i \in \mathbb{N},$$

and let

$$Y_n = \sum_{i=1}^{n} X_i, \qquad n \in \mathbb{N}.$$

1) *Find mean and variance of X_i.*
2) *Explain why Y_n only takes values in the interval $[-n, 2n]$.*
3) *Find by means of Chebyshev's inequality a positive constant a, such that*

$$P\left\{|Y_n| < 3\sqrt{n}\right\} \geq a \qquad \textit{for every } n \in \mathbb{N}.$$

4) *Find by means of the Central Limit Theorem*

$$\lim_{n\to\infty} P\left\{|Y_n| < 3\sqrt{n}\right\} \qquad (3 \textit{ decimals}).$$

5) *Find the distribution of Y_4.*

The simplest method is to introduce $Z_i \in B\left(1, \frac{1}{3}\right)$, which is Bernoulli distributed, and then note that $X_i = 3Z_i - 1$.

1) The mean is

$$E\{X_i\} = 3\,E\{Z_i\} - 1 = 3 \cdot \frac{1}{3} - 1 = 0.$$

Alternatively,

$$E\{X_i\} = 2 \cdot \frac{1}{3} - 1 \cdot \frac{2}{3} = 0.$$

2) The variance is

$$V\{X_i\} = 9\,V\{Z_i\} = 9 \cdot \frac{1}{3} \cdot \frac{2}{3} = 2.$$

Alternatively we get from $E\{X_i\} = 0$ that

$$V\{X_i\} = E\left\{X_i^2\right\} = 4 \cdot \frac{1}{3} + \frac{2}{3} = 2.$$

3) Since $-1 \leq X_i \leq 2$ for all i, we get

$$-n \leq \sum_{i=1}^{n} X_i = Y_n \leq 2n.$$

4) Since $E\{Y_n\} = 0$ and $\sigma_n^2 = V\{Y_n\} = 2n$, it follows by Chebyshev's inequality that

$$P\left\{|Y_n - 0| \geq 3\sqrt{n}\right\} \leq \frac{\sigma_n^2}{9n} = \frac{2}{9},$$

hence

$$P\left\{|Y_n| < 3\sqrt{n}\right\} \geq 1 - \frac{2}{9} = \frac{7}{9} = a.$$

5) According to the Central Limit Theorem,

$$\lim_{n\to\infty} P\left\{\frac{Y_n - 0}{\sqrt{2n}} \leq x\right\} = \lim_{n\to\infty} P\left\{Y_n \leq x \cdot \sqrt{2n}\right\} = \Phi(x).$$

If we put $x = \dfrac{3}{\sqrt{n}}$, then

$$\begin{aligned} P\left\{|Y_n| \leq 3\sqrt{n}\right\} &= P\left\{Y_n \leq \frac{3}{\sqrt{2}} \cdot \sqrt{2n}\right\} - P\left\{Y_n < -\frac{3}{\sqrt{2}} \cdot \sqrt{2n}\right\} \\ &\to 2\,\Phi\left(\frac{3}{\sqrt{2}}\right) - 1 \approx 2 \cdot \Phi(2,121) - 1 \\ &\approx 2 \cdot 0.983 - 1 = 0.966. \end{aligned}$$

6) Since

$$Y_4 = 3\sum_{i=1}^{4} Z_i - 4,$$

where

$$\sum_{i=1}^{4} Z_i \in B\left(4, \frac{1}{3}\right)$$

is binomially distributed, we get

$$P\left\{Y_4 = 3k-4\right\} = \binom{4}{k} \cdot \left(\frac{1}{3}\right)^{k} \cdot \left(\frac{2}{3}\right)^{4-k}, \qquad k = 0,\, 1,\, 2,\, 3,\, 4,$$

hence

$$\begin{aligned} P\{Y_4 = -4\} &= \frac{16}{81}, \\ P\{Y_4 = -1\} &= \frac{32}{81}, \\ P\{Y_4 = 2\} &= \frac{24}{81} = \frac{8}{27}, \\ P\{Y_4 = 5\} &= \frac{8}{81}, \\ P\{Y_4 = 8\} &= \frac{1}{81} \end{aligned}$$

Example 4.15 *Let $X_1, X_2, X_3, \dots$ be mutually independent random variables, all of the distribution given by*

$$P\{X_i = 1\} = \frac{3}{4}, \qquad P\{X_i = -3\} = \frac{1}{4}, \qquad i \in \mathbb{N},$$

and let

$$Y_n = \sum_{i=1}^{n} X_i, \qquad n \in \mathbb{N}.$$

1) Find mean and variance of X_i.

2) Find by means of the Central Limit Theorem,

$$\lim_{n \to \infty} P\left\{-2\sqrt{n} < Y_n < 2\sqrt{n}\right\} \qquad (3 \textit{ decimals}).$$

3) Explain why X_i has the same distribution as $4Z_i - 3$, where Z_i is Bernoulli distributed,

$$P\{Z_i = 1\} = \frac{3}{4}, \qquad P\{Z_i = 0\} = \frac{1}{4},$$

and find by means of this for every $n \in \mathbb{N}$ the probabilities

$$P\{Y_n = 4k - 3n\}, \qquad k = 0, 1, 2, \dots, n.$$

1) The mean is

$$E\{X_i\} = 1 \cdot \frac{3}{4} - 3 \cdot \frac{1}{4} = 0.$$

Hence the variance

$$V\{X_i\} = E\left\{X_i^2\right\} - (E\{X_i\})^2 ? 1 \cdot \frac{3}{4} + 9 \cdot \frac{1}{4} - 0^2 = \frac{3+9}{3} = 3.$$

2) The mean is

$$E\{Y_n\} = \sum_{i=1}^{n} E\{X_i\} = 0,$$

and the variance is $V\{Y_n\} = 3n$.
It follows from the Central Limit Theorem that

$$\begin{aligned}\lim_{n\to\infty} P\left\{-2\sqrt{n} < Y_n < 2\sqrt{n}\right\} &= \lim_{n\to\infty} P\left\{\left|\frac{Y_n - 0}{\sqrt{3n}}\right| < \frac{2}{\sqrt{3}}\right\} \\ &= 2\,\Phi\left(\frac{2}{\sqrt{3}}\right) - 1 = 2\,\Phi(1.1547) - 1 \\ &\approx 2 \cdot 0.876 - 1 = 0.752.\end{aligned}$$

3) This follows from **4.**, so we just indicate that

$$\begin{aligned}P\{Y_3 = -9\} &= \frac{1}{4^3} &&= \frac{1}{64}, \\ P\{Y_3 = -5\} &= \binom{3}{1} \cdot \frac{3}{4} \cdot \frac{1}{4^2} = \frac{9}{4^3} &&= \frac{9}{64}, \\ P\{Y_3 = -1\} &= \binom{3}{2} \cdot \left(\frac{3}{4}\right)^2 \cdot \frac{1}{4} = \frac{27}{4^3} &&= \frac{27}{64}, \\ P\{Y_3 = 3\} &= \left(\frac{3}{4}\right)^3 = \frac{27}{4^3} &&= \frac{27}{64}.\end{aligned}$$

4) It is obvious that if Z_i is Bernoulli distributed,

$$P\{Z_i = 1\} = \frac{3}{4}, \qquad P\{Z_i = 0\} = \frac{1}{4},$$

then

$$P\{X_i = 1\} = P\{4Z_i - 3 = 1\} = P\{Z_i = 1\} = \frac{3}{4},$$

and

$$P\{X_i = -3\} = P\{4Z_i - 3 = -3\} = P\{Z_i = 0\} = \frac{1}{4},$$

hence X_i and $4Z_i - 3$ have same distribution.
If we allow $X_i = 4Z_i - 3$, then

$$Y_n = \sum_{i=1}^{n} X_i = 4 \sum_{i=1}^{n} Z_i - 3n,$$

where

$$\sum_{i=1}^{n} Z_i \in B\left(n, \frac{3}{4}\right)$$

is binomially distributed.
Then

$$\begin{aligned}
P\left\{Y_n = 4k - 3n\right\} &= P\left\{4\sum_{i=1}^{n} Z_i - 3n = 4k - 3n\right\} = P\left\{\sum_{i=1}^{n} Z_i = k\right\} \\
&= \binom{n}{k}\left(\frac{3}{4}\right)^k\left(\frac{1}{4}\right)^{n-k} \qquad \text{for } k = 0,\, 1,\, \dots,\, n.
\end{aligned}$$

Putting $n = 3$ we get question **3.**, cf. the above.

5 The Maxwell distribution

Example 5.1 *1) Prove that the function*

$$f(x)=\begin{cases}\dfrac{4}{\sqrt{\pi}}\,x^2\,\exp\left(-x^2\right), & x\ge 0,\\[2ex] \qquad 0, & x<0,\end{cases}$$

is the frequency of a random variable X.

2) Find mean and variance of the random variable X.

3) Find the frequency of the random variable $Y=\dfrac{1}{2}\,X^2$.

4) Find the mean of the random variable Y.

1) Obviously, $f(x)\ge 0$ everywhere. Since

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\,dx &= \frac{4}{\sqrt{\pi}}\int_0^{\infty} x^2\,\exp\left(-x^2\right)\,dx=\frac{2}{\sqrt{\pi}}\int_0^{\infty} x\cdot\exp\left(-x^2\right)\cdot 2x\,dx\\ &= \frac{2}{\sqrt{\pi}}\left[x\left(-\exp\left(-x^2\right)\right)\right]_0^{\infty}+\frac{2}{\sqrt{\pi}}\int_0^{\infty}\exp\left(-x^2\right)\,dx=1,\end{aligned}$$

it follows that $f(x)$ is the frequency of a random variable X.

2) The mean is

$$E\{X\} = \frac{4}{\sqrt{\pi}} \int_0^{\infty} x \cdot x^2 \exp\left(-x^2\right)\, dx = \frac{2}{\sqrt{\pi}} \int_0^{\infty} y\, e^{-y}\, dy = \frac{1}{\sqrt{\pi}} \cdot 1! = \frac{2}{\sqrt{\pi}}.$$

Furthermore,

$$\begin{aligned} E\left\{X^2\right\} &= \frac{4}{\sqrt{\pi}} \int_0^{\infty} x^4 \exp\left(-x^2\right)\, dx = \frac{2}{\sqrt{\pi}} \int_0^{\infty} x^3 \exp\left(-x^2\right) \cdot 2x\, dx \\ &= \frac{2}{\sqrt{\pi}} \left[-x^3 \exp\left(-x^2\right)\right]_0^{\infty} + \frac{6}{\sqrt{\pi}} \int_0^{\infty} x^2 \exp\left(-x^2\right)\, dx \\ &= \frac{3}{2} \frac{4}{\sqrt{\pi}} \int_0^{\infty} x^2 \exp\left(-x^2\right)\, dx = \frac{3}{2} \cdot 1 = \frac{3}{2}, \end{aligned}$$

hence

$$V\{X\} = E\left\{X^2\right\} - (E\{X\})^2 = \frac{3}{2} - \frac{4}{\pi}.$$

3) Since $y = \dfrac{1}{2}\, x^2$ is a bijective map $\mathbb{R}_+ \to \mathbb{R}_+$ with the inverse $x = \sqrt{2y}$, the frequency of Y for $y > 0$ is given by

$$g(y) = f\left(\sqrt{2y}\right) \cdot \frac{1}{2} \cdot \frac{2}{\sqrt{2y}} = \frac{4}{\sqrt{\pi}} \cdot 2y \cdot e^{-2y} \cdot \frac{1}{\sqrt{2y}} = \frac{4}{\sqrt{\pi}} \sqrt{2y} \cdot e^{-2y},$$

hence

$$g(y) = \begin{cases} \dfrac{4}{\sqrt{\pi}} \sqrt{2y} \cdot e^{-2y} & \text{for } y > 0, \\ 0 & \text{for } y \leq 0. \end{cases}$$

4) Then by **2.**

$$E\{Y\} = E\left\{\frac{1}{2}\, X^2\right\} = \frac{1}{2}\, E\left\{X^2\right\} = \frac{1}{2} \cdot \frac{3}{2} = \frac{3}{4}.$$

Example 5.2 *Define a random variable Z by $Z = \dfrac{1}{2}\, m Y^2$, where Y is Maxwell distributed. Find the frequency of Z, and find mean and variance of Z. Which distribution has Z?*

The random variable $Z = \dfrac{1}{2}\, m Y^2$ represent the kinetic energy. Clearly, Z has only positive values. The frequency of Y is

$$f_Y(y) = \begin{cases} \sqrt{\dfrac{1}{\pi}} \cdot \dfrac{1}{\sigma^3}\, y^2 \exp\left(-\dfrac{y^2}{2\sigma^2}\right) & \text{for } y > 0, \\ 0 & \text{for } y \leq 0, \end{cases}$$

hence $Y \geq 0$. Therefore, $Y = \sqrt{\dfrac{2}{m}} \cdot \sqrt{Z}$, so when $z > 0$, the distribution function of Z is given by

$$F_Z(z) = P\{Z \leq z\} = P\left\{Y \leq \sqrt{\frac{2z}{m}}\right\} = F_Y\left(\sqrt{\frac{2z}{m}}\right).$$

By differentiation we get the frequency of Z for $z > 0$,

$$\begin{aligned} f_Z(z) &= f_Y\left(\sqrt{\frac{2z}{m}}\right) \cdot \sqrt{\frac{2}{m}} \cdot \frac{1}{2\sqrt{z}} = \frac{1}{\sqrt{2m}} \cdot \frac{1}{\sqrt{z}} \cdot \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sigma^3} \cdot \frac{2z}{m} \exp\left(-\frac{1}{2} \cdot \frac{2z}{\sigma^2 m}\right) \\ &= \frac{1}{\sqrt{\pi}} \cdot \frac{2}{m\sigma^3} \cdot \sqrt{z} \cdot \exp\left(-\frac{1}{2}\,\frac{2z}{\sigma^2 m}\right) = \frac{1}{\Gamma\left(\frac{3}{2}\right)} \left(\frac{2z}{m\sigma^2}\right)^{\frac{3}{2}-1} \cdot \frac{1}{2^{3/2}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{2z}{m\sigma^2}\right) \frac{2}{\sigma^2 m}. \end{aligned}$$

Since $f_Z(z) = 0$ for $z \leq 0$, it follows that we have a χ^2-distribution in the new random variable, $X = \dfrac{2Z}{m\sigma^2}$, or more precisely, $X \in \chi^2\left(\dfrac{3}{2}\right)$.

Notice that $\dfrac{dx}{dz} = \dfrac{2}{\sigma^2 m}$.

The mean is found by using the change of variable $t = \dfrac{y}{\sigma}$,

$$\begin{aligned} E\{Z\} &= \frac{m}{2}\, E\left\{Y^2\right\} = \frac{m}{2}\sqrt{\frac{2}{\pi}} \int_0^\infty \frac{1}{\sigma^3}\, y^2 y^2 \exp\left(-\frac{y^2}{2\sigma^2}\right) dy \\ &= \frac{m\sigma^2}{2} \cdot \sqrt{\frac{2}{\pi}} \int_0^\infty t^4 \exp\left(-\frac{t^2}{2}\right) dt = \frac{m\sigma^2}{2} \cdot \sqrt{\frac{2}{\pi}} \int_0^\infty t^3 \exp\left(-\frac{t^2}{2}\right) t\, dt \\ &= \frac{m\sigma^2}{2}\sqrt{\frac{2}{\pi}} \left\{\left[-t^3 \exp\left(-\frac{t^2}{2}\right)\right]_0^\infty + 3\int_0^\infty t \cdot \exp\left(-\frac{t^2}{2}\right) dt\right\} = \frac{3}{2}\, m\, \sigma^2. \end{aligned}$$

Furthermore,

$$\begin{aligned} E\left\{Z^2\right\} &= \frac{m^2}{4}\, E\left\{Y^4\right\} = \frac{m^2}{4}\sqrt{\frac{2}{\pi}} \int_0^\infty \frac{1}{\sigma^3}\, y^4 y^2 \exp\left(-\frac{y^2}{2\sigma^2}\right) dy, \qquad t = \frac{y}{\sigma} \\ &= \frac{m^2\sigma^4}{4}\sqrt{\frac{2}{\pi}} \int_0^\infty t^6 \exp\left(-\frac{t^2}{2}\right) dt = \frac{m^2\sigma^4}{4}\sqrt{\frac{2}{\pi}} \int_0^\infty t^5 \cdot \exp\left(-\frac{t^2}{2}\right) t\, dt \\ &= \frac{m^2\sigma^4}{4}\sqrt{\frac{2}{\pi}} \cdot 5 \int_0^\infty t^3 \cdot \exp\left(-\frac{t^2}{2}\right) t\, dt = \frac{m^2\sigma^4}{4}\sqrt{\frac{2}{\pi}} \cdot 5 \cdot 3 \int_0^\infty t \cdot \exp\left(-\frac{t^2}{2}\right) t\, dt \\ &= \frac{m^2\sigma^4}{4} \cdot 5 \cdot 3 \cdot 1 = \frac{15}{4}\left(m\sigma^2\right)^2, \end{aligned}$$

hence

$$V\{Z\} = E\left\{Z^2\right\} - (E\{Z\})^2 = \left(\frac{15}{4} - \frac{9}{4}\right) m^2\sigma^4 = \frac{3}{2}\, m^2\sigma^4.$$

6 The Gamma distribution

Example 6.1 *Let X_1 be a non-negative random variable of frequency $f_1(x)$, mean μ_1 and variance σ_1^2, and let the function f_2 be defined by*

$$f_2(x) = \begin{cases} \dfrac{1}{\mu_1}\, x\, f_1(x), & x > 0, \\[2mm] 0, & x \le 0. \end{cases}$$

1. *Prove that $f_2(x)$ is the frequency of a random variable X_2.*

2. *Compute the mean μ_2 of X_2 (expressed by μ_1 and σ_1^2).*

Let X_1 be exponentially distributed with parameter α and frequency $f_1(x)$, and let the functions $f_n(x)$, $n \in \mathbb{N}$, be defined recursively by

$$f_n(x) = \begin{cases} \dfrac{1}{\mu_{n-1}}\, x\, f_{n-1}(x), & x > 0, \\[2mm] 0, & x \le 0, \end{cases}$$

(where μ_{n-1} is the mean corresponding to the frequency $f_{n-1}(x)$).

3. *Prove that f_n is the frequency of a gamma distribution of form parameter n and scale parameter $\dfrac{1}{a}$.*

1) Since $f_1(x) \ge 0$ and $\mu_1 > 0$, we have $f_2(x) \ge 0$ and

$$\int_{-\infty}^{\infty} f_2(x)\, dx = \frac{1}{\mu_1} \int_0^{\infty} x\, f_1(x)\, dx = \frac{1}{\mu_1}\, E\{X_1\} = \frac{1}{\mu_1} \cdot \mu_1 = 1,$$

thus f_2 is the frequency of a random variable X_2.

2) Then by a straightforward computation,

$$E\{X_2\} = \int_0^{\infty} x\, f_2(x)\, dx = \frac{1}{\mu_1} \int_0^{\infty} x^2 f_1(x)\, dx = \frac{1}{\mu_1}\, E\left\{X^2\right\} = \frac{1}{\mu_1}\left(V\{X\} + \mu_1^2\right) = \frac{\sigma_1^2}{\mu_1} + \mu_1.$$

3) If

$$f_1(x) = \begin{cases} a\, e^{-ax} & \text{for } x \ge 0, \\[2mm] 0 & \text{for } x < 0, \end{cases}$$

then

$$f_2(x) = \begin{cases} \dfrac{a}{\mu_1}\, x\, e^{-ax} & \text{for } x \ge 0, \\[2mm] 0, & \text{for } x < 0. \end{cases}$$

This is a constant times the frequency of a $\Gamma\left(2, \dfrac{1}{a}\right)$-distribution. Since $f_2(x)$ itself is a frequency, the constant is 1.

Then assume that

$$f_n(x) = \begin{cases} \dfrac{a^n}{\Gamma(n)}\, x^{n-1} \exp(-ax) & \text{for } x > 0, \\ \\ 0 & \text{for } x < 0. \end{cases}$$

It follows by this construction that

$$f_{n+1}(x) = \begin{cases} \dfrac{1}{\mu_n} \cdot \dfrac{a^n}{\Gamma(n)}\, x^{n} \exp(-ax) & \text{for } x > 0, \\ \\ 0 & \text{for } x \le 0, \end{cases}$$

and we conclude that

$$\frac{1}{\mu_n} \cdot \frac{a^n}{\Gamma(n)} = \frac{a^{n+1}}{\Gamma(n+1)},$$

and that $X_{n+1} \in \Gamma\left(n+1, \dfrac{1}{a}\right)$, and the claim follows by induction.

Remark 6.1 It follows from the above that

$$\mu_n = a\,\frac{\Gamma(n+1)}{\Gamma(n)} = n\,a. \qquad \Diamond$$

Example 6.2 *A random variable X has the frequency*

$$f(x) = \begin{cases} \dfrac{1}{B(u,v)}\,x^{u-1}(1-x)^{v-1}, & 0<x<1, \\ \\ 0 & \textit{otherwise}, \end{cases}$$

where u and v denote positive constants, and where $B(u,v)$ denotes the Beta function.
Find mean and variance of the random variable X.
The distribution of X is called the Beta distribution of form parameters u and v *(this is written $X \in Be(u,v)$). If in particular $u = v = \frac{1}{2}$, then we get the* Arcussinus distribution*, and if $u = v = 1$, then we get the* uniform distribution *over* $]0,1[$.

The Beta function is defined by

$$B(u,v) = \int_0^1 z^{u-1}(1-z)^{v-1}\,dz = \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)}, \qquad u,\, v \in \mathbb{R}_+.$$

The mean is

$$E\{X\} = \frac{1}{B(u,v)}\int_0^1 x^u(1-x)^{v-1}\,dx = \frac{B(u+1,v)}{B(u,v)} = \frac{\Gamma(u+1)\Gamma(v)}{\Gamma(u+v+1)}\cdot\frac{\Gamma(u+v)}{\Gamma(u)\Gamma(v)} = \frac{u}{u+v}.$$

Analogously,

$$E\left\{X^2\right\} = \frac{B(u+2,v)}{B(u,v)} = \frac{(u+1)u}{(u+v+1)(u+v)},$$

so the variance is given by

$$\begin{aligned} V\{X\} &= \frac{u}{u+v}\left\{\frac{u+1}{u+v+1} - \frac{u}{u+v}\right\} = \frac{u}{u+v}\left\{\left(1-\frac{v}{u+v+1}\right) - \left(1-\frac{v}{u+v}\right)\right\} \\ &= \frac{uv}{u+v}\left\{\frac{1}{u+v} - \frac{1}{u+v+1}\right\} = \frac{uv}{(u+v)^2(u+v+1)}. \end{aligned}$$

If $u = v = \frac{1}{2}$, then

$$E\{X\} = \frac{1}{2} \qquad \text{and} \qquad V\{X\} = \frac{1}{8}.$$

If $u = v = 1$, then

$$E\{X\} = \frac{1}{2} \qquad \text{and} \qquad V\{X\} = \frac{1}{12}.$$

Example 6.3 *A two-dimensional random variable* (X,Y) *has the frequency*

$$f(x,y)=\begin{cases}\dfrac{1}{\Gamma(\mu)\Gamma(\nu)}\,x^{\mu-1}(y-x)^{\nu-1}e^{-y}, & 0<x<y,\\[2mm] 0, & \textit{otherwise},\end{cases}$$

where μ *and* ν *denote positive constants.*

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

2) Find mean and variance of the random variables X *and* Y*.*

3) Compute the correlation coefficient $\varrho(X,Y)$*.*

1) When $x>0$, then

$$f_X(x)=\frac{x^{\mu-1}}{\Gamma(\mu)\Gamma(\nu)}\int_y^{\infty}(y-x)^{\nu-1}e^{-y}\,dy=\frac{1}{\Gamma(\mu)}\,x^{\mu-1}\,\frac{1}{\Gamma(\nu)}\int_0^{\infty}t^{\nu-1}e^{-t-x}\,dt=\frac{1}{\Gamma(\mu)}\,x^{\mu-1}e^{-x},$$

so X is Gamma distributed, $X\in\Gamma(\mu,1)$.

When $y>0$, then

$$\begin{aligned}f_Y(y)&=\frac{e^{-y}}{\Gamma(\mu)\Gamma(\nu)}\int_0^{y}x^{\mu-1}(y-x)^{\nu-1}\,dx=\frac{1}{\Gamma(\mu)\Gamma(\nu)}\,e^{-y}\cdot y^{\mu+\nu-1}\int_0^1 t^{\mu-1}(1-t)^{\nu-1}\,dt\\&=\frac{1}{\Gamma(\mu+\nu)}\,y^{\mu+\nu-1}e^{-y},\end{aligned}$$

so Y is also Gamma distributed, $Y\in\Gamma(\mu+\nu,1)$.

2) Since $\alpha=1$ in both cases, it follows by a known formula that

$$E\{X\}=V\{X\}=\mu\quad\text{and}\quad E\{Y\}=V\{Y\}=\mu+\nu.$$

3) Compute,

$$\begin{aligned}E\{XY\}&=\frac{1}{\Gamma(\mu)\Gamma(\nu)}\int_0^{\infty}y\,e^{-y}\left\{\int_0^{y}x^{\mu}(y-x)^{\nu-1}\,dx\right\}dy\\&=\frac{1}{\Gamma(\mu)\Gamma(\nu)}\int_0^{\infty}y\,e^{-y}\cdot y^{\mu+\nu}\left\{\int_0^{1}t^{\mu}(1-t)^{\nu-1}\,dt\right\}dy\\&=\frac{1}{\Gamma(\mu)\Gamma(\nu)}\cdot\frac{\Gamma(\mu+1)\cdot\Gamma(\nu)}{\Gamma(\mu+\nu+1)}\int_0^{\infty}y^{\mu+\nu+1}e^{-y}\,dy\\&=\frac{\mu}{\Gamma(\mu+\nu+1)}\cdot\Gamma(\mu+\nu+2)=\mu(\mu+\nu+1).\end{aligned}$$

It follows that

$$\operatorname{Cov}(X,Y)=\mu(\mu+\nu+1)-\mu(\mu+\nu)=\mu,$$

hence

$$\varrho=\frac{\mu}{\sqrt{\mu(\mu+\nu)}}=\sqrt{\frac{\mu}{\mu+\nu}}.$$

Example 6.4 *Let X_1 and X_2 be mutually independent random variables,*

$$X_i \in \Gamma\left(\mu_i, \alpha\right), \qquad i = 1,\, 2,$$

and let

$$Z = \frac{X_1}{X_1 + X_2}.$$

Compute the frequency of Z

1) by finding the frequency of $Y = X_1/X_2$ and then use that

$$Z = \frac{Y}{1+Y}.$$

2) by finding the frequency of the two-dimensional random variable $(Z, X_1 + X_2)$.

1) In general, the frequency of $Y = X_1/X_2$ is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f_1(yx) f_2(x)\,|x|\,dx.$$

Since X_1 and X_2 are positive, Y is also positive. Then for $y > 0$,

$$\begin{aligned}
f_Y(y) &= \frac{1}{\Gamma\left(\mu_1\right)\alpha^{\mu_1}} \cdot \frac{1}{\Gamma\left(\mu_2\right)\alpha^{\mu_2}} \cdot \int_0^{\infty} (yx)^{\mu_1 - 1} \exp\left(-\frac{yx}{\alpha}\right) \cdot x^{\mu_2 - 1} \exp\left(-\frac{x}{\alpha}\right) x\,dx \\
&= \frac{y^{\mu_1 - 1}}{\Gamma\left(\mu_1\right)\Gamma\left(\mu_2\right)\alpha^{\mu_1+\mu_2}} \cdot \int_0^{\infty} x^{\mu_1+\mu_2-1} \exp\left(-\frac{y+1}{\alpha}\,x\right)\,dx \\
&= \frac{y^{\mu_1 - 1}}{\Gamma\left(\mu_1\right)\Gamma\left(\mu_2\right)\alpha^{\mu_1+\mu_2}} \left(\frac{\alpha}{y+1}\right)^{\mu_1+\mu_2-1} \cdot \frac{\alpha}{y+1} \int_0^{\infty} t^{\mu_1+\mu_2-1} e^{-t}\,dt \\
&= \frac{\Gamma\left(\mu_1+\mu_2\right)}{\Gamma\left(\mu_1\right)\Gamma\left(\mu_2\right)} \cdot \frac{y^{\mu_1-1}}{(y+1)^{\mu_1+\mu_2}}.
\end{aligned}$$

Now, consider the mapping

$$z = \varphi(y) = \frac{y}{1+y} = 1 - \frac{1}{1+y}, \qquad \varphi : \mathbb{R}_+ \to]0,1[.$$

Its inverse is given by

$$y = \varphi^{-1}(z) = \frac{z}{1-z} = -1 + \frac{1}{1-z},$$

from which we in particularly derive that

$$y + 1 = \frac{1}{1-z} \quad \text{and} \quad \frac{dy}{dz} = \frac{1}{(1-z)^2}, \qquad z \in]0,1[.$$

Hence, for $z \in]0,1[$,

$$\begin{aligned} f_Z(z) &= f_Y\left(\varphi^{-1}(z)\right) \cdot \left|\left(\varphi^{-1}\right)'(z)\right| = \frac{\Gamma(\mu_1+\mu_2)}{\Gamma(\mu_1)\,\Gamma(\mu_2)} \cdot \frac{\left(\dfrac{z}{1-z}\right)^{\mu_1-1}}{\left(\dfrac{1}{1-z}\right)^{\mu_1+\mu_2}} \cdot \frac{1}{(1-z)^2} \\ &= \frac{\Gamma(\mu_1+\mu_2)}{\Gamma(\mu_1)\,\Gamma(\mu_2)}\, z^{\mu_1-1}(1-z)^{\mu_1+\mu_2-(\mu_1-1)-2} = \frac{\Gamma(\mu_1+\mu_2)}{\Gamma(\mu_1)\,\Gamma(\mu_2)}\, z^{\mu_1-1}(1-z)^{\mu_2-1}, \end{aligned}$$

and $f_Z(z) = 0$ otherwise.

2) Alternatively, consider the mapping

$$\tau(x_1,x_2) = \left(\frac{x_1}{x_1+x_2}, x_1+x_2\right) = (z_1,z_2), \qquad (x_1,x_2) \in \mathbb{R}_+^2.$$

It follows that

$$x_1 = z_1(x_1+x_2) = z_1 z_2 \quad \text{and} \quad x_2 = z_2 - x_1 = z_2(1-z_1),$$

hence

$$x_1 = z_1 z_2 \qquad \text{and} \qquad x_2 = z_2(1-z_1).$$

The image of τ is $\tau\left(\mathbb{R}_+^2\right) =]0,1[\,\times\mathbb{R}_+$.
The Jacobian is

$$\frac{\partial\left(x_1,x_2\right)}{\partial\left(z_1,z_2\right)} = \left|\begin{array}{cc} z_2 & z_1 \\ -z_2 & 1-z_1 \end{array}\right| = \left|\begin{array}{cc} z_2 & z_1 \\ 0 & 1 \end{array}\right| = z_2 > 0,$$

thus the frequency of $\left(Z_1,Z_2\right) = \left(Z, X_1+X_2\right)$ for $\left(z_1,z_2\right) \in\,]0,1[\,\times\mathbb{R}_+$ is given by

$$\begin{aligned} f\left(z_1,z_2\right) &= f_1\left(z_1 z_2\right) f_2\left(z_2\left(1-z_1\right)\right)\cdot z_2 \\ &= \frac{1}{\Gamma\left(\mu_1\right)}\cdot\frac{1}{\alpha^{\mu_1}}\cdot\left(z_1 z_2\right)^{\mu_1-1}\exp\left(-\frac{z_1 z_2}{\alpha}\right)\times \\ &\qquad \times\frac{1}{\Gamma\left(\mu_2\right)}\cdot\frac{1}{\alpha^{\mu_2}}\,\left(z_2\left(1-z_1\right)\right)^{\mu_1-1}\exp\left(-\frac{z_2-z_1 z_2}{\alpha}\right) z_2 \\ &= \frac{1}{\Gamma\left(\mu_1\right)\Gamma\left(\mu_2\right)}\cdot\frac{1}{\alpha^{\mu_1+\mu_2}}\, z_1^{\mu_1-1}\left(1-z_1\right)^{\mu_1-1}\cdot z_2^{\mu_1-1+\mu_2-1+1}\exp\left(-\frac{z_2}{\alpha}\right) \\ &= \frac{\Gamma\left(\mu_1+\mu_2\right)}{\Gamma\left(\mu_1\right)\Gamma\left(\mu_2\right)}\, z_1^{\mu_1-1}\left(1-z_1\right)^{\mu_2-1}\cdot\frac{1}{\Gamma\left(\mu_1+\mu_2\right)\alpha^{\mu_1+\mu_2}}\, z^{\mu_1+\mu_2-1}\exp\left(-\frac{z_2}{\alpha}\right), \end{aligned}$$

hence Z_2 is $\Gamma\left(\mu_1+\mu_2,\alpha\right)$ distributed, and $Z = \dfrac{X_1}{X_1+X_2}$ has the frequency

$$f_Z(z) = \frac{\Gamma\left(\mu_1+\mu_2\right)}{\Gamma\left(\mu_1\right)\Gamma\left(\mu_2\right)}\, z^{\mu_1-1}(1-z)^{\mu_2-1} \qquad \text{for } z \in\,]0,1[,$$

which is seen by integrating with respect to z_2 and then putting $z_1 = z$, and of course $f_Z(z) = 0$ for $z \notin\,]0,1[$.

Example 6.5 *Let X_1, X_2, … be mutually independent random variables, all of the frequency*

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

and let $S_n = \sum_{i=1}^{n} X_i$, $n \in \mathbb{N}$, and $S_0 = 0$.

1. *Find the distribution of S_n.*

Let t be a fixed positive number. We define a random variable N of values in $\mathbb{N}_0$ by

$$N = n, \quad \text{if } S_n \le t \text{ and } S_{n+1} > t.$$

2. *Prove that*

$$P\{N = n\} = \frac{t^{2n+1}e^{-t}}{(2n+1)!} + \frac{t^{2n}e^{-t}}{(2n)!}, \qquad n \in \mathbb{N}_0.$$

3. *Compute the mean $E\{N\}$.*

1) Every X_i is $\Gamma(2,1)$ distributed, so S_n is $\Gamma(2n,1)$ distributed, and S_n has the frequency

$$f_n(s) = \begin{cases} \dfrac{1}{(2n-1)!}\, x^{2n-1}e^{-x} & \text{for } x > 0, \\ \\ 0 & \text{for } x \le 0, \end{cases} \qquad n \in \mathbb{N}.$$

If $n = 0$, then S_0 is causally distributed.

2) We see that

$$\begin{aligned} P\{N = 0\} &= P\{S_0 \le t,\, S_1 > t\} = P\{S_1 > t\} \\ &= \int_t^{\infty} x\,e^{-x}\,dx = \left[-x\,e^{-x} - e^{-x}\right]_t^{\infty} = t\,e^{-t} + e^{-t}, \end{aligned}$$

in agreement with the formula.

If $n \in \mathbb{N}$, then

$$\begin{aligned} P\{N = n\} &= P\{S_n \le t \wedge S_{n+1} > t\} = P\{S_n \le t\} - P\{S_n \le t \wedge S_{n+1} \le t\} \\ &= P\{S_n \le t\} - P\{S_{n+1} \le t\} = \int_0^t \frac{x^{2n-1}}{(2n-1)!}\,e^{-x}\,dx - \int_0^t \frac{x^{2n+1}}{(2n+1)!}\,e^{-x}\,dx, \end{aligned}$$

thus

$$\begin{aligned} P\{N = n\} &= \int_0^t \frac{x^{2n-1}}{(2n-1)!}\,e^{-x}\,dx - \int_0^t \frac{x^{2n+1}}{(2n+1)!}\,e^{-x}\,dx \\ &= \int_0^t \frac{x^{2n-1}}{(2n-1)!}\,e^{-x}\,dx + \left[\frac{x^{2n+1}}{(2n+1)!}\,e^{-x}\right]_0^t - \int_0^t \frac{x^{2n}}{(2n)!}\,e^{-x}\,dx \\ &= \int_0^t \frac{x^{2n-1}}{(2n-1)!}\,e^{-x}\,dt + \frac{t^{2n+1}}{(2n+1)!}\,e^{-t} + \left[\frac{x^{2n}}{(2n)!}\,e^{-x}\right]_0^t - \int_0^t \frac{x^{2n-1}}{(2n-1)!}\,e^{-x}\,dx \\ &= \frac{t^{2n+1}}{(2n+1)!}\,e^{-t} + \frac{t^{2n}}{(2n)!}\,e^{-t}, \end{aligned}$$

and the claim is proved.

3) The mean is

$$\begin{aligned}
E\{N\} &= e^{-t}\sum_{n=1}^{\infty}\frac{n\cdot t^{2n+1}}{(2n+1)!}+e^{-t}\sum_{n=1}^{\infty}\frac{n\cdot t^{2n}}{(2n)!}\\
&= \frac{1}{2}\,e^{-t}\left\{\sum_{n=1}^{\infty}(2n+1)\frac{t^{2n+1}}{(2n+1)!}-\sum_{n=1}^{\infty}\frac{t^{2n+1}}{(2n+1)!}+\sum_{n=1}^{\infty}\frac{t^{2n}}{(2n-1)!}\right\}\\
&= \frac{1}{2}\,e^{-t}\left\{t\sum_{n=1}^{\infty}\frac{t^{2n}}{(2n)!}+t\sum_{n=1}^{\infty}\frac{t^{2n-1}}{(2n-1)!}-\sum_{n=0}^{\infty}\frac{t^{2n+1}}{(2n+1)!}+t\right\}\\
&= \frac{1}{2}\,e^{-t}\left\{t\sum_{n=0}^{\infty}\frac{t^{n}}{n!}-\sum_{n=0}^{\infty}\frac{t^{2n+1}}{(2n+1)!}\right\}=\frac{1}{2}\,e^{-t}\left\{t\,e^{-t}-\sinh t\right\}=\frac{1}{2}\,t-\frac{1}{4}\left(1-e^{-2t}\right).
\end{aligned}$$

Example 6.6 *Given a sequence of random variables (X_n), where X_n has the frequency*

$$f_n(x)=\begin{cases}\dfrac{1}{(n-1)!}\,a^n x^{n-1}e^{-ax} & \text{for } x>0,\\[2mm] 0 & \text{for } x\le 0,\end{cases}$$

where a denotes a positive constant.

1) Find the mean $E\{X_n\}$ and the variance $V\{X_n\}$ of the random variable X_n.

2) Compute the probability $P\{X_4 > E\{X_4\}\}$.

3) Assuming that X_2 and X_n are independent, find the frequency of $X_n + X_2$.

4) Assuming that Y_1, Y_2, …, Y_n are mutually independent random variables, all of the same distribution as X_2, prove that the frequency of the random variable

$$Z_n = Y_1 + Y_2 + \cdots + Y_n$$

is given by f_{2n}.

5) Apply the Central Limit Theorem to prove that

$$P\{X_{2n} > E\{X_{2n}\}\} \to \frac{1}{2} \qquad \text{for } n\to\infty.$$

1) It follows from $X_n \in \Gamma\left(n,\dfrac{1}{a}\right)$ that

$$E\{X_n\}=\frac{n}{a} \qquad \text{and} \qquad V\{X_n\}=\frac{n}{a^2}.$$

2) By a computation,

$$\begin{aligned} P\{X_4 > E\{X_4\}\} &= P\left\{X_4 > \frac{4}{a}\right\} = \int_{\frac{4}{a}}^{\infty} \frac{1}{3!}\, a^4 x^3 e^{-ax}\, dx = \frac{1}{6}\int_4^{\infty} t^3 e^{-t}\, dt \\ &= \frac{1}{6}\left[-\left(t^3+3t^2+6t+6\right)e^{-t}\right]_4^{\infty} = \frac{64+48+24+6}{6}\, e^{-4} = \frac{71}{3}\, e^{-4}. \end{aligned}$$

3) Now, $X_n + X_2 \in \Gamma\left(n+2, \frac{1}{a}\right)$, so the frequency is

$$f(x) = f_{n+2}(x) = \begin{cases} \dfrac{1}{(n+1)!}\, a^{n+2} x^{n+1} e^{-ax} & \text{for } x > 0, \\ 0 & \text{for } x \le 0. \end{cases}$$

4) Since every $Y_k \in \Gamma\left(2, \frac{1}{a}\right)$, we have

$$Z_n = Y_1 + Y_2 + \cdots + Y_n \in \Gamma\left(2n, \frac{1}{a}\right),$$

and the frequency is again $f_{2n}(x)$, thus $Z_n = X_{2n}$.

5) Now,

$$X_{2n} = Z_n = Y_1 + Y_2 + \cdots + Y_n,$$

and all the Y_k have the same distribution and are mutually independent. Hence, we can apply the Central Limit Theorem,

$$\begin{aligned} &P\{X_{2n} > E\{X_{2n}\}\} = 1 - P\{X_{2n} - E\{X_{2n}\} \le 0\} \\ &\quad = 1 - P\left\{\frac{X_{2n} - E\{X_{2n}\}}{\sqrt{V\{X_{2n}\}}} \le 0\right\} \to 1 - \Phi(0) = 1 - \frac{1}{2} = \frac{1}{2}, \end{aligned}$$

for $n \to \infty$.

Example 6.7 *We assume that the lifetime of an instrument is exponentially distributed of parameter a. If the instrument is ruined, it is repaired, and the lifetime of the repaired instrument is assumed to have the same distribution as the lifetime of a new instrument. When the instrument is ruined for the k-th time, it is rejected. We assume that the lifetimes of the k periods of function are mutually independent random variables, thus they are all exponentially distributed of parameter a.*
Let Y_k denote the total lifetime of the instrument.

1) Find $\mu_k = E\{Y_k\}$ and $\sigma_k^2 = V\{Y_k\}$.

2) Prove that for every $\varepsilon > 0$,

$$P\left\{\left|\frac{Y_k}{\mu_k} - 1\right| \geq \varepsilon\right\} \to 0 \qquad \textit{for } k \to \infty.$$

3) What is the distribution of Y_k?

The Gamma distribution is reproductive in the form parameter, when the scale parameter is kept fixed. It therefore follows that $Y_k \in \Gamma\left(k, \frac{1}{a}\right)$. Hence

1)

$$\mu_k = E\{Y_k\} = \frac{k}{a} \qquad \text{and} \qquad \sigma_k^2 = V\{Y_k\} = \frac{k}{a^2}.$$

2) We get by Chebyshev's inequality,

$$P\left\{\left|\frac{Y_k}{\mu_k}-1\right|\ge\varepsilon\right\}=P\left\{|Y_k-\mu_k|\ge\varepsilon\,\mu_k\right\}\le\frac{\sigma_k^2}{\varepsilon^2\mu_k^2}=\frac{k\,a^2}{a^2\varepsilon^2k^2}=\frac{1}{\varepsilon^2k}\to 0\qquad\text{for } k\to\infty.$$

3) This has already been proved: $Y_k\in\Gamma\left(k,\dfrac{1}{a}\right)$.

Example 6.8 *Let X_1, X_2, …, X_n be mutually independent random variable, all $N\left(\mu,\sigma^2\right)$-distributed. As usual we introduce the random variables*

$$\overline{X}=\frac{1}{n}\sum_{i=1}^n X_i,\qquad S^2=\frac{1}{n-1}\sum_{i=1}^n\left(X_i-\overline{X}\right)^2.$$

Find the variance of S^2.
HINT: *Apply properties of the χ^2 distribution.*

Let

$$Y_n=\frac{1}{\sigma^2}\sum_{i=1}^n\left(X_i-\mu\right)^2.$$

Then

$$Y_n\in\Gamma\left(\frac{n}{2},2\right)=\chi^2(n),$$

hence

$$E\left\{Y_n\right\}=n\qquad\text{and}\qquad V\left\{Y_n\right\}=2n.$$

Then

$$\begin{aligned}S^2&=\frac{1}{n-1}\sum_{i=1}^n\left(X_i-\overline{X}\right)^2=\frac{1}{n-1}\sum_{i=1}^n\left(X_i-\mu+\mu-\overline{X}\right)^2\\&=\frac{1}{n-1}\sum_{i=1}^n\left(X_i-\mu\right)^2+\frac{2}{n-1}\sum_{i=1}^n\left(X_i-\mu\right)\left(\mu-\overline{X}\right)+\frac{1}{n-1}\sum_{i=1}^n\left(\overline{X}-\mu\right)^2\\&=\frac{\sigma^2}{n-1}Y_n-2\,\frac{n}{n-1}\,\left(\overline{X}-\mu\right)^2+\frac{n}{n-1}\,\left(\overline{X}-\mu\right)^2=\frac{\sigma^2}{n-1}Y_n-\frac{n}{n-1}\,\left(\overline{X}-\mu\right)^2.\end{aligned}$$

It follows that we can assume that $\mu=0$, which will simplify the setup. Thus

$$Y_n=\frac{1}{\sigma^2}\sum_{i=1}^n X_i^2,$$

and

$$\begin{aligned}
S^2 &= \frac{\sigma^2}{n-1}\,Y_n - \frac{n}{n-1}\,\overline{X}^2 = \frac{\sigma^2}{n-1}\,Y_n - \frac{n}{n-1}\left\{\frac{1}{n}\sum_{i=1}^{n} X_i\right\}^2 \\
&= \frac{\sigma^2}{n-1}\,Y_n - \frac{1}{n-1}\cdot\frac{1}{n}\left\{\sum_{i=1}^{n} X_i^2 + 2\sum_{1\le i<j\le n} X_iX_j\right\} \\
&= \frac{\sigma^2}{n-1}\,Y_n - \frac{\sigma^2}{n-1}\cdot\frac{1}{n}\,Y_n - \frac{1}{n-1}\cdot\frac{2}{n}\sum_{1\le i<j\le n} X_iX_j \\
&= \begin{cases} \dfrac{\sigma^2}{n}\,Y_n - \dfrac{2}{(n-1)n}\sum_{1\le i<j\le n} X_iX_j, \\[2ex] \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 - \dfrac{2}{(n-1)n}\sum_{1\le i<j\le n} X_iX_j, \end{cases}
\end{aligned}$$

where we shall use both expressions in the following.

First consider

$$\operatorname{Cov}\left(\sum_{i=1}^{n} X_i^2, \sum_{1\le j<k\le n} X_jX_k\right).$$

Since

$$\begin{aligned}
\operatorname{Cov}\left(X_i^2, X_jX_k\right) &= E\left\{X_i^2X_jX_k\right\} - \left(E\left\{X_i^2\right\}\,E\left\{X_jX_k\right\}\right)^2 \\
&= E\left\{X_\ell\right\}\cdots - \left(E\left\{X_i^2\right\}\cdot E\left\{X_j\right\}\cdot E\left\{X_k\right\}\right)^2 = 0,
\end{aligned}$$

where $\ell = j$ or $\ell = k$, and $E\left\{X_m\right\} = \mu = 0$, the covariance is zero.
This implies that

$$V\left\{S^2\right\} = \frac{\sigma^4}{n^2}\,V\left\{Y_n\right\} + \left(\frac{2}{(n-1)n}\right)^2 V\left\{\sum_{1\le i<j\le n} X_iX_j\right\}.$$

Analogously,

$$\begin{aligned}
V\left\{\sum_{1\le i<j\le n} X_iX_j\right\} &= \operatorname{Cov}\left(\sum_{j=2}^{n}\sum_{i=1}^{j-1} X_iX_j, \sum_{k=2}^{n}\sum_{\ell=1}^{k-1} X_kX_\ell\right) = \sum_{j=2}^{n}\sum_{i=1}^{j-1}\operatorname{Cov}\left(X_iX_j, \sum_{k=2}^{n}\sum_{\ell=1}^{k-1} X_kX_\ell\right) \\
&= \sum_{j=2}^{n}\sum_{i=1}^{j-1}\operatorname{Cov}\left(X_iX_j, X_iX_j\right) = \sum_{j=2}^{n}\sum_{i=1}^{j-1} V\left\{X_iX_j\right\},
\end{aligned}$$

where

$$\begin{aligned}
V\left\{X_iX_j\right\} &= E\left\{X_i^2X_j^2\right\} - \left(E\left\{X_i\right\}E\left\{X_j\right\}\right)^2 = E\left\{X_i^2\right\}\,E\left\{X_j^2\right\} - 0^2 \\
&= V\left\{X_i\right\}\,V\left\{X_j\right\} = \sigma^4.
\end{aligned}$$

The sum contains in total $\frac{1}{2}\,n(n-1)$ terms, hence by insertion,

$$V\left\{S^2\right\} = \frac{\sigma^4}{n^2}\cdot 2n + \left(\frac{2}{(n-1)n}\right)\cdot\frac{1}{2}\,(n-1)n\,\sigma^4 = \frac{2\sigma^4}{n}\left(1+\frac{1}{n-1}\right) = \frac{2\sigma^4}{n-1}.$$

ALTERNATIVELY (and somewhat easier) we see that

$$\frac{n-1}{\sigma^2}\,S^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \in \chi^2(n-1),$$

hence

$$V\left\{\frac{n-1}{\sigma^2}\,S^2\right\} = 2(n-1).$$

Then

$$V\left\{S^2\right\} = 2(n-1)\cdot\frac{\sigma^4}{(n-1)^2} = \frac{2\sigma^4}{n-1}.$$

Example 6.9 *Let X_1 and X_2 be independent random variables of the frequencies*

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

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

2. *Find the variances $V\{X_1\}$ and $V\{X_2\}$.*

Let the two-dimensional random variable $(Y_1,Y_2) = \tau(X_1,X_2)$ be given by

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

3. *Prove that τ maps $\mathbb{R}_+ \times \mathbb{R}_+$ bijectively onto $]0,1[\,\times\mathbb{R}_+$.*

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

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

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

7. *Find the means $E\{Y_1\}$ and $E\{Y_2\}$.*

8. *Find $P\{Y_1Y_2 < 1\}$.*

It is obvious that $X_1 \in \Gamma(2,1)$ and $X_2 \in \Gamma(3,1)$ are Gamma distributed.

1. and 2. It is immediately seen that

$$E\{X_1\} = 2, \quad E\{X_2\} = 3 \quad \text{and} \quad V\{X_1\} = 2, \quad V\{X_2\} = 3.$$

3. When the equations are solved with respect to (x_1, x_2), we get

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

The image is

$$\{(y_1, y_2) \mid y_1 y_2 > 0,\ y_1 > 0,\ (1 - y_1)\, y_2 > 0\},$$

i.e. $]0,1[\,\times\mathbb{R}_+$ after a reduction.
The Jacobian is

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

4. If $0 < y_1 < 1$ and $y_2 > 0$, then

$$\begin{aligned} k\,(y_1, y_2) &= f_{X_1}\,(y_1 y_2)\, f_{X_2}\,(y_2 - y_1 y_2) \cdot y_2 \\ &= y_1 y_2 \exp\,(-y_1 y_2) \cdot \frac{1}{2}\, y_2^2\, (1 - y_1)^2 \exp\,(-y_2 + y_1 y_2) \cdot y_2 = \frac{1}{2}\, y_1\, (1 - y_1)^2 \cdot y_2^4 e^{-y_2} \\ &= 12\, y_1\, (1 - y_1)^2 \cdot \frac{1}{24}\, y_2^4 e^{-y_2}, \end{aligned}$$

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

5., 6. and 7. Clearly, $Y_1 \in Be(2,3)$ is Beta distributed, and $Y_2 \in \Gamma(5,1)$ is Gamma distributed, and Y_1 and Y_2 are independent, and finally,

$$E\{Y_1\} = \frac{2}{2+3} = \frac{2}{5} \quad \text{and} \quad E\{Y_2\} = 5.$$

8. It follows from $Y_1 Y_2 = X_1$ that

$$P\{Y_1 Y_2 < 1\} = P\{X_1 < 1\} = \int_0^1 x_1 e^{-x_1}\, dx_1 = \left[-(x_1+1)\, e^{-x_1}\right]_0^1 = 1 - \frac{2}{e}.$$

Example 6.10 *Let X_1 and X_2 be independent random variables of the frequencies*

$$f_{X_1}(x_1) = \begin{cases} \dfrac{1}{6}\, x_1^3 e^{-x_1}, & x_1 > 0, \\[2ex] 0, & x_1 \le 0, \end{cases} \qquad f_{X_2}(x_2) = \begin{cases} e^{-x_2}, & x_2 > 0, \\[2ex] 0, & x_2 \le 0. \end{cases}$$

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

Let the two-dimensional random variable $(Y_1, Y_2) = \tau(X_1, X_2)$ be given by

$$Y_1 = X_1 + X_2, \qquad Y_2 = \frac{2X_2}{X_1 + X_2}.$$

2. *Prove that τ maps $\mathbb{R}_+ \times \mathbb{R}_+$ bijectively onto $\mathbb{R}_+ \times \left]0,2\right[$.*

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.*

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

6. *Find the means $E\{Y_1\}$ and $E\{Y_2\}$.*

1. Since $X_1 \in \Gamma(4,1)$ and $X_2 \in \Gamma(1,1)$, we get

$$E\{X_1\} = 4 \quad \text{and} \quad E\{X_2\} = 1.$$

2. The equations are solved,

$$\begin{cases} y_1 = x_1 + x_2, \\[2ex] y_2 = \dfrac{2x_2}{x_1 + x_2}, \end{cases} \qquad \text{fås} \qquad \begin{cases} x_1 = y_1 \left(1 - \dfrac{1}{2}\, y_2\right), \\[2ex] x_2 = \dfrac{1}{2}\, y_1 y_2, \end{cases}$$

and we see that τ is injective, and that τ maps $\mathbb{R}_+ \times \mathbb{R}_+$ onto

$$\left\{(y_1, y_2) \mid y_1 \left(1 - \frac{1}{2}\, y_2\right) > 0,\; y_1 y_2 > 0\right\} = \{(y_1, y_2) \mid y_1 > 0,\; 0 < y_2 < 2\},$$

thus

$$\tau : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+ \times]0, 2[\qquad \text{bijectively.}$$

The corresponding Jacobian is

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

3. It follows from the preparations in **2.** that the simultaneous frequency $k(y_1, y_2)$ of (Y_1, Y_2) for $(y_1, y_2) \in \mathbb{R}_+ \times]0, 2[$, is given by

$$\begin{aligned} k(y_1, y_2) &= f_{X_1}\left(y_1\left(1 - \frac{1}{2} y_2\right)\right) \cdot f_{X_2}\left(\frac{1}{2} y_1 y_2\right) \cdot \frac{1}{2} y_1 \\ &= \frac{1}{6} y_1^3 \left(1 - \frac{1}{2} y_2\right)^3 \cdot \exp\left(-y_1\left(1 - \frac{1}{2} y_2\right)\right) \cdot \exp\left(-\frac{1}{2} y_1 y_2\right) \cdot \frac{1}{2} y_1 \\ &= \left\{\frac{1}{24} y_1^4 e^{-y_1}\right\} \cdot 2\left\{1 - \frac{1}{2} y_2\right\}^3, \qquad y_1 > 0,\ 0 < y_2 < 2, \end{aligned}$$

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

4. and 5. Here **3.** immediately implies that Y_1 and Y_2 are independent random variables of the frequencies

$$k_{Y_1}(y_1) = \begin{cases} \frac{1}{24} y_1^4 e^{-y_1}, & y_1 > 0, \\ \\ 0, & y_1 \le 0, \end{cases} \qquad \text{i.e.} \quad Y_1 \in \Gamma(5, 1),$$

and

$$k_{Y_2}(y_2) = \begin{cases} 2\left\{1 - \frac{1}{2} y_2\right\} * 3, & 0 < y_2 < 2, \\ \\ 0, & \text{otherwise.} \end{cases}$$

6. Clearly, $E\{Y_1\} = 5$.
Furthermore,

$$\begin{aligned} E\{Y_2\} &= \int_0^2 y_2 k_{Y_2}(y_2)\, dy_2 = 2\int_0^2 y_2 \left(1 - \frac{1}{2} y_2\right)^3 dy_2 \\ &= 4\int_0^2 \left\{1 - \left(1 - \frac{1}{2} y_2\right)\right\} \cdot \left(1 - \frac{1}{2} y_2\right)^3 dy_2 \\ &= 8\int_0^2 \left\{\left(1 - \frac{1}{2} y_2\right)^3 - \left(1 - \frac{1}{2} y_2\right)^4\right\} \frac{1}{2}\, dy_2 \\ &= 8\left[-\frac{1}{4}\left(1 - \frac{1}{2} y_2\right)^4 + \frac{1}{5}\left(1 - \frac{1}{2} y_2\right)^5\right]_0^2 \\ &= 8\left\{\frac{1}{4} - \frac{1}{5}\right\} = \frac{8}{20} = \frac{2}{5}. \end{aligned}$$

ALTERNATIVELY, $2X_2 = Y_1 Y_2$. Since Y_1 and Y_2 are independent, we get

$$2\, E\left\{X_2\right\} = E\left\{Y_1\right\} \cdot E\left\{Y_2\right\},$$

hence

$$E\left\{Y_2\right\} = \frac{2\, E\left\{X_2\right\}}{E\left\{Y_1\right\}} = \frac{2}{5}.$$

Example 6.11 *Let (X_1, X_2) be a two-dimensional random variable of frequency*

$$h\left(x_1, x_2\right) = \begin{cases} \dfrac{1}{2}\left(x_1 + x_2\right) e^{-(x_1+x_2)}, & x_1 > 0 \text{ og } x_2 > 0, \\[2ex] 0, & \text{ellers.} \end{cases}$$

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

2. *Compute mean and variance of X_1.*

Define the random variables Y_1 and Y_2 by

$$\left(Y_1, Y_2\right) = \tau\left(X_1, X_2\right) = \left(X_1 + X_2\,,\, \frac{X_1 - X_2}{X_1 + X_2}\right).$$

One may use without proof that the vector function τ given by

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

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

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

3. *Find the simultaneous frequency $k\left(y_1, y_2\right)$ of (Y_1, Y_2).*

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

5. *Compute mean and variance of Y_1.*

6. *Compute $\operatorname{Cov}(X_1, X_2)$.*

1) We may without loss of generality restrict ourselves to the case where $x_1 > 0$ and $x_2 > 0$. Then

$$\begin{aligned} f_{X_1}\left(x_1\right) &= \frac{1}{2}\int_0^{\infty}\left(x_1 + x_2\right) e^{-(x_1+x_2)}\, dx_2 = \frac{1}{2}\, x_1 e^{-x_1}\int_0^{\infty} e^{-x_2}\, dx_2 + \frac{1}{2}\, e^{-x_1}\int_0^{\infty} x_2 e^{-x_2}\, dx_2 \\ &= \frac{1}{2}\left(x_1 + 1\right) e^{-x_1}, \end{aligned}$$

thus

$$f_{X_1}\left(x_1\right) = \begin{cases} \dfrac{1}{2}\left(x_1 + 1\right) e^{-x_1} & \text{for } x_1 > 0, \\[2ex] 0 & \text{for } x_1 \le 0. \end{cases}$$

By the symmetry,

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

2) The mean is

$$E\{X_1\} = \frac{1}{2}\int_0^\infty x_1^2 e^{-x_1}\,dx_1 + \frac{1}{2}\int_0^\infty x_1 e^{-x_1}\,dx_1 = 1 + \frac{1}{2} = \frac{3}{2}.$$

Now,

$$E\left\{X_1^2\right\} = \frac{1}{2}\int_0^\infty x_1^3 e^{-x_1} dx_1 + \frac{1}{2}\int_0^\infty x_1^2 e^{-x_1} dx_1 = \frac{3!}{2} + \frac{2!}{2} = 4,$$

so the variance is

$$V\{X_1\} = 4 - \left(\frac{3}{2}\right)^2 = \frac{16-9}{4} = \frac{7}{4}.$$

3) From

$$y_1 = x_1 + x_2 \quad \text{and} \quad y_2 = \frac{x_1 - x_2}{x_1 + x_2},$$

i.e. $y_1 y_2 = x_1 - x_2$, follows that

$$x_1 = \frac{1}{2}\,(y_1 + y_1 y_2) = \frac{1}{2}\,y_1\,(1 + y_2)$$

and

$$x_2 = \frac{1}{2}\,(y_1 - y_1 y_2) = \frac{1}{2}\,y_1\,(1 - y_2)\,.$$

The Jacobian is

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

thus if $y_1 > 0$ and $-1 < y_2 < 1$, then

$$k\,(y_1, y_2) = h\,(x_1(y), x_2(y)) \cdot \frac{1}{2}\,y_1 = \frac{1}{2}\,y_1^2 e^{-y_1} \cdot \frac{1}{2} = \frac{1}{4}\,y_1^2 e^{-y_1}.$$

4) It follows immediately from the result of **3.** that Y_1 and Y_2 are independent and that $Y_1 \in \Gamma(3,1)$, and Y_2 is uniformly distributed over $]-1,1[$.
The marginal frequencies are

$$g_{Y_1}\,(y_1) = \begin{cases} \frac{1}{2}\,y_1^2 e^{-y_1} & \text{for } y_1 > 0, \\[2ex] 0 & \text{for } y_1 \le 0, \end{cases}$$

and

$$g_{Y_2}\,(y_2) = \begin{cases} \frac{1}{2} & \text{for } -1 < y_2 < 1, \\[2ex] 0 & \text{ellers.} \end{cases}$$

5) From $Y_1 \in \Gamma(3,1)$ follows that $E\,\{Y_1\} = 3$ and $V\,\{Y_1\} = 3$.

6) The covariance can be found in several ways. We have e.g.

$$X_1 X_2 = \frac{1}{4}\left\{(X_1 + X_2)^2 - (X_1 - X_2)^2\right\} = \frac{1}{4}\,\left\{Y_1^2 - Y_1^2 Y_2^2\right\},$$

hence

$$\begin{aligned} \operatorname{Cov}\,(X_1 X_2) &= E\,\{X_1 X_2\} - E\,\{X_1\}\,E\,\{X_2\} = \frac{1}{4}\,E\left\{Y_1^2\right\}\left(1 - E\left\{Y_2^2\right\}\right) - \frac{3}{2}\cdot\frac{3}{2} \\ &= \frac{1}{4}\,\left(V\,\{Y_1\} + (E\,\{Y_1\})^2\right)\cdot\left(1 - V\,\{Y_2\} - (E\,\{Y_2\})^2\right) - \frac{9}{4} \\ &= \frac{1}{4}\,\left(3 + 3^3\right)\left\{1 - \frac{1}{2}\cdot 2^2 - 0^2\right\} - \frac{9}{4} = \frac{12}{4}\,\left(1 - \frac{1}{3}\right) - \frac{9}{4} = \frac{8}{4} - \frac{9}{4} = -\frac{1}{4}. \end{aligned}$$

ALTERNATIVELY,

$$\begin{aligned}
\operatorname{Cov}(X_1, X_2) &= \frac{1}{2}\int_0^{\infty}\int_0^{\infty} x_1 x_2 (x_1+x_2)\, e^{-(x_1+x_2)}\, dx_1\, dx_2 - \frac{9}{4} \\
&= \frac{1}{2}\int_0^{\infty}\int_0^{\infty} \left(x_1^2 x_2 + x_1 x_2^2\right) e^{-(x_1+x_2)} dx_1\, dx_2 - \frac{9}{4} \\
&= 2\cdot\frac{1}{2}\int_0^{\infty}\int_0^{\infty} x_1^2 x_2 e^{-x_1-x_2}\, dx_1 dx_2 - \frac{9}{4} \\
&= \int_0^{\infty} x_1^2 e^{-x_1}\, dx_1 \cdot \int_0^{\infty} x_2\, e^{-x_2}\, dx_2 - \frac{9}{4} = 2\cdot 1 - \frac{9}{4} = -\frac{1}{4}.
\end{aligned}$$

Example 6.12 *Let X_1 and X_2 be independent random variables of the frequencies*

$$f_{X_1}(x_1) = \begin{cases} \dfrac{1}{6}\, x_1^3 e^{-x_1}, & x_1 > 0, \\[2mm] 0, & \textit{otherwise.} \end{cases} \qquad f_{X_2}(x_2) = \begin{cases} \dfrac{2}{9}\,(3-x_2)\,, & 0 < x_2 < 3, \\[2mm] 0 & \textit{otherwise.} \end{cases}$$

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

Let the two-dimensional random variable $(Y_1, Y_2) = \tau(X_1, X_2)$ be given by

$$Y_1 = \frac{1}{3}\, X_1 X_2, \qquad Y_2 = \frac{1}{3}\, X_1 (3 - X_2)\,.$$

2. *Prove that τ maps $\mathbb{R}_+ \times\,]0,3[$ bijectively onto $\mathbb{R}_+ \times \mathbb{R}_+$.*

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.*

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

6. *Find the means $E\{Y_1\}$ and $E\{Y_2\}$.*

1) Clearly, X_1 is $\Gamma(4,1)$-distributed, so $E\{X_1\} = 4$. Furthermore,

$$E\{X_2\} = \frac{2}{9}\int_0^3 \left(3x_2 - x_2^2\right)\, dx_2 = \frac{2}{9}\left[\frac{3}{2}\, x_2^2 - \frac{1}{3}\, x_2^3\right]_0^3 = 3 - 2 = 1.$$

2) It follows from

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

that

$$x_1 = y_1 + y_2 \quad \text{and} \quad x_2 = \frac{3y_1}{y_1 + y_2} = 3 - \frac{3y_2}{y_1+y_2},$$

whence $x_1 = y_1 + y_2 > 0$, and thus $y_1 > 0$ (because $x_2 > 0$), and $y_2 > 0$ $\left(\text{because } \dfrac{y_1}{y_1+y_2} < 1\right)$,
so $\tau\left(\mathbb{R}_+ \times]0,3[\right) = \mathbb{R}_+ \times \mathbb{R}_+$.
The Jacobian is

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

3) The simultaneous frequency of (Y_1, Y_2) is for $y_1 > 0$ and $y_2 > 0$ given by

$$\begin{aligned} k\left(y_1,y_2\right) &= f_{X_1}\left(y_1+y_2\right)\, f_{X_2}\left(\frac{3y_1}{y_1+y_2}\right)\cdot\frac{3}{y_1+y_2} \\ &= \frac{1}{6}\,\left(y_1+y_2\right)^3\exp\left(-\left(y_1+y_2\right)\right)\cdot\frac{2}{9}\left(3-\frac{3y_1}{y_1+y_2}\right)\cdot\frac{3}{y_1+y_2} \\ &= \frac{1}{3}\,y_2\left(y_1+y_2\right)\exp\left(-\left(y_1+y_2\right)\right), \end{aligned}$$

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

4) When $y_1 > 0$ and $y_2 > 0$ the marginal frequencies are given by

$$k_{Y_1}\left(y_1\right) = \frac{1}{3}\,e^{-y_1}\int_0^\infty \left(y_2y_1+y_2^2\right)e^{-y_2}\,dy_2 = \frac{1}{3}\,\left(y_1+2\right)e^{-y_1},$$

and

$$k_{Y_2}\left(y_2\right) = \frac{1}{3}\,y_2e^{-y_2}\int_0^\infty \left(y_1+y_2\right)\,e^{-y_1}\,dy_1 = \frac{1}{3}\,y_2\,\left(1+y_2\right)\,e^{-y_2},$$

and $= 0$ otherwise.

5) Since $k_{Y_1}\left(y_1\right)k_{Y_2}\left(y_2\right) \neq k\left(y_1,y_2\right)$, it follows that Y_1 and Y_2 are not independent.

6) The means are

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

and

$$E\left\{Y_2\right\} = \frac{1}{3}\int_0^\infty \left(y_2^2+y_2^3\right)\,e^{-y_2}\,dy_2 = \frac{1}{3}\,(2!+3!) = \frac{8}{3}.$$

Example 6.13 *A shop is visited by both male and female customers, mutually independent of each other. The arrival times are measured from $t = 0$ (the opening time). Let*

$$X_1,\, X_1+X_2,\, X_1+X_2+X_3,\, \dots,$$

denote the arrival times of the first, the second, the third, ... of the male customers, and let analogously

$$Y_1,\, Y_1+Y_2,\, Y_1+Y_2+Y_3,\, \dots$$

denote the arrival times of the first, the second, the third ... of the female customers. We assume that the random variables X_i, $i \in \mathbb{N}$, are mutually independent and identically distributed of the frequency

$$f(x) = \begin{cases} \lambda\, e^{-\lambda x}, & x \geq 0, \\ \\ 0, & x < 0, \end{cases}$$

and that the random variables Y_i, $i \in \mathbb{N}$, are mutually independent and identically distributed of the frequency

$$g(y) = \begin{cases} \mu\, e^{-\mu x}, & y \geq 0, \\ \\ 0, & y < 0, \end{cases}$$

(here λ and μ are positive constants).
We furthermore assume that the Y_i are independent of the X_i.

1) *Find the frequency $f_2(x)$ of $X_1 + X_2$.*

2) *Find the frequency $f_3(x)$ of $X_1 + X_2 + X_3$.*

3) *Express by means of the random variables X_1, X_2 and Y_1 the event that at least two male customers arrive before the first female customer, and find the probability of this event.*

4) *Express by means of the random variables X_1, X_2, X_3 and Y_1 the event that at least three male customers arrive before the first female customer, and find the probability of this event.*

5) *Find the probability that precisely two male customers arrive before the first female customer.*

6) *Find the frequency of the random variable*

$$Z = \frac{X_1 + X_2}{Y_1}.$$

7) *Check if Z has a mean.*

1. and 2. Since $X_i \in \Gamma\left(1, \dfrac{1}{\lambda}\right)$ is exponential and Gamma distributed, we have

$$X_1 + X_2 \in \Gamma\left(2, \frac{1}{\lambda}\right) \quad \text{og} \quad X_1 + X_2 + X_3 \in \Gamma\left(3, \frac{1}{\lambda}\right),$$

and

$$f_2(x) = \begin{cases} \lambda^2 x\, e^{-\lambda x} & \text{for } x > 0, \\ 0 & \text{for } x \le 0, \end{cases}$$

and

$$f_3(x) = \begin{cases} \dfrac{1}{2}\,\lambda^3 x^2 e^{\lambda x} & \text{for } x > 0, \\ 0 & \text{for } x \le 0. \end{cases}$$

3. and 6. and 7. The event is expressed by

$$\{X_1 + X_2 < Y_1\} = \left\{\frac{X_1 + X_2}{Y_1} < 1\right\},$$

because X_1, X_2 and Y_1 only have positive values. When $z > 0$, then the frequency of

$$Z = \frac{X_1 + X_2}{Y_1}$$

is given by

$$\begin{aligned} f_Z(z) &= \int_0^{\infty} f_2(zx)\cdot g(x)|x|\,dx = \int_0^{\infty} \lambda^2 z\,x\,e^{-\lambda z x}\cdot \mu\,e^{-\mu x}\cdot x\,dx = \lambda^2\mu\,z\int_0^{\infty} x^2 e^{-(\lambda z+\mu)}\,dx \\ &= \frac{\lambda^2\mu z}{(\lambda z+\mu)^3}\int_0^{\infty} t^2 e^{-t}\,dt = \frac{2\lambda^2\mu z}{(\lambda z+\mu)^3}. \end{aligned}$$

Since $z\,f_Z(z)\sim \dfrac{2\mu}{\lambda}\cdot\dfrac{1}{z}$, and $\int_0^{\infty}\dfrac{1}{z}\,dz$ is divergent, Z does not have a mean.
Finally,

$$\begin{aligned} P\{X_1+X_2<Y_1\} &= P\{Z<1\} = \int_0^1 \frac{2\lambda^2\mu z}{(\lambda z+\mu)^3}\,dz = 2\lambda\mu\int_0^1 \frac{\lambda z+\mu-\mu}{(\lambda z+\mu)^3}\,dx \\ &= 2\lambda\mu\int_0^1\left\{\frac{1}{(\lambda z+\mu)^2}-\frac{\mu}{(\lambda z+\mu)^3}\right\}dz = 2\mu\left[-\frac{1}{\lambda z+\mu}+\frac{1}{2}\,\frac{\mu}{(\lambda z+\mu)^2}\right]_0^1 \\ &= 2\mu\left\{-\frac{1}{\lambda+\mu}+\frac{1}{2}\,\frac{\mu}{(\lambda+\mu)^2}+\frac{1}{\mu}-\frac{1}{2}\,\frac{\mu}{\mu^2}\right\} = -2\,\frac{\mu}{\lambda+\mu}+\left(\frac{\mu}{\lambda+\mu}\right)^2+1 \\ &= \left(1-\frac{\mu}{\lambda+\mu}\right)^2 = \left(\frac{\lambda}{\lambda+\mu}\right)^2. \end{aligned}$$

Analogously,

$$\{X_1+X_2+X_3\} = \left\{\frac{X_1+X_2+X_3}{Y_1}<1\right\}.$$

The frequency of

$$U = \frac{X_1+X_2+X_3}{Y_1}$$

for $u>0$ is given by

$$\begin{aligned} f_U(u) &= \int_0^{\infty} f_3(ux)\cdot g(x)\cdot x\,dx = \int_0^{\infty}\frac{1}{2}\,\lambda^3u^2x^2e^{-\lambda ux}\mu\,e^{-\mu x}x\,dx \\ &= \frac{1}{2}\,\lambda^3\mu\,u^2\int_0^{\infty} x^3e^{-(\lambda u+\mu)x}\,dx = \frac{1}{2}\,\frac{\lambda^3\mu u^2}{(\lambda u+\mu)^4}\int_0^{\infty} t^3e^{-t}\,dt = \frac{3\lambda^3\mu u^2}{(\lambda u+\mu)^4}. \end{aligned}$$

The probability of this event is

$$\begin{aligned}
P\left\{X_1+X_2+X_3<Y_1\right\} &= P\{U<1\} = \int_0^1 \frac{3\lambda^3\mu u^2}{(\lambda u+\mu)^4}\,du \\
&= 3\lambda\mu\int_0^1 \frac{\{(\lambda u+\mu)-\mu\}^2}{(\lambda u+\mu)^4}\,du = 3\lambda\mu\int_0^1 \frac{(\lambda u+\mu)^2-2\mu(\lambda u+\mu)+\mu^2}{(\lambda u+\mu)^4}\,du \\
&= 3\mu\int_0^1\left\{\frac{1}{(\lambda u+\mu)^2}-\frac{2\mu}{(\lambda u+\mu)^3}+\frac{\mu^2}{(\lambda u+\mu)^4}\right\}\lambda\,du \\
&= 3\mu\left[-\frac{1}{\lambda u+\mu}+\frac{\mu}{(\lambda u+\mu)^2}-\frac{1}{3}\,\frac{\mu^2}{(\lambda u+\mu)^3}\right]_0^1 \\
&= 3\mu\left\{-\frac{1}{\lambda+\mu}+\frac{\mu}{(\lambda+\mu)^2}-\frac{1}{3}\,\frac{\mu^2}{(\lambda+\mu)^3}+\frac{1}{\mu}-\frac{\mu}{\mu^2}+\frac{1}{3}\,\frac{\mu^2}{\mu^3}\right\} \\
&= 1-3\,\frac{\mu}{\lambda+\mu}+3\left(\frac{\mu}{\lambda+\mu}\right)^2-\left(\frac{\mu}{\lambda+\mu}\right)^3 = \left(1-\frac{\mu}{\lambda+\mu}\right)^3 = \left(\frac{\lambda}{\lambda+\mu}\right)^3.
\end{aligned}$$

The wanted probability is

$$\begin{aligned}
&P\left\{X_1+X_2<Y_1<X_1+X_2+X_3\right\} \\
&\quad= P\left\{X_1+X_2<Y_1\right\}-P\left\{X_1+X_2\le Y_1\,\wedge\,X_1+X_2+X_3\le Y_1\right\} \\
&\quad= P\left\{X_1+X_2<Y_1\right\}-P\left\{X_1+X_2+X_3\le Y_1\right\} \\
&\quad= \left(\frac{\lambda}{\lambda+\mu}\right)^2-\left(\frac{\lambda}{\lambda+\mu}\right)^3 = \left(\frac{\lambda}{\lambda+\mu}\right)\left\{1-\frac{\lambda}{\lambda+\mu}\right\} = \frac{\lambda^2\mu}{(\lambda+\mu)^3}.
\end{aligned}$$

Example 6.14 *There are two telephone booths in a waiting room. At a given time three persons A, B and C arrive, all of them wanting to telephone. A and B go immediately into each their booth and start their call, while C is waiting, until either A or B has finished his call. We shall in the following assume that the length of the telephone calls are mutually of each other, and that this length is a random variable, which is exponentially distributed of the frequency*

$$f(x)=\begin{cases}0, & x\le 0,\\[2mm] \lambda\, e^{-\lambda x}, & x>0,\end{cases}$$

where λ is a positive constant.
Let X_A, X_B and X_C denote the lengths of the telephone calls of A, B and C, respectively.

1) *Find the frequency and the distribution function of the random variable $Y = X_A/X_B$, and find $P\{X_A < X_B\}$.*

2) *Check if Y has a mean.*

3) *Find the distribution of the waiting time Z for C, and find the mean of Z.*

4) *Find the distribution of the random variable $X_A + X_C$.*

5) *Find the probability that A terminates his call before B, and that also C terminates his call before B, i.e. find*

$$P\{X_A+X_C<X_B\}.$$

6) *Find the probability that C is not the last one to terminate his call.*

1) The frequency of $Y = X_A/X_B$ for $y > 0$ is given by

$$\begin{aligned}g_Y(y) &= \int_0^\infty f_A(yx)f_B(x)\,dx=\int_0^\infty \lambda\, e^{-\lambda yx}\cdot \lambda\, e^{-\lambda x}\,x\,dx\\ &= \lambda^2\int_0^\infty x\,e^{-\lambda(1+y)x}\,dx=\frac{1}{(1+y)^2}, \qquad\qquad y>0,\end{aligned}$$

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

If $y \le 0$, we get the distribution function $G(y) = 0$. If $y > 0$, then

$$G(y)=\int_0^y g(u)\,du=\int_0^y\frac{du}{(1+u)^2}=\left[-\frac{1}{1+u}\right]_0^y=1-\frac{1}{1+y}\qquad\left(=\frac{y}{1+y}\right).$$

It follows by the symmetry that

$$P\{X_A<X_B\}=\frac12.$$

Alternatively,

$$P\{X_A<X_B\}=P\{Y<1\}=G(1)=\frac12.$$

2) Since

$$\int_0^{\infty} \frac{y}{(1+y)^2}\,dy = \infty,$$

it follows that the mean of Y does not exist.

ALTERNATIVELY,

$$E\left\{\frac{1}{X_B}\right\} = \int_0^{\infty} \frac{1}{x}\,\lambda\,e^{-\lambda x}\,dx = \infty,$$

hence

$$E\{Y\} = E\left\{X_A\right\} \cdot E\left\{\frac{1}{X_B}\right\} = \infty.$$

3) Clearly, $Z = \min\left\{X_A, X_B\right\}$. We get for $z > 0$,

$$P\{Z > z\} = P\left\{X_A > z\right\} \cdot P\left\{X_B > z\right\} = e^{-2\lambda z}.$$

This implies that Z is exponentially distributed of parameter 2λ, i.e.

$$f_Z(z) = \begin{cases} 2\lambda\,e^{-2\lambda z}, & z > 0, \\ 0, & z \le 0. \end{cases}$$

Consequently,

$$E\{Z\} = \frac{1}{2\lambda}.$$

4) According to the properties of the Gamma distribution, $X_A + X_C$ has the frequency

$$h(x) = \begin{cases} \lambda^2 x\,e^{-\lambda x}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

5) Since $X_A + X_C$ has the frequency $h(x)$, and X_B has the frequency $f_B(y)$, we get

$$\begin{aligned} P\left\{X_A + X_C < X_B\right\} &= \int_{x=0}^{\infty} \left\{\int_{y=x}^{\infty} h(x)\,f_B(y)\,dy\right\} dx = \int_0^{\infty} \lambda^2 x\,e^{-\lambda x} \left\{\int_{y=x}^{\infty} \lambda\,e^{-\lambda y}\,dy\right\} dx \\ &= \int_0^{\infty} \lambda^2 x\,e^{-\lambda x} \cdot e^{-\lambda x}\,dx = \int_0^{\infty} \lambda^2 x\,e^{-2\lambda x}\,dx = \frac{1}{4}. \end{aligned}$$

6) If C is not the last one to terminate his call, then

either $X_A + X_C < X_B$ or $X_B + X_C < X_A$,

hence

$$p = P\left\{X_A + X_C < X_B\right\} + P\left\{X_B + X_C < X_A\right\} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Example 6.15 *In a shop one first serves one customer with a serving time X_1 which is a random variable of frequency*

$$f_1(x_1) = \begin{cases} 2\,e^{-2x_1}, & x_1 > 0, \\ 0, & x_1 \le 0. \end{cases}$$

Then a family of 3 members is served, and the total serving time X_2 of this family is a random variable of frequency

$$f_2(x_2) = \begin{cases} 4x_2^2\,e^{-2x_2}, & x_2 > 0, \\ 0, & x_2 \le 0. \end{cases}$$

We assume that the random variables X_1 and X_2 are independent.

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

Define the random variables Y_1 and Y_2 by

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

Here, Y_1 is the total serving time of all 4 customers, and Y_2 is the quotient between the serving times of the single customer and the family. One may use without proof 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.

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. (This question may be answered with and without using the answer of* **2.***).*

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

5. *Find the means $E\{Y_1\}$ and $E\{Y_2\}$.*

6. *Find the median of Y_2.*

1. It follows from $X_1 \in \Gamma\left(1, \frac{1}{2}\right)$ and $X_2 \in \Gamma\left(3, \frac{1}{2}\right)$ that

$$E\{X_1\} = \frac{1}{2} \qquad \text{and} \qquad E\{X_2\} = \frac{3}{2}.$$

2. It follows from

$$y_1 = x_1 + x_2 \qquad \text{and} \qquad y_2 = \frac{x_1}{x_2}, \quad \text{i.e. } x_1 = y_2 x_2,$$

that $y_1 = y_2 x_2 + x_2 = (y_2+1)\,x_2$, thus $x_2 = y_1/(y_2+1)$, and hence

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

The Jacobian is

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

If therefore $(y_1,y_2) \in \mathbb{R}_+ \times \mathbb{R}_+$, then

$$\begin{aligned} k(y_1,y_2) &= f_1\left(\frac{y_1 y_2}{y_2+1}\right) f_2\left(\frac{y_1}{y_2+1}\right) \cdot \frac{y_1}{(y_2+1)^2} \\ &= 2\,\exp\left(-2y_1 + \frac{2y_1}{y_2+1}\right) \cdot 4\left(\frac{y_1}{y_2+1}\right)^2 \exp\left(-\frac{2y_1}{y_2+1}\right) \cdot \frac{y_1}{(y_2+1)^2} \\ &= 8\,y_1^3\,e^{-2y_1}\,\frac{1}{(y_2+1)^4} = \frac{1}{3}\,2^4\,y_1^3\,e^{-2y_1} \cdot \frac{3}{(y_2+1)^4}. \end{aligned}$$

3. and 4. It follows immediately from the expression of $k(y_1, y_2)$ that Y_1 and Y_2 are independent and that $Y_1 \in \Gamma\left(4, \frac{1}{2}\right)$, hence

$$k_{Y_1}(y_1) = \begin{cases} \dfrac{1}{3}\, 2^4\, y_1^3\, e^{-2y_1} & \text{for } y_1 > 0, \\ 0 & \text{for } y_1 \le 0, \end{cases}$$

and

$$k_{Y_2}(y_2) = \begin{cases} \dfrac{3}{(y_2+1)^4} & \text{for } y_2 > 0, \\ 0 & \text{for } y_2 \le 0. \end{cases}$$

5. Clearly,

$$E\{Y_1\} = 4 \cdot \frac{1}{2} = 2.$$

Furthermore,

$$\begin{aligned} E\{Y_2\} &= 3 \int_0^\infty \frac{y_2+1-1}{(y_2+1)^4}\, dy_2 = 3 \int_0^\infty \left\{ \frac{1}{(y_2+1)^3} - \frac{1}{(y_2+1)^4} \right\} dy_2 \\ &= 3 \left[-\frac{1}{2}\, \frac{1}{(y_2+1)^2} + \frac{1}{3}\, \frac{1}{(y_2+1)^3} \right]_0^\infty = 3\left(\frac{1}{2} - \frac{1}{3} \right) = \frac{1}{2}. \end{aligned}$$

6. The distribution function of Y_2 for $y_2 > 0$ is given by

$$F_{Y_2}(y_2) = \int_0^{y_2} \frac{3}{(t+1)^4}\, dt = \left[-\frac{1}{(t+1)^3} \right]_0^{y_2} = 1 - \frac{1}{(y_2+1)^3}.$$

Hence the median is given by

$$1 - \frac{1}{(y_2+1)^3} = \frac{1}{2},$$

i.e. $(y_2+1)^3 = 2$, so

$$(Y_2) = \sqrt[3]{2} - 1.$$

Example 6.16 *An instrument A contains two components, the lifetimes of which X_1 and X_2 are independent random variables, both of frequency*

$$f(x) = \begin{cases} 2a\,e^{-2ax}, & x > 0, \\ 0, & x \le 0, \end{cases}$$

where a is a positive constant.
We first use one of the components in the instrument A, and when this component is worn out, we shift immediately to the other component, thus the lifetime of A is $X = X_1 + X_2$.
Another instrument B only contains one component, the lifetime of which (which is also the lifetime of B) is a random variable Y of frequency

$$g(y) = \begin{cases} a\,e^{-ay}, & y > 0, \\ 0, & y \le 0. \end{cases}$$

1) Find mean and variance of Y.

2) Find frequency, mean and variance of X.

3) Let $U = \dfrac{X}{Y}$ denote the quotient between the lifetimes of the two instruments. Find the frequency of U.

4) Given 200 instruments of the same type as A. Let the random variable Z denote the sum of the lifetimes for all 200 instruments. Use the Central Limit Theorem to find an approximate expression of

$$P\left\{\frac{185}{a} < Z < \frac{215}{a}\right\}.$$

1) Since $Y \in \Gamma\left(1, \dfrac{1}{a}\right)$ is exponentially distributed, it follows by e.g. using a table that

$$E\{Y\} = \frac{1}{a} \qquad \text{and} \qquad V\{Y\} = \frac{1}{a^2}.$$

2) Since $X_i \in \Gamma\left(1, \dfrac{1}{2a}\right)$, $i = 1, 2$, are independent, the sum

$$X = X_1 + X_2 \in \Gamma\left(2, \frac{1}{2a}\right)$$

is Gamma distributed, hence (e.g. by using a table)

$$f_X(x) = \begin{cases} 4a^2 x\,e^{-2ax} & \text{for } x > 0, \\ 0, & \text{for } x \le 0, \end{cases}$$

and

$$E\{X\} = 2 \cdot \frac{1}{2a} = \frac{1}{a} \qquad \text{and} \qquad V\{X\} = 2 \cdot \left(\frac{1}{2a}\right)^2 = \frac{1}{2a^2}.$$

Alternatively, the frequency of X (for $x>0$) is

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X_1}(t) f_{X_2}(x-t)\,dt = \int_0^x 2a\,e^{-2at}\cdot 2a\,e^{-2a(x-t)}\,dt \\ &= 4a^2e^{-2ax}\int_0^x 1\,dt = 4a^2xe^{-2ax}. \end{aligned}$$

Then it follows that s

$$E\{X\} = \int_0^{\infty} x\,f_X(x)\,dx = \int_0^{\infty} 4a^2x^2e^{-2ax}\,dx = \frac{1}{2a}\int_0^{\infty} u^2e^{-u}\,du = \frac{1}{2a}\cdot 2 = \frac{1}{a},$$

and

$$E\left\{X^2\right\} = \int_0^{\infty} x^2 f_X(x)\,dx = \int_0^{\infty} 4a^2x^3e^{-2ax}dx = \frac{1}{4a^2}\int_0^{\infty} u^3e^{-u}du = \frac{1}{4a^2}\cdot 6 = \frac{3}{2a^2},$$

hence

$$V\{X\} = E\left\{X^2\right\} - (E\{X\})^2 = \frac{3}{2a^2} - \frac{1}{a^2} = \frac{1}{2a^2}.$$

3) The frequency of $U = \dfrac{X}{Y}$ is 0 for $u \le 0$. We get for $u>0$,

$$\begin{aligned} h_U(u) &= \int_0^{\infty} f_X(ux)g(x)|x|\,dx = \int_0^{\infty} 4a^2ux\,e^{-2aux}\cdot a\,e^{-ax}\cdot x\,dx \\ &= 4a^3u\int_0^{\infty} x^2e^{-a(1+2u)x}\,dx = \frac{4a^3u}{a^3(1+2u)^3}\int_0^{\infty} t^2e^{-t}\,dt = \frac{8u}{(1+2u)^3}. \end{aligned}$$

Summing up,

$$h_U(u) = \begin{cases} \dfrac{8u}{(1+2u)^3} & \text{for } u>0, \\ \\ 0, & \text{for } u \le 0. \end{cases}$$

Alternatively, one may for $u>0$ start with computing the distribution function $K(u)$ of U. This is given by

$$\begin{aligned} K(u) &= P\left\{\frac{X}{Y}\le u\right\} = \int_A f_X(x)f_Y(y)\,dx\,dx = \int_{x=a}^{\infty} f_X(x)\left\{\int_{y=\frac{x}{u}}^{\infty} f_Y(y)\,dy\right\}dx \\ &= \int_{x=0}^{\infty} 4a^2xe^{-2ax}\left\{\int_{y=\frac{x}{u}}^{\infty} a\,e^{-ay}dy\right\}dx = \int_{x=0}^{\infty} 4a^2x\,e^{-2ax}\cdot e^{-au}dx \\ &= \int_{x=0}^{\infty} 4a^2x\cdot\exp\left(-ax\left(2+\frac{1}{u}\right)\right)dx, \qquad v = a\left(2+\frac{1}{u}\right)x, \\ &= \int_{v=0}^{\infty} 4a^2v\cdot\frac{1}{a^2\left(2+\dfrac{1}{u}\right)^2}\cdot e^{-v}dv = \frac{4}{\left(2+\dfrac{1}{u}\right)^2} = \frac{4u^2}{(2u+1)^2} \\ &= 1-\frac{2}{2u+1}+\frac{1}{(2u+1)^2}. \end{aligned}$$

The we get the frequency (for $u > 0$) by differentiation,

$$f(u) = \frac{4}{(2u+1)^2} - \frac{4}{(2u+1)^3} = 4 \cdot \frac{(2u+1)-1}{(2u+1)^3} = \frac{8u}{(2u+1)^3}.$$

4) Since $Z = \sum_{i=1}^{200} \tilde{X}_i$, where the $\tilde{X}_i$ are mutually stochastically independent, we get

$$E\{Z\} = \frac{200}{a} \quad \text{and} \quad V\{Z\} = \frac{100}{a^2}, \quad \text{i.e.} \quad \sqrt{V\{Z\}} = \frac{10}{a}.$$

Then it follows by the Central Limit Theorem that

$$\begin{aligned}
P\left\{\frac{185}{a} < Z < \frac{215}{a}\right\} &= P\left\{\left|Z - \frac{200}{a}\right| < \frac{15}{a}\right\} \\
&= P\left\{\frac{\left|Z - \frac{200}{a}\right|}{\frac{10}{a}} < \frac{15}{a} \cdot \frac{1}{\frac{10}{a}}\right\} = P\left\{\frac{|Z - E\{Z\}|}{\sqrt{V\{Z\}}} < \frac{3}{2}\right\} \\
&\approx \Phi\left(\frac{3}{2}\right) - \Phi\left(-\frac{3}{2}\right) = 2\,\Phi\left(\frac{3}{2}\right) - 1 = 2\,\Phi\left(\frac{3}{2}\right) - 1 = 2 \cdot 0.9332 - 1 = 0.8664.
\end{aligned}$$

7 The normal distribution and the Gamma distribution

Example 7.1 *Assume that* (X_1, X_2) *has the frequency*

$$h(x_1, x_2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2}\left(x_1^2 + x_2^2\right)\right), \qquad (x_1, x_2) \in \mathbb{R}^2,$$

(i.e. X_1 *and* X_2 *are independent,* $N\left(0, \sigma^2\right)$ *distributed random variables), and let* $(X_1^\star, X_2^\star)$ *be given by*

$$X_1^\star = \min(X_1, X_2), \qquad X_2^\star = \max(X_1, X_2).$$

1. *Prove that* $(X_1^\star, X_2^\star)$ *has the frequency*

$$h^\star(x_1, x_2) = \begin{cases} 2h(x_1, x_2), & x_1 < x_2, \\ \\ 0, & x_1 \geq x_2. \end{cases}$$

Define random variables Y_1 *and* Y_2 *by*

$$Y_1 = \frac{1}{2}\left(X_1^\star + X_2^\star\right) \quad \left[= \frac{1}{2}\left(X_1 + X_2\right)\right],$$

$$Y_2 = \frac{1}{2}\left(X_1^\star - X_2^\star\right)^2 \quad \left[= \frac{1}{2}\left(X_2 - X_1\right)^2\right].$$

2. *Prove that the vector function* τ*, given by*

$$\tau(x_1, x_2) = \left(\frac{1}{2}\left(x_1 + x_2\right), \frac{1}{2}\left(x_2 - x_1\right)^2\right),$$

maps

$$D = \left\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 < x_2\right\}$$

bijectively onto

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

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.

5. *Are* Y_1 *and* Y_2 *independent?*

6. *Find the means and variances of* Y_1 *and* Y_2.

1. If $x_1 < x_2$, then

$$P\left\{X_1^\star > x_1 \wedge X_2^\star \leq x_2\right\} = P\left\{x_1 < X_1 \leq x_2 \wedge x_1 < X_2 \leq x_2\right\} = \left\{F(x_2) - F(x_1)\right\}^2,$$

where F is the distribution function of an $N\left(0,\sigma^2\right)$-distribution. This implies that

$$\begin{aligned} H^\star\left(x_1,x_2\right) &= P\left\{X_1^\star \le x_1 \,\wedge\, X_2^\star \le x_2\right\} = P\left\{X_2^\star \le x_2\right\} - P\left\{X_1^\star > x_1 \,\wedge\, X_2^\star \le x_2\right\} \\ &= \left\{F\left(x_2\right)\right\}^2 - \left\{F\left(x_2\right) - F\left(x_1\right)\right\}^2 = -F\left(x_1\right)^2 + 2F\left(x_1\right)F\left(x_2\right). \end{aligned}$$

Then by differentiation,

$$h^\star\left(x_1,x_2\right) = \frac{\partial^2}{\partial x_1 \partial x_2}\left\{-F\left(x_1\right)^2 + 2F\left(x_1\right)\cdot F\left(x_2\right)\right\} = 2h\left(x_1,x_2\right).$$

If $x_2 \le x_1$, then of course $h^\star\left(x_1,x_2\right) = 0$.

2. Clearly, $Y_2 \ge 0$. If $y_1 \in \mathbb{R}$ and $y_2 > 0$, then it follows from

$$y_1 = \frac{1}{2}\left(x_1+x_2\right) \quad \text{and} \quad y_2 = \frac{1}{2}\left(x_2-x_1\right)^2 \quad \text{and} \quad x_2 > x_1,$$

that

$$y_1 = \frac{1}{2}\left(x_1+x_2\right) \quad \text{and} \quad \sqrt{\frac{y_2}{2}} = \frac{1}{2}\left(x_2-x_1\right) \quad [\ge 0],$$

hence

$$x_1 = y_1 - \sqrt{\frac{y_2}{2}} \quad \text{and} \quad x_2 = y_1 + \sqrt{\frac{y_2}{2}}.$$

The solution is unique, and since the equation can always be solved, we find that

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

The Jacobian is

$$\frac{\partial\left(x_1,x_2\right)}{\partial\left(y_1,y_2\right)} = \begin{vmatrix} 1 & -\dfrac{1}{2}\,\dfrac{1}{\sqrt{2y_2}} \\ & \\ 1 & \dfrac{1}{2}\,\dfrac{1}{\sqrt{2y_2}} \end{vmatrix} = \frac{1}{\sqrt{2y_2}} > 0.$$

3. The simultaneous frequency of $\left(Y_1,Y_2\right)$ for $y_2 > 0$ is given by

$$\begin{aligned} k\left(y_1,y_2\right) &= h^\star\left(y_1 - \sqrt{\frac{y_2}{2}}\,,\, y_1 + \sqrt{\frac{y_2}{2}}\right)\cdot\frac{1}{\sqrt{2y_2}} \\ &= 2h\left(y_1 - \sqrt{\frac{y_2}{2}}\,,\, y_1 + \sqrt{\frac{y_2}{2}}\right)\cdot\frac{1}{\sqrt{2y_2}} \\ &= \frac{2}{2\pi\sigma^2}\exp\left(-\frac{1}{2\sigma^2}\left(2y_1^2+y_2\right)\right)\cdot\frac{1}{\sqrt{2y_2}} \\ &= \frac{1}{\sqrt{\pi\sigma^2}}\exp\left(-\frac{y_1}{2\sigma^2}\right)\cdot\frac{1}{\sqrt{\pi}}\cdot\frac{1}{\sqrt{2\sigma^2}}\cdot\frac{1}{\sqrt{y_2}}\cdot\exp\left(-\frac{y_2}{2\sigma^2}\right). \end{aligned}$$

If $y_2 \le 0$, then $k\left(y_1,y_2\right) = 0$.

4. and 5. It follows immediately from **3.** that Y_1 and Y_2 are independent and that

$$k_{Y_1}(y_1) = \frac{1}{\sqrt{\pi\sigma^2}} \cdot \exp\left(-\frac{1}{2\sigma^2}\, y_1^2\right), \qquad y_1 \in \mathbb{R}, \quad Y_1 \in N\left(0\,,\,\frac{\sigma^2}{2}\right),$$

and

$$k_{Y_2}(y_2) = \begin{cases} \dfrac{1}{\sqrt{\pi}} \cdot \dfrac{1}{\sqrt{2\sigma^2}} \cdot \dfrac{1}{\sqrt{y_2}}\, \exp\left(-\dfrac{y_2}{2\sigma^2}\right), & y_2 > 0, \\ 0, & y_2 \le 0, \end{cases} \qquad Y_2 \in \Gamma\left(\frac{1}{2}\,,\,2\sigma^2\right).$$

6. Then obviously,

$$E\{Y_1\} = 0, \qquad V\{Y_1\} = \frac{\sigma^2}{2},$$

$$E\{Y_2\} = \sigma^2, \qquad V\{Y_2\} = \frac{1}{2} \cdot \left(2\sigma^2\right) = 2\sigma^4.$$

Remark 7.1 We shall in the following Example 7.2 treat this problem in an alternative way. ◊

Example 7.2 *Let X_1 and X_2 be independent $N\left(0,\sigma^2\right)$-distributed random variables, and let the random variables Y_1 and Y_2 be given by*

$$Y_1 = \frac{1}{2}\left(X_1 + X_2\right), \qquad Y_2 = \frac{1}{2}\left(X_2 - X_1\right)^2.$$

1) *Prove that the random variables $X_1 + X_2$ and $X_2 - X_1$ are independent, e.g. by first finding the simultaneous frequency.*

2) *Prove that the random variables Y_1 and Y_2 are independent.*

3) *Find the frequencies of Y_1 and Y_2.*

4) *Find the simultaneous frequency of (Y_1, Y_2).*

5) *Find the means and variances of Y_1 and Y_2.*

1) From $z_1 = x_1 + x_2$ and $z_2 = x_2 - x_1$ follows that

$$x_1 = \frac{1}{2}\left(z_1 - z_2\right) \quad \text{and} \quad x_2 = \frac{1}{2}\left(z_1 + z_2\right)$$

of the Jacobian

$$\frac{\partial(x_1, x_2)}{\partial(z_1, z_2)} = \begin{vmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{vmatrix} = \frac{1}{2}.$$

Since

$$x_1^2 + x_2^2 = \frac{1}{4}\left(z_1 - z_2\right)^2 + \frac{1}{4}\left(z_1 + z_2\right)^2 = \frac{1}{2}\left(z_1^2 + z_2^2\right),$$

the simultaneous frequency of $(Z_1, Z_2) = (X_1 + X_2, X_2 - X_1)$ is given by

$$\begin{aligned} f_Z(z_1, z_2) &= \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2} \cdot \left\{\left(\frac{z_1}{\sqrt{2}}\right)^2 + \left(\frac{z_2}{\sqrt{2}}\right)^2\right\}\right) \cdot \frac{1}{2} \\ &= \frac{1}{2\sqrt{\pi}\,\sigma} \exp\left(-\frac{1}{2}\left\{\frac{z_1}{\sqrt{2}\,\sigma}\right\}^2\right) \cdot \frac{1}{2\sqrt{\pi}\,\sigma} \exp\left(-\frac{1}{2}\left\{\frac{z_2}{\sqrt{2}\,\sigma}\right\}^2\right). \end{aligned}$$

It follows immediately that Z_1 and Z_2 are independent and that

$$f_{Z_i}(z_i) = \frac{1}{2\sqrt{\pi}\,\sigma} \exp\left(-\frac{1}{2}\left\{\frac{z_i}{\sqrt{2}\,\sigma}\right\}^2\right), \qquad z_i \in \mathbb{R}, \quad i = 1, 2.$$

2) Since $Y_1 = \frac{1}{2} Z_1$ and $Y_2 = \frac{1}{2} Z_2^2$, and Z_1 and Z_2 are independent, we conclude that Y_1 and Y_2 are also independent.

3) It follows from $y_1 = \frac{1}{2} z_1$, that $z_2 = 2y_1$, and $\frac{dz_1}{dy_1} = 2$, so

$$f_{Y_1}(y_1) = f_{Z_1}(2y_1) \cdot 2 = \frac{1}{\sqrt{\pi} \cdot \sigma} \exp\left(-\frac{1}{2}\left\{\frac{\sqrt{2}\,y_1}{\sigma}\right\}^2\right), \qquad y_1 \in \mathbb{R}.$$

From $y_2 = \dfrac{1}{2}\, z_2^2 \geq 0$ follows that $|z_2| = \sqrt{2y_2}$, thus

$$\begin{aligned} F_{Y_2}(y_2) &= P\left\{|Z_2| \leq \sqrt{2y_2}\right\} = P\left\{\left|\frac{Z_2}{\sqrt{2}\,\sigma}\right| \leq \frac{\sqrt{y_2}}{\sigma}\right\} \\ &= \Phi\left(\frac{\sqrt{y_2}}{\sigma}\right) - \Phi\left(-\frac{\sqrt{y_2}}{\sigma}\right) = 2\,\Phi\left(\frac{\sqrt{y_2}}{\sigma}\right) - 1. \end{aligned}$$

If $y_2 > 0$, then by differentiation,

$$f_{Y_2}(y_2) = 2 \cdot \varphi\left(\frac{\sqrt{y_2}}{\sigma}\right) \cdot \frac{1}{2\sqrt{y_2}\,\sigma} = \frac{1}{\sigma\sqrt{y_2}} \cdot \frac{1}{\sqrt{2\pi}}\,\exp\left(-\frac{y_2}{2\sigma^2}\right),$$

and $f_{Y_2}(y_2) = 0$ for $y_2 \leq 0$, i.e.

$$f_{Y_2}(y_2) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\,\sigma} \cdot \dfrac{1}{\sqrt{y_2}}\,\exp\left(-\dfrac{y_2}{2\sigma^2}\right) & \text{for } y_2 > 0, \\[2ex] 0 & \text{for } y_2 \leq 0. \end{cases}$$

4) The simultaneous frequency is given by

$$h(y_1, y_2) = f_{Y_1}(y_1) \cdot f_{Y_2}(y_2) = \frac{1}{\sqrt{2\pi}\,\sigma^2} \cdot \frac{1}{\sqrt{y_2}}\,\exp\left(-\frac{y_1^2}{\sigma^2} - \frac{y_2}{2\sigma^2}\right)$$

for $y_1 \in \mathbb{R}$ and $y_2 > 0$, and $h(y_1, y_2) = 0$ for $y_2 \leq 0$.

5) The mean and variance of Y_1 are

$$E\{Y_1\} = 0 \qquad \text{and} \qquad V\{Y_1\} = \frac{\sigma^2}{2}.$$

For Y_2 we get by the substitution $t^2 = \dfrac{y_2}{\sigma^2}$ that

$$E\{Y_2\} = \frac{1}{\sqrt{2\pi}\,\sigma}\int_0^\infty y_2\,\exp\left(-\frac{y_2}{2\sigma^2}\right)\frac{1}{\sqrt{y_2}}\,dy_2 = \sqrt{\frac{2}{\pi}}\int_0^\infty \sigma^2 t^2\exp\left(-\frac{1}{2}\,t^2\right)dt = \sigma^2,$$

and

$$E\left\{Y_2^2\right\} = \sqrt{\frac{2}{\pi}}\int_0^\infty \sigma^4 t^4\exp\left(-\frac{1}{2}\,t^2\right)dt = 3 \cdot 1 \cdot \sigma^4,$$

så

$$V\{Y_2\} = E\left\{Y_2^2\right\} - (E\{Y_2\})^2 = 3\sigma^4 - \sigma^4 = 2\sigma^4.$$

8 Convergence in distribution

Example 8.1 *Given a sequence of random variables $(U_n)_{n=1}^{\infty}$, which converges in distribution towards a random variable U of distribution function $F_U(u)$ and frequency $f_U(u)$.*

1. *Prove that the sequence $\left(U_n^2\right)$ converges in distribution towards U^2 of distribution function*

$$F_U\left(\sqrt{u}\right) - F_U\left(-\sqrt{u}\right)$$

and frequency

$$\frac{1}{2\sqrt{u}}\left\{f_U\left(\sqrt{u}\right) + f_U\left(-\sqrt{u}\right)\right\}, \qquad u > 0.$$

We perform a series of throws with an (honest) coin, where we assume that the throws are mutually independent. Define the random variables X_1, X_2, ... by

$$X_i = \begin{cases} 1, & \text{if the } i\text{-th throw results in } i \text{ heads,} \\ 0, & \text{if the } i\text{-th throw results in } i \text{ tails,} \end{cases}$$

and the random variables K_1, K_2, ... by

$$K_n = \text{ number of heads in the first } n \text{ throws.}$$

2. *Express K_n by means of the X_i and find mean and variance of K_n.*

3. *Prove that the sequence $\left(\frac{2}{\sqrt{n}}\left\{K_n - \frac{n}{2}\right\}\right)$ converges in distribution towards a normally distributed random variable.*

4. *Define the random variables Y_n by*

$$Y_n = \frac{4}{n}\left(K_n - \frac{n}{2}\right)^2, \qquad n \in \mathbb{N}.$$

Prove that the sequence (Y_n) converges in distribution towards a random variable Y, and find the distribution of Y.

1) It is given that

$$F_{U_n}(u) \to F_U(u) \qquad \text{for } n \to \infty,$$

and that for $u > 0$,

$$F_{U^2}(u) = P\left\{U^2 \le u\right\} = P\left\{-\sqrt{u} \le U \le \sqrt{u}\right\} = F\left(\sqrt{u}\right) - F\left(-\sqrt{u}\right).$$

If $u > 0$, then

$$\begin{aligned} F_{U_n^2}(u) &= P\left\{U_n^2 \le u\right\} = P\left\{-\sqrt{u} \le U_n \le \sqrt{u}\right\} \\ &= F_{U_n}\left(\sqrt{u}\right) - F_{U_n}\left(-\sqrt{u}\right) \to F_U\left(\sqrt{u}\right) - F_U\left(-\sqrt{u}\right) \\ &= F_{U^2}(u) \qquad \text{for } n \to \infty, \end{aligned}$$

so $\left(U_n^2\right)$ converges in distribution towards U^2.

The frequency is found by differentiation,

$$f_{U^2}(u) = \frac{d}{du}\left\{F_U\left(\sqrt{u}\right) - F_U\left(-\sqrt{u}\right)\right\} = \frac{1}{2\sqrt{u}}\left\{f_U\left(\sqrt{u}\right) + f_U\left(-\sqrt{u}\right)\right\}.$$

2) Clearly, $K_n = \sum_{i=1}^n X_i$, so

$$E\{K_n\} = \sum_{i=1}^n E\{X_i\} = n \cdot \frac{1}{2} = \frac{n}{2} = \mu_n.$$

Since the X_i are mutually independent, we get

$$V\{K_n\} = \sum_{i=1}^n V\{X_i\} = n \cdot \left\{\frac{1}{2} - \left(\frac{1}{2}\right)^2\right\} = \frac{n}{4} = s_n^2.$$

3) It follows from the Central Limit theorem that

$$P\left\{\frac{2}{\sqrt{n}}\left(K_n - \frac{n}{2}\right) \le x\right\} = P\left\{\frac{K_n - \mu_n}{s_n} \le x\right\} \to \Phi(x) \qquad \text{for } n \to \infty.$$

4) According to **1.** and **3.** the sequence (Y_n) converges for $x > 0$ in distribution towards a random variable Y of the distribution function

$$\frac{1}{2\sqrt{x}}\{\Phi(x)-\Phi(-x)\} = \frac{1}{2\sqrt{x}}\{2\,\Phi(x)-1\} = \frac{1}{\sqrt{x}}\,\Phi(x) - \frac{1}{2\sqrt{x}},$$

and $= 0$ for $x \leq 0$.
The corresponding frequency for $x > 0$ is given by

$$\frac{1}{2\sqrt{x}}\cdot\left\{\varphi\left(\sqrt{x}\right)+\varphi\left(-\sqrt{x}\right)\right\} = \frac{\varphi\left(\sqrt{x}\right)}{\sqrt{x}},$$

and $= 0$ for $x \leq 0$.

Example 8.2 *Given a series of throws with an (honest) coin, where we assume that the throws are mutually independent. Define the random variables X_1, X_2, … and K_n and P_n by*

$$\begin{aligned} X_i &= \begin{cases} 1, & \textit{if the } i\textit{-th throw results in a head,} \\ \\ 0, & \textit{if the } i\textit{-th throw results in a tail,} \end{cases} \\ K_n &= \textit{number of heads in the first } n \textit{ throws,} \\ P_n &= \textit{number of tails in the first } n \textit{ throws.} \end{aligned}$$

1. *Prove that K_n and P_n both have the mean $\frac{n}{2}$.*

Define Z_n by

$$Z_n = \frac{2}{n}\left\{\left(K_n - \frac{n}{2}\right)^2 + \left(P_n - \frac{n}{2}\right)^2\right\}, \qquad n \in \mathbb{N}.$$

2. *Express Z_n by the X_i.*

3. *Find the mean of Z_n.*

4. *Prove that the sequence (Z_n) converges in distribution towards a random variable Z, and find the distribution of Z.*

1) We put

$$Y_i = 1 - X_i = \begin{cases} 0, & \text{if the } i\text{-th throw results in a head,} \\ \\ 1, & \text{if the } i\text{-th throw results in a tail.} \end{cases}$$

Then

$$K_n = \sum_{i=1}^n X_i \qquad \text{and} \qquad P_n = \sum_{i=1}^n Y_i = n - \sum_{i=1}^n X_i,$$

hence

$$E\{K_n\} = \sum_{i=1}^n E\{X_i\} = \frac{n}{2} \quad \text{and} \quad E\{P_n\} = n - \sum_{i=1}^n E\{X_i\} = \frac{n}{2}.$$

2) It follows from the above that

$$
\begin{aligned}
Z_n &= \frac{2}{n}\left\{\left(K_n-\frac{n}{2}\right)^2+\left(P_n-\frac{n}{2}\right)^2\right\}=\frac{2}{n}\left\{\left(\sum_{i=1}^n X_i-\frac{n}{2}\right)^2+\left(n-\sum_{i=1}^n X_i-\frac{n}{2}\right)^2\right\}\\
&= \frac{4}{n}\left(\sum_{i=1}^n\left\{X_i-\frac{1}{2}\right\}\right)^2=\frac{4}{n}\sum_{i=1}\left(X_i-\frac{1}{2}\right)^2+\frac{2}{n}\sum_{1\le i<j\le n}\left(X_i-\frac{1}{2}\right)\left(X_j-\frac{1}{2}\right).
\end{aligned}
$$

3) Now, $E\{X_i\}=\frac{1}{2}$, and the X_i are mutually independent. Hence

$$
\begin{aligned}
E\{Z_n\} &= \frac{4}{n}E\left\{\sum_{i=1}^n\left(X_i-\frac{1}{2}\right)^2\right\}+\frac{8}{n}\sum_{1\le i<j\le n}E\left\{X_i-\frac{1}{2}\right\}\cdot E\left\{X_j-\frac{1}{2}\right\}\\
&= \frac{4}{n}\sum_{i=1}^n E\left\{\left(X_i-\frac{1}{2}\right)^2\right\}+0=\frac{4}{n}\sum_{i=1}^n V\{X_i\}=\frac{4}{n}\cdot n\cdot\frac{1}{4}=1.
\end{aligned}
$$

4) Since

$$
V\left\{\sum_{i=1}^n\left(X_i-\frac{1}{2}\right)\right\}=\sum_{i=1}^n V\{X_i\}=\frac{n}{4}=s_n^2,
$$

we get for $x>0$ be the Central Limit Theorem that

$$
\begin{aligned}
P\{Z_n\le x\} &= P\left\{\frac{4}{n}\left(\sum_{i=1}^n\left\{X_i-\frac{1}{2}\right\}\right)^2\le x\right\}=P\left\{-\sqrt{x}\le\frac{\sum_{i=1}^n\left(X_i-\frac{1}{2}\right)}{\frac{\sqrt{n}}{2}}\le\sqrt{x}\right\}\\
&\to \Phi\left(\sqrt{x}\right)-\Phi\left(-\sqrt{x}\right)=2\,\Phi\left(\sqrt{x}\right)-1\qquad\text{for } n\to\infty.
\end{aligned}
$$

When instead $x\le 0$, then

$$
P\{Z_n\le x\}=0\to 0\qquad\text{for } n\to\infty.
$$

Consequently, (Z_n) converges in distribution towards a random variable Z, where

$$
F_Z(z)=\begin{cases}2\,\Phi\left(\sqrt{z}\right)-1 & \text{for } z>0,\\[2mm] 0 & \text{for } z\le 0.\end{cases}
$$

If $z>0$, then the corresponding frequency is given by

$$
f_Z(z)=F_Z'(z)=\frac{1}{\sqrt{z}}\,\varphi\left(\sqrt{z}\right)=\frac{1}{\sqrt{2\pi}}\,\frac{1}{\sqrt{z}}\,\exp\left(-\frac{z}{2}\right),
$$

and Z is Gamma distributed of form parameter $\mu=\frac{1}{2}$ and scale parameter $\alpha=2$.

9 The χ^2 distribution

Example 9.1 *Let $X \in \chi^2(n)$. Find the frequency of $\sqrt{X}$, and find the mean and variance of $\sqrt{X}$.*

From $X \in \chi^2(n) = \Gamma\left(\dfrac{n}{2}, 2\right)$, follows that X has the frequency

$$f_X(x) = \begin{cases} \dfrac{1}{\Gamma\left(\frac{n}{2}\right)} \cdot \dfrac{1}{2^{n/2}} \cdot x^{\frac{n}{2}-1}\, e^{-\frac{x}{2}} & \text{for } x > 0, \\[2ex] 0 & \text{for } x \le 0. \end{cases}$$

The map $y = \sqrt{x}$ is bijective $\mathbb{R}_+ \to \mathbb{R}$ with the inverse $x = y^2$, so if $Y = \sqrt{X}$, then the frequency $f_Y(y)$ for Y, when $y > 0$, is given by

$$f_Y(y) = f_X(x(y)) \cdot \left|\frac{dx(y)}{dy}\right| = \frac{1}{\Gamma\left(\frac{n}{2}\right)} \cdot \frac{1}{2^{n/2}}\, e^{-\frac{y^2}{2}} \cdot 2y = \frac{1}{\Gamma\left(\dfrac{n}{2}\right) 2^{\frac{n}{2}-1}}\, y^{n-1} \exp\left(-\frac{y^2}{2}\right),$$

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

The mean is given by

$$\begin{aligned} E\left\{\sqrt{X}\right\} &= \int_0^\infty \sqrt{x}\, f_X(x)\, dx = \frac{1}{\Gamma\left(\dfrac{n}{2}\right) \cdot 2^{\frac{n}{2}}} \int_0^\infty x^{\frac{n}{2}-1+\frac{1}{2}} e^{-\frac{x}{2}}\, dx \\ &= \frac{\Gamma\left(\frac{n+1}{2}\right) 2^{\frac{n+1}{2}}}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} \int_0^\infty \frac{1}{\Gamma\left(\frac{n+1}{2}\right) 2^{\frac{n+1}{2}}}\, x^{\frac{n+1}{2}-1} e^{-\frac{x}{2}}\, dx = \sqrt{2} \cdot \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)}. \end{aligned}$$

Using

$$E\left\{\left(\sqrt{X}\right)^2\right\} = E\{X\} = n,$$

we get

$$V\left\{\sqrt{X}\right\} = E\{X\} - \left(E\left\{\sqrt{X}\right\}\right)^2 = n - 2\left\{\frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)}\right\}^2.$$

10 The F distribution

Example 10.1 *Let* $X \in F(n_1, n_2)$. *Prove that* $\dfrac{1}{X} \in F(n_2, n_1)$.

If $X \in F(n_1, n_2)$, then $f_X(x) = 0$ for $x \le 0$, and if $x > 0$, then the frequency is

$$f_X(x) = \frac{\left(\dfrac{n_1}{n_2}\right)^{\frac{1}{2}n_1}}{B\left(\dfrac{n_1}{2}, \dfrac{n_2}{2}\right)} \cdot \frac{x^{\frac{1}{2}n_1 - 1}}{\left(1 + \dfrac{n_1}{n_2}\,x\right)^{\frac{1}{2}(n_1+n_2)}}.$$

The map $w = \dfrac{1}{x}$, $x = \dfrac{1}{w}$, is bijective, $\mathbb{R}_+ \to \mathbb{R}_+$, with $\dfrac{dx}{dw} = -\dfrac{1}{w^2}$. If we put $W = \dfrac{1}{X}$, then $f_W(w) = 0$ for $w \le 0$. If $w > 0$, then

$$\begin{aligned} f_W(w) &= \frac{\left(\dfrac{n_1}{n_2}\right)^{\frac{1}{2}n_1}}{B\left(\dfrac{n_1}{2}, \dfrac{n_2}{2}\right)} \cdot \frac{\left(\dfrac{1}{w}\right)^{\frac{1}{2}n_1 - 1}}{\left(1 + \dfrac{n_1}{n_2} \cdot \dfrac{1}{w}\right)^{\frac{1}{2}(n_1+n_2)}} \cdot \frac{1}{w^2} \\ &= \frac{\left(\dfrac{n_1}{n_2}\right)^{\frac{1}{2}n_1}}{B\left(\dfrac{n_1}{2}, \dfrac{n_1}{2}\right)} \cdot \frac{w^{\frac{1}{2}(n_1+n_2)}}{\left(\dfrac{n_1}{n_2}\right)^{\frac{1}{2}n_1 + \frac{1}{2}n_2}} \cdot \frac{w^{-\frac{1}{2}n_1 + 1 - 2}}{\left(1 + \dfrac{n_2}{n_1}\,w\right)^{\frac{1}{2}(n_1+n_2)}} \\ &= \frac{\left(\dfrac{n_2}{n_1}\right)^{\frac{1}{2}n_2}}{B\left(\dfrac{n_2}{2}, \dfrac{n_1}{2}\right)} \cdot \frac{w^{\frac{1}{2}n_2 - 1}}{\left(1 + \dfrac{n_2}{n_1}\,w\right)^{\frac{1}{2}(n_1+n_2)}}, \end{aligned}$$

and we conclude that $W = \dfrac{1}{X} \in F(n_2, n_1)$.

Alternatively, an $F(n_1, n_2)$-distribution occurs as the distribution of

$$X = \frac{\dfrac{1}{n_1}\,X_1}{\dfrac{1}{n_2}\,X_2}, \qquad \text{where } X_1 \in \chi^2(n_1) \text{ and } X_2 \in \chi^2(n_2),$$

and where X_1, X_2 are independent. Then we immediately get

$$\frac{1}{X} = \frac{\dfrac{1}{n_2}\,X_2}{\dfrac{1}{n_1}\,X_1} \in F(n_2, n_1).$$

Example 10.2 *Let $X \in F(n_1, n_2)$. Prove that $E\{X\} = \dfrac{n_2}{n_2 - 2}$ for $n_2 > 2$.*

Let $n_2 > 2$. Introduce the change of variable

$$x = \frac{n_2}{n_1}\left(\frac{1}{y} - 1\right), \qquad y \in]0, 1[,$$

i.e.

$$y = \frac{1}{1 + \dfrac{n_1}{n_2}\,x} \quad \text{og} \quad 1 - y = \frac{\dfrac{n_1}{n_2}\,x}{1 + \dfrac{n_1}{n_2}\,x}.$$

We get

$$
\begin{aligned}
E\{X\} &= \int_0^{\infty} \frac{\left(\dfrac{n_1}{n_2}\right)^{\frac{1}{2}\,n_1}}{B\left(\dfrac{n_1}{2},\dfrac{n_2}{2}\right)\cdot \dfrac{x^{\frac{1}{2}\,n_1}}{\left(1+\dfrac{n_1}{n_2}\,x\right)^{\frac{1}{2}\,(n_1+n_2)}}}\,dx \\
&= \frac{1}{B\left(\dfrac{n_1}{2},\dfrac{n_2}{2}\right)}\int_0^{\infty}\left\{\frac{\dfrac{n_1}{n_2}\,x}{1+\dfrac{n_1}{n_2}\,x}\right\}^{\frac{1}{2}\,n_1}\cdot\left\{\frac{1}{1+\dfrac{n_1}{n_2}\,x}\right\}^{\frac{1}{2}\,n_2}dx \\
&= \frac{1}{B\left(\dfrac{n_1}{2},\dfrac{n_2}{2}\right)}\int_0^1 (1-y)^{\frac{1}{2}\,n_1}y^{\frac{1}{2}\,n_2}\cdot\frac{n_2}{n_1}\cdot\frac{1}{y^2}\,dy \\
&= \frac{1}{B\left(\dfrac{n_1}{2},\dfrac{n_2}{2}\right)}\cdot\frac{n_2}{n_1}\int_0^1 y^{\frac{1}{2}(n_2-2)-1}(1-y)^{\frac{1}{2}(n_1+2)-1}\,dy \\
&= \frac{n_2}{n_1}\cdot\frac{B\left(\dfrac{n_1}{2}+1,\dfrac{n_1}{2}-1\right)}{B\left(\dfrac{n_1}{2},\dfrac{n_2}{2}\right)} = \frac{n_2}{n_1}\cdot\frac{\Gamma\left(\dfrac{n_1}{2}+1\right)\Gamma\left(\dfrac{n_2}{2}-1\right)}{\Gamma\left(\dfrac{n_1}{2}+1+\dfrac{n_2}{2}-1\right)}\cdot\frac{\Gamma\left(\dfrac{n_1}{2}+\dfrac{n_2}{2}\right)}{\Gamma\left(\dfrac{n_1}{2}\right)\Gamma\left(\dfrac{n_2}{2}\right)} \\
&= \frac{n_2}{n_1}\cdot\frac{\dfrac{n_1}{2}\,\Gamma\left(\dfrac{n_1}{2}\right)}{\Gamma\left(\dfrac{n_1}{2}\right)}\cdot\frac{\Gamma\left(\dfrac{n_2}{2}-1\right)}{\left(\dfrac{n_2}{2}-1\right)\Gamma\left(\dfrac{n_2}{2}-1\right)} = \frac{n_2}{n_1}\cdot\frac{n_1}{2}\cdot\frac{1}{\dfrac{n_2}{2}-1} = \frac{n_2}{n_2-2}.
\end{aligned}
$$

Alternatively,

$$X = \frac{n_2}{n_1}\,\frac{X_1}{X_2}, \qquad \text{where } X_1 \in \chi^2(n_1) \text{ and } X_2 \in \chi^2(n_2),$$

and where X_1 and X_2 are independent. Then for $n_2 > 2$,

$$
\begin{aligned}
E\{X\} &= \frac{n_2}{n_1}\cdot E\{X_1\}\cdot E\left\{\frac{1}{X_2}\right\} = n_2\cdot E\left\{\frac{1}{X_2}\right\} \\
&= n_2\int_0^{\infty}\frac{1}{x}\cdot x^{(n_2/2)-1}e^{-(x/2)}dx\cdot\frac{1}{\Gamma\left(\dfrac{n_2}{2}\right)2^{n_2/2}} \\
&= n_2\int_0^{\infty}x^{(n_2/2)-2}e^{-(x/2)}dx\cdot\frac{1}{\Gamma\left(\dfrac{n_2}{2}\right)2^{n_2/2}} \qquad (x=2y) \\
&= n_2\int_0^{\infty}y^{(n_2/2)-2}e^{y}dy\cdot\frac{2^{(n_2/2)-1}}{\Gamma\left(\dfrac{n_2}{2}\right)\cdot 2^{n_2/2}} \\
&= n_2\Gamma\left(\frac{n_2}{2}-1\right)\cdot\frac{1}{\Gamma\left(\dfrac{n_2}{2}\right)\cdot 2} = n_2\cdot\frac{1}{\left(\dfrac{n_2}{2}-1\right)2} = \frac{n_2}{n_2-2}.
\end{aligned}
$$

It follows form the computations above that $E\{X\}$ does not exist for $n_2 \le 2$.

11 The F distribution and the t distribution

Example 11.1 *Let $X \in t(n)$. Prove that $X^2 \in F(1,n)$. Prove that the mean $E\{X\}$ exists, if and only if $n > 1$, and find $E\{X\}$ for $n > 1$.*
Prove that the variance $V\{X\}$ exists, if and only if $n > 2$, and find $V\{X\}$ for $n > 2$.

The random variable $X \in t(n)$ has the frequency

$$f(x) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{n\,\pi}\,\Gamma\left(\dfrac{n}{2}\right)} \cdot \left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \qquad x \in \mathbb{R}.$$

Let $Y = X^2$ have the frequency $g(y)$. Then $g(y) = 0$ for $y \le 0$. For $y > 0$ we obtain the distribution function

$$\begin{aligned} G(y) &= P\{Y < leqy\} = P\left\{X^2 \le y\right\} = P\left\{-\sqrt{y} < leqX \le \sqrt{y}\right\} \\ &= P\left\{X \le \sqrt{y}\right\} - P\left\{X \le -\sqrt{y}\right\} = F\left(\sqrt{y}\right) - F\left(-\sqrt{y}\right), \end{aligned}$$

hence

$$\begin{aligned} g(y) &= G'(y) = f\left(\sqrt{y}\right) \cdot \frac{1}{2\sqrt{y}} - f\left(-\sqrt{y}\right) \cdot \frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{y}}\, f\left(\sqrt{y}\right) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{n\,\pi}\,\Gamma\left(\dfrac{n}{2}\right)} \cdot \frac{1}{\left(1+\dfrac{y}{n}\right)^{\frac{n+1}{2}}} \cdot \frac{1}{\sqrt{y}} \\ &= \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{1}{2}\right)\Gamma\left(\dfrac{n}{2}\right)} \left(\frac{1}{n}\right)^{\frac{1}{2}-1} \cdot \frac{y^{\frac{1}{2}\cdot 1-1}}{\left(1+\dfrac{1}{n}\,y\right)^{\frac{1}{2}\,(1+n)}} \in B(1,n), \end{aligned}$$

as required.

If $n = 1$, then

$$\int_{-\infty}^{\infty} \frac{\Gamma(1)}{\sqrt{\pi}\,\Gamma\left(\dfrac{1}{2}\right)}\, \frac{|x|}{1+x^2}\, dx = \infty,$$

and the mean does not exist.

If $n > 1$, then

$$\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}} \cdot |x| \sim |x|^{-n},$$

and $\int |x|^{-n}\, dx$ is convergent. The integrand $\left(1+\dfrac{x^2}{n}\right)^{-\frac{n+1}{2}} x$ is an odd function, hence the mean is

$$E\{X\} = 0.$$

12 Estimation of parameters

Example 12.1 *A random variable X has its distribution function $F(x)$ given by*

$$F(x)=\begin{cases} 1-\left(\dfrac{2}{x}\right)^{\alpha} & \text{for } x \geq 2,\\[2mm] 0 & \text{for } x<2,\end{cases}$$

where α is a positive constant.

1) Find all values of α, for which $E\{X\}$ exists, and find $E\{X\}$ in these cases.

2) Find all values of α, for which $V\{X\}$ exists and find $V\{X\}$ in these cases.

3) Given n observations $x_1,\, x_2,\, \dots,\, x_n$ of the mutually independent random variables $X_1,\, X_2,\, \dots,\, X_n$, all distributed as X above. Use the maximum method to find the best estimate of the parameter α.

The distribution is called a Pareto distribution.

1) The corresponding frequency is found by differentiation,

$$f(x)=\begin{cases} \dfrac{\alpha\cdot 2^{\alpha}}{x^{\alpha+1}} & \text{for } x\geq 2,\\[2mm] 0 & \text{for } x<2.\end{cases}$$

Since $x\cdot f(x)\sim x^{-\alpha}$, the mean exists, if and only if $\alpha>1$. When this is the case, then

$$E\{X\}=\int_2^{\infty} x\cdot\alpha\cdot\frac{2^{\alpha}}{x^{\alpha+1}}\,dx=\frac{2\alpha}{\alpha-1}\int_2^{\infty}(\alpha-1)\cdot\frac{2^{\alpha-1}}{x^{(\alpha-1)+1}}\,dx=\frac{2\alpha}{\alpha-1}.$$

2) Analogously, $V\{X\}$ exists if and only if $\alpha>2$. If so, we first compute

$$E\left\{X^2\right\}=\int_2^{\infty} x^2\alpha\cdot\frac{2^{\alpha}}{x^{\alpha-1}}\,dx=\frac{4\alpha}{\alpha-2}\int_2^{\infty}(\alpha-2)\cdot\frac{2^{\alpha-2}}{x^{(\alpha-2)+1}}\,dx=\frac{4\alpha}{\alpha-2},$$

hence

$$\begin{aligned} V\{X\} &= E\left\{X^2\right\}-(E\{X\})^2=\frac{4\alpha}{\alpha-2}-\frac{4\alpha^2}{(\alpha-1)^2}\\ &= \frac{4\alpha}{(\alpha-2)(\alpha-1)^2}\left\{(\alpha-1)^2-\alpha(\alpha-2)\right\}=\frac{4\alpha}{(\alpha-1)^2(\alpha-2)}.\end{aligned}$$

3) We shall find the maximum in α of the function

$$h(\alpha)=\prod_{i=1}^{n}\alpha\cdot\frac{2^{\alpha}}{x_i^{\alpha+1}},\qquad \alpha>0,\quad x_i>2,$$

i.e. for

$$\begin{aligned} h_1(\alpha) &= \ln h(\alpha) = \sum_{i=1}^{n} \{\ln \alpha + \alpha \ln 2 - (\alpha+1) \ln x_i\} \\ &= n \ln \alpha + \alpha \cdot n \ln 2 - (\alpha+1) \sum_{i=1}^{n} \ln x_i . \end{aligned}$$

Since

$$h_1'(\alpha) = \frac{n}{\alpha} + n \ln 2 - \sum_{i=1}^{n} \ln x_i = 0$$

for

$$\frac{1}{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \ln x_i - \ln 2 = \frac{1}{n} \sum_{i=1}^{n} \ln \left(\frac{x_i}{2} \right),$$

the estimate of α is given by

$$\alpha^\star = \frac{1}{\sum_{i=1}^{n} \ln \left(\frac{x_i}{2} \right)} = \frac{n}{\ln \prod_{i=1}^{n} \left(\frac{x_i}{2} \right)},$$

where we check that $h(\alpha) \to 0$ for $\alpha \to 0+$ and for $\alpha \to \infty$, if $x_i > 2$.

Example 12.2 *1) Let $a > 0$ be a positive constant. Prove that the function*

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

can be considered as a frequency of a random variable X.

2) Find mean and variance of the random variable X.

3) Assuming that n observations of the mutually independent random variables $X_1, X_2, \dots, X_n$, all distributed like X above, have given the results $x_1, x_2, \dots, x_n$, we shall apply the maximum method to find the best estimate of a and thus the best estimate of $V\{X\}$.

4) Prove that the estimator corresponding to the estimate of $V\{X\}$ is a central estimator, and then prove that the corresponding sequence of estimators is a consistent sequence.

1) Since $a > 0$, we have $f(x) \ge 0$, and

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^{\infty} \frac{x}{a}\exp\left(-\frac{x^2}{2a}\right)dx = \left[-\exp\left(-\frac{x^2}{2a}\right)\right]_0^{\infty} = 1,$$

hence $f(x)$ is the frequency of a random variable X.

2) The mean is

$$\begin{aligned} E\{X\} &= \int_0^{\infty} x\cdot\frac{x}{a}\exp\left(-\frac{x^2}{2a}\right)dx = \left[-x\exp\left(-\frac{x^2}{2a}\right)\right]_0^{\infty} + \int_0^{\infty}\exp\left(-\frac{s^2}{2a}\right)dx \\ &= 0 + \sqrt{2\pi\,a}\cdot\frac{1}{\sqrt{2\pi\,a}}\int_{-\infty}^{0}\exp\left(-\frac{x^2}{2a}\right)dx = \sqrt{2\pi\,a}\cdot\Phi(0) = \sqrt{\frac{\pi\,a}{2}}. \end{aligned}$$

Using that

$$E\left\{X^2\right\} = \int_0^{\infty} x^2\cdot\frac{x}{a}\exp\left(-\frac{x^2}{2a}\right)dx = 2a\int_0^{\infty} y\,e^{-y}\,dy = 2a,$$

we get

$$V\{X\} = E\left\{X^2\right\} - (E\{X\})^2 = 2a - \frac{\pi\,a}{2} = \left(2-\frac{\pi}{2}\right)a.$$

3) We shall find the maximum of the function (in a)

$$h(a) = \prod_{i=1}^{n}\frac{x_i}{a}\exp\left(-\frac{x_i^2}{2a}\right) = \frac{1}{a^n}\cdot\prod_{i=1}^{n}x_i\cdot\exp\left(-\frac{1}{2a}\sum_{i=1}^{n}x_i^2\right), \qquad a > 0.$$

Since $x_i > 0$, we have $h(a) \to 0$ for $a \to 0+$ and for $a \to \infty$, so we shall find the maximum for $a > 0$ of the auxiliary function

$$h_1(a) = \ln h(a) = -n\,\ln a + \sum_{i=1}^{n}\ln x_i - \frac{1}{2a}\sum_{i=1}^{n}x_i^2.$$

Now,

$$h_1'(a) = -\frac{n}{1} + \frac{1}{2a^2} \sum_{i=1}^{n} x_i^2 = \frac{n}{a^2} \left\{ -a + \frac{1}{2n} \sum_{i=1}^{n} x_i^2 \right\}.$$

this maximum is attained for the estimate

$$a^{\star} = \frac{1}{2n} \sum_{i=1}^{n} x_i^2.$$

The best estimate of $V\{X\}$ is

$$\left(2 - \frac{\pi}{2}\right) a = \left(2 - \frac{\pi}{2}\right) \cdot \frac{1}{2n} \sum_{i=1}^{n} x_i^2 = \frac{4-\pi}{4n} \sum_{i=1}^{n} x_i^2,$$

corresponding to the estimator

$$Y_n = \left(1 - \frac{\pi}{4}\right) \frac{1}{n} \sum_{i=1}^{n} X_i^2$$

for $V\{X\}$.

4) We shall finally prove that $E\{Y_n\} = \left(2 - \frac{\pi}{2}\right) a$. We get

$$E\{Y_n\} = \left(1 - \frac{\pi}{4}\right) \frac{1}{n} \sum_{i=1}^{n} E\left\{C_i^2\right\} = \left(1 - \frac{\pi}{4}\right) E\left\{X^2\right\} = \left(1 - \frac{\pi}{4}\right) \cdot 2a = \left(2 - \frac{\pi}{2}\right) a,$$

from which we conclude that Y_n is a central estimator.

Since

$$V\{Y_n\} = \left(1 - \frac{\pi}{4}\right)^2 \cdot \frac{1}{n} V\left\{X^2\right\} = \frac{c}{n} = s_n^2,$$

where $c > 0$ is some constant, which we do not need to find, it follows by Chebyshevs inequality that

$$P\left\{ \left| Y_n - \left(2 - \frac{\pi}{2}\right) a \right| \geq \varepsilon \right\} \leq \frac{s_n^2}{\varepsilon^2} = \frac{c}{\varepsilon^2} \cdot \frac{1}{n} \to 0 \qquad \text{for } n \to \infty,$$

proving that (Y_n) is a consistent sequence of estimators.

Example 12.3 *A random variable X has its distribution given by*

(6) $P\{X=k\} = \left(\begin{array}{c} k+4 \\ 4 \end{array}\right) p^5 q^k, \qquad k \in \mathbb{N}_0,$

where $p>0$, $q>0$ and $p+q=1$. The mean of X is given by

(7) $E\{X\} = \mu = \dfrac{5q}{p}.$

(The proof is not required.)

1) Assuming that n observations of the mutually independent random variables $X_1, X_2, \dots, X_n$, all distributed like X above, have given the results $x_1, x_2, \dots, x_n$, apply the maximum method to find the best estimates of p and q, by also using (7) an estimate of μ.

2) Prove that the estimator corresponding to the estimate of μ is a central estimator.

3) Find for a distribution of the type (6) above with $\mu = 5$,

$$P\{X<5\}, \qquad P\{X=5\} \quad \textit{and} \quad P\{X>5\}.$$

1) We shall maximize the function (in p)

$$h(p) = \prod_{i=1}^{n} \left(\begin{array}{c} x_i+4 \\ 4 \end{array}\right) p^5 (1-p)^{x_i}, \qquad p \in]0,1[.$$

Since $h(0) = h(1) = 0$, this corresponds to an investigation of the solutions of the equation $h_1'(p) = 0$, $0<p<1$, where

$$h_1(p) = \ln h(p) = \sum_{i=1}^{n} \ln \left(\begin{array}{c} x_i+4 \\ 4 \end{array}\right) + 5n \ln p + \sum_{i=1}^{n} x_i \cdot \ln(1-p).$$

By differentiation,

$$\begin{aligned} h_1'(p) &= \frac{5n}{p} - \frac{1}{1-p}\sum_{i=1}^{n} x_i = \frac{1}{p(1-p)}\left\{5n - 5np - p\sum_{i=1}^{n} x_i\right\} \\ &= \frac{1}{p(1-p)}\left\{5n - p\left(5n + \sum_{i=1}^{n} x_i\right)\right\}, \end{aligned}$$

thus

$$p^\star = \frac{5n}{5n+\sum_{i=1}^{n} x_i}, \quad \text{corresponding to} \quad q^\star = \frac{\sum_{i=1}^{n} x_i}{5n+\sum_{i=1}^{n} x_i}.$$

Using (7) we get the estimate of the mean

$$\mu^\star = 5\,\frac{q^\star}{p^\star} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

2) Clearly,

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

is a central estimator, because

$$E\left\{Y_n\right\} = \frac{1}{n}\sum_{i=1}^{m} E\left\{X_i\right\} = E\{X\} = \mu.$$

3) If $\mu = 5\,\dfrac{q}{p} = 5$, er $p = q = \dfrac{1}{2}$, then

$$P\{X = k\} = \binom{k+4}{4}\left\{\frac{1}{2}\right\}^{5+k}, \qquad k \in \mathbb{N}_0.$$

Then

$$\begin{aligned}
P\{X<5\} &= P\{X=0\}+P\{X=1\}+P\{X=2\}+P\{X=3\} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad +P\{X=4\} \\
&= \binom{4}{4}\frac{1}{2^5}+\binom{5}{4}\frac{1}{2^6}+\binom{6}{4}\frac{1}{2^7}+\binom{7}{4}\frac{1}{2^8}+\binom{8}{4}\frac{1}{2^9} \\
&= \frac{1}{2^9}\left\{2^4+5\cdot 2^3+\frac{6\cdot 5}{1\cdot 2}\cdot 2^2+\frac{7\cdot 6\cdot 5}{1\cdot 2\cdot 3}\cdot 2+\frac{8\cdot 7\cdot 6\cdot 5}{1\cdot 2\cdot 3\cdot 4}\right\} \\
&= \frac{1}{2^9}\left\{2^4+5\cdot 2^3+15\cdot 2^2+35\cdot 2+35\cdot 2\right\} \\
&= \frac{1}{2^9}\left\{2^4+5\cdot 2^3+15\cdot 2^2+35\cdot 2^2\right\} \\
&= \frac{1}{2^7}\left\{2^2+5\cdot 2+15+35\right\}=\frac{64}{128}=\frac{1}{2},
\end{aligned}$$

and

$$P\{X=5\}=\binom{9}{4}\frac{1}{2^{10}}=\frac{9\cdot 8\cdot 7\cdot 6}{1\cdot 2\cdot 3\cdot 4}\cdot\frac{1}{1024}=\frac{63}{512},$$

and

$$P\{X>5\}=1-P\{X<5\}-P\{X=5\}=1-\frac{1}{2}-\frac{63}{512}=\frac{193}{512}.$$

Example 12.4 *A random variable X has the frequency*

$$f_X(x)=\begin{cases} k\cdot x^{a-1}\,\exp\left(-b\,x^a\right) & \text{for } x>0, \\[2ex] 0 & \text{for } x>leq0, \end{cases}$$

where a and b are positive constants.

1. *Find, expressed by a and b, the constant k.*

2. *Find the median of X.*

We assume in the following two questions that $a=4$.

3. *Find the mean $E\left\{X^4\right\}$ and the variance $V\left\{X^4\right\}$.*

4. *Assuming that n observations of the mutually independent random variables $X_1, X_2, \dots, X_n$, all distributed like X above, have given the results $x_1, x_2, \dots, x_n$, apply the maximum method the best estimate of the parameter b.*

1) It is obvious that $f_X(x)\ge 0$ for $k>0$. Then we get the condition

$$1=\int_0^{\infty} f_X(x)\,dx=k\int_0^{\infty} x^{a-1}\,\exp\left(-b\,x^a\right)\,dx=\frac{k}{a}\int_0^{\infty}\exp\left(-b\,x^a\right)\,d\left(x^a\right)=\frac{k}{ab},$$

from which we derive that $k=ab$.

2) Clearly, $x > 0$. The condition of the median x is

$$\frac{1}{2} = F_X(x) = \int_0^x f_X(t)\,dt = ab\int_0^x t^{a-1}\,\exp\left(-b\,t^a\right)\,dt = 1 - \exp\left(-b\,x^a\right),$$

hence

$$\exp\left(b\,x^a\right) = 2,$$

so $b\,x^a = \ln 2$, and the median is

$$(X) = \left\{\frac{\ln 2}{b}\right\}^{\frac{1}{a}}.$$

3) If $a = 4$, then by the change of variable $t = x^4$,

$$E\left\{X^4\right\} = 4b\int_0^\infty x^4 x^3 \exp\left(-b\,x^4\right)\,dx = b\int_0^\infty t\,e^{-bt}\,dt = \frac{1}{b}\int_0^\infty u\,e^{-u}\,du = \frac{1}{b},$$

and

$$E\left\{X^8\right\} = b\int_0^\infty t^2 e^{-bt}\,dt = \frac{1}{b^2}\int_0^\infty u^2 e^{-u}\,du = \frac{2}{b^2},$$

hence

$$V\left\{X^4\right\} = E\left\{X^8\right\} - \left(E\left\{X^4\right\}\right)^2 = \frac{2}{b^2} - \frac{1}{b^2} = \frac{1}{b^2}.$$

4) We shall maximize the function

$$h(b) = \prod_{k=1}^n f_X\left(x_k\right) = 4^n b^n \prod_{k=1}^n x_k^3 \exp\left(-b\,x_k^4\right) = 4^n\left(\prod_{k=1}^n x_k^3\right) b^n \exp\left(-b\sum_{k=1}^n x_k^4\right),$$

i.e.

$$h_1(b) = \ln h(b) = \ln\left\{4^n\prod_{k=1}^n x_k^3\right\} + n\,\ln b - b\sum_{k=1}^n x_k^4,$$

where

$$h_1'(b) = \frac{h'(b)}{h(b)} = \frac{n}{b} - \sum_{k=1}^n x_k^4 = 0 \qquad \text{for } b = \frac{n}{\sum_{k=1}^n x_k^n}.$$

Since $h(b) > 0$ for $b > 0$, and $h(b) \to 0$ for $b \to 0+$ and for $b \to \infty$, we conclude that

$$b^\star = \frac{n}{\sum_{k=1}^n x_k^4}$$

is the best estimat of b.

Example 12.5 *Let (X,Y) be a two-dimensional random variable, where X and Y are independent, and where X and Y have the same distribution, given by*

$$P\{X=n\}=P\{Y=n\}=pq^{n-1}, \qquad n\in\mathbb{N},$$

where $p>0$, $q>0$ and $p+q=1$.

1) *Find everyone of the following probabilities*

$$P\{Y=X\}, \quad P\{Y<X\}, \quad P\{Y>X\} \quad \text{and} \quad P\{Y=2X\}.$$

2) *The random variables U and V are given by*

$$(U,V)=(X+aY,X-aY),$$

where a is a real constant.
Find the correlation coefficient $\varrho(U,V)$.

3) *Assuming that m observations of the mutually independent random variables X_1, X_2, $\dots$, X_m, all distributed like X above, have given the results x_1, x_2, $\dots$, x_m, apply the maximum method to find the best estimate of the parameter p.*

4) *Assuming that m observations of the mutually independent two-dimensional random variables (X_1,Y_1), (X_2,Y_2), $\dots$, (X_m,Y_m), all distributed like (X,Y) above, have given the results (x_1,y_1), (x_2,y_2), $\dots$, (x_m,y_m).*
What is a reasonable estimate of p?

1) A straightforward summation gives

$$\begin{aligned}P\{Y=X\} &= \sum_{n=1}^{\infty} P\{X=n\}\cdot P\{Y=n\} = \sum_{n=1}^{\infty} p^2 q^{2n-2} = \frac{p^2}{1-q^2}\sum_{n=1}^{\infty}\left(1-q^2\right)\left(q^2\right)^{n-1} \\ &= \frac{p^2}{1-q^2} = \frac{p^2}{(1+q)(1-q)} = \frac{p}{1+q} = \frac{p}{2-p}.\end{aligned}$$

By the symmetry, $P\{Y<X\}=P\{X<Y\}$. We then conclude from

$$1=P\{Y=X\}+P\{Y<X\}+P\{X<Y\}=P\{Y=X\}+2\,P\{Y<X\},$$

that

$$\begin{aligned}P\{Y<X\} &= P\{X<Y\}=\frac{1}{2}(1-P\{Y=X\})=\frac{1}{2}\left\{1-\frac{p}{1+q}\right\}=\frac{1}{2}\left\{1-\frac{1-q}{1+q}\right\} \\ &= \frac{1}{2}\cdot\frac{1+q-1+q}{1+q}=\frac{q}{1+q}.\end{aligned}$$

Finally,

$$\begin{aligned}P\{Y=2X\} &= \sum_{n=1}^{\infty} P\{X=n\}\cdot P\{Y=2n\} = \sum_{n=1}^{\infty} pa^{n-1}\cdot pq^{2n-1} = p^2\sum_{n=1}^{\infty} q^{3n-2} \\ &= p^2 q\sum_{n=1}^{\infty}\left(q^3\right)^{n-1} = \frac{p^2 q}{1-q^3} = \frac{pq}{q^2+q+1}.\end{aligned}$$

2) It follows from

$$E\{X\} = \sum_{n=1}^{\infty} pnq^{n-1} = p \cdot \frac{1}{(1-q)^2} = \frac{1}{p},$$

that

$$E\{U\} = E\{X\} + a\,E\{Y\} = (1+a)E\{X\} = \frac{1+a}{p}$$

and

$$E\{V\} = E\{X\} - a\,E\{Y\} = (1-a)E\{X\} = \frac{1-a}{p}.$$

Since X and Y are independent, and $V\{X\} = \dfrac{q}{p^2}$, we get

$$V\{U\} = V\{V\} = V\{X\} + a^2 V\{Y\} = \left(1+a^2\right) V\{X\} = \left(1+a^2\right) \cdot \frac{p}{q^2}.$$

Finally,

$$
\begin{aligned}
\operatorname{Cov}(U,V) &= E\{UV\}-E\{U\}\cdot E\{V\} = \left(E\left\{X^2-a^2Y\right\}-\frac{1-a^2}{p^2}\right) \\
&= \left[V\{X\}+(E\{X\})^2-a^2\left(V\{Y\}+(E\{Y\})^2\right)-\frac{1-a^2}{p^2}\right] \\
&= \left(1-a^2\right)\left(V\{X\}+(E\{X\})^2-\frac{1}{p^2}\right)=\left(1-a^2\right)\left(\frac{q}{p^2}+\frac{1}{p^2}-\frac{1}{p^2}\right) \\
&= \left(1-a^2\right)\frac{q}{p^2},
\end{aligned}
$$

hence

$$
\varrho(U,V)=\frac{\operatorname{Cov}(U,V)}{\sqrt{V\{U\}\cdot V\{V\}}}=\frac{\operatorname{Cov}(U,V)}{V\{U\}}=\frac{q}{p^2}\cdot\left(1-a^2\right)\cdot\frac{p^2}{q\left(1+a^2\right)}=\frac{1-a^2}{1+a^2}.
$$

3) Using that $q=1-p$, it follows that we shall maximize the function

$$
h(p)=\prod_{k=1}^{m} p\,q^{x_k-1}=p^m\prod_{k=1}^{m}(1-p)^{x_k-1},
$$

or equivalently,

$$
\ln h(p)=m\,\ln p+\left\{\sum_{k=1}^{m}x_k-m\right\}\ln(1-p).
$$

Now

$$
\frac{d}{dp}\,\ln h(p)=\frac{h'(p)}{h(p)}=\frac{m}{p}-\frac{1}{1-p}\left\{\sum_{k=1}^{m}x_k-m\right\}
$$

is zero for

$$
m(1-p)=p\left\{\sum_{k=1}^{m}x_k-m\right\},
$$

i.e. for

$$
p\sum_{k=1}^{m}x_k=m,
$$

and $h(p)\to 0$ for $p\to 0$ or for $p\to 1$ [provided that at least one $x_k>1$]. We therefore conclude that the estimate of p is

$$
p^\star=\frac{m}{\sum_{k=1}^{m}x_k}=\frac{1}{\overline{x}},\qquad \text{with } \overline{x}=\frac{1}{m}\sum_{k=1}^{m}x_k.
$$

4) Using the same method as in **3.**, it follows from the independency of X and Y that (x_1,y_1), (x_2,y_2), $\dots$, (x_m,y_m) can be considered as $2m$ observations of X_1, X_2, $\dots$, X_m, M_{m+1}, $\dots$, X_{2m}, where $Y_j=X_{m+j}$. Then by **3.** the estimate becomes

$$
p^{\star\star}=\frac{2m}{\sum_{k=1}^{m}x_k+\sum_{k=1}^{m}y_k}=\frac{2m}{\sum_{k=1}^{m}(x_k+y_k)}=\frac{2}{\overline{x}+\overline{y}}.
$$

Example 12.6 *A random variable X has the frequency*

$$f(x)=\begin{cases} a\,e^{-a\sqrt{x}}\cdot\dfrac{1}{2\sqrt{x}}, & x>0,\\[2mm] 0, & x\le 0,\end{cases}$$

where a is a positive constant.

1) Find $E\{X\}$ and $E\left\{\sqrt{X}\right\}$.

2) Assuming that n observations of the mutually independent random variables X_1, X_2, ..., X_n, all distributed like X above, have given the results x_1, x_2, ..., x_n, apply the maximum method to find the best estimate of a and hence the best estimate of $E\left\{\sqrt{X}\right\}$.

3) Prove that the estimator corresponding to the estimate of $E\left\{\sqrt{X}\right\}$ is a central estimator.

1) We get by the substitution $t=\sqrt{x}$,

$$E\{X\}=a\int_0^{\infty}x\cdot e^{-a\sqrt{x}}\cdot\frac{1}{2\sqrt{x}}\,dx=a\int_0^{\infty}t^2e^{-at}\,dt=\frac{1}{a^2}\int_0^{\infty}u^2e^{-u}\,du=\frac{2}{a^2}$$

and

$$E\left\{\sqrt{X}\right\}=a\int_0^{\infty}\sqrt{x}\cdot e^{-a\sqrt{x}}\cdot\frac{1}{2\sqrt{x}}\,dx=a\int_0^{\infty}t\,e^{-at}\,dt=\frac{1}{a}\int_0^{\infty}u\,e^{-u}\,du=\frac{1}{a}.$$

2) We shall maximize the function

$$h(a)=a^n\cdot\exp\left(-a\sum_{i=1}^n\sqrt{x_i}\right)\cdot\frac{1}{2^n\sqrt{\prod_{i=1}^n x_i}},\qquad a>0,$$

where we see that it suffices to maximize the function

$$h_1(a)=\ln h(a)=n\cdot\ln a-a\sum_{i=1}^n\sqrt{x_i}-\ln\left(2^n\sqrt{\prod_{i=1}^n x_i}\right),\qquad a>0.$$

Since $x_i>0$, we have $h(a)\to 0$ for $a\to 0+$ and for $a\to\infty$, and $h(a)>0$ otherwise. From

$$h_1'(a)=\frac{h'(a)}{h(a)}=\frac{n}{a}-\sum_{i=1}^n\sqrt{x_i}=0,\qquad \text{only for } a=\frac{n}{\sum_{i=1}^n\sqrt{x_i}}=\frac{1}{\overline{\sqrt{x}}},$$

follows that the best estimate of a is given by

$$a^{\star}=\frac{n}{\sum_{i=1}^n\sqrt{x_i}}=\frac{1}{\overline{\sqrt{x}}}.$$

3) The estimator of $E\left\{\sqrt{X}\right\}$ is

$$Y_n = \frac{1}{n}\sum_{i=1}^{n}\sqrt{X_i}.$$

Its mean is

$$E\left\{Y_n\right\} = \frac{1}{n}\sum_{i=1}^{n} E\left\{\sqrt{X_i}\right\} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{a} = \frac{1}{a} = E\left\{\sqrt{X}\right\},$$

hence Y_n is a central estimator of $E\left\{\sqrt{X}\right\}$.

Example 12.7 *A random variable X has the frequency*

$$f(x) = \begin{cases} \lambda^2 x\, e^{-\lambda x}, & x > 0, \\ 0, & x \le 0, \end{cases}$$

where λ is a positive constant.

1. *Compute the mean μ.*

A Geiger counter is only recording every second particle, which arrives to the counter. The particles arrive according to a Poisson process of an (unknown) intensity λ. The difference in time between two successive recorded *arrivals has the frequency $f(x)$ given above (this shall* not *be proved), and the difference in time between the first and the second recorded particle, between the second and the third recorded particle … are mutually independent random variables $X_1, X_2, \dots, $ all of frequency $f(x)$.*

2. *Given the observed differences in time $x_1, x_2, \dots, x_n$, find by means of the maximum method the best estimate $\lambda_n^\star$ of the parameter λ, and hence the best estimate $\mu_n^\star$ of the mean μ.*

3. *Prove that the estimator Y_n corresponding to the estimate $\mu_n^\star$ of μ is a central estimator.*

4. *Allowing n to vary, prove that the sequence of estimators $(Y_n)_{n=1}^{\infty}$ of μ is a consistent sequence.*

1) Since $X \in \Gamma\left(2, \frac{1}{\lambda}\right)$, the mean is $\mu = \frac{2}{\lambda}$.

2) We shall find the maximum of the function

$$h(\lambda) = \prod_{i=1}^{n}\lambda^2 x_i e^{-\lambda x_i} = \left(\prod_{i=1}^{n} x_i\right)\cdot\lambda^{2n}\exp\left(-\lambda\sum_{i=1}^{n} x_i\right),$$

which is equivalent to finding the maximum of the function

$$h_1(\lambda) = \lambda h(\lambda) = 2n\,\ln\lambda - \lambda\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\ln x_i.$$

It follows from $x_i > 0$ that $h(\lambda) \to 0$ for $\lambda \to 0+$ and for $\lambda \to \infty$. Since

$$h_1'(\lambda) = \frac{h'(\lambda)}{h(\lambda)} = \frac{2n}{\lambda} - \sum_{i=1}^{n} x_i$$

is only zero for

$$\lambda = \frac{2n}{\sum_{i=1}^{n} x_i} = \frac{2}{\overline{x}},$$

where $\overline{x}$ as usual denotes the mean, this corresponds to our maximum. Hence the best estimate of λ is given by

$$\lambda_n^\star = \frac{2n}{\sum_{i=1}^{n} x_i} = \frac{2}{\overline{x}},$$

and the best estimate of $\mu = \dfrac{2}{\lambda}$ is

$$\mu_n^\star = \frac{2}{\lambda_n^\star} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}.$$

3) The estimator Y_n corresponding to $\mu_n^\star$ is

$$Y_n = \frac{1}{n}\sum_{i=1}^n X_i.$$

Since

$$E\{Y_n\} = \frac{1}{n}\sum_{i=1}^n E\{X_i\} = \frac{2}{\lambda} = \mu,$$

the estimator is central.

4) Finally,

$$V\{Y_n\} = \frac{1}{n}\,V\{X\} = \frac{2}{n\lambda^2} = s_n^2,$$

hence by Chebyshev's inequality,

$$P\{|Y_n - \mu| \ge \varepsilon\} \le \frac{s_n^2}{\varepsilon^2} = \frac{2}{\varepsilon^2\lambda^2}\cdot\frac{1}{n} \to 0 \qquad \text{for } n \to \infty.$$

It follows that the sequence of estimators is consistent.

Example 12.8 *A random variable X has its distribution given by*

$$P\{X=k\}=(k-1)p^2q^{k-2}, \qquad k=2,\,3,\,4,\,\dots,$$

where $p>0$, $q>0$ and $p+q=1$, thus $X\in \mathrm{Pas}(2,p)$.

1. *Find mean μ and variance σ^2 of X.*

2. *Assuming that n observations of the mutually independent random variables $X_1,\,X_2,\,\dots,\,X_n$, all distributed like X above, have given the results $x_1,\,x_2,\,\dots,\,x_n$ (where not all x_i are equal to 2), apply the maximum method the best estimate of p and the best estimate of μ.*

3. *Prove that the corresponding estimator of μ is*

$$\tilde{X}_n=\frac{1}{n}\sum_{i=1}^n X_i,$$

and that this is a central estimator.

We now perform m observations of the mutually independent random variables $Y_1,\,Y_2,\,\dots,\,Y_m$, all distributed like X above and furthermore all independent of the X_i. The results are called $y_1,\,y_2,\,\dots,\,y_m$ (where not all y_j are equal to 2). We denote by $\tilde{Y}_m$ the estimator of μ which has been found by the maximum method from $Y_1,\,Y_2,\,\dots,\,Y_m$. Based on the two estimators $\tilde{X}_n$ and $\tilde{Y}_m$ we form another two estimators of μ, namely

$$Z=\frac{1}{2}\left(\tilde{X}_n+\tilde{Y}_m\right) \quad og \quad U=\frac{n}{n+m}\,\tilde{X}_n+\frac{m}{n+m}\,\tilde{Y}_m.$$

4. *Prove that Z and U are both central estimators of μ.*

5. *Prove for $m\neq n$ that the estimator U is more efficient than the estimator Z.*

1) When $X\in\mathrm{Pas}(2,\mu)$, then

$$E\{X\}=\mu=\frac{2}{p} \quad\text{and}\quad V\{X\}=\sigma^2=\frac{2(1-p)}{p^2}.$$

2) We shall maximize the function

$$h(p)=\prod_{i=1}^n (x_i-1)\,p^2q^{x_i-2}=\left\{\prod_{j=1}^n (x_j-1)\right\}\cdot p^{2n}\cdot(1-p)^{\sum_{i=1}^n x_i-2n},$$

which is equivalent to maximizing the function

$$h_1(p)=\ln h(p)=\sum_{j=1}^n \ln(x_j-1)+2n\cdot\ln p+\left\{\sum_{i=1}^n x_i-2n\right\}\ln(1-p).$$

Since not every x_i is 2, we have $h(0)=h(1)=0$, and $h(p)>0$ for $0<p<1$. Since

$$h_p'(p)=\frac{2n}{p}-\frac{1}{1-p}\left\{\sum_{i=1}^n x_i-2n\right\}=0$$

for

$$(1-p)\cdot 2n - p\left\{\sum_{i=1}^{n} x_i - 2n\right\} = (1-p)\cdot 2n - pn\,(\overline{x}-2) = 0,$$

i.e. for

$$2n - p\sum_{i=1}^{n} x_i = n\,(2-p\,\overline{x}) = 0,$$

we obtain the maximum (i.e. the best estimate for p),

$$p^\star = \frac{2n}{\sum_{i=1}^{n} x_i} = \frac{2}{\overline{x}}.$$

Since $\mu = \dfrac{2}{p}$, the best estimate of μ is given by

$$\mu^\star = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x},$$

i.e. by the mean of the observations.

3) The estimator corresponding to $\mu^\star$ is precisely

$$\tilde{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

It follows from

$$E\left\{\tilde{X}_n\right\} = \frac{1}{n}\sum_{i=1}^{n} E\left\{X_i\right\} = \frac{1}{n}\cdot n\cdot\frac{2}{p} = \mu,$$

that $\tilde{X}_n$ is a central estimator.

4) Since both $E\left\{\tilde{X}_n\right\} = \mu$ and $E\left\{\tilde{Y}_m\right\} = \mu$, we get

$$E\left\{r\,\tilde{X}_n + (1-r)\,\tilde{Y}_m\right\} = \mu, \qquad r\in\mathbb{R},$$

hence $r\,\tilde{X}_n + (1-r)\tilde{Y}_m$ is central for every $r\in\mathbb{R}$. By choosing $r=2$, we get Z. By choosing $r = \dfrac{n}{n+m}$, we get U.

5) Since

$$V\left\{\tilde{X}_n\right\} = \frac{\sigma^2}{n} \quad \text{og} \quad V\left\{\tilde{Y}_m\right\} = \frac{\sigma^2}{m},$$

and since $\tilde{X}_n$ and $\tilde{Y}_m$ are independent, we get

$$V\left\{r\,\tilde{X}_n + (1-r)\,\tilde{Y}_m\right\} = \sigma^2\left(\frac{r^2}{n} + \frac{(1-r)^2}{m}\right).$$

If we put

$$f(r)=\frac{r^2}{n}+\frac{(1-r)^2}{m},$$

then

$$f'(r)=\frac{2r}{n}-\frac{2(1-r)}{m}=\frac{2}{mn}\,(rm-n(1-r))=\frac{2}{nm}\,(r(m+n)-n)=0$$

for $r=\dfrac{n}{n+m}$, corresponding to a minimum. Hence the variance er is smallest for U, i.e. U is the most efficient estimator in the family.

Example 12.9 *A random variable $X \in N(0,a)$ is normally distributed of mean 0 and variance a.*

1) Compute $E\left\{X^4\right\}$.

2) Assuming that n observations of the mutually independent random variables $X_1,\,X_2,\,\dots,\,X_n$, all distributed like X above, have given the results $x_1,\,x_2,\,\dots,\,x_n$, apply the maximum method in order to find the best estimate of a.

3) Prove that the corresponding estimator Y_n of a is a central estimator.

4) Compute the variance of Y_n.

5) Prove that the corresponding sequence of estimators $(Y_n)_{n=1}^{\infty}$ is a consistent sequence.

1) By a small computation,

$$\begin{aligned}
E\left\{X^4\right\} &= \frac{1}{\sqrt{2\pi\,a}}\int_{-\infty}^{\infty} x^4\exp\left(-\frac{x^2}{2a}\right)\,dt = a^2\,\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t^4\exp\left(-\frac{1}{2}\,t^2\right)\,dt\\
&= a^2\,\frac{2}{\sqrt{2\pi}}\int_0^{\infty} t^3\cdot\exp\left(-\frac{1}{2}\,t^2\right)t\,dt = 3a^2\cdot\frac{2}{\sqrt{2\pi}}\int_0^{\infty} t^2\exp\left(-\frac{1}{2}\,t^2\right)\,dt\\
&= 3a^2.
\end{aligned}$$

2) Since

$$\varphi_a(x)=\frac{1}{\sqrt{2\pi\,a}}\,\exp\left(-\frac{x^2}{2a}\right),$$

we shall find the maximum of the function

$$h(a)=\prod_{i=1}^n \varphi_a\left(x_i\right)=\left\{\frac{1}{\sqrt{2\pi}}\right\}^n\cdot a^{-\frac{n}{2}}\cdot\exp\left(-\frac{1}{2}\cdot\frac{1}{a}\sum_{i=1}^n x_i^2\right),$$

which is equivalent to finding the maximum of the function

$$h_1(a)=\ln h(a)=\ln\left(\left\{\frac{1}{\sqrt{2\pi}}\right\}^n\right)-\frac{n}{2}\,\ln a-\frac{1}{2}\cdot\frac{1}{a}\sum_{i=1}^n x_i^2.$$

Clearly, $h(a) \to 0$ for $a \to \infty$. If not all the x_i are 0, then $h(a) \to 0$ for $a \to 0+$ due to to rule of magnitudes. From

$$h_1'(a) = -\frac{n}{2a} + \frac{1}{2}\,\frac{1}{a^2}\sum_{i=1}^{n} x_i^2 = 0 \qquad \text{for } a = \frac{1}{n}\sum_{i=1}^{n} x_i^n,$$

follows that the maximum is attained at the estimate

$$a^\star = \frac{1}{n}\sum_{i=1}^{n} x_i^2.$$

3) The estimator corresponding to $a^\star$ is

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

From $X_i^2 \in \Gamma\left(\frac{a}{2}, 2\right)$, follows that

$$E\left\{Y_n\right\} = \frac{1}{n}\sum_{i=1}^{n} \frac{a}{2}\cdot 2 = a,$$

and Y_n is a central estimator of a.

4) Since the X_i are mutually independent, the variance is

$$V\left\{Y_n\right\} = \frac{1}{n^2}\sum_{i=1}^{n} V\left\{X_i\right\} = \frac{a}{n} = s_n^2.$$

5) By Chebyshev's inequality,

$$P\left\{|Y_n - a| \geq \varepsilon\right\} \leq \frac{s_n^2}{\varepsilon^2} = \frac{a}{\varepsilon^2}\cdot\frac{1}{n} \to 0 \qquad \text{for } n \to \infty,$$

and the sequence is consistent.

Example 12.10 *A random variable X has the frequency*

$$f(x) = \begin{cases} \lambda\, e^{-\lambda(x-1)}, & x \geq 1, \\ \\ 0, & x < 1, \end{cases}$$

where λ is a positive constant.

1. *Find the mean μ of X.*

2. *Assuming n observations of the mutually independent random variables X_1, X_2, $\dots$, X_n, all distributed like X above, of the results x_1, x_2, $\dots$, x_n, use the maximum method to find the best estimate $\lambda_n^\star$ of λ and hence also the best estimate $\mu_n^\star$ of μ.*

The estimator corresponding to the estimat $\lambda_n^\star$ is denoted by Y_n, and the estimator corresponding to the estimate $\mu_n^\star$ is denoted by Z_n.

3. *Prove that Z_n is a central estimator of μ.*

4. *Find the frequency of Z_n.*
Hint: *Start by computing the frequency of*

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

5. *Check if Y_n is a central estimator of λ.*

1) One may either translate X, or compute directly,

$$E\{X\} = \lambda\int_1^\infty x\, e^{-\lambda(x-1)}\, dx = \lambda\int_0^\infty (x+1)\, e^{-\lambda x}\, dx = \frac{1}{\lambda}+1.$$

2) We shall maximize the function

$$h(\lambda) = \prod_{i=1}^{n} f\left(x_i\right) = \lambda^n \exp\left(-\lambda\left(\sum_{i=1}^{n} x_i - n\right)\right),$$

where $h(\lambda) \to 0$ for $\lambda \to 0+$ and for $\lambda \to \infty$. We get from the equation

$$h'(\lambda) = h(\lambda) \left\{ \frac{n}{\lambda} - \left(\sum_{i=1}^{n} x_i - n \right) \right\} = 0$$

the estimate

$$\lambda_n^\star = \frac{n}{\sum_{i=1}^n x_i - n} = \frac{1}{\frac{1}{n}\sum_{i=1}^n x_i - 1} = \frac{1}{\overline{x} - 1}.$$

Hence

$$\mu_n^\star = \frac{1}{\lambda_n^\star} + 1 = \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{x}.$$

3) If

$$Y_n = \frac{1}{\frac{1}{n}\sum_{i=1}^n X_i - 1} \quad \text{and} \quad Z_n = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

then

$$E\{Z_n\} = \frac{1}{n} \sum_{i=1}^{n} E\{X_i\} = E\{X\} = \frac{1}{\lambda} + 1 = \mu,$$

and Z_n is central.

4) From $U_i = X_i - 1 \in \Gamma\left(1, \frac{1}{\lambda}\right)$, follows that

$$U = \sum_{i=1}^{n} U_i \in \Gamma\left(n, \frac{1}{\lambda}\right),$$

thus

$$f_U(u) = \begin{cases} \dfrac{\lambda^n}{(n-1)!}\, u^{n-1} \, \exp(-\lambda\, u) & \text{for } u > 0, \\[2ex] 0 & \text{for } u \le 0. \end{cases}$$

From

$$Z_n = \frac{1}{n} \sum_{i=1}^{n} X_i = \frac{1}{n} U + 1,$$

and $a = \dfrac{1}{n}$ and $b = 1$ follows for $z > 1$ that the frequency of Z_n is given by

$$\begin{aligned} f_{Z_n}(z) &= n \cdot f_U(n(z-1)) = n \cdot \frac{\lambda^n}{(n-1)!} \cdot n^{n-1}(z-1)^{n-1} \cdot e^{-\lambda n(z-1)} \\ &= \frac{n^2}{(n-1)!} \cdot \lambda^n (z-1)^{n-1} e^{-\lambda n(z-1)}, \end{aligned}$$

and of course $f_{Z_n}(z) = 0$ for $z \le 1$.

5) Since $Y_n = \frac{n}{U}$, where $U \in \Gamma\left(n, \frac{1}{\lambda}\right)$, we get for $n > 1$,

$$
\begin{aligned}
E\{Y_n\} &= n\,E\left\{\frac{1}{U}\right\} = n\int_0^{\infty} \frac{1}{u}\cdot\frac{1}{(n-1)!}\,\lambda^n u^{n-1}e^{-\lambda u}du \\
&= \frac{n}{(n-1)!}\int_0^{\infty} u^{n-2}\lambda^n e^{-\lambda u}du \qquad\qquad (t=\lambda u) \\
&= \frac{n}{(n-1)!}\,\lambda\int_0^{\infty} t^{n-2}e^{-t}dt = \frac{n}{(n-1)!}\cdot\lambda(n-2)! \\
&= \frac{n}{n-1}\,\lambda \neq \lambda,
\end{aligned}
$$

proving that Y_n is not a central estimator of λ.

Example 12.11 *A random variable X has the frequency*

$$f(x) = \begin{cases} \frac{1}{2}\,\lambda^3 x^2 e^{-\lambda x}, & x > 0, \\ \\ \quad 0, & x \le 0, \end{cases}$$

where λ is a positive constant.

1. *Compute the mean μ.*

Customers arrive to a shop according to a Poisson process of (unknown) intensity λ. One day a shop assistant has been asked to write down all the arrival times of the customers, but due to his laziness he is only recording the arrival times of every third customer. The time difference between two successive recorded *arrivals has the frequency $f(x)$ given above (a proof of this claim is* not *required), and the time differences between the first and the second recorded arrival, between the second and the third recorded arrival etc. are mutually independent random variables $X_1, X_2, \dots,$ all of frequency $f(x)$.*

2. *Assuming that the time differences $x_1, x_2, \dots, x_n$, have been recorded, use the maximum method based on these observations to find the best estimate $\lambda_n^\star$ of λ, and hence the best estimate of the mean μ.*

3. *Prove that the estimator Y_n corresponding to the estimate $\mu_n^\star$ of μ is a central estimator.*

4. *Allowing n to vary, prove that the sequence of estimators $(Y_n)_{n=1}^{\infty}$ is a consistent sequence.*

5. *Find for every value of n the distribution of Y_n.*

1) Since $X \in \Gamma\left(3, \frac{1}{\lambda}\right)$, we have $\mu = \frac{3}{\lambda}$.

2) We shall maximize the function

$$h(\lambda) = \prod_{i=1}^{n} f(x_i) = \frac{1}{2^n} \prod_{i=1}^{n} x_i^2 \cdot \lambda^{3n} \exp\left(-\lambda \sum_{i=1}^{n} x_i\right),$$

where clearly $h(\lambda) \to 0$ for $\lambda \to 0+$ and for $\lambda \to \infty$. Since

$$h'(\lambda) = h(\lambda) \cdot \left\{ \frac{3n}{\lambda} - \sum_{i=1}^{n} x_i \right\} = 0$$

for

$$\lambda = \frac{3n}{\sum_{i=1}^{n} x_i} = \frac{3}{\overline{x}},$$

we get

$$\lambda_n^\star = \frac{3n}{\sum_{i=1}^{n} x_i} = \frac{3}{\overline{x}} \quad \text{and} \quad \mu_n^\star = \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{x}.$$

3) It follows from

$$Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

that

$$E\{Y_n\} = \frac{1}{2} \sum_{i=1}^{n} E\{X_i\} = E\{X\} = \mu = \frac{3}{\lambda},$$

hence Y_n is a central estimator.

4) Now,

$$V\{Y_n\} = \frac{1}{n} V\{X\} = \frac{9}{\lambda n} = s_n^2,$$

so it follows by Chebyshev's inequality that

$$P\{|Y_n - \mu| \geq \varepsilon\} \leq \frac{s_n^2}{\varepsilon^2} = \frac{9}{\lambda \varepsilon^2} \cdot \frac{1}{n} \to 0 \qquad \text{for } n \to \infty,$$

and the sequence (Y_n) is consistent.

5) Since $X \in \Gamma\left(3, \frac{1}{\lambda}\right)$, We get

$$Z_n = \sum_{i=1}^{n} X_i \in \Gamma\left(3n, \frac{1}{\lambda}\right).$$

Since $Y_n = \frac{1}{n} Z_n$ and $\frac{dz}{dy} = n$ for $z = ny$, we get for $y > 0$ that

$$f_{Y_n}(y) = f_{Z_n}(ny) \cdot n = \frac{\lambda^{3n}}{\Gamma(3n)} n^{3n-1} y^{3n-1} \cdot \exp(-\lambda n y) \cdot n = \frac{\lambda n)^{3n}}{\Gamma(3n)} y^{3n-1} \exp(-\lambda n y),$$

thus $Y_n \in \Gamma\left(3n, \frac{1}{\lambda n}\right)$.

Example 12.12 *A random variable X has the frequency*

$$f(x)=\frac{1}{2a}\,\exp\left(-\frac{|x|}{a}\right),\qquad x\in\mathbb{R},$$

where a is a positive constant.

1) Compute $E\{X\}$, $E\{|X|\}$ and $E\left\{X^2\right\}$.

2) Assuming that n observations of the mutually independent random variables $X_1, X_2, \dots, X_n$, all distributed like X above, have given the results $x_1, x_2, \dots, x_n$ (where not all the x_i are equal to 0), apply the maximal method to give the best estimate $a_n^\star$ of a and the corresponding estimator Y_n of a.

3) Prove that Y_n is a central estimator of a.

4) Check if the corresponding sequence of estimators $(Y_n)_{n=1}^{\infty}$ is a consistent sequence.

1) Clearly, $E\{X\}$, $E\{|X|\}$ and $E\left\{X^2\right\}$ all exist, and $E\{X\}=0$. It follows by the symmetry that

$$E\{|X\|\}=\frac{1}{2a}\cdot 2\int_0^{\infty} x\cdot\exp\left(-\frac{x}{a}\right)\,dx=a\int_0^{\infty} t\,e^{-t}\,dt=a$$

and

$$E\left\{X^2\right\}=\frac{1}{2a}\cdot 2\int_0^{\infty} x^2\exp\left(-\frac{x}{a}\right)\,dx=a^2\int_0^{\infty} t^2e^{-t}\,dt=2a^2.$$

2) We shall maximize the function

$$h(a)=\prod_{i=1}^{n}\frac{1}{2a}\,\exp\left(-\frac{|x_i|}{a}\right),\qquad a>0.$$

Since not all x_i are 0, we have $h(a)\to 0$ for $a\to 0$ and for $a\to\infty$. It therefore suffices to find the maximum of

$$h_1(a)=\ln h(a)=-n\,\ln 2-n\,\ln a-\frac{1}{a}\sum_{i=1}^{n}|x_i|\,.$$

Since

$$h_1'(a)=-\frac{n}{a}+\frac{1}{a^2}\sum_{i=1}^{n}|x_i|=0$$

only for $a=\dfrac{1}{n}\sum_{i=1}^{n}|x_i|$, we get

$$a_n^\star=\frac{1}{n}\sum_{i=1}^{n}|x_i|\,,$$

with the corresponding estimator

$$Y_n=\frac{1}{n}\sum_{i=1}^{n}|X_i|\,.$$

3) Since

$$E\{Y_n\} = \frac{1}{n}\sum_{i=1}^{n} E\{|X_i|\} = \frac{1}{n}\sum_{i=1}^{n} a = a,$$

it follows that Y_n is a central estimator.

4) Since

$$V\{|X|\} = E\left\{X^2\right\} - (E\{|X|\})^2 = 2a^2 - a^2 = a^2,$$

and all the X_i are mutually independent, we get

$$V\{Y_n\} = \frac{1}{n^2}\sum_{i=1}^{n} V\{|X_i|\} = \frac{a^2}{n} = s_n^2.$$

It follows from Chebyshev's inequality that

$$P\{|Y_n - a| \geq \varepsilon\} \leq \frac{s_n^2}{\varepsilon^2} = \frac{a^2}{\varepsilon^2}\cdot\frac{1}{n} \to 0 \qquad \text{for } n \to \infty,$$

thus the sequence of estimators (Y_n) is consistent.

Example 12.13 *Let $X_1, X_2, \dots, X_{2n+1}$ be mutually independent random variables of the same distribution function $F(x)$ and frequency $f(x)$, $x \in \mathbb{R}$. The random variables $X_1, X_2, \dots, X_{2n+1}$ are sorted according to size, defining new random variables $X_1^\star, X_2^\star, \dots, X_{2n+1}^\star$, which satisfy*

$$X_1^\star \le X_2^\star \le \cdots \le X_{2n+1}^\star$$

(i.e. $X_1^\star$ is the smallest one, $X_2^\star$ the second smallest one, etc.).

1. *Find, expressed by means of $F(x)$ and $f(x)$, the distribution functions and frequencies of the random variables $X_1^\star$ and $X_{2n+1}^\star$.*

2. *Find, for $u < v$,*

$$P\left\{X_1^\star > u \,\wedge\, X_{2n+1}^\star \le v\right\},$$

and then derive the simultaneous frequency of $\left(X_1^\star, X_{2n+1}^\star\right)$.

3. *Prove that $X_{n+1}^\star$ (the "middle one") has the frequency*

$$f_{n+1}^\star = (2n+1)\binom{2n}{n}\{F(x)\}^n\{1-F(x)\}^n f(x), \qquad x \in \mathbb{R}.$$

We assume in the following that $X_1, X_2, \dots, X_{2n+1}$ are mutually independent and rectangularly distributed over the interval $]0,a[$ (where $a > 0$).

4. *Prove that the three random variables*

$$Y_{2n+1} = X_{n+1}^\star, \quad Z_{2n+1} = \frac{1}{2n+1}\sum_{i=1}^{2n+1} X_i, \quad U_{2n+1} = \frac{1}{2}\left\{X_1^\star + X_{2n+1}^\star\right\}$$

are all central estimators of the mean $\dfrac{a}{2}$.

5. *Prove that Z_{2n+1} is more efficient than Y_{2n+1}, and that U_{2n+1} is more efficient than Z_{n+1}.*

1) It follows from

$$\begin{aligned} F_{X_1^\star}(x) &= P\{X_1^\star \le x\} = P\left\{\min_{i=1,\dots,2n+1} X_i \le x\right\} = 1 - P\{X_1 > x, \dots, X_{2n+1} > x\} \\ &= 1 - \{1 - F_X(x)\}^{2n+1}, \end{aligned}$$

that the frequency is given by

$$f_{X_1^\star}(x) = (2n+1)\{1 - F_X(x)\}^{2n} f_X(x).$$

Analogously,

$$F_{X_{2n+1}^\star}(x) = P\{X_1 \le x, \dots, X_{2n+1} \le x\} = \{F_X(x)\}^{2n+1},$$

thus

$$f_{X_{2n+1}^\star}(x) = (2n+1) \cdot \{F_X(x)\}^{2n} f_X(x).$$

2) If $u < v$, then

$$
\begin{aligned}
P\left\{X_1^\star > u \,\wedge\, X_{2n+1}^\star \le v\right\} &= P\left\{u < \min X_i \,\wedge\, \max X_i \le v\right\} \\
&= P\left\{u < X_1 \le v,\, u < X_2 \le v,\, \dots,\, u < X_{2n+1} \le v\right\} \\
&= P\left\{u < X_1 \le v\right\} \cdot P\left\{u < X_2 \le v\right\} \cdots P\left\{u < X_{2n+1} \le v\right\} \\
&= \left\{F_X(v) - F_X(u)\right\}^{2n+1},
\end{aligned}
$$

and the distribution function of $\left(X_1^\star, X_{2n+1}^\star\right)$ for $u < v$ is given by

$$
\begin{aligned}
G(u,v) &= P\left\{X_1^\star \le u \,\wedge\, X_{2n+1}^\star \le v\right\} \\
&= P\left\{X_{2n+1}^\star \le v\right\} - P\left\{X_1^\star > u \,\wedge\, X_{2n+1}^\star \le v\right\} \\
&= \left\{F_X(v)\right\}^{2n+1} - \left\{F_X(v) - F_X(u)\right\}^{2n+1},
\end{aligned}
$$

and $G(u,v) = 0$ for $u \ge v$.
The simultaneous frequency $g(u,v)$ for $u < v$ is given by

$$
\begin{aligned}
g(u,v) &= \frac{\partial^2 G}{\partial u \partial v} = \frac{\partial}{\partial v}\left\{+(2n+1)\left\{F_X(v) - F_X(u)\right\}^{2n} \cdot f_X(u)\right\} \\
&= (2n+1)\cdot 2n\left\{F_X(v) - F_X(u)\right\}^{2n-1} f_X(u)\, f_X(v),
\end{aligned}
$$

and $g(u,v) = 0$ for $u \ge v$.

3) The distribution function of the “middle” random variable is

$$F_{X_{n+1}^\star}(x) = P\left\{n+1 \text{ of the } X_i \le x \text{ and the rest } > x\right\}.$$

We can choose $n+1$ variables $X_i \le x$ in $\binom{2n+1}{n+1}$ ways. If we consider a fixed set

$$\mathbf{U} = \left(X_{i(1)}, \dots, X_{i(n+1)}\right)$$

among these without caring about the remaining n variables, then we get by **1.** that

$$F_{\mathbf{U}}(x) = \left\{F_X(x)\right\}^{n+1} \qquad \text{with } f_{\mathbf{U}}(x) = (n+1)\left\{F_X(x)\right\}^n f_X(x).$$

For the full system we get a conditional frequency (conditional, because the remaining n variables are $> x$),

$$f_{\mathbf{U}}^\star(x) = f_{\mathbf{U}}(x) \cdot \{1 - F(x)\}^n.$$

Notice that $\mathbf{U}$ and $\complement\mathbf{U}$ play different roles; only $\mathbf{U}$ is referring directly to $\le x$.

When we collect all contributions, we get

$$
\begin{aligned}
f_{X_{n+1}^\star}^\star(x) &= \binom{2n+1}{n+1} \cdot (n+1) \cdot \left\{F_X(x)\right\}^n f_X(x) \cdot \left\{1 - F_X(x)\right\}^n \\
&= (2n+1)\binom{2n}{n} \cdot \left\{F_X(x)\right\}^n \left\{1 - F_X(x)\right\}^n f_X(x), \qquad x \in \mathbb{R}.
\end{aligned}
$$

Alternatively, start by isolating $X_{n+1}^{\star}$. This random variable can be chosen in $2n+1$ ways. There are $\binom{2n}{n}$ ways to choose $X_1^{\star}, \dots, X_n^{\star}$, and then $X_{n+2}^{\star}, \dots, X_{2n+1}^{\star}$ are given. Thus

$$f_{X_{n+1}^{\star}}(x)\,dx \approx P\left\{X_{n+1}^{\star} \in]x, x+dx]\right\} \approx (2n+1)\binom{2n}{n}\{F_X(x)\}^n\{1-F_X(x)\}^n f_X(x)\,dx,$$

and the result follows.
By isolating $x < X_{n+1}^{\star} \le x+dx$ it becomes more clear what in the first version is meant by "conditional probability".

4) When X is rectangularly distributed over $]0,a[$, then

$$f_X(x) = \begin{cases} \dfrac{1}{a} & \text{for } x \in]0,a[, \\ 0 & \text{ellers,} \end{cases}$$

and

$$F_X(x) = \begin{cases} 0 & \text{for } x \le 0, \\ \dfrac{x}{a} & \text{for } 0 < x < a, \\ 1 & \text{for } x > a. \end{cases}$$

Let $0 < x < a$. By insertion,

$$\begin{aligned} f_{Y_{2n+1}}(x) &= f_{X_{n+1}^{\star}}(x) = (2n+1)\binom{2n}{n}\left\{\frac{x}{a}\right\}^n\left\{1-\frac{x}{a}\right\}^n \cdot \frac{1}{a} \\ &= \frac{1}{B(n+1,n+1)} \cdot \frac{1}{a}\left\{\frac{x}{a}\right\}^n\left\{1-\frac{x}{a}\right\}^n, \end{aligned}$$

(a Beta distribution), so

$$\begin{aligned} E\{Y_{2n+1}\} &= \frac{1}{B(n+1,n+1)}\int_0^a x\left\{\frac{x}{a}\right\}^n\left\{1-\frac{x}{a}\right\}^n \cdot \frac{1}{a}\,dx \\ &= \frac{a}{B(n+1,n+1)}\int_0^1 t^{n+1}(1-t)^n dt \\ &= a \cdot \frac{B(n+2,n+1)}{B(n+1,n+1)} = a \cdot \frac{(n+1)!\,n!}{(2n+2)!} \cdot \frac{(2n+1)!}{n!\,n!} = \frac{a}{2}, \end{aligned}$$

proving that Y_{n+1} is central.
Then we get that Z_{n+1} is central, because

$$E\{Z_{2n+1}\} = \frac{1}{2n+1}\sum_{i=1}^{2n+1} E\{X_i\} = \frac{a}{2}.$$

Finally,

$$\begin{aligned}
E\left\{U_{2n+1}\right\} &= \frac{1}{2}\,E\left\{X_1^{\star}\right\}+\frac{1}{2}\,E\left\{X_{2n+1}^{\star}\right\}\\
&= \frac{1}{2}\,(2n+1)\int_0^a x\left\{1-\frac{x}{a}\right\}^{2n}\,\frac{1}{a}\,dx+\frac{1}{2}\,(2n+1)\int_0^a x\left\{\frac{x}{a}\right\}^{2n}\,\frac{1}{a}\,dx\\
&= \frac{a}{2}\,(2n+1)\left\{\int_0^1 t(1-t)^{2n}\,dt+\int_0^1 t^{2n+1}\,dt\right\}\\
&= \frac{a}{2}\,(2n+1)\left\{\int_0^1 (1-u)u^{2n}\,du+\int_0^1 t^{2n+1}\,dt\right\}\\
&= \frac{a}{2}\,(2n+1)\int_0^1 u^{2n}\,du=\frac{a}{2},
\end{aligned}$$

proving that U_{2n+1} is also central.

5) We shall compute $V\left\{Y_{2n+1}\right\}$, $V\left\{Z_{2n+1}\right\}$ and $V\left\{U_{2n+1}\right\}$, and compare these expressions. The easiest computation is

$$V\left\{Z_{2n+1}\right\}=V\left\{\frac{1}{2n+1}\sum_{i=1}^{2n+1}X_i\right\}=\frac{1}{2n+1}\cdot\frac{a^2}{12}.$$

Then

$$\begin{aligned}E\left\{Y_{2n+1}^2\right\} &= (2n+1)\left(\begin{array}{c}2n\\ n\end{array}\right)\int_0^a x^2\left\{\frac{x}{a}\right\}^n\left\{1-\frac{x}{a}\right\}^n\frac{1}{a}\,dx\\ &= (2n+1)\left(\begin{array}{c}2n\\ n\end{array}\right)a^2\int_0^1 t^{n+2}(1-t)^n\,dt\\ &= a^2\cdot\frac{(2n+1)!}{n!\,n!}\cdot B(n+3,n+1)=a^2\cdot\frac{(2n+1)!}{n!\,n!}\cdot\frac{(n+2)!n!}{(2n+3)!}\\ &= a^2\,\frac{(n+2)(n+1)}{(2n+3)(2n+2)}=\frac{a^2}{2}\cdot\frac{n+2}{2n+3}\end{aligned}$$

implies that

$$V\left\{Y_{2n+1}\right\}=\frac{a^2}{2}\cdot\frac{n+2}{2n+3}-\frac{a^2}{4}=\frac{a^2}{4}\cdot\frac{1}{2n+3}.$$

From

$$V\left\{Z_{2n+1}\right\}=\frac{1}{2n+1}\cdot\frac{a^2}{12}=\frac{1}{6n+3}\cdot\frac{a^2}{4}<\frac{1}{2n+3}\cdot\frac{a^2}{4}=V\left\{Y_{2n+1}\right\},$$

follows that Z_{2n+1} is more efficient than Y_{2n+1}.

We still have to compute $V\left\{U_{2n+1}\right\}$. It follows from

$$U_{2n+1}^2=\frac{1}{4}\left\{X_1^{\star 2}+X_{2n+1}^{\star 2}+2X_1^{\star}X_{2n+1}^{\star}\right\},$$

and

$$\left(E\left\{U_{2n+1}\right\}\right)=\frac{a^2}{4},$$

that

$$V\left\{U_{2n+1}\right\}=\frac{1}{4}\left(E\left\{X_1^{\star 2}\right\}+E\left\{X_{2n+1}^{\star 2}\right\}+2\,E\left\{X_1^{\star}X_{2n+1}^{\star 2}\right\}-a^2\right).$$

The former two terms are simplest, cf. **1.**,

$$\begin{aligned}E\left\{X_1^{\star 2}\right\} &= (2n+1)\int_0^a x^2\left\{1-\frac{x}{a}\right\}^{2n}\cdot\frac{1}{a}\,dx\\ &= a^2(2n+1)\int_0^1 t^2(1-t)^{2n}\,dt=a^2(2n+1)B(3,2n+1)\\ &= a^2(2n+1)\cdot\frac{2!(2n)!}{(2n+3)!}=\frac{a^2}{(2n+3)(n+1)!},\end{aligned}$$

and

$$E\left\{X_{2n+1}^{\star 2}\right\} = (2n+1)\int_0^a x^2 \left\{\frac{x}{a}\right\}^{2n} \frac{1}{a}\,dx = a^2(2n+1)\int_0^1 t^{2n+2}\,dt = \frac{2n+1}{2n+3}\,a^2.$$

According to **2.**, the simultaneous frequency of $\left(X_1^\star, X_{2n+1}^\star\right)$ is

$$g(u,v) = \begin{cases} 2n(2n+1)\left\{\dfrac{v}{a}-\dfrac{u}{a}\right\}^{2n-1}\cdot\dfrac{1}{a^2} & \text{for } 0<u<v<a, \\[2mm] 0 & \text{otherwise.}\end{cases}$$

Then

$$\begin{aligned}
E\left\{X_1^\star X_{2n+1}^\star\right\} &= \int_0^a \left\{\int_0^v 2n(2n+1)\left\{\frac{v-u}{a}\right\}^{2n-1}\cdot uv\cdot\frac{1}{a^2}\,du\right\}dv \\
&= a^2\int_0^1\left\{\int_0^v 2n(2n+1)(v-u)^{2n-1}\cdot uv\,du\right\}dv \\
&= a^2 2n(2n+1)\int_0^1 v\left\{\int_0^v \left\{v(v-u)^{2n-1}-(v-u)^{2n}\right\}\,du\right\}dv \\
&= a^2 2n(2n+1)\int_0^1 v\left\{\int_0^v \left(vt^{2n-1}-t^{2n}\right)\,dt\right\}dv \\
&= a^2 2n(2n+1)\int_0^1 v\,\left[\frac{v^{2n+1}}{2n}-\frac{v^{2n+1}}{2n+1}\right]\,dv \\
&= a^2\int_0^1 v^{2n+2}\,dv = \frac{a^2}{2n+3}.
\end{aligned}$$

By insertion,

$$\begin{aligned}
V\left\{U_{2n+1}\right\} &= \frac{a^2}{4}\cdot\left\{\frac{1}{(2n+3)(n+1)}+\frac{2n+1}{2n+3}+\frac{2}{2n+3}-1\right\} \\
&= \frac{1}{(2n+3)(n+1)}\cdot\frac{a^2}{4}.
\end{aligned}$$

Since

$$(2n+3)(n+1) = 2n^2+5n+3 > 6n+3 = 3(2n+1),$$

we get $V\left\{U_{2n+1}\right\} < V\left\{Z_{2n+1}\right\}$, hence U_{2n+1} is more efficient than Z_{2n+1}.

Example 12.14 *A random variable X has the frequency*

$$f(x)=\frac{|x|}{2a^2}\,\exp\left(-\frac{|x|}{a}\right),\qquad x\in\mathbb{R},$$

where a is a positive constant.

1) *Compute $E\{X\}$ and $E\{|X|\}$.*

2) *Assuming that n observations of the mutually independent random variables $X_1, X_2, \dots, X_n$, all distributed like X above, have given the results $x_1, x_2, \dots, x_n$ (where all x_i are different from 0), apply the maximum method to find the best estimate $a_n^\star$ of a and hence also the best estimate $b_n^\star$ of $E\{|X|\}$.*

3) *Prove that the estimator Y_n corresponding to the estimate $b_n^\star$ is a central estimator.*

4) *We perform 10 observations. The results are*

$$-3.3,\quad -2.6,\quad -3.6,\quad 3.0,\quad 3.2,\quad -3.1,\quad 3.5,\quad -2.7,\quad 2.6,\quad -2.4.$$

What is the best estimate of $E\{|X|\}$ in this case?

1) It follows by the symmetry that $E\{X\}=0$, and

$$E\{|X|\}=2\cdot\frac{1}{2a^2}\int_0^\infty x^2\exp\left(-\frac{x}{a}\right)\,dx=a\int_0^\infty t^2e^{-t}\,dt=2a.$$

2) We shall maximize

$$h_n(a)=\prod_{i=1}^n\frac{|x_i|}{2a^2}\,\exp\left(-\frac{|x_i|}{a}\right)=\frac{1}{2^n}\cdot\frac{1}{a^{2n}}\prod_{i=1}^n|x_i|\cdot\exp\left(-\frac{1}{a}\sum_{i=1}^n|x_i|\right).$$

Clearly, $h_n(a)>0$, and $h_n(a)\to 0$ for $a\to 0+$, and for $a\to\infty$.
We conclude from

$$\ln h_n(a)=-n\,\ln 2-\sum_{i=1}^n\ln|x_i|-2n\,\ln a-\frac{1}{a}\sum_{i=1}^n|x_i|,$$

that

$$\frac{h_n'(a)}{h_n(a)}=-\frac{2n}{a}+\frac{1}{a^2}\sum_{i=1}^n|x_i|=\frac{1}{a^2}\left\{\sum_{i=1}^n|x_i|-2n\,a\right\},$$

so the unique maximum (the best estimate of a) is

$$a_n^\star=\frac{1}{2n}\sum_{i=1}^n|x_i|.$$

The best estimate $b_n^\star$ of $E\{|X|\}=2a$ is

$$b_n^\star=\frac{1}{n}\sum_{i=1}^n|x_i|.$$

3) The estimator corresponding to $b_n^\star$ for $E\{|X|\}$ is

$$Y_n = \frac{1}{n}\sum_{i=1}^n |X_i|\,.$$

We get from

$$E\{Y_n\} = \frac{1}{n}\sum_{i=1}^n E\{|X_i|\} = \frac{1}{n}\cdot n\cdot E\{|X|\} = E\{|X|\},$$

that Y_n is a central estimator.

4) With the given measurements, the best estimate of $E\{|X|\}$ is

$$b_{10}^\star = \frac{1}{10}\{3.3+2.6+3.6+3.0+3.2+3.1+3.5+2.7+2.6+2.4\} = \frac{30.0}{10} = 3.$$

Example 12.15 *A random variable X has its distribution given by*

$$P\{X=k\}=(k+1)p^2(1-p)^k, \qquad k\in\mathbb{N}_0,$$

where p is a constant, $0<p<1$, i.e. $X\in$ NB$(2,p)$.

1) Find the mean μ and variance σ^2 of the random variable X.

2) Assuming that n observations of the mutually independent random variables $X_1, X_2, \dots, X_n$, all distributed like X above, have given the results $x_1, x_2, \dots, x_n$ (where not all x_i are equal to 0), apply the maximal method to find the best estimate of p and the best estimate $\mu_n^\star$ of μ.

3) Prove that the estimator Y_n corresponding to the estimate $\mu_n^\star$ is a central estimator of μ.

4) We perform 10 observations of the results

$$4, \quad 5, \quad 1, \quad 7, \quad 0, \quad 4, \quad 2, \quad 10, \quad 3, \quad 4.$$

What is the best estimate of p in this case?

1) Since $X\in\text{NB}(2,p)$, we get from a table that

$$\mu=2\cdot\frac{q}{p}=2\left(\frac{1}{p}-1\right) \quad\text{and}\quad \sigma^2=2\,\frac{q}{p^2} \quad \left(=2\left(\frac{1}{p^2}-\frac{1}{p}\right)\right).$$

2) We shall maximize the function

$$h_n(p)=\prod_{i=1}^n P\left\{X_i=x_i\right\}=\prod_{i=1}^n\left(x_i+1\right)p^2(1-p)^{x_i}.$$

Clearly,

$$h_n(p)>0 \qquad \text{for } 0<p<1,$$

and $h_n(p)\to 0$ for $p\to 0+$, and for $p\to 1-$, where we have assumed that not every x_i is 0.
It follows from

$$\ln h_n(p)=\sum_{i=1}^n \ln\left(x_i+1\right)+2n\,\ln p+\ln(1-p)\cdot\sum_{i=1}^n x_i,$$

that

$$\frac{h_n'(p)}{h_n(p)}=\frac{2n}{p}-\frac{1}{1-p}\sum_{i=1}^n x_i.$$

This expression is equal to 0 for

$$2n\left(1-p_n^\star\right)=p_n^\star\sum_{i=1}^n x_i,$$

thus for

$$p_n^\star = \frac{2n}{2n + \sum_{i=1}^n x_i} = \frac{2}{2+\overline{x}},$$

where $\overline{x} = \dfrac{1}{n}\sum_{i=1}^n x_i$.
The best estimate of μ is

$$\mu_n^\star = 2\left(\frac{1}{p_n^\star} - 1\right) = 2 + \frac{1}{n}\sum_{i=1}^n x_i - 2 = \overline{x}.$$

3) The estimator Y_n is

$$Y_n = \frac{1}{n}\sum_{i=1}^n X_i,$$

and it is obvious that

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

hence the estimator is central.

4) Since

$$\frac{1}{10}\sum_{i=1}^{10} x_i = \frac{1}{10}\left(4+5+1+7+0+4+2+10+3+4\right) = \frac{40}{10} = 4,$$

the best estimate of p is given by

$$p_n^\star = \frac{2}{2 + \frac{1}{10}\sum_{i=1}^{10} x_i} = \frac{2}{2+4} = \frac{1}{3}.$$

Index