Mathematical Appendix

We present in this concise appendix the main probabilistic notions used in the previous chapters. For more details, the diligent reader may refer to Jacod and Protter (2004) for the classical Probability theory, to Williams (1991) for martingales and to Lapeyre et al. (2013) and Robert and Casella (2010) for Monte Carlo methods.

Let $(\Omega, \mathcal{A}, P)$ be a probability space where Ω is a set, $\mathcal{A}$ a σ-algebra and P a probability defined on $\mathcal{A}$. In this section we denote by E the expectation under the probability P. For $p \in \mathbb{N}^*$, we classically denote by $L^p(\Omega, \mathcal{A}, P)$ the set of the real random variables $Z : (\Omega, \mathcal{A}) \to \mathbb{R}$ such that $E[|Z|^p] < \infty$.

Gaussian Random Variables

A real random variable X follows a Gaussian distribution of mean m and variance σ^2 (the standard notation is $\mathcal{N}(m, \sigma^2)$) if its density function is equal to

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

The associated moment generating function is given by

$$E[e^{tX}] = e^{tm+\frac{\sigma^2 t^2}{2}}, \ \forall t \in \mathbb{R}.$$

There are no explicit expressions for the distribution function

$$N(x) = P(X \leq x)$$

of a $\mathcal{N}(m, \sigma^2)$, but due to the importance of this quantity in mathematical finance (in particular in the Black and Scholes formula (4.3)) we mention that very accurate approximations for N and N^{-1} exist (see Cody 1969) and are implemented in R via the *pnorm* and *qnorm* functions of the *stats* package.

The following result gives a simple method (known as the Box-Muller method) to generate samples of Gaussian distributions from classical random number

C. Chorro et al., *A Time Series Approach to Option Pricing*,
DOI 10.1007/978-3-662-45037-6

generators. In R, this method may be used via the *rnorm* command selecting *RNGkind(normal.kind =Box-Muller)* in the *stats* package.

Proposition A.1 *If U_1 and U_2 are two independent uniform random variables on* $[0, 1]$ *then*

$$G_1 = \sqrt{-2log(U_1)}cos(2\pi U_2) \text{ and } G_2 = \sqrt{-2log(U_1)}sin(2\pi U_2)$$

are two independent $\mathcal{N}(0, 1)$.

Proof Let us define

$$\Phi(x, y) = (u = \sqrt{-2log(x)}cos(2\pi y), v = \sqrt{-2log(x)}sin(2\pi y)).$$

It is easy to see that Φ is a C^1-diffeomorphism[1] from $]0, 1[^2$ into $\mathbb{R}^2-(\mathbb{R}_+\times\{0\})$ with a Jacobian determinant fulfilling $|J(\Phi)(x, y)| = \frac{2\pi}{x}$. Since $u^2+v^2 = -2log(x)$, we have $|J(\Phi^{-1})(u, v)| = \frac{1}{2\pi}e^{-\frac{u^2+v^2}{2}}$. According to the change of variables theorem, for any $F : \mathbb{R}^2 \to \mathbb{R}$ continuous and bounded

$$\int_{]0,1[^2} F(\Phi(x, y))dxdy = \int_{\mathbb{R}^2-(\mathbb{R}_+\times\{0\})} F(u, v)\frac{1}{2\pi}e^{-\frac{u^2+v^2}{2}}dudv$$

and the result follows. □

Conditional Expectation

Let X be a random variable in $L^1(\Omega, \mathcal{A}, P)$ and let $\mathcal{G}$ denote a sub σ-algebra of $\mathcal{A}$.

Definition A.1 There exists a random variable $Z \in L^1(\Omega, \mathcal{A}, P)$ such that

$$i)\ Z \text{ is } \mathcal{G}\text{-measurable}$$

$$ii)\ E[XU] = E[ZU],\ \forall U\ \mathcal{G}\text{-measurable and bounded.}$$

[1]If $\mathcal{O}_1$ and $\mathcal{O}_2$ are two open sets of $\mathbb{R}^n$, we say that $\Phi : \mathcal{O}_1 \to \mathcal{O}_2$ is a C^1-diffeomorphism if it is a bijection that is C^1 and if Φ^{-1} is C^1 as well. In this case, we define by $J(\Phi)$ the determinant of the matrix of the partial derivatives of Φ called the Jacobian and we have the change of variables theorem: $\forall F : \mathcal{O}_2 \to \mathbb{R}$ continuous and bounded,

$$\int_{\mathcal{O}_2} F(y)dy = \int_{\mathcal{O}_1} F(\Phi(x))|J(\Phi)(x)|dx.$$

Z is denoted by $E[X|\mathcal{G}]$ and is called the conditional expectation of X given $\mathcal{G}$. Moreover, Z is unique up to almost-sure equality.

Remark A.1 When Y is a random variable, $E[X|Y]$ is simply defined as $E[X|\sigma(Y)]$ where $\sigma(Y)$ is the smallest σ-algebra that makes Y measurable. In particular, when the pair (X, Y) owns a density $f_{(X,Y)}$ with respect to the Lebesgue measure on $\mathbb{R}^2$, the marginal densities of X and Y are given by

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x, y)dy \text{ and } f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x, y)dx.$$

If X and Y are independent, we have $f_{(X,Y)} = f_X f_Y$ and if the condition of independence is relaxed, we obtain the following disintegration formula:

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

where

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

if $f_Y(y) \neq 0$ and $f_{X|Y}(x, y) = 0$ otherwise. The function $f_{X|Y}$ is called the conditional density of X given Y. In fact, if ϕ satisfies $\phi(X) \in L^1(\Omega, \mathcal{A}, P)$,

$$E[\phi(X)|Y] = \Phi(Y) \text{ with } \Phi(y) = \int_{\mathbb{R}} \phi(x) f_{X|Y}(x, y)dx.$$

The practical computations involving conditional expectations are simplified by the following properties that are used all along this book:

Proposition A.2 *Let X and Y be two random variables in $L^1(\Omega, \mathcal{A}, P)$ and let $\mathcal{G}$ be a sub σ-algebra of $\mathcal{A}$, then,*

(a) (Positivity) If $X \geq 0$ P-a.s, $E[X|\mathcal{G}] \geq 0$ P-a.s.
(b) (Linearity) If $(\alpha, \beta) \in \mathbb{R}^2$, $E[\alpha X + \beta Y|\mathcal{G}] = \alpha E[X|\mathcal{G}] + \beta E[Y|\mathcal{G}]$ P-a.s.
(c) (Monotony) If $X \geq Y$ P-a.s, $E[X|\mathcal{G}] \geq E[Y|\mathcal{G}]$ P-a.s.
(d) (Jensen inequality) If $\psi : \mathbb{R} \to \mathbb{R}$ is a convex function such that $\psi(X) \in L^1(\Omega, \mathcal{A}, P)$, then, $\psi(E[X|\mathcal{G}]) \leq E[\psi(X)|\mathcal{G}]$ P-a.s.
(e) $E[E[X|\mathcal{G}]] = E[X]$.
(f) If X is $\mathcal{G}$-measurable, $E[X|\mathcal{G}] = X$ P-a.s.
(g) (Taking out what is known) If Y is $\mathcal{G}$-measurable and bounded, $E[XY|\mathcal{G}] = YE[X|\mathcal{G}]$ P-a.s.
(h) (Role of independence) If X is independent of $\mathcal{G}$, $E[X|\mathcal{G}] = E[X]$ P-a.s.
(i) (Tower property) If $\mathcal{G}'$ is a sub σ-algebra of $\mathcal{A}$ such that $\mathcal{G}' \subset \mathcal{G}$, then, $E[E[X|\mathcal{G}]|\mathcal{G}'] = E[X|\mathcal{G}']$ P-a.s.

The preceding properties are often sufficient to compute conditional expectations and we only come back to the definition for more difficult cases. In particular, the following result is useful for the study of financial models:

Proposition A.3 *If X is independent of $\mathcal{G}$ and if Y is $\mathcal{G}$-measurable, then, for all measurable function $\Phi : \mathbb{R}^2 \to \mathbb{R}$ such that $E[|\Phi(X,Y)|] < \infty$, we have*

$$E[\Phi(X,Y)|\mathcal{G}] = \psi(Y) \text{ where } \psi(y) = E[\Phi(X,y)].$$

Proof We have $\psi(y) = \int_{\mathbb{R}} \Phi(x,y)dP_X(x)$ and the measurability of ψ is a classical consequence of Fubini's theorem. For $G \in \mathcal{G}$, we set $Z = 1_G$. We deduce from the hypotheses that $P_{(X,Y,Z)} = P_X \otimes P_{(Y,Z)}$, thus,

$$E[\Phi(X,Y)1_G] = \int_{\mathbb{R}} \int_{\mathbb{R}^2} \Phi(x,y) z dP_{(Y,Z)}(y,z) dP_X(x).$$

By Fubini's theorem,

$$E[\Phi(X,Y)1_G] = \int_{\mathbb{R}^2} \psi(y) z dP_{(Y,Z)}(y,z)$$

so

$$E[\Phi(X,Y)1_G] = E[\psi(Y)1_G]$$

which completes the proof. □

The notion of filtration is used in mathematical finance to represent the evolution of financial information along time. It is presented below in its discrete time version:

Definition A.2

(a) A non decreasing family $(\mathcal{F}_t)_{t\in\{0,\ldots,T\}}$ of sub σ-algebras of $\mathcal{A}$ is called a filtration. This filtration is said to be complete when it contains negligible sets that is $\forall t \in \{0,\ldots,T\}, \mathcal{N} \subset \mathcal{F}_t$ where

$$\mathcal{N} = \{N \subset \Omega; \exists A \in \mathcal{A}, N \subset A, P(A) = 0\}.$$

(b) A family of random variables $(X_t)_{t\in\{0,1,\ldots,T\}}$ is adapted to the filtration $(\mathcal{F}_t)_{t\in\{0,\ldots,T\}}$ if $\forall t \in \{0,\ldots,T\}$, X_t is $\mathcal{F}_t$ measurable.
(c) A family of random variables $(X_t)_{t\in\{0,\ldots,T\}}$ is predictable with respect to the filtration $(\mathcal{F}_t)_{t\in\{0,1,\ldots,T\}}$ if $\forall t \in \{0,\ldots,T-1\}$, X_{t+1} is $\mathcal{F}_t$ measurable and if X_0 is $\mathcal{F}_0$ measurable.

We have seen in Sect. 3.2 that the absence of arbitrage opportunities in financial models is intrinsically linked to the notion of martingales that is defined below:

Definition A.3 A family of random variables $(M_t)_{t\in\{0,\dots,T\}}$ in $L^1(\Omega, \mathcal{A}, P)$ that is adapted to the filtration $(\mathcal{F}_t)_{t\in\{0,\dots,T\}}$ is called:

– a martingale if for $T \geq t \geq s \geq 0$, $E[M_t|\mathcal{F}_s] = M_s$.
– a supermartingale if for $T \geq t \geq s \geq 0$, $E[M_t|\mathcal{F}_s] \leq M_s$.
– a submartingale if for $T \geq t \geq s \geq 0$, $E[M_t|\mathcal{F}_s] \geq M_s$.

The following proposition is the key stone of the definition of the stochastic discount factor associated to an equivalent martingale measure (see Sect. 3.2.2):

Proposition A.4 *Let $(M_t)_{t\in\{0,\dots,T\}}$ be a non negative martingale such that $E[M_T] = 1$. If X is a random variable in $L^1(\Omega, \mathcal{A}, P)$ that is $\mathcal{F}_t$ measurable, then, for $t \geq s \geq 0$*

$$E_Q[X|\mathcal{F}_s] = \frac{1}{M_s} E[XM_t|\mathcal{F}_s]$$

where Q is the probability defined by the density function $\frac{dQ}{dP} = M_T$ with respect to P.

Proof First, when $s \leq T$, for an $\mathcal{F}_s$ measurable and bounded random variable Z we may deduce from the martingale property of (M_t) that

$$E_Q[Z] = E[ZM_T] = E[M_s Z].$$

Moreover, for $t \geq s \geq 0$, we have from Proposition A.2

$$E_Q[XZ] = E[M_T XZ] = E[E[XM_T|\mathcal{F}_s]Z] = E[E[XE[M_T|\mathcal{F}_t]|\mathcal{F}_s]Z]$$

thus

$$E_Q[XZ] = E[E[M_t X|\mathcal{F}_s]Z] = E\left[\frac{M_s}{M_s} E[XM_t|\mathcal{F}_s]Z\right] = E_Q\left[\frac{1}{M_s} E[XM_t|\mathcal{F}_s]Z\right].$$

Hence, from Definition A.1,

$$E_Q[X|\mathcal{F}_s] = \frac{1}{M_s} E[XM_t|\mathcal{F}_s].$$

□

Monte Carlo Methods

In this section, $\underset{a.s}{\rightarrow}$, $\underset{L^1}{\rightarrow}$ and $\underset{\mathcal{D}}{\rightarrow}$ denote the almost sure convergence, the convergence in $L^1(\Omega, \mathcal{A}, P)$ and the convergence in distribution for sequences of random variables. Theoretical foundations of Monte Carlo Methods are mainly based on two fundamental asymptotic results: The Strong Law of Large Numbers and the Central Limit Theorem. The Strong Law of Large Numbers ensures that, under integrability conditions, the mean of a sequence of i.i.d random variables is an approximation of the expectation:

Theorem A.1 *Let (X_n) be a sequence of i.i.d random variables.*

(a) Suppose that $X \in L^1(\Omega, \mathcal{A}, P)$. Denoting $S_n = X_1 + \ldots + X_n$, we have

$$\frac{S_n}{n} \underset{a.s \text{ and } L^1}{\rightarrow} E[X_1].$$

(b) If $E[|X_1|] = +\infty$, the sequence S_n diverges almost surely.

The Central Limit Theorem gives some precisions concerning the speed of convergence in the Strong Law of Large Numbers:

Theorem A.2 *Let (X_n) be a sequence of i.i.d random variables. Suppose that $X_1 \in L^2(\Omega, \mathcal{A}, P)$, then,*

$$\frac{S_{n-nm}}{\sqrt{n}\sigma} \underset{\mathcal{D}}{\rightarrow} \mathcal{N}(0,1)$$

where $m = E[X_1]$ and $\sigma^2 = Var[X_1]$.

The preceding theorem is classically used to build asymptotic confidence intervals. In fact, we obtain $\forall a \in \mathbb{R}_+$,

$$P\left(\frac{-a\sigma}{\sqrt{n}} \leq \frac{S_n}{n} - E[X_1] \leq \frac{a\sigma}{\sqrt{n}}\right) \underset{n\to\infty}{\rightarrow} \int_{-a}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

In practice we know from tables that

$$P(|\mathcal{N}(0,1)| \leq 1.96) = 0.95$$

thus when n is large enough, with a confidence of 95 %,

$$E[X_1] \in \left[\frac{S_n}{n} - \frac{1.96\sigma}{\sqrt{n}}, \frac{S_n}{n} + \frac{1.96\sigma}{\sqrt{n}}\right].$$

The magnitude of the error is given by $\frac{3.92\sigma}{\sqrt{n}}$: the magnitude of n and σ is fundamental to control the length of the confidence interval. When σ is unknown, the so-called empirical variance gives us an unbiased and consistent estimator:

Proposition A.5 *Let $(X_n)_{n\in\mathbb{N}}$ be i.i.d random variables in $L^2(\Omega, \mathcal{A}, P)$. Then, if we define*

$$\hat{\sigma_n}^2 = \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\frac{1}{n}\sum_{i=1}^{n} X_i)^2\right)$$

we have $E[\hat{\sigma_n}^2] = \sigma^2$ and $\hat{\sigma_n}^2 \underset{a.s}{\to} \sigma^2$.

According to the following result, that is a direct consequence of the Slutsky's lemma,[2] the preceding estimator is used in practice to deduce confidence intervals from observations:

Proposition A.6 *Let $(X_n)_{n\in\mathbb{N}}$ be i.i.d random variables in $L^2(\Omega, \mathcal{A}, P)$, then,*

$$\frac{S_n - nE[X_1]}{\sqrt{n}\hat{\sigma_n}} \underset{\mathcal{D}}{\to} \mathcal{N}(0, 1).$$

Basically, to approximate the quantity $E[f(X)]$ by Monte Carlo Methods we have to

- Generate a n-sample of the distribution of X,
- Compute $\frac{1}{n}\sum_{k=1}^{n} f(X_k)$ for large n,
- Compute the confidence interval $\left[\frac{S_n}{n} - \frac{1.96\hat{\sigma_n}}{\sqrt{n}}, \frac{S_n}{n} + \frac{1.96\hat{\sigma_n}}{\sqrt{n}}\right]$ coming from the Central Limit Theorem.

This method is easy to implement on any software once we are able to generate samples of particular distributions. Moreover, only integrability conditions are required for f. Nevertheless, it is important to keep in mind that the precision of the method (measured by the size of the confidence interval) is a random variable depending on the magnitude of σ. From Proposition A.1, we have seen how to generate independent Gaussian random variables from uniform ones (that may

[2] If (X_n) and (Y_n) are two sequences of random variables such that $X_n \underset{\mathcal{D}}{\to} X$ and $Y_n \underset{\text{a.s}}{\to} c$ where c is a constant, then, $X_n Y_n \underset{\mathcal{D}}{\to} Xc$.

be obtained using R with the *runif* command). We generalize this idea to the distributions involved in the empirical Chap. 4 namely the Poisson, the Mixture of two Gaussians and the Generalized Hyperbolic distributions:

Poisson distribution: A Poisson random variable X of parameter $\lambda \in \mathbb{R}_+^*$ is a random variable with values in $\mathbb{N}$ such that $\forall k \in \mathbb{N}$,

$$P(X = k) = p_k = \frac{e^{-\lambda}\lambda^k}{k!}.$$

As all discrete random variables, the Poisson distribution may be generated by the inversion method that is based on the following result:

Proposition A.7 *Let X be a random variable and F_X its distribution function. If U follows a uniform distribution on* $[0, 1]$*, then, $F_X^-(U)$ and X have the same distribution where F_X^- is the generalized inverse of F_X given $\forall u \in]0, 1[$ by*

$$F_X^-(u) = inf\{x \mid F_X(x) \geq u\}.$$

The preceding proposition is based on the elementary relation $F_X^-(u) \leq x \Leftrightarrow u \leq F_X(x)$ and in the case of the Poisson distribution

$$F_X^-(u) = \sum_{k=1}^{\infty} k 1_{\left\{\sum_{j=0}^{k-1} p_j < u \leq \sum_{j=0}^{k} p_j\right\}}.$$

For $\lambda \leq 10$ this method is implemented in R via the command *rpois* of the *stats* package. For $\lambda > 10$, this command uses the more complex and efficient methodology proposed in Ahrens and Dieter (1982).

Mixture of two Gaussian distributions: Generate mixture of two Gaussian distributions is an easy task remarking that

$$\mathbf{1}_{U \leq \phi} G_1 + \mathbf{1}_{U > \phi} G_2 \hookrightarrow MN(\phi, \mu_1, \mu_2, \sigma_1, \sigma_2)$$

when U, G_1 and G_2 are independent random variables such that U follows a uniform distribution on [0, 1], $G_1 \hookrightarrow \mathcal{N}(\mu_1, \sigma_1^2)$ and $G_2 \hookrightarrow \mathcal{N}(\mu_2, \sigma_2^2)$.

Generalized Hyperbolic distributions: In Barndorff-Nielsen (1977) we find the normal variance-mean mixture representation of the $GH(\lambda, \alpha, \beta, \delta, \mu)$ distribution: We say that W follows an Generalized Inverse Gaussian distribution of parameter $(\lambda, \chi, \psi) \in \mathbb{R} \times (\mathbb{R}_+^*)^2$ (and we use the notation $W \hookrightarrow GIG(\lambda, \chi, \psi)$) if its density function is given $\forall x \in \mathbb{R}_+^*$ by

$$\frac{\left(\frac{\psi}{\chi}\right)^{\frac{\lambda}{2}}}{2K_\lambda(\sqrt{\psi\chi})} x^{\lambda-1} e^{-\frac{1}{2}\left(\chi x^{-1} + \psi x\right)}$$

where K_λ is the Bessel function of the third kind. If W and Z are two independent random variables such that $Z \hookrightarrow \mathcal{N}(0,1)$ and $W \hookrightarrow GIG(\lambda, \delta^2, \alpha^2 - \delta^2)$, then,

$$X = \mu + W\beta + \sqrt{W}Z \hookrightarrow GH(\lambda, \alpha, \beta, \delta, \mu).$$

Thus, we deduce from the preceding relation that GH distributions may easily be generated from *GIG* ones. In Dagpunar (1989), such an algorithm, based on the rejection method, is proposed for the *GIG* distribution and its numerical performances are discussed in Hörmann and Leydold (2014) where the authors prove its efficiency for $|\lambda| > 1$ and $\sqrt{\delta^2(\alpha^2 - \delta^2)} > 0.5$ (these conditions are compatible with the estimated parameters provided in Table 4.7). This method is implemented in R via the *rgh* command of the *fBasics* package.

Convergence of Discrete Time Markov Processes to Diffusions

We present here the basic techniques developed in Stroock and Varadhan (1979) that are classically applied to the study of the convergence of stochastic difference equations to diffusions.

For $T \in \mathbb{R}_+^*$ and $n \in \mathbb{N}^*$, we consider a Markov chain[3] indexed by the time unit $\tau = \frac{1}{n}$, $Z^{(n)} = (Y_{k\tau}^{(n)}, h_{k\tau}^{(n)})_{k \in \{0,\ldots,nT\}}$, with values in $\mathbb{R}^2$ and starting from $(y_0, h_0) \in \mathbb{R}^2$. Then, $Z^{(n)}$ is embedded into a continuous time process $(Z_t^{(n)})_{t \in [0,T]}$ by defining

$$Z_t^{(n)} = Z_{k\tau}^{(n)} \text{ if } k\tau \le t < (k+1)\tau. \tag{A.1}$$

The sample paths of the latter process are by construction right continuous with left limit (cadlag). The next theorem gives general conditions to ensure the weak convergence[4] of $(Z_t^{(n)})_{t \in [0,T]}$ toward a bivariate diffusion.

Theorem A.3 *Let $p \in \mathbb{N}^*$ and μ (resp. Σ) be a continuous function from $\mathbb{R}^2$ into $\mathbb{R}^2$ (resp. into the set of the real matrix of size $2 \times p$). Suppose that for all $(r_1, r_2) \in (\mathbb{R}_+^*)^2$ and for some $s > 0$,*

[3]Let $(\Omega, \mathcal{A}, P)$ be a probability space equipped with a filtration $(\mathcal{F}_t)_{t \in \mathcal{T}}$, for some (totally ordered) index set $\mathcal{T}$ and let $(S, \mathcal{S})$ be a measurable space. An S-valued and adapted stochastic process $(X_t)_{t \in \mathcal{T}}$ is said to possess the Markov property with respect to the filtration $(\mathcal{F}_t)_{t \in \mathcal{T}}$ if, for each $A \in \mathcal{S}$ and each $s, t \in \mathcal{T}$ with $s < t$, $P(X_t \in A | \mathcal{F}_s) = P(X_t \in A | X_s)$. A Markov process is a stochastic process which satisfies the Markov property with respect to its natural filtration $(\mathcal{F}_t = \sigma(X_u; u \le t))$. When $\mathcal{T}$ is a discrete space, we say that $(X_t)_{t \in \mathcal{T}}$ is a Markov chain.

[4]Here, by weak convergence we mean weak convergence in the Skorokhod space of cadlag functions with values in $\mathbb{R}^2$ endowed with the Skorokhod topology (see e.g. Jacod and Shiryaev 2003, Section 6).

$$\lim_{\tau\to 0} \sup_{|y|\le r_1,|h|\le r_2} \left| \frac{1}{\tau} E\left[Z^{(n)}_{(k+1)\tau} - Z^{(n)}_{k\tau} | Z^{(n)}_{k\tau} = (y,h) \right] - \mu(y,h) \right| = 0, \tag{A.2}$$

$$\lim_{\tau\to 0} \sup_{|y|\le r_1,|h|\le r_2} \left| \frac{1}{\tau} Var\left[Z^{(n)}_{(k+1)\tau} - Z^{(n)}_{k\tau} | Z^{(n)}_{k\tau} = (y,h) \right] - \Sigma(y,h)\Sigma^t(y,h) \right| = 0, \tag{A.3}$$

$$\limsup_{\tau\to 0} \sup_{|y|\le r_1,|h|\le r_2} \tau^{\frac{-(2+s)}{2}} E\left[\left\| Z^{(n)}_{(k+1)\tau} - Z^{(n)}_{k\tau} \right\|^{2+s} | Z^{(n)}_{k\tau} = (y,h) \right] < \infty. \tag{A.4}$$

Then, if the stochastic differential equation

$$dZ_t = \mu(Z_t) + \Sigma(Z_t)dW_t, Z_0 = (y_0, h_0) \tag{A.5}$$

(where W is a p-dimensional standard Brownian motion) admits a unique weak solution[5] *on* $[0,T]$*, then the process* $(Z_t^{(n)})_{t\in[0,T]}$ *weakly converges toward* $(Z_t)_{t\in[0,T]}$ *when τ goes to zero.*

From Moment Generating Functions to Option Prices

Here, using the notations of Sect. 3.7, we briefly remind how to obtain, up to numerical integration, option prices from the moment generating function of the logarithm of the risky asset.

Let us denote by $d_{t,T}$ the density function, under an EMM $\mathbb{Q}$, of $log(S_T)$ given $\mathcal{F}_t$. Thus, the arbitrage-free price, at time t, of a European call option with strike K and maturity T is given by

$$e^{-r(T-t)} E_{\mathbb{Q}}[(S_T - K)_+ \mid \mathcal{F}_t] = e^{-r(T-t)} \int_{log(K)}^{+\infty} (e^{\phi} - K) d_{t,T}(\phi) d\phi.$$

Exploiting the fact that $d_{t,T}$ and

$$d^*_{t,T} = \frac{e^x d_{t,T}}{\mathbb{G}^{\mathbb{Q}}_{log(S_T)|\mathcal{F}_t}(1)} = \frac{e^x d_{t,T}}{e^{r(T-t)} S_t}$$

are densities of probability, we use twice the classical inversion formula (see Stuart and Ord 1994, Chapter 4)

$$\int_{log(K)}^{+\infty} d(\phi)d\phi = \frac{1}{2} + \frac{1}{\pi}\int_0^{+\infty} \mathcal{R}e\left[\frac{K^{-i\phi}\mathbb{G}(i\phi)}{i\phi}\right] d\phi \tag{A.6}$$

[5]See for example Nelson (1990) for classical conditions.

(where d is a density function and $\mathbb{G}$ the associated moment generating function) to obtain

$$
\begin{aligned}
e^{-r(T-t)}E_{\mathbb{Q}}[(S_T-K)_+ \mid \mathcal{F}_t] = \frac{S_t}{2} + \frac{e^{-r(T-t)}}{\pi} \\
\int_0^{+\infty} \mathcal{R}e\left[\frac{K^{-i\phi}\mathbb{G}^{\mathbb{Q}}_{log(S_T)|\mathcal{F}_t}(i\phi+1)}{i\phi}\right]d\phi \\
-Ke^{-r(T-t)}(\frac{1}{2}+\frac{1}{\pi} \\
\int_0^{+\infty} \mathcal{R}e\left[\frac{K^{-i\phi}\mathbb{G}^{\mathbb{Q}}_{log(S_T)|\mathcal{F}_t}(i\phi)}{i\phi}\right]d\phi).
\end{aligned}
$$

References

Ahrens JH, Dieter U (1982) Computer generation of poisson deviates from modified normal distributions. ACM Trans Math Softw 8:163–179

Barndorff-Nielsen OE (1977) Exponentially decreasing distributions for the logarithm of particle size. Proc R Soc Lond Ser A 353:401–419

Cody WJ (1969) Rational chebyshev approximations for the error function. Math Comput 23:631–637

Dagpunar JS (1989) An easily implemented generalized inverse gaussian generator. Commun Stat B Simul Comput 18:703–710

Hörmann W, Leydold J (2014) Generating generalized inverse gaussian random variates. Stat Comput 24(4):547–557

Jacod J, Protter P (2004) Probability essentials. Springer, Berlin

Jacod J, Shiryaev AN (2003) Limit theorems for stochastic processes, 2nd edn. Springer, Berlin

Lapeyre B, Sulem A, Talay D (2013) Understanding numerical analysis for financial models. Cambridge University Press, Cambridge

Nelson DB (1990) ARCH models as diffusion approximations. J Econ 45:7–38

Robert CP, Casella G (2010) Introducing monte carlo methods with R. Springer, New York

Stroock DW, Varadhan SRS (1979) Multidimensional diffusion processes. Springer, New York

Stuart A, Ord JK (1994) Kendall's advanced theory of statistics. Vol. 1: Distribution theory, 6th edn. Wiley-Blackwell, New York

Williams D (1991) Probability with martingales. Cambridge University Press, Cambridge

Index

C. Chorro et al., *A Time Series Approach to Option Pricing*,
DOI 10.1007/978-3-662-45037-6