APPENDIX A

Probability and Statistics

The basics of probability and statistics are presented in this appendix, providing the reader with a reference source for the main body of this book without interrupting the flow of argumentation there with unnecessary excursions. Only the basic definitions and results have been included in this appendix. More complicated concepts (such as stochastic processes, Ito Calculus, Girsanov Theorem, etc.) are discussed directly in the main body of the book as the need arises.

A.1 PROBABILITY, EXPECTATION, AND VARIANCE

A variable whose value is dependent on *random events*, is referred to as a *random variable* or a *random number*. An example of a random variable is the number observed when throwing a dice. The distribution of this random number is an example of a *discrete* distribution. A random number is discretely distributed if it can take on only discrete values such as whole numbers, for instance. A random variable has a *continuous* distribution if it can assume arbitrary values in some interval (which can by all means be the infinite interval from $-\infty$ to ∞). Intuitively, the values which can be assumed by the random variable lie together *densely*, i.e., arbitrarily close to one another.

An example of a discrete distribution on the interval from 0 to ∞ (infinity) might be a random variable which could take on the values 0.00, 0.01, 0.02, 0.03, . . ., 99.98, 99.99, 100.00, 100.01, 100.02, . . ., etc., like a Bund future with a tick-size of 0.01% corresponding to a value change of 10 euros per tick on a nominal of 100,000 euros. In contrast, a continuously distributed random variable can take on arbitrary values in the pertinent interval, for example, $\sqrt{5}$, π, or $7/3$.

The *probability* $P(x < a)$ that a random variable x will be less than some arbitrary number a is the sum of the probabilities of all events in which

x takes on values less than a. For continuous distributions, for which the possible values of x lie "infinitely dense" in some interval, the sum takes on the form of an integral:

$$P(x < a) = \sum_{i \text{ where } x_i < a} p(x_i) \longrightarrow \int_{-\infty}^{a} p(x)dx \tag{A.1}$$

The function p is called the *probability density* of the random variable x. It is often referred to as the *probability distribution*, *distribution density*, or simply *the distribution*. In this text, the abbreviation pdf for *probability density function* will frequently be used in reference to the function p.

The function P is called the *cumulative probability*. It is often referred to as the *cumulative probability distribution* or simply *cumulative distribution*. We will frequently use the abbreviation cpf for *cumulative probability function* in reference to P.

It is certain that a random number will take on *some* value. The probability of something at all happening is thus equal to one. This property is called the *normalization to one* and holds for all probability distributions:

$$1 = \sum_i p(x_i) \longrightarrow \int_{-\infty}^{\infty} p(x)dx = 1 \tag{A.2}$$

The *expectation* of x is computed by taking the weighted sum of all possible values taken on by the random variable where the weights are the corresponding probabilities belonging to those values

$$\mathrm{E}[x] = \sum_i x_i p(x_i) \longrightarrow \int_{-\infty}^{\infty} xp(x)dx \tag{A.3}$$

A function f of a random variable x is again a random variable. The expectation of this function is calculated analogously by taking the weighted sum of the values of the function $f(x)$ evaluated at all possible values of x. The weights are again the probabilities belonging to the values of x:

$$\mathrm{E}\left[f(x)\right] = \sum_i f(x_i)p(x_i) \longrightarrow \int_{-\infty}^{\infty} f(x)p(x)dx \tag{A.4}$$

A special case is particularly interesting: setting $f(x)$ equal to the square of the deviation of x from its expected value, i.e., $f(x) = (x-\mathrm{E}[x])^2$, measures how strongly x fluctuates around its expected value. This measure is called the *variance*.

$$\begin{aligned}\operatorname{var}[x] &= \mathrm{E}\left[(x-\mathrm{E}[x])^2\right] \\ &= \sum_i (x_i - \mathrm{E}[x])^2 p(x_i) \longrightarrow \int_{-\infty}^{\infty} (x-\mathrm{E}[x])^2 p(x)dx \end{aligned} \tag{A.5}$$

The square root of the variance is called the *standard deviation*, abbreviated as std.

$$\operatorname{std}[x] := \sqrt{\operatorname{var}[x]} = \sqrt{\mathrm{E}\left[(x-\mathrm{E}[x])^2\right]} \tag{A.6}$$

For both discrete and continuous distributions, there exists a simple connection between the variance and the expectation: the variance is equal to the difference between the *expectation of the square* of the random variable and the *square of the expectation* of the same random variable:

$$\operatorname{var}[x] = \mathrm{E}\left[x^2\right] - \mathrm{E}[x]^2 \tag{A.7}$$

The derivation is presented here for discrete distributions; the proof for the continuous case is completely analogous:

$$\begin{aligned}\mathrm{E}\left[(x-\mathrm{E}[x])^2\right] &= \sum_i (x_i - \mathrm{E}[x])^2 p(x_i) \\ &= \underbrace{\sum_i x_i^2 p(x_i)}_{\mathrm{E}[x^2]} - 2\mathrm{E}[x] \underbrace{\sum_i x_i p(x_i)}_{\mathrm{E}[x]} + \mathrm{E}[x]^2 \underbrace{\sum_i p(x_i)}_{1} \\ &= \mathrm{E}\left[x^2\right] - 2\mathrm{E}[x]^2 + \mathrm{E}[x]^2 = \mathrm{E}\left[x^2\right] - \mathrm{E}[x]^2 .\end{aligned}$$

A.2 MULTIVARIATE DISTRIBUTIONS, COVARIANCE, CORRELATION, AND BETA

Two random variables x and y which are not statistically independent (for example, the price of a Siemens share and the DAX) are said to be *correlated*. The probability $P(x < a, y < b)$ that a random variable x will be less than some value a and *simultaneously* that the second random variable y will be less than a value b equals the sum of the probabilities of all events in which $x < a$ and $y < b$. For continuous distributions, i.e., for infinitely dense values of the random variables, the sums converge to integrals:

$$P(x < a, y < b) = \sum_{x_i < a} \sum_{y_j < b} p(x_i, y_j) \longrightarrow \int_{-\infty}^{a} dx \int_{-\infty}^{b} dy\, p(x,y) \tag{A.8}$$

The function p is in this case the *joint probability density* for a pair of random variables x and y. Since such density functions refer to more than one (in this case two) random variables, they are referred to as *multivariate probability densities* or *multivariate probability distributions*, *multivariate distribution densities*, or simply *multivariate distributions*.

Just as for a single random variable, the expectation of an arbitrary function $f(x,y)$ is calculated by taking the weighted sum of the values of the function $f(x,y)$ evaluated at all possible values of x and y. The weights are the joint probabilities belonging to the value pairs of (x,y):

$$\mathrm{E}\left[f(x,y)\right] = \sum_i \sum_j p(x_i,y_j) f(x_i,y_j) \longrightarrow \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, p(x,y) f(x,y) \tag{A.9}$$

A special case is particularly interesting: setting the arbitrary function $f(x,y)$ equal to the product of the deviation of the random variables from their respective expectations, i.e., $f(x,y) = (x - \mathrm{E}[x])(y - \mathrm{E}[y])$, yields a measure for the fluctuation of each random variable from its expectation as well as the degree of statistical dependence between the two random variables; or in other words, a "simultaneous" measure for the variance and the correlation. This expectation is called the *covariance* between x and y.

$$\begin{aligned} \mathrm{cov}[x,y] &= \mathrm{E}\left[(x - \mathrm{E}[x])\,(y - \mathrm{E}[y])\right] \\ &= \sum_{i,j} (x_i - \mathrm{E}[x])\left(y_j - \mathrm{E}[y]\right) p(x_i,y_j) \\ &\longrightarrow \int\int dx dy p(x,y)\,(x - \mathrm{E}[x])\,(y - \mathrm{E}[y]) \end{aligned} \tag{A.10}$$

For both discrete and continuous distributions, there exists a natural extension (which is just as easily shown) of the relation in Equation A.7 between the variance and the expectations: the covariance is the difference between the expectation of the product and the product of the expectations of the random variables under consideration:

$$\mathrm{cov}[x,y] = \mathrm{E}[xy] - \mathrm{E}[x]\,\mathrm{E}[y] \tag{A.11}$$

The symmetry property of the covariance can be seen immediately from this equation; the covariance of x with y is the same as the covariance of y with x. The covariance of a random variable with itself is its variance.

$$\mathrm{cov}[x,y] = \mathrm{cov}[y,x], \quad \mathrm{cov}[x,x] = \mathrm{var}[x] \tag{A.12}$$

The covariance has an additional useful property, it is bilinear: let a, b, c, d be constants and x, y, u, z random variables. Then

$$\begin{aligned}\mathrm{cov}[ax + by, cu + dz] = ac\mathrm{cov}[x, u] + ad\mathrm{cov}[x, z] \\ + bc\mathrm{cov}[y, u] + bd\mathrm{cov}[y, z]\end{aligned} \tag{A.13}$$

As often mention in the main body of this book, the covariances between n random variables can be represented in the form of an n by n matrix, called the *covariance matrix*. Because of the symmetry in Equation A.12, the information in the matrix entries appearing above the diagonal is the same as that below the diagonal; the entries in the diagonal itself are the variances of each of the n respective variables.

Dividing the covariance by the standard deviations of the two respective random variables yields the *correlation coefficient* ρ, also called the *correlation* between x and y. This value always lies between -1 and $+1$:

$$\rho(x, y) := \frac{\mathrm{cov}[x, y]}{\sqrt{\mathrm{var}[x]\mathrm{var}[y]}} = \frac{\mathrm{E}[xy] - \mathrm{E}[x]\,\mathrm{E}[y]}{\sqrt{(\mathrm{E}\left[x^2\right] - \mathrm{E}[x]^2)(\mathrm{E}\left[y^2\right] - \mathrm{E}[y]^2)}} \tag{A.14}$$

The symmetry property of the correlation can be seen immediately from this equation: The correlation of x with y is the same as the correlation of y with x. The correlation of a random variable with itself equals one.

$$\rho(x, y) = \rho(y, x), \quad \rho(x, x) = 1 \tag{A.15}$$

Two random numbers x_1 and x_2 are called *uncorrelated* if the correlation between the two is zero, i.e., if

$$\rho(x_i, x_j) = \delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}, \quad i, j \in \{1, 2\} \tag{A.16}$$

where δ_{ij} denotes the well-known *Kronecker delta*.

The *variance of a sum* of random variables is just the sum of the *covariances* of these random variables (in the same way as the expectation of a sum equals the sum of the expectations):

$$\begin{aligned}\mathrm{E}\left[\sum_i x_i\right] &= \sum_i \mathrm{E}[x] \\ \mathrm{var}\left[\sum_i x_i\right] &= \sum_{i,j} \mathrm{cov}[x_i, x_j] = \sum_{i,j} \rho\left(x_i, x_j\right)\sqrt{\mathrm{var}[x_i]\mathrm{var}[x_j]}\end{aligned} \tag{A.17}$$

The first of these equations, i.e., the linearity of the expectation, follows directly from the definition, Equation A.3. This linearity of the expectation is the only thing needed to show the second equation, i.e., the equation for the variance of a sum:

$$\begin{aligned}
&\mathrm{var}\left[\sum_i x_i\right] \\
&= \mathrm{E}\left[\left(\sum_i x_i - \mathrm{E}\left[\sum_j x_j\right]\right)^2\right] = \mathrm{E}\left[\left(\sum_i x_i - \sum_j \mathrm{E}\left[x_j\right]\right)^2\right] \\
&= \mathrm{E}\left[\left(\sum_i (x_i - \mathrm{E}\left[x_i\right])\right)^2\right] = \mathrm{E}\left[\sum_{i,j} (x_i - \mathrm{E}\left[x_i\right])\left(x_j - \mathrm{E}\left[x_j\right]\right)\right] \\
&= \sum_{i,j} \mathrm{E}\left[(x_i - \mathrm{E}\left[x_i\right])\left(x_j - \mathrm{E}\left[x_j\right]\right)\right] = \sum_{i,j} \mathrm{cov}[x_i, x_j].
\end{aligned}$$

The first step is just the definition of the variance, Equation A.5. The linearity of the expectation is used in the second step for the inner expectation and in the fifth step for the outer expectation.

A direct application of Equation A.17 is the fact that the variance of a sum of *uncorrelated* random numbers (see Equation A.16) is just the sum of the individual variances:

$$\mathrm{var}\left[\sum_i x_i\right] = \sum_{i,j} \underbrace{\rho\left(x_i, x_j\right)}_{\delta_{ij}} \sqrt{\mathrm{var}[x_i]\mathrm{var}[x_j]} = \sum_i \mathrm{var}[x_i] \tag{A.18}$$

Thus, the standard deviation, Equation A.6, of such a sum of *uncorrelated* random numbers is obtained by adding the squares of each individual standard deviation and then taking the square root of this sum:

$$\mathrm{std}\left[\sum_i x_i\right] = \sqrt{\sum_i \mathrm{var}[x_i]} = \sqrt{\sum_i \mathrm{std}[x_i]^2} \tag{A.19}$$

Besides the above two symmetric quantities (covariance and correlation), a further *a*symmetric quantity is quite useful as well. This is the *beta* of y with respect to x defined as the covariance of x and y divided by the variance of x:

$$\beta(x, y) = \frac{\mathrm{cov}[x, y]}{\mathrm{var}[y]} = \sqrt{\frac{\mathrm{var}[x]}{\mathrm{var}[y]}}\rho\left(x, y\right) = \frac{\mathrm{E}[xy] - \mathrm{E}\left[x\right]\mathrm{E}[y]}{\mathrm{E}\left[y^2\right] - \mathrm{E}[y]^2} \tag{A.20}$$

A symmetry property can be seen immediately in this definition: the beta of a random variable with itself is indeed equal to one, the beta of x with respect to y is however not the same as the beta of y with respect to x. The conversion can be accomplished as follows:

$$\beta(y,x) = \frac{\text{var}[y]}{\text{var}[x]}\beta(x,y), \quad \beta(x,x) = 1 \tag{A.21}$$

Note that all of the above equations (in particular the very useful Equation A.17 and the properties A.18 and A.19 of uncorrelated variables) have been derived directly from first principles and are therefore valid no matter what probability distribution the random variables may have.

A.3 MOMENTS AND CHARACTERISTIC FUNCTIONS

The expectation and the variance are examples of the *moments* of a distribution. In general, the n-th moment of the distribution of the random variable x is defined as the expectation of the n^{th} power of the random variable:

$$\mathrm{E}\left[x^n\right] = \sum_i x_i^n p(x_i) \longrightarrow \int_{-\infty}^{\infty} x^n p(x)dx \tag{A.22}$$

The *central moments* μ_i of a distribution are defined as the "expectation of the powers of the difference between a random variable and its expectation":

$$\mu_j := \mathrm{E}[(x - \mathrm{E}[x])^j] \tag{A.23}$$

The first central moment ($j = 1$) is thus by definition equal to zero. The *expectation* of x is the first moment, the *variance* the *second central* moment of the distribution. With the third central moment $\mathrm{E}[(x - \mathrm{E}[x])^3]$ we can calculate the *skewness*, a measure for the asymmetry of the density. With the fourth central moment $\mathrm{E}[(x - \mathrm{E}[x])^4]$ we can calculate the *kurtosis*, a measure for the weight of the distribution at the tail ends of its range. The exact definitions of the skewness and kurtosis are

$$\begin{aligned} \text{Skewness} &:= \frac{\mu_3}{\sqrt{\mu_2^3}} = \frac{\mathrm{E}[(x - \mathrm{E}[x])^3]}{\mathrm{E}[(x - \mathrm{E}[x])^2]^{3/2}} \\ \text{kurtosis} &:= \frac{\mu_4}{\mu_2^2} = \frac{\mathrm{E}[(x - \mathrm{E}[x])^4]}{\mathrm{E}[(x - \mathrm{E}[x])^2]^2} \end{aligned} \tag{A.24}$$

A.3.1 Moment generating functions

The moment generating function is a very useful tool for the explicit computation of moments. The *moment generating function* (in short *MGF*) of a random variable x with density function pdf(x) is defined as the expectation of e^{sx} for an arbitrary real value s

$$G_x(s) = \mathrm{E}[e^{sx}] = \int_{-\infty}^{\infty} e^{sx}\,\mathrm{pdf}(x)dx \tag{A.25}$$

if this integral exists. This corresponds to the *Laplace transformation* of the pdf. Expanding the exponential function e^{sx} in its Taylor series, we see that the coefficient of s^n is determined by the n^{th} moment of the distribution:

$$\begin{aligned} G_x(s) &= \int_{-\infty}^{\infty} \underbrace{\sum_{n=0}^{\infty} \frac{1}{n!} s^n x^n}_{e^{sx}} \mathrm{pdf}(x)dx \\ &= \sum_{n=0}^{\infty} \frac{s^n}{n!} \int_{-\infty}^{\infty} x^n\,\mathrm{pdf}(x)dx \\ &= \sum_{n=0}^{\infty} \frac{s^n}{n!} \mathrm{E}[x^n] \end{aligned} \tag{A.26}$$

Differentiating the moment generating function with respect to s at the point $s = 0$ yields all moments of the distribution (and thus the name):

$$\left.\frac{\partial^n G_x(s)}{\partial s^n}\right|_{s=0} = \mathrm{E}[x^n] \tag{A.27}$$

This extraordinarily useful fact can be shown as follows:

$$\begin{aligned} \left.\frac{\partial^n G_x(s)}{\partial s^n}\right|_{s=0} &= \left.\frac{\partial^n}{\partial s^n} \sum_{i=0}^{\infty} \frac{1}{i!} s^i \mathrm{E}[x^i]\right|_{s=0} \\ &= \left.\frac{1}{i!} \frac{\partial^{n-1}}{\partial s^{n-1}} \sum_{i=0}^{\infty} \mathrm{E}[x^i] \frac{\partial s^i}{\partial s}\right|_{s=0} \\ &= \left.\frac{1}{i!} \frac{\partial^{n-1}}{\partial s^{n-1}} \sum_{i=1}^{\infty} \mathrm{E}[x^i] i s^{i-1}\right|_{s=0} \\ &= \left.\frac{1}{i!} \frac{\partial^{n-2}}{\partial s^{n-2}} \sum_{i=2}^{\infty} \mathrm{E}[x^i] i(i-1) s^{i-2}\right|_{s=0} \\ &= \cdots \\ &= \left.\frac{1}{i!} \sum_{i=n}^{\infty} \mathrm{E}[x^i] i(i-1)\cdots(i-n+1) s^{i-n}\right|_{s=0} . \end{aligned}$$

Notice that the lower limit in the sum increases by one each time a derivative is taken. The first term of the sum is always independent of the differentiating variable. For $i=0$, for example, $\partial s^i/\partial s = \partial s^0/\partial s = \partial 1/\partial s = 0$, and so on. In the last step, all derivatives have been performed. The expression can now be evaluated at $s=0$. Naturally, $s^{i-n}=0$ for $s=0$ and $i>n$. Thus, *only* the first summand where $i = n$ makes a contribution to the sum since in this term we have $s^{i-n} = s^0 = 1$. Setting $i = n$ in this term immediately yields Equation A.27.

The *central* moments defined in Equation A.23 can likewise be calculated using the moment generating function: If a random number x has a distribution pdf(x) with expectation $\mu = \mathrm{E}[x]$, then the moments of the random number $\widetilde{x} := x - \mu$ are exactly equal to the *central* moments of x. But for the moments of $\widetilde{x}$ we can use the MGF of the distribution of $\widetilde{x}$.

$$\begin{aligned} G_{\widetilde{x}}(s) &= \int_{-\infty}^{\infty} e^{s\widetilde{x}}\, \mathrm{pdf}(\widetilde{x}) d\widetilde{x} = \int_{-\infty}^{\infty} e^{s\widetilde{x}}\, \mathrm{pdf}(x) d\widetilde{x} \\ &= \int_{-\infty}^{\infty} e^{s(x-\mu)}\, \mathrm{pdf}(x) dx = e^{-s\mu} \int_{-\infty}^{\infty} e^{sx}\, \mathrm{pdf}(x) dx. \end{aligned}$$

The first step is just Definition A.25. For the second step we made use of the fact that if the difference between two random numbers is just a constant they must have the same distribution.[1] In the third step we used the fact that $\widetilde{x} = x - \mu$ for the differentials $d\widetilde{x} = dx$. Thus the MGF for the central moments is simply $e^{-s\mu}$ times the MGF for the (ordinary) moments where μ denotes the first (ordinary) moment:

$$G_{x-\mu}(s) = e^{-s\mu} G_x(s) \tag{A.28}$$

With this MGF the central moments defined in Equation A.23 can be calculated completely analogously to Equation A.27:

$$\mathrm{E}[(x - \mathrm{E}[x])^n] = \left. \frac{\partial^n}{\partial s^n} \exp\left(-s\mathrm{E}[x]\right) G_x(s) \right|_{s=0} \tag{A.29}$$

The general procedure for calculating central moments is thus: first calculate the expectation using Equation A.27. Then insert the result into Equation A.29 for the central moments.

For many distributions an explicit analytical expression for the MGF can be obtained using the integral representation in Equation A.25. This will be demonstrated below for several important distributions. Having obtained

1 This is trivial: if $\widetilde{x} = x - \mu$ then the probability for $\widetilde{x} < a - \mu$ is of course the same as the probability for $x < a$.

such an expression, the moments can be calculated by simply differentiating this function as indicated in Equation A.27.

The MGF has another very useful property: If two random variables x and y are independent then

$$G_{x+y}(s) = G_x(s)G_y(s) \tag{A.30}$$

The *distribution* of a *sum* of independent random variables is generally very difficult to determine, even when the distributions of the individual random variables in the sum are known.[2] The *MGF* of such a sum, in contrast, can be calculated quite easily by taking the product of the MGFs of each of the distributions! This is the most useful property of the moment generating function. In Equation A.30, each of the random variables in the sum can by all means be governed by completely *different* distributions. The only condition which needs to be satisfied for Equation A.30 to hold is the statistical *independence* of the random variables under consideration. Equation A.30 is quite simple to prove:

$$G_{x+y}(s) \equiv \mathrm{E}[e^{s(x+s)}] = \mathrm{E}[e^{sx}e^{sy}] = \mathrm{E}[e^{sx}]\mathrm{E}[e^{sy}] = G_x(s)G_Y(s).$$

This is a consequence of the fact that $\mathrm{E}[f(x)g(y)] = \mathrm{E}[f(x)]\mathrm{E}[g(y)]$ holds for arbitrary functions f, g of independent random variables x, y.

A further property of the MGF in connection with Equation A.30 is that for all non-stochastic values a, b, and random variables x we have

$$G_{ax+b}(s) = e^{sb}G_x(as) \tag{A.31}$$

The proof of this result is also quite simple:

$$G_{ax+b}(s) = \mathrm{E}[e^{s(ax+b)}] = e^{bs}\mathrm{E}[e^{(as)x}] = e^{sb}G_x(as).$$

We have already encountered a special case of this in Equation A.28.

A.3.2 Characteristic functions

Similar to the moment generating function, the *characteristic function* of a random variable x with probability density function $\mathrm{pdf}(x)$ is defined as the expectation of e^{isx} for an arbitrary real value s

$$\Phi_x(s) := \mathrm{E}[e^{isx}] = \int_{-\infty}^{\infty} e^{isx}\,\mathrm{pdf}(x)dx \tag{A.32}$$

2 Only in a few special cases, for example when each of the random variables is normally distributed, can the distribution of the sum be easily specified.

Here i denotes the imaginary number satisfying $i^2 = -1$. This is just the *Fourier transformation* of the pdf .

The *Fourier transformation* of the cpf, the *cumulative* distribution function, is sometimes used as well. The results derived below hold for these functions as well. Thus, we will formulate the characteristic function more generally by writing

$$\Phi_x(s) = \int_{-\infty}^{\infty} e^{isx} f(x)dx$$

where the function f can be taken to be either $f(x) = \text{pdf}(x)$ or $f(x) = \text{cpf}(x)$.

The advantage of the characteristic function is that its inverse function, the *inverse Fourier transformation* always exists:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-isx} \Phi_x(s)ds \tag{A.33}$$

Thus, if Φ_x is known, then the *distribution* (pdf or cpf) can be computed directly (and not only its moments as was the case with the moment generating function). The validity of Equation A.33 can be shown quite easily:

$$\begin{aligned}
\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-isx} \Phi_x(s)ds &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-isx} \int_{-\infty}^{\infty} e^{isx'} f(x')dx'ds \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-is(x-x')} f(x')dx'ds \\
&= \int_{-\infty}^{\infty} \delta(x-x') f(x')dx' \\
&= f(x)
\end{aligned}$$

where the *Dirac delta function* was used in the above derivation:

$$\delta(x-x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-is(x-x')} ds.$$

This is not a function in the strict sense but a so-called *distribution* with the defining property

$$\int_{-\infty}^{\infty} \delta(x-x') f(x')dx' := f(x).$$

It is precisely this property of the delta function that yields the invertibility of the Fourier transformation.

Analogously to the moment generating function, Equation A.30 holds for the characteristic function as well, i.e.,

$$\Phi_{x+y}(s) = \Phi_x(s)\Phi_y(s) \tag{A.34}$$

for independent random variables x and y. Likewise, for non-stochastic values a, b, and a random variable x

$$\Phi_{ax+b}(s) = e^{ibs}\Phi_x(as) \tag{A.35}$$

holds. The proof is completely analogous to that of Equation A.31.

A.4 A COLLECTION OF IMPORTANT DISTRIBUTIONS

A.4.1 The uniform distribution

The simplest of all distributions is the *uniform distribution*. This distribution has both a continuous and a discrete version. A random variable is said to be *uniformly distributed* if its distribution density is constant. The normalizing equation A.2 implies immediately that the distribution density for a discrete random variable with n possible outcomes (or for a continuously distributed random variable taking on values in an interval $[a, b]$) is given by:

$$\begin{aligned}
1 &= \sum_{i=1}^{n} \underbrace{p(x_i)}_{\text{Constant}} = np \Longrightarrow p = \frac{1}{n} \\
1 &= \int_{-\infty}^{\infty} \underbrace{p(x)}_{\text{Constant}} dx = p\int_a^b dx = p(b-a) \Longrightarrow p = \frac{1}{b-a}
\end{aligned} \tag{A.36}$$

The expectation and the variance of the continuous form of the uniform distribution are

$$\begin{aligned}
\mathrm{E}[x] &= \int_{-\infty}^{\infty} xp(x)dx = \frac{1}{b-a}\int_a^b xdx = \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{a+b}{2} \\
\mathrm{var}[x] &= \int_{-\infty}^{\infty} (x-\mathrm{E}[x])^2 p(x)dx = \frac{1}{b-a}\int_a^b \left(x - \frac{a+b}{2}\right)^2 dx \\
&= \frac{(b-a)^2}{12}
\end{aligned} \tag{A.37}$$

Most random number generators generate uniformly distributed random numbers between 0 and 1. For $a=0$ and $b=1$, the expectation and the variance of the uniform distribution is given by $1/2$ and $1/12$, respectively.

The moment generating function of the uniform distribution is by definition A.25

$$G_x(s) \equiv \int_{-\infty}^{\infty} e^{sx} p(x)dx = \frac{1}{b-a}\int_a^b e^{sx}dx = \frac{1}{b-a}\left[\frac{1}{s}e^{sx}\right]_{x=a}^{x=b}.$$

Thus

$$G_x(s) = \frac{e^{bs}-e^{as}}{s(b-a)} \tag{A.38}$$

Naive differentiation of G_x with respect to s does not lead us directly to the desired result since the factor s appears in the denominator forbidding us from setting $s=0$. The function can, however, be written in such a way that it is well defined for this value since as s approaches zero the numerator converges toward zero faster than the denominator (since the exponential function converges toward 1 faster than s converges to zero). This can be seen by expanding the exponential function in its Taylor series as in Equation A.38:

$$\frac{1}{s}e^{bs} = \frac{1}{s}\sum_{n=0}^{\infty}\frac{1}{n!}b^n s^n = \sum_{n=0}^{\infty}\frac{1}{n!}b^n s^{n-1} = \frac{1}{s}+\sum_{i=0}^{\infty}\frac{b^{i+1}}{(i+1)}\frac{s^i}{i!}$$

and analogously for e^{as}/s. Taking the difference of these two expressions, the $1/s$ term disappears and we are left with an equivalent form of the MGF in Equation A.38. This is given by

$$G_x(s) = \sum_{n=0}^{\infty}\frac{s^n}{n!}\frac{b^{n+1}-a^{n+1}}{(n+1)(b-a)}.$$

Comparison of the coefficients of s^n with Equation A.26 yields *all* of the moments immediately

$$\mathrm{E}[x^n] = \frac{b^{n+1}-a^{n+1}}{(n+1)(b-a)}.$$

For example, the first moment ($n=1$) is simply

$$\mathrm{E}[x] = \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2}$$

which agrees with the result in Equation A.37.

The characteristic function, A.32, of the continuous uniform distribution function is given by

$$\Phi_x(s) \equiv \int_{-\infty}^{\infty} e^{isx} \frac{1}{b-a} dx = \frac{e^{ibs} - e^{ias}}{is(b-a)}.$$

This equation can be found in every table of Fourier-transformed functions since it is just the Fourier transform of a constant. Note that the characteristic function can also be found by replacing s with is in the moment generating function Equation A.38.

A.4.2 The binomial distribution and the Bernoulli experiment

Suppose that an experiment has only two possible outcomes (for example, heads or tails when tossing a coin or up or down in one time step of a binomial model) and the probability of one of the two outcomes (for example, "heads" or "up") of the experiment is p. Then the normalizing equation A.2 implies that the probability of the alternate outcome ("tails" or "down") is $1 - p$. In mathematics, such an experiment is referred to as a *Bernoulli experiment*.

If p is independent of the number of experimental trials then the probability of the observing exactly j outcomes associated with the probability p in n trials is given by the *binomial distribution*.

$$p_n(j) = \binom{n}{j} p^j (1-p)^{n-j} \qquad \text{(A.39)}$$

Obviously, this is a discrete distribution taking on $n + 1$ possible values. The term $p^j(1-p)^{n-j}$ is the probability that the result of the n Bernoulli trials will occur *in a certain order*, for instance, up, up, down, up, down, down, etc., with precisely j "ups." However, since the number of "up" terms does not depend on the order in which they appear, this probability must be multiplied by the number of all *permutations* having j "ups." The number of these permutations is given by the *binomial coefficient*:[3]

$$\binom{n}{j} := \frac{n!}{j!(n-j)!} = \frac{1 \times 2 \times \cdots \times n}{(1 \times 2 \times \cdots \times j)(1 \times 2 \times \cdots \times n-j)} \qquad \text{(A.40)}$$

This is a result from the theory of combinations which has long been known.

The probability that "up" will be observed *at least* k times is naturally the sum from k to n of the density defined in Equation A.39. This yields

3 This reads "n choose j" or "n over j."

precisely the binomial probability arising in the recombinant binomial trees for European options:

$$B_{n,p}(j \geq k) = \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} \tag{A.41}$$

The moment generating function of the binomial distribution is, by the definition in Equation A.25,

$$\begin{aligned} G_{B_{n,p}}(s) &\equiv \mathrm{E}\left[e^{sx}\right] = \sum_{j=0}^{n} e^{sj} p_n(j) \\ &= \sum_{j=0}^{n} e^{sj} \binom{n}{j} p^j (1-p)^{n-j} \\ &= \sum_{j=0}^{n} \binom{n}{j} \left[pe^s\right]^j \left[1-p\right]^{n-j}. \end{aligned}$$

Note that this is precisely the function appearing in the *binomial formula* for the n^{th} power of a sum $(a+b)^n$ where $a = pe^s$ and $b = 1-p$. Thus the moment generating function of the binomial distribution is simply

$$G_{B_{n,p}}(s) = \left(pe^s + 1 - p\right)^n \tag{A.42}$$

According to Equation A.27 all of the moments can now be calculated directly by differentiating the MGF with respect to s. Doing so yields, for example, the expectation as

$$\mathrm{E}[j] = \left.\frac{\partial G_{B_{n,p}}(s)}{\partial s}\right|_{s=0} = n\left(pe^s + 1 - p\right)^{n-1} pe^s\Big|_{s=0} = np.$$

The second moment can be computed analogously:

$$\begin{aligned} \mathrm{E}[j^2] = \left.\frac{\partial^2 G_{B_{n,p}}(s)}{\partial s^2}\right|_{s=0} &= np \frac{\partial}{\partial s}\left(pe^s + 1 - p\right)^{n-1} e^s\Big|_{s=0} \\ &= np\,(n-1)\left(pe^s + 1 - p\right)^{n-2} pe^s e^s\Big|_{s=0} \\ &\quad + np\left(pe^s + 1 - p\right)^{n-1} e^s\Big|_{s=0} \\ &= n(n-1)p^2 + np. \end{aligned}$$

Thus, according to Equation A.7, the second *central* moment, the variance of the binomial distribution, becomes

$$\mathrm{var}\left[j\right] = \mathrm{E}[j^2] - (\mathrm{E}[j])^2 = np(1-p).$$

Summarizing, the expectation and variance of the binomial distribution are given by

$$
\begin{aligned}
\mathrm{E}\left[j\right] &\equiv \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} j = np \\
\mathrm{var}\left[j\right] &\equiv \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} (j-np)^2 = np(1-p)
\end{aligned}
\tag{A.43}
$$

Analogously to Equation A.42, the characteristic function of the binomial distribution, can be found from Equation A.32 to be

$$
\Phi_{B_{n,p}}(s) = \left(pe^{is} + 1 - p\right)^n \tag{A.44}
$$

A.4.3 The normal distribution and the central limit theorem

It is well known that the sum of random variables is itself a random variable. One of the most important theorems of mathematical statistics, the *central limit theorem*, makes a broad statement on the sum of random variables. The intuitive content of this theorem can be expressed as follows:

Theorem 10 (Central Limit) *The sum of a large number of independent random variables is approximately normally distributed, regardless of how the individual random variables are distributed, if the contribution of each random variable to the sum is almost negligible.*

This theorem is the reason for the extraordinary importance of the normal distribution above all others. For instances, the *mean* of some random variables is defined as the *sum* of these random variables divided by the number of variables. Thus, such means are *always* approximately normally distributed according to the central limit theorem, *regardless of the distribution of the random variables*. The fact that such sums are often *not* normally distributed can only mean that the assumptions in the statement of the theorem are not satisfied. Either the measured variables are not (purely) random variables and/or they are not completely independent (uncorrelated). The most common reason, however, is an insufficient number of available trial results. Because only in the *limit*, i.e., for an *infinite* number of random variables, does the central limit theorem hold exactly and not merely as an approximation (hence the name central *limit* theorem).

The *normal distribution*, also called *Gaussian distribution*, is a continuous distribution which is completely determined by two parameters (denoted by

μ and σ). The density has the explicit form

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{A.45}$$

The expectation and the variance of the normal distribution according to the definition in Equation A.4 and Equation A.5 are given by

$$\begin{aligned} \mathrm{E}[x] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu \\ \mathrm{var}[x] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sigma^2 \end{aligned} \tag{A.46}$$

Hence, the two parameters μ and σ are equal to the expectation and the standard deviation of the distribution, respectively. This is quite practical, since through observing the results of random trials x, the expectation and the variance can be approximated by measuring the mean of x and x^2 and applying Equation A.7, thus obtaining an approximation of the entire distribution. The density of the distribution can be written as

$$p(x) = \frac{1}{\sqrt{2\pi \mathrm{E}\left[(x - \mathrm{E}[x])^2\right]}} \exp\left\{ -\frac{1}{2} \frac{(x-\mathrm{E}[x])^2}{\mathrm{E}\left[(x-\mathrm{E}[x])^2\right]} \right\} \tag{A.47}$$

The ratio of the square of the deviation of the random variable from its expectation to the *expectation* of this same factor appears in the argument of the exponential function.

The probability $P(x \leq a)$ of the event that a random variable x, distributed according to Equation A.45 with parameter values μ and σ, will be less than or equal to a given value a is the cumulative probability cpf of the normal distribution evaluated at a, frequently denoted by $\mathrm{N}_{\mu,\sigma}(a)$. Thus, according to Equation A.1 we have

$$\mathrm{N}_{\mu,\sigma}(a) = \mathrm{cpf}(a) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{a} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \tag{A.48}$$

The following notation is also often used to express the fact that a random variable x is normally distributed with expectation μ and variance σ^2

$$x \sim \mathrm{N}(\mu, \sigma^2) \tag{A.49}$$

The reader should be careful not to confuse this expression with that for the cumulative distribution function $\mathrm{N}_{\mu,\sigma}(a)$ in Equation A.48.

The integral in the definition of the *cumulative normal distribution* cannot be calculated as a closed form expression. But through a simple change of variable

$$y := \frac{x-\mu}{\sigma} \Longrightarrow dx = \sigma dy \tag{A.50}$$

$$y(x=-\infty) = -\infty, \quad y(x=a) = \frac{a-\mu}{\sigma}$$

the normal distribution can be transformed into the *standard* normal distribution. This is the normal distribution with expectation 0 and variance 1, i.e., with the density function

$$p(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \tag{A.51}$$

In the notation introduced in Equation A.49 we write

$$y \sim \mathrm{N}(0, 1).$$

Random variables which are standard normally distributed with density function A.51 form the basis of many random walk models applied in this book. The probability $P(x \le a)$ that a standard normally distributed random variable x will be less than or equal to a given number a is called the *cumulative standard normal distribution* $\mathrm{N}(a)$:

$$\mathrm{N}(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx \tag{A.52}$$

This distribution appears for instance in the Black-Scholes formula.

The density Equation A.51 of the standard normal distribution is symmetric about zero and thus $p(x) = p(-x)$. From this we derive a very useful *symmetry* relation holding for the standard normal distribution, namely

$$\mathrm{N}(-a) = 1 - \mathrm{N}(a) \tag{A.53}$$

This symmetry is so important in practical applications that we will provide the proof here. We first write

$$\mathrm{N}(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx = \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx}_{1} - \frac{1}{\sqrt{2\pi}} \int_{a}^{\infty} e^{-\frac{x^2}{2}} dx.$$

The first integral equals 1 because of Equation A.2. In the second integral, we make the substitution

$$u := -x \Rightarrow du = -dx$$
$$u(x = a) = -a, \quad u(x = \infty) = -\infty$$

thus obtaining

$$\mathrm{N}(a) = 1 + \frac{1}{\sqrt{2\pi}} \int_{-a}^{-\infty} e^{-\frac{x^2}{2}} du = 1 - \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-a} e^{-\frac{x^2}{2}} du}_{\mathrm{N}(-a)}$$

which proves the symmetry relation in Equation A.53.

The integral in Equation A.52 cannot be computed explicitly. However, there exist tables and numerical routines for the computation of the cumulative standard normal distribution. A very good polynomial approximation, exact up to six decimal places, is (see for example [1])

$$\mathrm{N}(x) = 1 - \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \sum_{i=1}^{5} a_i y^i \quad \text{for} \quad x \geq 0 \tag{A.54}$$

where

$$y = \frac{1}{1 + 0.2316419x}$$
$$a_1 = 0.319381530, \quad a_2 = -0.356563782$$
$$a_3 = 1.781477937, \quad a_4 = -1.821255978$$
$$a_5 = 1.330274429$$

This approximation holds only for nonnegative values $x \geq 0$. But the values for $x < 0$ can simply be obtained by applying the symmetry relation A.53.

Calculations with the *standard* normal distribution are thus quite simple. A frequently used method is therefore to transform normally distributed random variables into *standard* normal random variables via a transformation as in Equation A.50. Then all necessary calculations are performed using tools like Equations A.54 and A.53. Having completed the calculation, an inverse transformation can be performed to determine the original variables. For example, a 99% confidence interval for a standard normal distribution is

$$0,99 = P(y \leq a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{y^2}{2}} dy \Rightarrow a \approx 2,326.$$

This is the upper bound for the *standard* normally distributed random variable y. Performing the inverse transformation back to the variable x according to Equation A.50 yields the boundary of the confidence interval for the original random variable: the expectation plus 2.326 standard deviations.

The moment generating function of the standard normal distribution is, by Definition A.25

$$\begin{aligned} G_{\mathrm{N}(0,1)}(s) &= \mathrm{E}[e^{sx}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sx} e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{x^2 - 2sx}{2}\right\} dx \\ &= \exp\left(\frac{1}{2}s^2\right) \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{(x-s)^2}{2}\right\} dx}_{1} \end{aligned}$$

where we have used the method of completing the squares in the last step.[4] The remaining integral is precisely the probability that a normally distributed random variable $x \sim \mathrm{N}(s, 1)$ will take on any arbitrary value. Equation A.2 implies that this probability is 1. The moment generating function of the *standard* normal distribution is thus simply

$$G_{\mathrm{N}(0,1)}(s) = \exp\left(\frac{1}{2}s^2\right) \tag{A.55}$$

The moment generating function of the normal distribution with expectation μ and variance σ now follows immediately from the transformation A.50 with Equation A.31

$$G_{\mathrm{N}(\mu,\sigma^2)}(s) = e^{\mu s} G_{\mathrm{N}(0,1)}(\sigma s) = \exp\left(\mu s + \frac{1}{2}\sigma^2 s^2\right) \tag{A.56}$$

From this equation, all *moments* of a normal distribution can be calculated by means of Equation A.27. The *central* moments are given by Equation A.29:

$$\begin{aligned} \mathrm{E}[(x - \mathrm{E}[x])^n] &= \frac{\partial^n}{\partial s^n} \exp(-s\mu) \exp\left(\mu s + \frac{1}{2}\sigma^2 s^2\right)\bigg|_{s=0} \\ &= \frac{\partial^n}{\partial s^n} \exp\left(\frac{1}{2}\sigma^2 s^2\right)\bigg|_{s=0} . \end{aligned}$$

4 $x^2 - 2sx = (x-s)^2 - s^2$.

The second central moment, i.e., the variance is

$$\begin{aligned}
\mathrm{E}[(x-\mathrm{E}[x])^2] &= \frac{\partial^2}{\partial s^2} e^{\sigma^2 s^2/2}\Big|_{s=0} = \frac{\partial}{\partial s}\sigma^2 s e^{\sigma^2 s^2/2}\Big|_{s=0} \\
&= \sigma^2 e^{\sigma^2 s^2/2} + \left(\sigma^2 s\right)^2 e^{\sigma^2 s^2/2}\Big|_{s=0} \\
&= \left(1+\sigma^2 s^2\right)\sigma^2 e^{\sigma^2 s^2/2}\Big|_{s=0} = \sigma^2.
\end{aligned}$$

Differentiating the result in the penultimate step yields the third central moment

$$\begin{aligned}
\mathrm{E}[(x-\mathrm{E}[x])^3] &= \frac{\partial^3}{\partial s^3} e^{\sigma^2 s^2/2}\Big|_{s=0} = \frac{\partial}{\partial s}\left(1+\sigma^2 s^2\right)\sigma^2 e^{\sigma^2 s^2/2}\Big|_{s=0} \\
&= 2s\sigma^4 e^{\sigma^2 s^2/2} + \left(1+\sigma^2 s^2\right) s\sigma^4 e^{\sigma^2 s^2/2}\Big|_{s=0} \\
&= \left(3s+\sigma^2 s^3\right)\sigma^4 e^{\sigma^2 s^2/2}\Big|_{s=0} = 0.
\end{aligned}$$

Again, differentiating the result in the penultimate step yields the next (i.e., fourth) central moment

$$\begin{aligned}
\mathrm{E}[(x-\mathrm{E}[x])^4] &= \frac{\partial^4}{\partial s^4} e^{\sigma^2 s^2/2}\Big|_{s=0} = \frac{\partial}{\partial s}\left(3s+\sigma^2 s^3\right)\sigma^4 e^{\sigma^2 s^2/2}\Big|_{s=0} \\
&= \left(3+3\sigma^2 s^2\right)\sigma^4 e^{\sigma^2 s^2/2} + \left(3s+\sigma^2 s^3\right)\sigma^6 s e^{\sigma^2 s^2/2}\Big|_{s=0} \\
&= \left(3+3\sigma^2 s^2+3\sigma^2 s+\sigma^4 s^3\right)\sigma^4 e^{\sigma^2 s^2/2}\Big|_{s=0} = 3\sigma^4.
\end{aligned}$$

We summarize these results for reference:

$$\begin{aligned}
\mathrm{E}[x] &= \mu \\
\mathrm{E}[(x-\mathrm{E}[x])^2] &= \sigma^2 \\
\mathrm{E}[(x-\mathrm{E}[x])^4] &= 3\sigma^4 \\
\mathrm{E}[(x-\mathrm{E}[x])^n] &= 0 \quad \text{for all odd } n > 2
\end{aligned} \tag{A.57}$$

From these moments the *skewness* and the *kurtosis* of the normal distribution follow directly from their respective definitions, Equation A.24:

$$\begin{aligned}
\text{Skewness} &\equiv \frac{\mathrm{E}[(x-\mathrm{E}[x])^3]}{\mathrm{E}[(x-\mathrm{E}[x])^2]^{3/2}} = 0 \\
\text{kurtosis} &\equiv \frac{\mathrm{E}[(x-\mathrm{E}[x])^4]}{\mathrm{E}[(x-\mathrm{E}[x])^2]^2} = 3
\end{aligned} \tag{A.58}$$

The characteristic function A.32 of the standard normal density function can be found by replacing s with is in Equation A.55

$$\Phi_{N(0,1)}(s) = \exp\left(-\frac{1}{2}s^2\right) \tag{A.59}$$

The characteristic function of a normal distribution with expectation μ and variance σ follows immediately from Equation A.35

$$\Phi_{N(\mu,\sigma^2)}(s) = e^{i\mu s}\Phi_{N(0,1)}(\sigma s) = \exp\left(i\mu s - \frac{1}{2}\sigma^2 s^2\right) \tag{A.60}$$

The multivariate normal distribution

If two random variables $x_i, i=1,2$ are both normally distributed with expectation μ_i and covariance[5]

$$\mathrm{cov}\left[x_i, x_j\right] = \sigma_i \rho_{ij} \sigma_j$$

then their *joint probability distribution*, i.e., the probability that both random numbers simultaneously will have certain values, is given by the *bivariate normal distribution* with the density

$$p(x_1, x_2) = \frac{\exp\left\{-\frac{1}{1-\rho_{1,2}^2}\frac{1}{2}\sum_{i,j=1,2}\frac{x_i-\mu_i}{\sigma_i}\rho_{ij}\frac{x_j-\mu_j}{\sigma_j}\right\}}{\sqrt{1-\rho_{1,2}^2}\prod_{i=1,2}\sqrt{2\pi\sigma_i^2}}.$$

With this density, the joint (cumulative) probability, as defined in Equation A.8, for $x_1 < a$ and $x_2 < b$ is

$$P(x_1 < a, x_2 < b) = \frac{\int_{-\infty}^{a} dx_1 \int_{-\infty}^{b} dx_2 \exp\left\{-\frac{1}{1-\rho_{1,2}^2}\frac{1}{2}\sum_{i,j=1}^{2}\frac{x_i-\mu_i}{\sigma_i}\rho_{ij}\frac{x_j-\mu_j}{\sigma_j}\right\}}{\sqrt{1-\rho_{12}^2}\prod_{i=1}^{2}\sqrt{2\pi\sigma_i^2}}.$$

If the two random variables are *uncorrelated*, i.e., if

$$\rho_{ij} = \delta_{ij} = \begin{cases} 1 & \text{for } i=j \\ 0 & \text{for } i \neq j \end{cases}$$

5 This means that they have variances σ_1^2 and σ_2^2 and the correlation between the two is ρ_{12} (see Equation A.14).

where δ_{ij} denotes the *Kronecker delta*, then the joint distribution of the variables is equal to the product of the distributions of each individual variable:

$$
\begin{aligned}
p(x_1, x_2) &= \frac{\exp\left\{-\frac{1}{1-0}\frac{1}{2}\sum_{i,j=1,2}\frac{x_i-\mu_i}{\sigma_i}\delta_{ij}\frac{x_j-\mu_j}{\sigma_j}\right\}}{\sqrt{1-0}\prod_{i=1,2}\sqrt{2\pi\sigma_i^2}} \\
&= \frac{\exp\left\{-\frac{1}{2}\sum_{i=1,2}\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2\right\}}{\prod_{i=1,2}\sqrt{2\pi\sigma_i^2}} \\
&= \prod_{i=1,2}\frac{1}{\sqrt{2\pi\sigma_i^2}}\exp\left\{-\frac{1}{2}\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2\right\} \\
&= \prod_{i=1,2} p(x_i)
\end{aligned}
\qquad \text{(A.61)}
$$

If the joint probability density of two random variables x_1 and x_2 fulfills

$$p(x_1, x_2) = p(x_1)p(x_2) \qquad \text{(A.62)}$$

then the two random variables are said to be *independent*. As we have just seen, a *necessary* condition for independence is that the two variables are uncorrelated. However, this is generally not a *sufficient* condition, i.e., uncorrelated random variables are not always independent. But – as we have just seen – in the special case of normally distributed random variables independence and uncorrelation are equivalent.

A.4.4 The lognormal distribution

A random variable is *lognormally distributed* if its logarithm is normally distributed. Since the lognormal distribution is defined through the normal distribution, it is also continuous and completely determined by the two parameters μ and σ. The precise definition of the *lognormal distribution* can be stated as follows: let x be a lognormally distributed random variable with parameters μ and σ. Then the probability $\mathrm{H}_{\mu,\sigma}(a)$ that x will be less than a given value a is equal to the cumulative distribution function of the normal distribution, Equation A.48, with the same parameters, evaluated at $\ln(a)$:

$$\mathrm{H}_{\mu,\sigma}(a) := \mathrm{N}_{\mu,\sigma}(\ln(a)) = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\ln(a)} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx \qquad \text{(A.63)}$$

The density function of the lognormal distribution can be derived from the above definition. It has the explicit form

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\frac{1}{x}\exp\left\{-\frac{1}{2}\frac{(\ln(x)-\mu)^2}{\sigma^2}\right\} \tag{A.64}$$

Note the factor $1/x$. The density is thus not merely obtained by substituting x with $\ln(x)$ in Equation A.45. With this density function we can express $H_{\mu,\sigma}$ explicitly as

$$H_{\mu,\sigma}(a) = \frac{1}{\sqrt{2\pi\sigma^2}}\int_0^a \exp\left\{-\frac{1}{2}\frac{(\ln(x)-\mu)^2}{\sigma^2}\right\}\frac{1}{x}dx \tag{A.65}$$

Note that the lower limit in the integral is zero. The range of such a lognormally distributed random number lies only between zero and infinity.

With the simple change in variable given by $u = \ln(x)$ we can immediately verify that the cumulative distribution function $H_{\mu,\sigma}(a)$ actually satisfies the definition in Equation A.63:

$$\begin{aligned}
H_{\mu,\sigma}(a) &= \frac{1}{\sqrt{2\pi\sigma^2}}\int_0^a \exp\left\{-\frac{1}{2}\frac{(\ln(x)-\mu)^2}{\sigma^2}\right\}\frac{1}{x}dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\ln(0)=-\infty}^{\ln(a)} \exp\left\{-\frac{1}{2}\frac{(u-\mu)^2}{\sigma^2}\right\}du = N_{\mu,\sigma}(\ln(a))
\end{aligned}$$

$$\text{where} \quad u := \ln(x) \Longrightarrow du = \frac{1}{x}dx, \quad x = e^u.$$

The expectation and variance of the lognormal distribution can be calculated using Equations A.4 and A.5 as

$$\begin{aligned}
E[x] &= \int_0^\infty xp(x)dx = \frac{1}{\sqrt{2\pi\sigma^2}}\int_0^\infty e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}dx = e^{\mu+\sigma^2/2} \\
\operatorname{var}[x] &= \int_0^\infty (x - E[x])^2 p(x)dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\int_0^\infty (x - e^{\mu+\sigma^2/2})^2\frac{1}{x}e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}dx \\
&= e^{2\mu}(e^{2\sigma^2} - e^{\sigma^2})
\end{aligned} \tag{A.66}$$

In general, the moments of the lognormal distribution are given by

$$\mathrm{E}[x^n] = \exp\left(n\mu + n^2\sigma^2/2\right) \quad \text{for } n = 1, 2, \ldots \tag{A.67}$$

A.4.5 The gamma distribution

The *gamma distribution* is an important distribution because it encompasses a whole class of different distributions (which includes the exponential distribution and the χ^2-distribution, for example). Like the lognormal distribution, the range of the gamma distribution consists solely of the nonnegative real numbers. A random variable x has a gamma distribution with parameters λ and t if it is governed by the following distribution density

$$\mathrm{pdf}(x) = \frac{1}{\Gamma(t)}\lambda^t x^{t-1} e^{-\lambda x} \quad \text{where } \lambda > 0, t > 0 \text{ and } x \in [0, \infty[\tag{A.68}$$

where $\Gamma(t)$ denotes the *gamma function*. This function is defined by

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx \tag{A.69}$$

A description of its properties can be found in any mathematical collection of special mathematical functions (see for instance [1]). An important property of the gamma function is the recursion

$$\Gamma(t + 1) = t\Gamma(t) \tag{A.70}$$

This allows the gamma function to be interpreted as a generalization of the factorial operation. Two special function values which often serve as the initial values in the above recursion relation are given by

$$\Gamma(1) = 1, \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \tag{A.71}$$

It follows that the gamma function for whole numbers t is in fact nothing other than the factorial:

$$\Gamma(n) = (n - 1)! \quad \text{for } n = 1, 2, \ldots$$

Furthermore, this function has a symmetry property which allows it to be evaluated for negative values if the function evaluated at the corresponding positive value is known:

$$\Gamma(-x) = -\frac{\pi}{x\sin(\pi x)}\frac{1}{\Gamma(x)}.$$

From this symmetry relation, we see immediately that the gamma function has a simple pole at each negative integer value of x. It is well defined for all positive values. Fortunately, in our consideration of the gamma distribution, Equation A.68, our needs are restricted to the positive arguments of the gamma function only.

The explicit form of the moment generating function of the gamma distribution can be calculated quite easily. Consider the definition A.25. Then for the gamma distribution, we have

$$G_{\Gamma(t,\lambda)}(s) = \mathrm{E}[e^{sx}] = \frac{\lambda^t}{\Gamma(t)} \int_0^\infty x^{t-1} e^{-(\lambda-s)x} dx.$$

The integral appearing in this expression is quite similar to that in the definition of the gamma function, Equation A.69. We make the substitution

$$u := (\lambda - s)x \Longrightarrow du = (\lambda - s)dx,$$
$$u(x = 0) = 0, \quad u(x = \infty) = \infty \quad \text{for } s < \lambda.$$

The condition $s < \lambda$ is required in order for the upper limit of the integral to remain equal to $+\infty$. From Equation A.27 we see that the s of interest are in a small neighborhood of zero. Since λ in the distribution Equation A.68 is strictly greater than zero, the condition $s < \lambda$ is no obstacle for our purposes here. A simple substitution yields

$$\begin{aligned} G_{\Gamma(t,\lambda)}(s) &= \frac{\lambda^t}{\Gamma(t)} \int_0^\infty \frac{1}{(\lambda - s)^{t-1}} u^{t-1} e^{-u} \frac{1}{\lambda - s} du \\ &= \frac{\lambda^t}{\Gamma(t)} \frac{1}{(\lambda - s)^t} \underbrace{\int_0^\infty u^{t-1} e^{-u} du.}_{\Gamma(t)} \end{aligned}$$

The moment generating function of the gamma distribution thus has the following simple form:

$$G_{\Gamma(t,\lambda)}(s) = \left(\frac{\lambda}{\lambda - s}\right)^t \quad \text{with} \quad \lambda > 0, t > 0, s < \lambda \tag{A.72}$$

Replacing s with is in Equation A.72 yields the characteristic function given in Equation A.32 for the gamma distribution

$$\Phi_{\Gamma(t,\lambda)}(s) = \left(\frac{\lambda}{\lambda - is}\right)^t \tag{A.73}$$

All *moments* can be obtained as indicated in Equation A.27 directly by differentiating the expression in Equation A.72 with respect to s. For example, the expectation becomes

$$\mathrm{E}[x] = \left.\frac{\partial G_{\Gamma(t,\lambda)}(s)}{\partial s}\right|_{s=0} = \lambda^t t \left.\frac{1}{(\lambda - s)^{t+1}}\right|_{s=0} = \frac{t}{\lambda} \tag{A.74}$$

The second moment is

$$\mathrm{E}[x^2] = \left.\frac{\partial^2 G_{\Gamma(t,\lambda)}(s)}{\partial s^2}\right|_{s=0} = \lambda^t t(t+1) \left.\frac{1}{(\lambda - s)^{t+2}}\right|_{s=0} = \frac{t(t+1)}{\lambda^2}.$$

Proceeding analogously, *all* moments can be explicitly calculated for the gamma distribution:

$$\mathrm{E}[x^n] = \frac{t(t+1)\cdots(t+n-1)}{\lambda^n} \quad \text{for} \quad n = 1, 2, \ldots \tag{A.75}$$

The variance is then

$$\mathrm{var}\,[x] = \mathrm{E}[x^2] - \mathrm{E}[x]^2 = \frac{t}{\lambda^2} \tag{A.76}$$

A.4.6 The χ^2-distribution

As was shown in Section A.4.3, a sum of normally distributed random variables is itself normally distributed. A situation frequently encountered (for example, in a value at risk computation or the determination of variances from historical data) involves taking sums of the *squares* of random variables. If the random variables whose squares are added have a standard normal distribution, the distribution of the resulting random variable is easily determined: the sum of *n squared, independent, standard normally distributed random variables,* x_i, $(i = 1, \ldots, n)$ has a distribution known as the *χ^2-distribution with n degrees of freedom*

$$x_i \sim \mathrm{N}(0, 1),\; i = 1, \ldots, n, \quad x_i \;\; iid \Longrightarrow \sum_{i=1}^{n} x_i^2 =: y \sim \chi^2(n) \tag{A.77}$$

It is essential that the random variables x_i in the above definition be independent. The *degree of freedom*, n, of the χ^2-distribution $\chi^2(n)$ can be intuitively thought of as the number of independent (standard normal) random variables which "make up" the random variable with the $\chi^2(n)$-distribution; thus the name "degree of freedom."

It is often the case that the above sum is taken over only one single element. If a random variable x has a standard normal distribution, then this random variable *squared* is governed by the χ^2*-distribution with one degree of freedom.* We write

$$x \sim \mathrm{N}(0,1) \Longrightarrow x^2 =: y \sim \chi^2(1) \tag{A.78}$$

In fact, $\chi^2(1)$ (or its noncentral counterpart, see below) is sufficient for almost all our needs if we are working with moment generating functions since via Equation A.30 and A.31, the MGFs of sums of independent random variables can be written as products of MGFs of $\chi^2(1)$.

The MGF also proves to be a helpful tool in calculating the density function of the χ^2-distribution. We first generalize the statement of the problem somewhat, keeping in mind that our goal is to arrive at the MGF of $\chi^2(1)$.

Suppose a random variable x is distributed according to some known distribution whose distribution density we denote by pdf_x. Then, for any arbitrary function f, the random variable $y = f(x)$ is distributed according to another distribution function pdf_y. This distribution is in general unknown but from the construction it is immediately obvious that the probability for the function f to assume a value y is just the probability of the original stochastic variable to assume the value $x = f^{-1}(y)$:

$$\mathrm{pdf}_y(y) = \mathrm{pdf}_x(x = f^{-1}(y)) \tag{A.79}$$

Let's now look at the MGF. The MGF of the unknown distribution pdf_y is the expectation of a function g of y (specifically, $g(y) = e^{sy}$). According to Equation A.4, the expectation of a function g of y is given by

$$\mathrm{E}\left[g(y)\right]_{\mathrm{pdf}_y} = \int g(y)\,\mathrm{pdf}_y(y)dy.$$

Here, the expectation is computed with respect to the (unknown) distribution pdf_y (this fact is emphasized by the subscript in the expectation). Since $y = f(x)$, the function

$$g(y) = g(f(x)) = h(x) \tag{A.80}$$

is also a function of x and therefore the expectation of h is according to Equation A.4

$$\mathrm{E}\left[h(x)\right]_{\mathrm{pdf}_x} = \int h(x)\,\mathrm{pdf}_x(x)dx.$$

Here, the expectation is computed by means of the *known* distribution pdf_x.

Computing the expectation of $h(x)$ means integrating over *all* possible values of x where each x is weighted with $\mathrm{pdf}_x(x)$. Since by definition $y = f(x)$ this operation can also be viewed as integrating over *all possible* values of y where each random number $y = f(x)$ is weighted with the weight $\text{pdf}_x(x = f^{-1}(y))$:

$$\begin{aligned}\text{E}\,[h(x)]_{\text{pdf}_x} &= \int h(x)\,\text{pdf}_x(x)dx \\ &= \int h(x)\,\text{pdf}_x(f^{-1}(y))dy.\end{aligned}$$

But this weight is just the density $\text{pdf}_y(y)$, see Equation A.79. We can therefore write

$$\begin{aligned}\text{E}\,[h(x)]_{\text{pdf}_x} &= \int h(x)\,\text{pdf}_y(y)dy \\ &= \int g(y)\,\text{pdf}_y(y)dy\end{aligned}$$

where we have used Equation A.80 in the last step. The expression we have now arrived at is simply the expectation of the function g with respect to the distribution pdf_y. In summary we thus have the quite general result

$$\text{E}\left[g(y)\right]_{\text{pdf}_y} = \text{E}\left[g(f(x))\right]_{\text{pdf}_x} \quad \text{for } y = f(x) \tag{A.81}$$

Choosing the function $g(y) = e^{sy}$, the left-hand side becomes the MGF of the unknown distribution pdf_y

$$G_{\text{pdf}_y}(s) = \text{E}[e^{sy}]_{\text{pdf}_y} = \text{E}[e^{sf(x)}]_{\text{pdf}_x}$$

Thus, for every function f of a random variable x, the moment generating function of the distribution of $f(x)$ can be expressed through the expectation with respect to the distribution of x:

$$G_{\text{pdf}_{f(x)}}(s) = \text{E}[e^{sf(x)}]_{\text{pdf}_x} \tag{A.82}$$

For the χ^2-distribution with one degree of freedom, we have $f(x) = x^2$ and obtain[6] the moment generating function of $\chi^2(1)$ as

$$G_{\chi^2(1)}(s) = \text{E}[e^{sx^2}]_{\text{N}(0,1)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sx^2} e^{-x^2/2} dx \tag{A.83}$$

6 with the notation: $\text{pdf}_{f(x)} = \text{pdf}_{x^2} = \chi^2(1)$, $\text{pdf}_x = \text{N}(0, 1)$

This integral can be solved explicitly using the substitution

$$
\begin{aligned}
u &= x\sqrt{1-2s} \Rightarrow dx = \frac{du}{\sqrt{1-2s}} \\
u(x=-\infty) &= -\infty \quad \text{für } s < 1/2 \\
u(x=+\infty) &= +\infty \quad \text{für } s < 1/2
\end{aligned}
\tag{A.84}
$$

Note that the integration limits remain the same only for $s < 1/2$.

$$
\begin{aligned}
\mathrm{E}[e^{sx^2}]_{\mathrm{N}(0,1)} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(1-2s)x^2/2} dx \\
&= \frac{1}{\sqrt{1-2s}} \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2} du}_{1}.
\end{aligned}
$$

The remaining integral is the probability that a standard normal random variable will assume *any* arbitrary value and, according to Equation A.2, equals one. Thus, the MGF is simply

$$
G_{\chi^2(1)}(s) = \frac{1}{\sqrt{1-2s}} \tag{A.85}
$$

Since the random variables in the sum appearing in Definition A.77 are all *independent* we can use Equation A.30 to immediately obtain the MGF for a χ^2-distribution with n degrees of freedom

$$
G_{\chi^2(n)}(s) = \frac{1}{(1-2s)^{n/2}} \quad \text{for } n = 1, 2, \ldots \tag{A.86}
$$

The *moments* can be derived directly from Equation A.27 by differentiating the function in Equation A.86 with respect to s or simply from Equation A.75 with $\lambda = 1/2$ and $t = n/2$:

$$
\begin{aligned}
\mathrm{E}[x^k]_{\chi^2(n)} &= \frac{t(t+1)\cdots(t+k-1)}{\lambda^k} \\
&= \frac{\frac{n}{2}\left(\frac{n}{2}+1\right)\cdots\left(\frac{n}{2}+k-1\right)}{\left(\frac{1}{2}\right)^k} \\
&= 2^k \frac{n}{2}\left(\frac{n}{2}+1\right)\cdots\left(\frac{n}{2}+k-1\right) \\
&= n(n+2)\cdots(n+2(k-1)).
\end{aligned}
$$

Therefore

$$
\mathrm{E}[x^k]_{\chi^2(n)} = \prod_{i=0}^{k-1}(n+2i) \tag{A.87}
$$

For example, the expectation and variance are given by

$$\mathrm{E}[x]_{\chi^2(n)} = n, \quad \mathrm{var}\,[x]_{\chi^2(n)} = 2n.$$

From the definition in Equation A.32, the characteristic function of the χ^2-distribution can be obtained by replacing s with is in Equation A.86 or from Equation A.73 with $\lambda = 1/2$ and $t = n/2$.

$$\Phi_{\chi^2(n)}(s) = \frac{1}{(1-2is)^{n/2}} \quad \text{for } n = 1, 2, \ldots \tag{A.88}$$

The density of the χ^2-distribution with one degree of freedom

The MGF Equation A.83 has been derived based on the general Transformation A.81. It is also possible to calculate the MGF without using Transformation A.81. This has the advantage that an explicit expression for the density of the χ^2-Distribution emerges.

So let $y := x^2$ with $x \sim \mathrm{N}(0, 1)$ as in Equation A.78. Denote the desired density of the χ^2-Distribution with p and the Standardnormal density with q.

$$p(y) := \mathrm{pdf}_{\chi^2(1)}(y)$$

$$q(x) := \mathrm{pdf}_{\mathrm{N}(0,1)}(x) = \frac{1}{\sqrt{2\pi}} e^{-x/2}.$$

Then, the cumulative probability is

$$\begin{aligned}\int_{-\infty}^{a} p(y)dy &= P(y < a) \\ &= P(x^2 < a) \\ &= P(-\sqrt{a} < x < \sqrt{a}) \\ &= 2P(0 < x < \sqrt{a}) \\ &= 2\int_0^{\sqrt{a}} q(x)dx.\end{aligned}$$

We now differentiate with respect to the upper integration limit a. The left side becomes

$$\frac{\partial}{\partial a}\int_{-\infty}^{a} p(y)dy = p(a).$$

And with the chain rule the right side reads

$$\frac{\partial}{\partial a} 2\int_0^{\sqrt{a}} q(x)dx = \frac{1}{2\sqrt{a}} 2q(\sqrt{a}) = \frac{q(\sqrt{a})}{\sqrt{a}}.$$

Comparing both sides yields the relation between the normal and the χ^2-distribution

$$\text{pdf}_{\chi^2(1)}(a) = \frac{1}{\sqrt{a}}\,\text{pdf}_{\text{N}(0,1)}(\sqrt{a}) \tag{A.89}$$

Finally, inserting the standard normal density we arrive at the explicit form of the χ^2-distribution

$$\text{pdf}_{\chi^2(1)}(a) = \frac{1}{\sqrt{2\pi a}}e^{-a/2} = \frac{1}{\sqrt{2\pi a \exp(a)}} \tag{A.90}$$

With Equation A.89, the MGF Gl. A.83 can now be derived directly without referring to the general Transformation A.81:

$$\begin{aligned}
\text{E}[e^{sy}]_{\chi^2(1)} &= \int_{-\infty}^{\infty} e^{sy} \underbrace{p(y)}_{=0 \text{ for } y<0} dy \\
&= \int_{0}^{\infty} e^{sy} p(y) dy \\
&= \int_{0}^{\infty} e^{sx^2} \frac{\text{pdf}_{\text{N}(0,1)}(\sqrt{y})}{\sqrt{y}} \underbrace{dy}_{=2xdx} \\
&= \int_{0}^{\infty} e^{sx^2} \frac{\text{pdf}_{\text{N}(0,1)}(x)}{x} 2xdx \\
&= 2\int_{0}^{\infty} e^{sx^2}\, \text{pdf}_{\text{N}(0,1)}(x) dx \\
&= \int_{-\infty}^{\infty} e^{sx^2}\, \text{pdf}_{\text{N}(0,1)}(x) dx \\
&= \text{E}[e^{sx^2}]_{\text{N}(0,1)}.
\end{aligned}$$

Here we used Equation A.89 for $p(y)$ in the 3rd row and made the substitution $x = \sqrt{y}$ in the 4th row.

An explicit expression for the density function of the χ^2-distribution can also be obtained by taking the following approach: if we succeed in transforming Equation A.83 into the form $\int e^{sx}p(x)dx$, this integral can be interpreted as the expectation of e^{sx} taken with respect to the probability density $p(x)$. This expectation is then by definition the MGF of $p(x)$. On the other hand, Equation A.83 is the MGF of $\chi^2(1)$-distribution. Since two distributions are exactly equal if they have the same MGF[7] we could then conclude that p is the density function of $\chi^2(1)$. In this way, knowledge of p would give us an explicit expression for the density function of $\chi^2(1)$. We wish to take this approach now.

First, we observe that in the integrand in Equation A.83, x appears only in the form x^2, i.e., the integral is symmetric about zero; thus we can write

$$G_{\chi^2(1)}(s) = \frac{2}{\sqrt{2\pi}} \int_0^\infty e^{sx^2} e^{-x^2/2} dx.$$

We then perform the following substitution in the integral

$$\begin{aligned} y &:= x^2 \Longrightarrow y = +\sqrt{y} \quad \text{since } x \geq 0 \\ y(x = 0) &= 0, \quad y(x = \infty) = \infty \\ \frac{dy}{dx} &= 2x \Longrightarrow dx = \frac{1}{2\sqrt{y}} dy \end{aligned}$$

and obtain

$$\begin{aligned} G_{\chi^2(1)}(s) &= \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{sy} e^{-y/2} \frac{1}{\sqrt{y}} dy \\ &= \int_0^\infty e^{sy} p(y) dy = \mathrm{E}[e^{sy}]_p. \end{aligned}$$

We have thus attained our goal. The MGF of $\chi^2(1)$ is expressed as the expectation of e^{sy} with respect to a density function p. This density is the explicit expression for the density function of $\chi^2(1)$

$$\mathrm{pdf}_{\chi^2(1)}(x) = \frac{1}{\sqrt{2\pi}} x^{-1/2} e^{-x/2} \quad \text{with } x \in [0, \infty[\qquad \text{(A.91)}$$

in complete agreement with Equation A.90.

The range of an χ^2-distributed random variable is thus restricted to the positive real numbers which makes sense since it is the *square* of a standard normally distributed random variable. Comparing this density to the density of the gamma distribution in Equation A.68 shows that $\chi^2(1)$ is equal to the gamma distribution for the parameters $\lambda = 1/2$ and $t = 1/2$. This can also be seen immediately by comparing their moment generating function in Equations A.85 and A.72; they are equal for $\lambda = 1/2$ and $t = 1/2$ and thus the associated density functions must be the same for these parameter

---

7 Since the MGF defines all (infinitely many) moments of a distribution, the distribution is completely determined by the MGF. This implies that if two MGFs are equal then they are the MGF of the *same* distribution.

values as well. Likewise, a comparison of the moment generating functions in Equations A.86 and A.72 shows that the χ^2-ditribution with n degrees of freedom equals the gamma distribution with parameters $\lambda = 1/2$ and $t = n/2$:

$$\text{pdf}_{\chi^2(n)}(x) = \frac{1}{\Gamma(n/2)}\left(\frac{1}{2}\right)^{n/2} x^{\frac{n}{2}-1}e^{-x/2} \text{ with } x \in [0, \infty[, \quad n = 1, 2, \ldots \tag{A.92}$$

Due to the recursion relation in Equation A.70 with the initial values A.71, the gamma functions appearing here can be given explicitly as:

$$\Gamma(n/2) = \begin{cases} (n/2-1)! & \text{for even values of } n \\ (n/2-1)(n/2-2)(n/2-3)\cdots(1/2)\sqrt{\pi} & \text{for odd values of } n \end{cases}.$$

The Noncentral χ^2-distribution

The χ^2-distribution described above is the distribution of a sum of n squared independent *standard* normal random numbers x_i, $(i = 1, \ldots, n)$. A slight but often needed generalization of this is the situation in which the random numbers x_i have expectations $\mu_i \neq 0$. The distribution of a sum of n squared random numbers of this type is called the *noncentral* χ^2-distribution with n degrees of freedom and with *noncentral parameter* θ, where θ denotes the sum of the squared expectations μ_i:

$$x_i \sim \text{N}(\mu_i, 1), \quad i = 1, \ldots, n, \quad x_i \ iid$$
$$\Longrightarrow$$
$$\sum_{i=1}^{n} x_i^2 =: y \sim \chi^2(n, \theta) \quad \text{with} \quad \theta = \sum_{i=1}^{n} \mu_i^2 \tag{A.93}$$

The square of a single random number $x \sim \text{N}(0, \mu)$ thus has the noncentral χ^2-distribution with one degree of freedom:

$$x \sim \text{N}(\mu, 1) \Longrightarrow x^2 =: y \sim \chi^2(1, \mu^2).$$

To determine the moment generating function of the noncentral χ^2-distribution we again start from the general Equation A.82:

$$\begin{aligned} G_{\chi^2(1,\mu^2)}(s) &= \text{E}[e^{sx^2}]_{\text{N}(\mu,1)} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sx^2} e^{-(x-\mu)^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{sx^2 - \frac{1}{2}(x-\mu)^2\right\} dx. \end{aligned}$$

The decisive step to calculate this integral is to complete the square in the argument of the exp-function:

$$\begin{aligned} sx^2 - \frac{1}{2}(x-\mu)^2 &= -\frac{1}{2}\left[(1-2s)x^2 - 2\mu x + \mu^2\right] \\ &= -\frac{1}{2}\left[\left(\sqrt{1-2s}x - \frac{1}{\sqrt{1-2s}}\mu\right)^2 - \frac{1}{1-2s}\mu^2 + \mu^2\right] \\ &= -\frac{1}{2}\left(\sqrt{1-2s}x - \frac{1}{\sqrt{1-2s}}\mu\right)^2 + \frac{s\mu^2}{1-2s}. \end{aligned}$$

Thus, the expectation becomes

$$\begin{aligned} &\mathrm{E}[e^{sx^2}]_{\mathrm{N}(\mu,1)} \\ &= \exp\left\{\frac{s\mu^2}{1-2s}\right\}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2}\left(\sqrt{1-2s}x - \frac{1}{\sqrt{1-2s}}\mu\right)^2\right\}dx \\ &= \frac{1}{\sqrt{1-2s}}\exp\left\{\frac{s\mu^2}{1-2s}\right\}\underbrace{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2}\left(u - \frac{\mu}{\sqrt{1-2s}}\right)^2\right\}du}_{1} \end{aligned}$$

where in the last step we have used the substitution A.84 (again with the condition $s < 1/2$). The remaining integral is the probability that a normally distributed random number with expectation $\mu/\sqrt{1-2s}$ will assume *any* arbitrary value and, according to Equation A.2, equals one. Thus, the MGF of the noncentral χ^2-distribution with one degree of freedom is

$$G_{\chi^2(1,\mu^2)}(s) = \frac{1}{\sqrt{1-2s}}\exp\left\{\frac{s\mu^2}{1-2s}\right\} \tag{A.94}$$

Since the random variables in the sum appearing in the Definition A.93 are all *independent* we can use Equation A.30 to immediately obtain the MGF for a noncentral χ^2-distribution with n degrees of freedom:

$$G_{\chi^2(n,\theta)}(s) = \frac{1}{(1-2s)^{n/2}}\exp\left\{\frac{s\theta}{1-2s}\right\} \quad \text{with} \quad \theta = \sum_{j=1}^{n}\mu_j^2 \tag{A.95}$$

The characteristic function, Equation A.32, of the noncentral χ^2-distribution follows again by replacing s by is in Equation A.95.

$$\Phi_{\chi^2(n,\theta)}(s) = \frac{1}{(1-2is)^{n/2}}\exp\left\{\frac{is\theta}{1-2is}\right\} \quad \text{with} \quad \theta = \sum_{j=1}^{n}\mu_j^2 \tag{A.96}$$

A.5 TRANSFORMATIONS BETWEEN DISTRIBUTIONS

It is possible to transform random variables with a certain distribution into random variables which have *another* distribution. This is particularly useful, for instance, when simulating random walks with the Monte Carlo method. Most *random number generators* generate uniformly distributed random numbers. These can be transformed into random numbers distributed according to a more suitable (for example, normal) distribution function. Two examples of transformations into normally distributed random variables and one transformation into random variables governed by any desired distribution will be introduced below.

A.5.1 Summations

By directly applying the central limit theorem, a sufficient number of independent, identically distributed, random variables are added to obtain an (approximately) normally distributed random variable:

z_i iid, uniformly distributed between 0 and 1

$$\Rightarrow \sqrt{\frac{12}{n}}\sum_{i=1}^{n}\left(z_i - \frac{1}{2}\right) =: x \sim \mathrm{N}(0,1) \tag{A.97}$$

According to Equation A.37, uniformly distributed random variables z_i on [0,1] have an expectation of 1/2 and a variance of 1/12. The random numbers $(Z_i - 1/2)$ are uniformly distributed between $-1/2$ and $+1/2$, thus having zero expectation. However, the variance of these variables remains 1/12. According to Equation A.18 the variance of the sum of these variables is

$$\mathrm{var}\left[\sum_{i=1}^{n}\left(z_i - \frac{1}{2}\right)\right] = \sum_{i=1}^{n}\mathrm{var}\left[\left(z_i - \frac{1}{2}\right)\right] = \sum_{i=1}^{n}\frac{1}{12} = \frac{n}{12}.$$

Thus, this sum has zero expectation and a variance of $n/12$, or equivalently a standard deviation of $\sqrt{n/12}$. Dividing the sum by the factor $\sqrt{12/n}$ compensates for this standard deviation. The resulting random variable x in Equation A.97 has a variance of 1 and is thus in consequence of the central limit theorem (approximately) a *standard* normally distributed random variable.

n uniformly distributed *iid* random numbers are required in order to generate a single, approximately normally distributed random variable. For most applications, it is sufficient to take $n = 12$. Figures A.1 and A.2 show the effect of the transformation for $n = 12$.

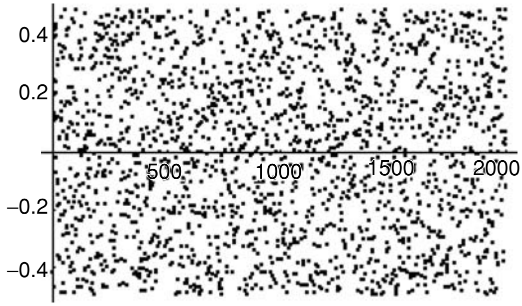

Figure A.1 2000 *iid* random numbers uniformly distributed between −1/2 and +1/2

Figure A.2 The random numbers from Figure A.1 after Transformation A.97 with $n = 12$. The random numbers are now approximately N(0,1)-distributed

Note that the transformed random variables can only take on values between $-\sqrt{3n}$ and $+\sqrt{3n}$. Events lying more than $\sqrt{3n}$ standard deviations from the expectation will never occur when using this transformation. The probability of such an event is so small, however, that it does not play a role in most practical applications in finance. For example, for $n = 12$, the probability is approximately 1 to half a billion:

$$1 - \frac{1}{\sqrt{2\pi}} \int_{-6}^{6} e^{-x^2/2} dx \approx 1,973210^{-9}.$$

A.5.2 Box-Muller transformations

The transformation in Equation A.97 requires a significant amount of effort, since n random numbers must be generated in order to obtain one single normally distributed random number. The transformation found by *Box* and *Muller* [14] is much more effective. From two independent uniformly distributed random numbers, two independent, normally distributed random variables are generated as follows:

$$x_1 = \sqrt{-2\ln(z_1)}\cos(2\pi z_2), \quad x_2 = \sqrt{-2\ln(z_1)}\sin(2\pi z_2) \qquad \text{(A.98)}$$

with uniformly distributed z_1, z_2 between 0 and 1 $\Rightarrow$

$$x_1 \sim \mathrm{N}(0,1), \quad x_2 \sim \mathrm{N}(0,1), \quad \mathrm{cov}\,[x_1, x_2] = 0$$

A.5.3 Inversion of cumulative distribution functions

A very simple method for generating random numbers obeying any desired distribution from a set of random numbers uniformly distributed on the interval between 0 and 1 is to evaluate the inverse cumulative probability function of the desired distribution with the uniformly distributed random numbers as its arguments. Or more precisely: let z be a *uniformly* distributed random variables taking on values in the interval between 0 and 1. Let $f(z)$ be an arbitrary distribution density function with $f(z) > 0\ \forall z$. Then

$$x := F^{-1}(z) \quad \text{with} \quad F(z) = \int_{-\infty}^{z} f(u)du \qquad \text{(A.99)}$$

is distributed according to the distribution associated with the density f. Intuitively, since the uniformly distributed random number z lies between 0 and 1, it can be interpreted as a "probability." The inverse of the cumulative distribution function is then nothing other than the *percentile* of this "probability" z with respect to the desired distribution. The random, uniformly distributed numbers are thus interpreted as probabilities. The percentiles associated with these "probabilities" with respect to the *desired* distribution are then random variables distributed according to the *desired* distribution.

This surprisingly simple method for generating random numbers distributed according to an arbitrary distribution is so important in practice that we will provide the reader with a proof. We begin by proving the existence of an inverse of the cumulative distribution by showing that the cumulative distribution is strictly increasing if the associated density function is strictly positive:

$$\begin{aligned} F(x+dx) &\equiv \int_{-\infty}^{x+dx} f(u)du = \int_{-\infty}^{x} f(u)du + \int_{x}^{x+dx} f(u)du \\ &= F(x) + f(x)dx > F(x) \quad \text{since } f(x) > 0. \end{aligned}$$

Generally $f(x)dx$ is the probability that a random number governed by a distribution density f will take on a value between x and $x + dx$. This is just the definition of the distribution density, i.e.,

$$\begin{aligned} f(x)dx &= F(x+dx) - F(x) \\ &= z + dz - z = dz \end{aligned}$$

where in the second step, we have made use of the fact that Equation A.99 implies $F(x) = z$ and, because F is monotone, $F(x+dx) = z+dz$ holds as well. In summary, we may write

$$z = F(x) \Longleftrightarrow dz = f(x)dx \tag{A.100}$$

A random variable, uniformly distributed on the interval $[a, b]$, has a constant density function $p(z) = 1/(b-a)$ for all z in the interval $[a, b]$ as was shown in Equation A.36. In the interval $[0, 1]$ the density of the uniform distribution is thus $p(z) = 1$. The probability that such a uniformly distributed random variable will lie between z and $z+dz$ is therefore simply $p(z)dz = dz$. It follows that Equation A.100 can be interpreted as follows: if the random variable x is governed by a distribution with density function $f(x)$ then $z = F(x)$ is *uniformly* distributed with density function 1, i.e., uniformly distributed on an interval of length 1. Since equality holds in each step of the above derivation, the conclusion holds in both directions: if z is uniformly distributed on an interval of length 1 (and thus has a distribution density of 1) then $x = F^{-1}(z)$ is distributed according to the density $f(x)$. qed.

One application of Equation A.99 is the generation of standard normally distributed random numbers:

$$z \text{ uniformly distributed between 0 and 1} \Rightarrow \mathrm{N}^{-1}(z) =: x \sim \mathrm{N}(0, 1) \tag{A.101}$$

The cumulative distribution function of the normal distribution has the disadvantage that it and its inverse function cannot be computed analytically, although they can be computed with applications commonly offered in many widely available software packages. For demonstration purposes, however, it is instructive to apply this method to a distribution which *can* be treated analytically. Consider therefore what is known as the *Cauchy distribution.* The Cauchy distribution has a parameter $\lambda > 0$ and the density function

$$f_\lambda(x) = \frac{\lambda}{\pi\left(\lambda^2 + x^2\right)} \tag{A.102}$$

The cumulative distribution function of a Cauchy distributed random variable can be computed analytically with little difficulty. The necessary integral can be found in any collection of mathematical formulas (see for instance [22]):

$$\begin{aligned} F_\lambda(x) &\equiv \int_{-\infty}^{x} f_\lambda(u)du = \frac{\lambda}{\pi}\int_{-\infty}^{x} \frac{1}{\left(\lambda^2 + u^2\right)}du \\ &= \frac{\lambda}{\pi}\left[\frac{1}{\lambda}\arctan\left(\frac{u}{\lambda}\right)\right]_{u=-\infty}^{u=x} = \frac{1}{\pi}\arctan\left(\frac{x}{\lambda}\right) + \frac{1}{2} \end{aligned}$$

where several well-known properties of the *arctangent function* have been employed. We thus have an analytic expression of the cumulative Cauchy distribution. The inverse function can now be obtained by solving the equation $z = F_\lambda(x)$ for x:

$$\begin{aligned}
F_\lambda(x) &= z \\
\frac{1}{\pi}\arctan\left(\frac{x}{\lambda}\right) + \frac{1}{2} &= z \\
\arctan\left(\frac{x}{\lambda}\right) &= \pi\left(z - \frac{1}{2}\right) \\
\frac{x}{\lambda} &= \tan\left[\pi\left(z - \frac{1}{2}\right)\right] \\
x &= \lambda\tan\left[\pi z - \frac{\pi}{2}\right] \\
&= -\lambda\cot(\pi Z).
\end{aligned}$$

Thus, the inverse of the cumulative distribution function of the Cauchy distribution is

$$F_\lambda^{-1}(z) = x = -\lambda\cot(\pi z).$$

This implies that if z is a uniformly distributed random variable on $[0, 1]$, then $x = -\lambda\cot(\pi z)$ is a Cauchy distributed random variable taking on values between $-\infty$ and ∞ and having a distribution density function as given in Equation A.102.

Bibliography

[1] Abramowitz, M., Stegun, I., *Handbook of Mathematical Functions*, Dover Publications, New York, 1972.

[2] Alexander, C. (Ed.), *The Handbook of Risk Management and Analysis*, John Wiley & Sons, Chichester, 1996.

[3] Anderson, N., Breedon, F., Deacon, M., Derry, A., Murphy, G., *Estimating and Interpreting the Yield Curve*, John Wiley & Sons, New York, 1996.

[4] Babbel, D. F., Merrill, C. B., *Valuation of Interest-Sensitive Instruments*, Society of Actuaries, Schaumburg, Illinois, 1996.

[5] Baschnagel, J., Paul, W., *Stochastic Processes – From Physics to Finance*, Springer Verlag, Berlin, 2000.

[6] Basler Ausschuß für Bankenaufsicht, *Richtlinien für das Risikomanagement im Derivativgeschäft*, Juli, Bank for International Settlements, Basel, 1994.

[7] Baxter, M., Rennie, A., *Financial Calculus*, Cambridge University Press, Cambridge, 1997.

[8] Beau, C. L., Lucas, D., *Technical Traders Guide to Computer Analyses of the Futures Market*, Hoppenstedt, Darmstadt, 1992.

[9] Becker, U., *Lexikon Terminhandel*, Gabler Verlag, Wiesbaden, 1994.

[10] Best, P., *Implementing Value at Risk*, John Wiley & Sons, New York, 1998.

[11] Binder, K., *Monte Carlo Simulations in Statistical Physics*, John Wiley & Sons, New York, 1986.

[12] Bittner, J., Goddard, R., *Successful Bank Asset-Liability Management*, John Wiley & Sons, New York, 1992.

[13] Black, F., Karasinski, P., *Bond and Option Pricing when Short Rates are Lognormal*, Financial Analysts Journal, July/August 1991, 52–59.

[14] Box, G. E. P., Muller, M. E., A Note on the Generation of Random Normal Deviates, *Annals of Mathematical Statistics*, 29, 1958, 610–611.

[15] Panjer, Harry H., Boyle, Phelim P., Cox, Samuel H., Gerber, Hans U., Mueller, Heinz H., Pedersen, Hal W., Pliska, Stanley R., Sherris, Michael, Shiu, Elias S., Tan, Ken Seng, *Financial Economics*, The Actuarial Foundation, Schaumburg, Illinois, 1998.

[16] Brace, A., Gatarek, D., Musiela, M., The Market Model of Interest Rate Dynamics, *Mathematical Finance*, 7 (2) (April, 1997), 127–147.

[17] Brigham, E. O., *The Fast Fourier Transform*, Prentice Hall, Englewood Cliffs, New Jersey, 1974.

[18] Broadie, M., Glasserman, P. (Eds), *Hedging with Trees*, Risk Publications, London, 1998.

[19] Brockhaus, O., Farkas, M., Ferraris, A., Long, D., Overhaus, M., *Equity Derivatives and Market Risk Models*, Risk Books, London, 2000.

[20] Brockhaus, Oliver, Ferraris, A., Gallus, C., Long, D., Martin, R., Overhaus, M., Gallus, C., *Modelling and Hedging Equity Derivatives*, Risk Books, London, 1999.

[21] Brockwell, P. J., Davis, R. A., *Introduction to Time Series and Forecasting*, Springer Verlag, Heidelberg, 1998.

[22] Bronstein, S., *Taschenbuch der Mathematik*, Verlag Harri Deutsch, Frankfurt, 1979.

[23] Bundesaufsichtsamt für das Kreditwesen, *Neufassung des Grundsatzes I, Erläuterungen zum Vorentwurf*, Berlin, 10.5.96, Geschäftsnummer 17-A223-2/93.

[24] Bundesaufsichtsamt für das Kreditwesen, *Verlautbarung über Mindestanforderungen an das Betreiben von Handelsgeschäften der Kreditinstitute*, Berlin, 23.10.95, Geschäftsnummer I4-42-3/86.
[25] Byron, J. T. M., *Elements of Simulation*, Chapman & Hall, London, 1984.
[26] Calamos, J. P., *Convertible Securities*, McGraw-Hill, New York, 1998.
[27] Campbell, J. Y., Lo, A. W., Mackinlay, A. G., *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey, 1997.
[28] Chiang, A. C., *Fundamental Methods of Mathematical Economics*, 2nd edition, McGraw-Hill, Singapore, 1984.
[29] Chriss, N. A., *Black Scholes and Beyond*, McGraw-Hill, New York, 1997.
[30] Clarke R., *Estimating and Using Volatility: Part 1*, Derivatives Quarterly, Fall 1994, S. 40–44.
[31] Clewlow, L., Strickland, C., *Implementing Derivatives Models*, John Wiley & Sons, New York, 1998.
[32] Collinson, F., Giddings, M., Sullivan, M., *Financial Products: A Survival Guide*, Euromoney Books, 1996.
[33] Connolly, K. B., *Pricing Convertible Bonds*, John Wiley & Sons, New York, 1998.
[34] Cornish, E. A., Fisher, R. A., *Moments and Cumulants in the Specification of Distributions*, Review of the International Statistical Institute, 5, 1937, 307–320.
[35] Cox, J. C., Rubinstein, M., *Options Markets*, Prentice Hall, Englewood Cliffs, New Jersey, 1985.
[36] Cox, J. C., Ingersoll, J. E., Ross, S. A., *A Theory of the Term Structure of Interest Rates*, Econometrica, 53, 1985, 385–407.
[37] Culp, L., Miller, M. H. (Eds), *Corporate Hedging in Theory and Practice*, Risk Publications, London, 1999.
[38] Das, S., *Credit Derivatives*, John Wiley & Sons, New York, 1998.
[39] Das, S., *Swap and Derivative Financing*, revised ed., McGraw-Hill, New York, 1994.
[40] Dattatreya, R. E., Fabozzi, F. J., *Active Total Return Management of Fixed Income Portfolios*, Probus, Chicago, 1989.
[41] Davenport, W. B., *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, New York, 1958, chap. 6.
[42] Deutsch, H. -P., *Quantitative Portfoliosteuerung*, Schäffer-Poeschel Verlag, Stuttgart, 2005.
[43] Dixit, A. K., Pindick, R. S., *Investment under Uncertainty*, Princeton University Press, Princeton, New Jersey, 1994.
[44] Doi, M., Edwards, S. F., *The Theory of Polymer Dynamics, Clarendon Press*, Oxford University Press, Oxford, 1986.
[45] Dothan, M., *Prices in Financial Markets*, Oxford University Press, New York, 1990.
[46] Dowd, K., *Beyond Value-At-Risk*, John Wiley & Sons, New York, 1998.
[47] Duffie, D., *Security Markets – Stochastic Models*, Academic Press, San Diego, 1988.
[48] Duffie, D., *Dynamic Asset Pricing Theory*, 2nd edition, Princeton University Press, Princeton, New Jersey, 1996.
[49] Dunis, C., Zhou, B., *Nonlinear Modelling of High Frequency Financial Time Series*, John Wiley & Sons, New York, 1998.
[50] Dupire, B., *Monte Carlo – Methodologies and Applications*, Risk Books, London, 1998.
[51] Eller, R. (Ed.), *Handbuch Derivativer Instrumente*, Schäffer-Poeschel Verlag, Stuttgart, 1996.
[52] Eller, R. (Ed.), *Handbuch des Risikomanagements*, Schäffer-Poeschel Verlag, Stuttgart, 1997.
[53] Eller, R., Gruber, W., Reif, M. (Eds), *Handbuch Bankenaufsicht und Interne Risikosteuerungsmodelle*, Schäffer-Poeschel Verlag, Stuttgart, 1998.
[54] Elton, E., Gruber, M., *Modern Portfolio Theory and Investment Analysis*, 5th edition, John Wiley & Sons, New York, 1995.
[55] Embrechts, P., Klüppelberg, C., Mikosch, T., *Modelling Extremal Events for Insurance and Finance*, Springer Verlag, Heidelberg, 1999.
[56] Fabozzi, F. J., *The Handbook of Fixed Income Securities*, 5th edition, McGraw-Hill, New York, 2001.
[57] Fabozzi, F. J., *Valuation of Fixed Income Securities and Derivatives*, 3rd edition, John Wiley and Sons, New York, 1998.
[58] Feller, W., *Probability Theory and Its Applications*, vols. 1 and 2, John Wiley & Sons, New York, 1950.

[59] Ferris, K., Jones, M., *The Reuters Guide to Official Interest Rates*, 2nd edition, Probus, Chicago, 1995.
[60] Fishman, G. S., *Monte Carlo Concepts, Algorithms and Applications*, Springer Verlag, New York, 1996.
[61] Flury, B., *A First Course in Multivariate Statistics*, Springer Verlag, Heidelberg, 1997.
[62] Francis, J., Clark, F., Joyce A., Whittaker, J. Gregg, *The Handbook of Credit Derivatives*, McGraw-Hill, New York, 1998.
[63] Gardiner, C. W., *Handbook of Stochastic Methods*, 2nd edition, Springer Verlag, Heidelberg, 1985.
[64] Gardner, M. J., Mills, D. L., *Managing Financial Institutions*, 3rd edition, Dryden, Fort Worth, 1994.
[65] Gershenfeld, N., *The Nature of Mathematical Modelling*, Cambridge University Press, Cambridge, 1999.
[66] Gibson, R. (Ed.), *Model Risk – Concepts, Calibration and Pricing*, Risk Publications, London, 1999.
[67] Greene, W. H., *Econometric Analysis*, 3rd edition, Prentice Hall, London, 1997.
[68] Grill, W., Gramlich, L., Eller, R., *Gabler Bank Lexikon*, 11th edition, Gabler Verlag, Wiesbaden, 1995.
[69] Grimmett, G. R., Stirzaker, D. R., *Probability and Random Processes*, 2nd edition, Oxford University Press, Oxford, 1998.
[70] Grinold, R., Kahn, R., *Active Portfolio Management*, Probus, Chicago, 1995.
[71] Group of Thirty, *Derivatives, Practices and Principles*, Washington D.C., 1993.
[72] Gruber, W., Overbeck, L., Working Paper, Deutsche Bundesbank, Frankfurt, 1996.
[73] Gwilym, O., Sutcliffe, C. (Eds), *High-Frequency Financial Market Data*, Risk Publications, London, 1999.
[74] Hamilton, J. D., *Time Series Analysis*, Princeton University Press, Princeton, New Jersey, 1994.
[75] Hammer D., *Dynamic Asset Allocation*, John Wiley & Sons, New York, 1991.
[76] Hanenberg, L., *Zur Verlautbarung über Mindestanforderungen an das Betreiben von Handelsgeschäften der Kreditinstitute des Bundesaufsichtsamtes für das Kreditwesen*, Die Wirtschaftsprüfung, S. 637–648, Heft 18/1996, Jahrgang 49.
[77] Harrison, M., Pliska, S., *Martingales and Stochastic Integrals In the Theory of Continuous Trading, Stochastic Processes and their Applications*, 11, 1981, 215–260.
[78] Haug, Gaarder, *The Comlete Guide to Option Pricing Formulas*, McGraw Hill, New York, 1997.
[79] Haugen, R. A., *Modern Investment Theory*, 4th edition, Prentice Hall, Upper Saddle River, New Jersey, 1997.
[80] Heath, D., Jarrow, R., Morton, A., *Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation*, Econometrica, 60, 1992, 77–105.
[81] Heidorn, T., *Vom Zins zur Option – Finanzmathematik in der Bankpraxis*, Gabler Verlag, Wiesbaden, 1994.
[82] Hess, D. E., *Die Dynamik der Zinsstruktur*, Gabler Edition Wissenschaft, Wiesbaden, 1995.
[83] Hicks, A., *Foreign Exchange Options*, 2nd edition, Woodhead Publishing, Cambridge, 1998.
[84] Higgins, R. C., *Analysis for Financial Management*, 5th edition, Irwin/ McGraw-Hill, Boston, 1998.
[85] Ho, S. Y., Lee, S. -B., *Term Structure Movements and Pricing Interest Rate Contingent Claims*, Journal of Finance, 41, 1986, 1011–1029.
[86] Ho, Thomas, *Fixed Income Solutions*, Irwin, Chicago, 1996.
[87] Hughston, L. (Ed.), *Vasicek and Beyond: Approaches to Building and Applying IR-Models*, Risk Publications/Euro Brokers, London, 1996.
[88] Hughston, L. (Ed.), *Options*, Risk Publications, London, 1999.
[89] Hughston, L. (Ed.), *The New Interest Rate Models: Recent Developments in the Theory and Application of Yield Curve Dynamics*, Risk Books, London, 2000.
[90] Hull, J. C., *Options, Futures and other Derivatives*, 4th edition, Prentice Hall, Upper Saddle River, New Jersey, 2000.
[91] Hull, J. C., White, A., *Hull-White on Derivatives*, Risk Publications, London, 1996.
[92] Hull, J. C., White, A., *Pricing Interest Rate Derivative Securities*, The Review of Financial Studies, 3, 1990, 573–592.

[93] Hunt, P. J., Kennedy, J. E., *Financial Derivatives in Theory and Practice*, John Wiley & Sons, Chichester, 2000.

[94] Ingersoll Jr., Jonathan E., *Theory of Financial Decision Making*, Rowman & Littlefield, Bollman Place, Savage, Maryland, 1987.

[95] James, J., Webber, N., *Interest Rate Modelling*, John Wiley & Sons, Chichester, 2000.

[96] Jamshidian, F., Option and Futures Evaluation with Deterministic Volatilities?, *Mathematical Finance*, 3 (2), 1993, 149–159.

[97] Jarrow, R., Turnball S., *Derivative Securities*, Southwestern Publishers, Cincinnati 1996.

[98] Jarrow, R., *Modelling Fixed Income Securities and Interest Rate Options*, McGraw Hill, Boston, 1996.

[99] Jarrow, R. (Ed.), *Over the Rainbow*, Risk Publications, London, 1995.

[100] Jarrow, R. (Ed.), *Volatility – New Estimation Techniques for Pricing Derivatives*, Risk Publications, London, 1998.

[101] Johnson, N. L., Kotz, S., Balakrishan, N., *Continuous Univariate Distributions*, vol. 1, John Wiley & Sons, New York, 1994.

[102] Johnson, R. A., Wichern, D. W., *Applied Multivariate Statistical Analysis*, 4th edition, Prentice Hall, Upper Saddle River, New Jersey, 1999.

[103] Jollife, I. T., *Principal Component Analysis*, Springer Verlag, Heidelberg, 1986.

[104] Jordan, D. W., Smith, P., *Mathematical Techniques*, 2nd edition, Oxford University Press, Oxford, 1998.

[105] Jorion, P., *Value at Risk: The New Benchmark for Controlling Derivatives Risk*, McGraw-Hill, New York, 1997.

[106] Kalos, M. H., Whitlock, P. A., *Monte Carlo Methods*, John Wiley & Sons, New York, 1986.

[107] Karatzas, I., Shreve, S. E., *Brownian Motion and Stochastic Calculus*, 2nd edition, Springer Verlag, Heidelberg, 1999.

[108] Karlin, H. M., *A First Course in Stochastic Processes*, Academic Press, New York, 1975.

[109] Klein, R. A., Lederman, J., *Derivatives Risk and Responsibility*, Irwin, Chicago, 1996.

[110] Kloeden, P. E., Platen, E., Schurz, H., *Numerical Solutions of SDE Through Computer Experiments*, Springer Verlag, Heidelberg, 1994.

[111] Kofler, M., *VBA-Programmierung mit Excel 97*, Pearson Education, München, 1997.

[112] Korajzykk, R. A. (Ed.), *Asset Pricing and Portfolio Performance*, Risk Publications, London, 1999.

[113] Krzanowski, W. J., *Principles of Multivariate Analysis*, Oxford Science Publications, Oxford, reprinted 1998.

[114] Kwok, Y. K., *Mathematical Models of Financial Derivatives*, Springer Verlag, Heidelberg, 1998.

[115] Lamberton, D., Lapeyere, B., *Introduction to Stochastic Calculus Applied to Finance*, Chapman & Hall, London, 1996.

[116] Leibowitz, M., Kogelman, S., Klaffky, T., *A Shortfall Approach to Duration Management*, Salomon Brothers, New York, 1990.

[117] Lewis, A. L., *Option Valuation under Stochastic Volatility*, Financepress, London, 2000.

[118] Lewis, E., *An Introduction to Credit Scoring*, Athena Press, San Rafael, California, 1994.

[119] Loistl, O., *Kapitalmarkttheorie*, Oldenbourg-Verlag, München, 1994.

[120] Longerstaey, J., *Introduction to RiskMetrics*, Morgan Guaranty Trust Company, New York, 1998.

[121] Mantegna, R. N., Stanley, H. E., *An Introduction to Econophysics*, Cambridge University Press, New York, 1999.

[122] Markowitz, H. M., *Mean Variance*, Blackwell, Berlin, 1987.

[123] Markowitz, H. M., *Portfolio Selection*, 2nd edition, Blackwell, Berlin, 1998.

[124] Matten, C., *Managing Bank Capital*, John Wiley & Sons, Chichester, 1996.

[125] Merton, R. C., *Continuous Time Finance*, revised edition, Blackwell, Malden, Massachusetts, 1998.

[126] Mills, T. C., *The Econometric Modelling of Financial Time Series*, Cambridge University Press, Cambridge, 1997.

[127] Miron, P., Swannel, P., *Pricing and Hedging of Swaps*, Euromoney Books, London, 1991.

[128] Musiela, M., Rutkowski, M., *Martingale Methods in Financial Modelling*, Springer Verlag, Heidelberg, 1997.

[129] Neftci, Salih N., *An Introduction to the Mathematics of Financial Derivatives*, 2nd edition, Academic Press, San Diego, 2000.
[130] Nelken, Israel, *The Handbook of Exotic Options*, Probus, Chicago, 1995.
[131] Niederreiter, H., *Random Number Generation and Quasi Monte Carlo Methods*, SIAM, Philadelphia, 1992.
[132] Nielsen, L. T., *Pricing and Hedging of Derivative Securities*, Oxford University Press, Oxford, 1999.
[133] Oksendal, B., *Stochastic Differntial Equations*, 5th edition, Springer Verlag, Heidelberg, 1998.
[134] Pindyck, R. S., Rubinfeld, D. L., *Econometric Models and Economic Forecasts*, 4th edition, McGraw Hill, Boston, 1997.
[135] Press, W. H., Teukolsky, S.A., Vetterling, W. T., Flannery, B. P., *Numerical Recipes in C – The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1994.
[136] Prudenzi, P. S., *VBA mit Excel 97*, Addison-Wesley, Bonn, 1997.
[137] Questa, G. S., *Fixed Income Analysis for the Global Financial Market*, John Wiley & Sons, New York, 1999.
[138] Rebonato, R., *Interest-Rate Option Models*, John Wiley & Sons, Chichester, 1996.
[139] Rebonato, R., *Volatility & Correlation*, John Wiley & Sons, Chichester, 1999.
[140] Rice, J. A., *Mathematical Statistics and Data Analysis*, 2nd edition, Duxbury Press, Boston, 1994.
[141] Ripley, B. D., *Stochastic Simulation*, John Wiley & Sons, New York, 1987.
[142] RiskMetrics, *Technical Document*, 4th edition, JP Morgan Global Research, New York, 1996.
[143] Rönz, B., Förster, E., *Regressions – und Korrelationsanalyse, Grundlagen, Methoden, Beispiele*, Gabler Verlag, Wiesbaden, 1992.
[144] Rouvinez, C., *Going Greek with VaR*, RISK Magazine, 10 (2), February 1994.
[145] Rubinstein, M., *Rubinstein on Derivatives*, Risk Publications, London, 1999.
[146] Rubinstein, M., *Return to Oz*, RISK Magazine, 10 (11), November 1994.
[147] Rubinstein, R. Y., *Simulation and the Monte Carlo Method*, John Wiley & Sons, New York, 1981.
[148] Schierenbeck, H., *Bank – und Versicherungslexikon*, Oldenbourg Verlag, Munich, 1990.
[149] Schierenbeck, H., *Ertragsorientiertes Bankmanagement*, 5th edition, Gabler Verlag, Wiesbaden, 1997.
[150] Sharpe, W. F., Alexander, G. J., Bailey, J. V., *Investments*, 5th edition, Prentice Hall, Upper Saddle River, New Jersey, 1995.
[151] Smith C. W., Smithson C. W., *The Handbook of Financial Engineering, New Financial Product Innovations, Applications, and Analyses*, Harper Business Books, New York, 1990.
[152] Smith, C. D., *Option Strategies*, John Wiley & Sons, New York, 1996.
[153] Smith, G. D., *Numerical Solution of Partial Differential Equations*, 3rd edition, Oxford University Press, Oxford, 1998.
[154] Taleb, N., *Dynamic Hedging*, John Wiley & Sons, New York, 1997.
[155] Tavakoli, J. M., *Credit Derivatives*, John Wiley & Sons, New York, 1998.
[156] Thomas, J. W., *Numerical Partial Differential Equations*, Springer Verlag, Heidelberg, 1995.
[157] Tavella, D., Randall, C., *Pricing Financial Instruments: The Finite Difference Method*, John Wiley & Sons, New York, 2000.
[158] Tiner, J. I., Conneely, J. M., *Accounting for Treasury Products*, 2nd edition, Woodhead-Faulkner, Cambridge, 1989.
[159] Tomkins, R., *Options Explained*, MacMillan Business, London, 1994.
[160] Tong, Howell, *Non-linear Time Series*, Oxford Science Publications, Oxford, reprinted 1999.
[161] Uyemura, D. G., van Deventer, D. R., *Financial Risk Management in Banking*, McGraw Hill, New York, 1993.
[162] Vasicek, O. A., *An Equilibrium Characterization of the Term Structure*, Journal of Finance, 5, 1977, 177–188.
[163] Walmsley, J., *The New Financial Instruments*, John Wiley & Sons, New York, 1988.
[164] Welcker J., Kloy, J. W., Schindler, K., *Professionelles Optionsgeschäft*, MI Verlag, Zürich, 1992.
[165] Williams, D., *Probability with Martingales*, Cambridge Mathematical Textbooks, Cambridge, 1998.
[166] Wilmott, P. J., *Derivatives – The Theory and Practice of Financial Engineering*, John Wiley & Sons, New York, 1998.

[167] Wilmott, P. J., Dewynne, J., Howison, S., *The Mathematics of Financial Derivatives*, Cambridge University Press, Cambridge, 1998.

[168] Wilmott, P. J., Dewynne, J., Howison, S., *Option Pricing – Mathematical Models and Computation*, Oxford Financial Press, Oxford, 1997.

[169] Wilson, T. C., Calculating Risk Capital, in C. Alexander (Ed.), *Risk Management and Analysis*, John Wiley & Sons, New York, 1998.

[170] Zangari P., A VaR Methodology for Portfolios that Include Options, *RiskMetrics Monitor*, 1st quarter, 1996.

[171] Zhang, P. G., *Exotic Options: a Guide to Second Generation Options*, World Scientific, Singapore, 1997.

[172] Zimmermann, H., Jaeger, S., Staub, Z., *Asset – and Liability-Management*, 2nd edition, Verlag neue Zürcher Zeitung, Zurich, 1996.

Index