MFE 230D

Derivatives: Quantitative Methods

Exercises

Alberto Gennaro

https://github.com/GennaroAlberto/mfe230d-derivatives

Built October 31, 2026

Contents

I Analytical pricing 3

1 The standard model, revisited 4

2 The Black–Scholes formula three ways 8

3 Arbitrage in the strike and the smile 11

4 Static replication and the Kolmogorov equations 14

II Monte Carlo methods 18

5 Random numbers 19

6 Monte Carlo pricing and its error 22

7 Simulating paths 26

8 Antithetic variates, moment matching and control variates 30

9 Importance sampling 33

10 Stratified sampling and conditional Monte Carlo 36

11 Quasi-Monte Carlo 40

12 Multilevel Monte Carlo 44

13 Monte Carlo Greeks 48

14 American options by simulation 51

III Path-dependent and multi-asset payoffs **54**

15 Asian options **55**

16 Barriers, lookbacks and the Brownian bridge **58**

17 FX and multi-asset options **62**

IV Beyond Black–Scholes **65**

18 Hedging in practice **66**

19 Stochastic volatility **69**

20 Trees and finite differences **72**

Part I

Analytical pricing

Chapter 1

The standard model, revisited

Exercise 1.1 *(A trinomial Brownian step)* [⋆] A tree replaces the Brownian increment over a step of length h by a variable ζ that takes the values $\beta\sqrt{h}$, 0 and $-\beta\sqrt{h}$ with probabilities p_u, p_0 and p_d, where $\beta > 0$.

(a) Show that $\mathbb{E}[\zeta] = 0$ and $\mathbb{E}[\zeta^2] = h$ hold if and only if $p_u = p_d = 1/(2\beta^2)$ and $p_0 = 1 - 1/\beta^2$. For which β is this a probability distribution, and which β gives back the binomial step?

(b) Find the β for which, in addition, $\mathbb{E}[\zeta^4]$ equals the fourth moment $3h^2$ of $\mathcal{N}(0, h)$, and the corresponding probabilities.

(c) With these values, compare $\mathbb{E}[\zeta^6]$ with the sixth moment of $\mathcal{N}(0, h)$.

(d) Let f have bounded derivatives up to order six. Expand $\mathbb{E}[f(x + \zeta)] - \mathbb{E}[f(x + \sqrt{h}Z)]$ in powers of h, and explain why the choice of (b) makes one step of the tree more accurate for smooth payoffs than any other β.

Exercise 1.2 *(Probabilities under the two measures)* [⋆] In the standard model, let $K > 0$ and $\theta = (\mu - r)/\sigma$.

(a) Show that $\mathbb{Q}(S_T > K) = N(d_-)$ and $\mathbb{P}(S_T > K) = N(d_- + \theta\sqrt{T})$, with $d_- = \big(\log(S_0/K) + \nu T\big)/(\sigma\sqrt{T})$. Which of the two enters the price of the digital call paying $\mathbf{1}_{\{S_T > K\}}$, and why?

(b) Show that the density $Z_T = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}$ on $\mathcal{F}_T$ is a function of S_T alone, $Z_T = c\,(S_T/S_0)^{-\theta/\sigma}$, and find the constant c. For $\mu > r$, which terminal prices does $\mathbb{Q}$ weight more heavily than $\mathbb{P}$?

(c) Take the test case $S_0 = K = 100$, $T = 1$, $r = 1\%$, $y = 2\%$, $\sigma = 20\%$, with $\mu = 6\%$. Compute both probabilities and the price of the digital call. A colleague prices it as $e^{-\mu T}\mathbb{P}(S_T > K)$; how far off is she?

(d) Define the rate $\rho(K)$ by $e^{-\rho(K)T}\mathbb{P}(S_T > K) =$ the price of the digital. Compute $\rho(80)$, $\rho(100)$ and $\rho(120)$, and find the limits of $\rho(K)$ as $K \downarrow 0$ and as $K \to \infty$ (for $\mu > r$). Can a single risk-adjusted discount rate price all digitals?

Exercise 1.3 *(The dividend stream)* [⋆] In the standard model, the holder of one share receives the dividends $yS_s\,\mathrm{d}s$. Let $u(t,x)$ be the value at time t, when $S_t = x$, of the dividends paid between t and T: $u(t,x) = \mathbb{E}^{\mathbb{Q}}\big[\int_t^T e^{-r(s-t)} yS_s\,\mathrm{d}s \,\big|\, S_t = x\big]$.

(a) Write the equation Equation (1.10) and the terminal condition that Theorem 1.6 associates with this expectation (identify b, σ, r, c and g).

(b) Look for a solution of the form $u(t,x) = a(T-t)\,x$, and find a.

(c) Check hypothesis (b) of Theorem 1.6, and conclude that $u(t,x) = x\big(1 - e^{-y(T-t)}\big)$.

(d) Recover the same value directly, by exchanging the expectation and the time integral. Justify the exchange.

(e) Interpret $x = u(t,x) + xe^{-y(T-t)}$ with Theorem 1.4. What happens when $y < 0$?

Exercise 1.4 *(A Gaussian payoff)* [⋆⋆] In the standard model, let $X_t = \log(S_t/K)$ for a constant $K > 0$, and consider the claim that pays $f_\alpha(X_T) = (2\pi\alpha)^{-1/2} e^{-X_T^2/(2\alpha)}$ at T, for a constant $\alpha > 0$: a bet on S_T ending close to K.

(a) Find the dynamics of X under $\mathbb{Q}$, and the equation that the price $F(t,x)$, written as a function of $x = \log(S_t/K)$, must satisfy.

(b) Compute $F(t,x) = e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[f_\alpha(X_T) \mid X_t = x]$, $\tau = T - t$, in closed form, using that a sum of independent normal variables is normal. Check directly that it solves the equation of (a), and that $F(t,x) \to f_\alpha(x)$ as $t \uparrow T$.

(c) Show that $f_\alpha(X_T) \to \delta(X_T)$ as $\alpha \downarrow 0$, in the sense of Section 1.4, and compute $F_0(t,x) = \lim_{\alpha\downarrow 0} F(t,x)$. Show that $F_0(t, \log(S/K)) = K\,G(t,S;T,K)$, and explain the factor K by a change of variables in the integral $\int \delta(\log(\xi/K))\,\varphi(\xi)\,\mathrm{d}\xi$.

(d) Compare F with F_0, using the heat kernel Equation (1.1). In what sense is the Gaussian payoff "an Arrow–Debreu security that has already been diffusing for a while"?

Exercise 1.5 *(Moments from the forward equation)* [⋆⋆] In the standard model, fix t and $x > 0$, let $q \in \mathbb{R}$, and for $\tau > 0$ let $m_q(\tau) = \int_0^\infty \xi^q\, G(t,x;t+\tau,\xi)\,\mathrm{d}\xi$, the price at t of the claim paying $S_{t+\tau}^q$ at $t+\tau$.

(a) Using the forward equation Equation (1.19) and two integrations by parts, show that $m_q'(\tau) = \lambda_q\, m_q(\tau)$ with $\lambda_q = q(r-y) + \frac{1}{2}q(q-1)\sigma^2 - r$. Justify the differentiation under the integral sign and the vanishing of the boundary terms. (*Hint:* in $\eta = \log\xi$, G and its derivatives are a polynomial times a Gaussian.)

(b) Use Theorem 1.12(iii) to find $m_q(0+)$, solve the equation, and compare with Equation (1.5).

(c) For which q is $m_q(\tau) = x^q$ for every τ? Find these q for the test case ($r = 1\%$, $y = 2\%$, $\sigma = 20\%$) and for $y = 0$. What does the answer say about the process $e^{-rt}S_t^q$ under $\mathbb{Q}$?

Exercise 1.6 *(Butterflies and the Arrow–Debreu price)* [⋆⋆] For $\varepsilon > 0$ let $\Lambda_\varepsilon(x) = c_\varepsilon(\varepsilon - |x|)^+$, a tent of width 2ε.

(a) Find c_ε such that $\int \Lambda_\varepsilon = 1$, and show that $\Lambda_\varepsilon(x-K) = c_\varepsilon\big[(x-K+\varepsilon)^+ - 2(x-K)^+ + (x-K-\varepsilon)^+\big]$: the payoff of a butterfly. Show that $\Lambda_\varepsilon(\cdot - K) \to \delta_K$ as $\varepsilon \downarrow 0$.

(b) Let $C(K)$ be the price at time 0 of the call with strike K in the standard model, and $B_\varepsilon(K) = c_\varepsilon\big[C(K-\varepsilon) - 2C(K) + C(K+\varepsilon)\big]$ the price of the butterfly. Show that $B_\varepsilon(K) \to G(0,S_0;T,K)$ and that $B_\varepsilon(K) - G(0,S_0;T,K) = \frac{\varepsilon^2}{12}\,\partial_{KK}G(0,S_0;T,K) + O(\varepsilon^4)$. You may use Theorem 1.14.

(c) *Computational.* For the test case and $K = 100$, compute $B_\varepsilon(100)$ in double precision with the Black–Scholes call of `mfe230d.bs`, for $\varepsilon = 10, 1, 10^{-1}, 10^{-2}, \dots, 10^{-7}$, and compare with $G(0,100;1,100)$. Explain the errors you observe for large and for small ε, and estimate the best ε.

Exercise 1.7 *(A dipole)* [⋆⋆⋆] Let k be the heat kernel Equation (1.1) and, for $0 \le t < T$ and $x \in \mathbb{R}$,

$$u(t,x) = -\partial_x k(T-t,x) = \frac{x}{\sqrt{2\pi}\,(T-t)^{3/2}}\, e^{-x^2/(2(T-t))}.$$

(a) Show that u solves $\partial_t u + \frac12 \partial_{xx} u = 0$ on $[0,T)\times\mathbb{R}$ and that $u(t,x) \to 0$ as $(t,x) \to (T,x_0)$ for every $x_0 \neq 0$. Does u have a limit at $(T,0)$?

(b) Let W be a Brownian motion, $x_0 \neq 0$ and $X = x_0 + W$; take $b = 0$, $\sigma = 1$, $r = c = 0$, $g = 0$ and $E = \{0\}$ in Theorem 1.6. Show that $u(t,X_t)$ is a martingale on $[0,T']$ for every $T' < T$ and that $u(t,X_t) \to 0$ almost surely as $t \uparrow T$. Deduce that the Itô integral $\int_0^t u_x(s,X_s)\,\mathrm{d}W_s$ is defined on $[0,T]$ but is not a martingale there, so hypothesis (a) of the theorem fails. Which property of the family $(u(t,X_t))_{t<T}$ is missing, and on which paths is the expectation carried just before T?

(c) Show that $\int_{\mathbb{R}} u(t,x)\,\varphi(x)\,\mathrm{d}x \to \varphi'(0)$ as $t \uparrow T$, for every $\varphi \in C^1(\mathbb{R})$ with φ and φ' bounded. So the terminal value of u is the distribution $-\delta_0'$, and not 0.

(d) With $r = 0$ and W as the underlying, the price of the Arrow–Debreu security paying $\delta_a(x_0 + W_T)$ is $k(T, a - x_0)$. Show that $u(0,x_0)$ is the limit as $\varepsilon \downarrow 0$ of the price of a portfolio of $1/(2\varepsilon)$ Arrow–Debreu securities at ε and $-1/(2\varepsilon)$ at $-\varepsilon$, and relate this portfolio to (c).

Exercise 1.8 *(Discounting at a quadratic rate)* [⋆⋆⋆] Let W be a Brownian motion, $\lambda > 0$, and $u(t,x) = \mathbb{E}\big[\exp\big(-\frac{\lambda^2}{2}\int_t^T (x + W_s - W_t)^2\,\mathrm{d}s\big)\big]$: the Feynman–Kac expectation with $X = x + W$, discount rate $r(t,x) = \frac12\lambda^2 x^2$, $c = 0$ and $g = 1$.

(a) Look for a solution of the equation of Theorem 1.6 of the form $u(t,x) = \exp\big(a(T-t) + b(T-t)x^2\big)$. Show that $a' = b$ and $b' = 2b^2 - \frac12\lambda^2$ with $a(0) = b(0) = 0$, and solve: $b(\tau) = -\frac{\lambda}{2}\tanh(\lambda\tau)$, $a(\tau) = -\frac12 \log\cosh(\lambda\tau)$.

(b) Check the hypotheses of Theorem 1.6, and deduce that $\mathbb{E}\big[\exp\big(-\frac{\lambda^2}{2}\int_0^T W_s^2\,\mathrm{d}s\big)\big] = (\cosh \lambda T)^{-1/2}$.

(c) What goes wrong in parts (a) and (b) if $\frac12\lambda^2 x^2$ is replaced by $-\frac12\lambda^2 x^2$?

(d) *Computational.* For $\lambda = 2$, $T = 1$ and $x = 0$, estimate the expectation by simulating W on a uniform grid of m steps (`mfe230d.paths`) and approximating the integral by the trapezoid rule, for $m = 4$, 16 and 64 with 10^5 paths each. Report the estimates with

their standard errors. Is the discretisation bias visible? How does it scale with the step $h = T/m$? (*Hint:* the discretised expectation can be computed exactly from the covariance matrix of $(W_{t_1}, \ldots, W_{t_m})$, as a determinant.)

Chapter 2

The Black–Scholes formula three ways

Throughout, the model is the standard model of the notes, with spot S, strike K, maturity T, rate r, yield y and volatility σ; $d_\pm$, $m = \log(F/K)$, $w = \sigma\sqrt{T}$ and the normalised price $c(m, w)$ are as in Equation (2.1) and Equation (2.16). The test case is $S = K = 100$, $T = 1$, $r = 1\%$, $y = 2\%$, $\sigma = 20\%$; the course's index parameters are $S = K = 3900$, $T = 1$, $r = 3\%$, $y = 1.3\%$, $\sigma = 25\%$. You may use the results of the notes, quoting them, and the closed forms of `bs.py`.

Exercise 2.1 *(Gap options)* [⋆] A *gap call* with trigger $K_2 > 0$ and strike $K_1 \geq 0$ pays $(S_T - K_1)\mathbf{1}_{\{S_T > K_2\}}$.

(a) Price it at time 0 by direct integration, with $d_\pm$ computed at the strike K_2.

(b) In the test case, with $K_2 = 100$, compute the price for $K_1 = 90, 100, 110, 120$. Explain why a gap call can have a negative price.

(c) Show that its vega is $e^{-rT}N'(d_-)(K_2\sqrt{T} + (K_1 - K_2)d_+/\sigma)$, and find where it is negative. For $K_1 = 120$, $K_2 = 100$ and the other test-case inputs, show that the price is not monotone in σ, and find two volatilities that give the price -1. Which part of Theorem 2.9 fails for gap options?

Exercise 2.2 *(Strike Greeks and homogeneity)* [⋆] Write $C = C(S, K)$ for the call price at time 0, $\Delta_S = \partial_S C$, $\Delta_K = \partial_K C$, $\Gamma_{SS} = \partial_{SS} C$, $\Gamma_{SK} = \partial_{SK} C$ and $\Gamma_{KK} = \partial_{KK} C$.

(a) Compute the cross gamma Γ_{SK} in closed form ($\Gamma_{KK} = \partial_{KK} C$ is in the notes).

(b) Show that $S\Gamma_{SS} + K\Gamma_{SK} = 0$ and $S\Gamma_{SK} + K\Gamma_{KK} = 0$, and deduce $S^2\Gamma_{SS} = K^2\Gamma_{KK}$. Interpret them: which functions of (S, K) are homogeneous of degree zero?

(c) Three of the ratios $\rho_y/(S\Delta_S)$, $\rho_r/(K\Delta_K)$, $\mathcal{V}/(\sigma S^2\Gamma_{SS})$, $\mathcal{V}/(\sigma SK\Gamma_{SK})$ and $\mathcal{V}/(\sigma K^2\Gamma_{KK})$ follow from Theorem 2.7; derive the other two.

(d) State the units of Δ_S, Δ_K, Γ_{SS}, Γ_{SK}, Γ_{KK}, ρ_r, ρ_y and $\mathcal{V}$ when prices are in dollars and time in years, and compute them, and the five ratios, for the one-year at-the-money call on the index.

Exercise 2.3 *(At-the-money rules of thumb)* [⋆⋆] Let $D = e^{-rT}$.

(a) For the call struck at the forward, $K = F$, whose price is given in Theorem 2.2(a), show that $DFw/\sqrt{2\pi}$ overestimates it with relative error $w^2/24 + O(w^4)$. Find, numerically, the total volatilities w at which the relative error equals 10% and 1%, and compare with the leading term.

(b) Show that $\partial_m c(m, w) = e^m N(d_+)$ and deduce the approximation $C \approx D\big(Kw/\sqrt{2\pi} + (F - K)/2\big)$ for strikes near the forward. What is the order of its error in m and w?

(c) For the one-year at-the-money (spot-strike) call on the index, compare the exact price with the rule $0.4\,\sigma\sqrt{T}\,K$, with $DFw/\sqrt{2\pi}$, and with the approximation of (b). Which error dominates the rule of thumb here?

Exercise 2.4 *(Put–call parity by residues)* [⋆⋆] Let $u_0^{\mathrm{c}}(\xi) = (e^\xi - 1)^+$ and $u_0^{\mathrm{p}}(\xi) = (1 - e^\xi)^+$, and let u^{c}, u^{p} be the corresponding heat solutions $u(s, \xi) = \mathbb{E}[u_0(\xi + \sqrt{s}Z)]$.

(a) Show that for $\operatorname{Im} z > 0$ the generalised transform $\int_{\mathbb{R}} e^{-iz\xi} u_0^{\mathrm{p}}(\xi)\,\mathrm{d}\xi$ exists and equals $-1/(z(z + i))$: the same function as for the call (Equation (2.8)), on another half-plane.

(b) Write $u^{\mathrm{p}}(s, \xi)$, for $s > 0$, as an integral along a line $\operatorname{Im} z = b > 0$, as in the proof of Theorem 2.5. Then show, with the residue theorem, that $u^{\mathrm{c}}(s, \xi) - u^{\mathrm{p}}(s, \xi) = e^{\xi + s/2} - 1$; justify the passage to the limit on the vertical sides of the rectangle. Translate the result into a statement about prices.

(c) Which payoff has the transform $-1/(z(z + i))$ on the strip $-1 < \operatorname{Im} z < 0$? Write the corresponding representation of the call price along $\operatorname{Im} z = -\tfrac{1}{2}$.

Exercise 2.5 *(Inverting the Laplace transform on the branch cut)* [⋆⋆⋆] Let $\xi \in \mathbb{R}$ and $s > 0$, and let $\bar g(\lambda) = e^{-q|\xi|}/q$ with $q = \sqrt{2\lambda}$ (principal branch), the Laplace transform of the heat kernel (Theorem 2.6). Show, without using Equation (2.12), that the Bromwich integral

$$\lim_{R \to \infty} \frac{1}{2\pi i} \int_{c - iR}^{c + iR} e^{\lambda s}\, \bar g(\lambda)\,\mathrm{d}\lambda, \qquad c > 0,$$

equals the heat kernel $k(s, \xi) = (2\pi s)^{-1/2} e^{-\xi^2/(2s)}$, as follows.

(a) Close the line with two arcs of radius R in the left half-plane and a keyhole around the cut $(-\infty, 0]$, with a small circle of radius ε around 0. Show that the integrals over the arcs and over the small circle tend to 0 as $R \to \infty$ and $\varepsilon \to 0$. *Hint:* $|\bar g(\lambda)| \le |2\lambda|^{-1/2}$ when $\operatorname{Re} q \ge 0$; for the arcs, bound $\int_{\pi/2}^{\pi} e^{sR\cos\theta}\,\mathrm{d}\theta$.

(b) Show that the two sides of the cut contribute $\frac{1}{\pi}\int_0^\infty e^{-\rho s} \cos\big(\sqrt{2\rho}\,|\xi|\big)(2\rho)^{-1/2}\,\mathrm{d}\rho$.

(c) Evaluate this integral. *Hint:* substitute $\rho = v^2/2$ and use the characteristic function of a normal variable.

Exercise 2.6 *(Equity as a call on the firm)* [⋆⋆] In Merton's model of a firm, the value V of the firm's assets follows a geometric Brownian motion with volatility σ_V and no payouts, and the firm has issued one zero-coupon bond with face value N_d and maturity T. At T the bondholders

receive $\min(V_T, N_d)$ and the shareholders $(V_T - N_d)^+$. Assume that V is the price of a traded asset, so that the notes apply with $S = V$ and $y = 0$.

(a) Give the time-0 values E_0 of the equity and B_0 of the debt, the risk-neutral probability of default, and the credit spread $-\frac{1}{T}\log(B_0/N_d) - r$.

(b) Show that the equity value $E_t = e(t, V_t)$ has instantaneous volatility $\sigma_E = \sigma_V \, V_t \, \partial_V e / E_t$, and that $\sigma_E > \sigma_V$. Interpret.

(c) Compute (a) and σ_E for $V_0 = 100$, $N_d = 80$, $T = 1$, $r = 3\%$, $\sigma_V = 25\%$.

(d) Only E_0 and σ_E are observed. With $E_0 = 30$, $\sigma_E = 60\%$ and the other inputs of (c), find V_0 and σ_V, the default probability and the spread. Describe your numerical method, prove that for fixed σ_V the equation for V_0 has a unique root, and check numerically that the remaining equation in σ_V has a single root on $[0.01, 2]$.

Exercise 2.7 *(How well is implied volatility determined?)* [⋆⋆]

(a) Show that the map from a call price to its implied volatility is continuous and increasing on the interval Equation (2.17), tends to 0 and to ∞ at its ends, and has derivative $1/\mathcal{V}$.

(b) Let $\kappa(w) = c(m, w)/(w \, \partial_w c(m, w))$, the relative change in the implied volatility per relative change in the price. For fixed $m < 0$, show that $\kappa(w) \to 0$ as $w \downarrow 0$ and $\kappa(w) \to \infty$ as $w \to \infty$; find the limit as $w \downarrow 0$ when $m = 0$. *Hint:* below the inflection point, combine $c \, \partial_{ww} c < (\partial_w c)^2$ from Theorem 2.9(iv) with the formula for $\partial_{ww} c$ in part (ii) of that proposition.

(c) Take $S = 100$, $r = y = 0$, $T = 0.25$, $\sigma = 20\%$ and $K = 100, 110, 120, 130, 140$. For each strike compute the price, the vega and κ, and the change in implied volatility caused by an error of half a cent in the price. Which of these quotes, rounded to the cent, still determine the volatility to one volatility point? Reconcile your answer with (b).

Exercise 2.8 *(The course's Newton experiment)* [⋆⋆] This is a computational exercise. Let $\hat{C} = C/(Ke^{-rT})$ and $\hat{F} = F/K$.

(a) By Theorem 2.11(2) it suffices to treat calls with $\hat{F} \le 1$ and $0 < \hat{C} < \hat{F}$. Show that then $\hat{w}_{\text{imp}} \ge 2N^{-1}((1 + \hat{C})/2)$, with equality if and only if $\hat{F} = 1$. *Hint:* how does $c(m, w)$ depend on m?

(b) Implement Newton's method on the price, $\hat{w} \leftarrow \hat{w} - (c(\log \hat{F}, \hat{w}) - \hat{C})/N'(d_-)$, started at $\hat{w}_0 = 0.9951$ and stopped when $|c - \hat{C}| \le 10^{-8}$. On the grid $\hat{F} \in \{0.1, 0.25, 0.5, 0.75, 1\}$, $\hat{C} \in \{0.01, 0.1, 0.5, 0.9, 0.99\} \times \hat{F}$, report the number of iterations and the limit, or the failure, in each cell.

(c) Repeat with the starting point of Theorem 2.9(iii), and compare. Which cells explain the curious value 0.9951?

(d) Propose a starting function $\hat{w}_0(\hat{C}, \hat{F})$ of your own that uses only elementary functions and N^{-1}, and report the iteration counts it gives on the grid.

Chapter 3

Arbitrage in the strike and the smile

Throughout, calls and puts have one maturity T, the bond paying 1 at T costs e^{-rT}, the prepaid forward paying S_T costs S_0e^{-yT}, $F = S_0e^{(r-y)T}$, and $S_T \geq 0$. Hypotheses (A), (B) and (P) are those of Section 3.1. You may use the results of the notes, quoting them, and the closed forms and the implied-volatility solver of `bs.py`. In computational parts report your numbers to the precision asked; if you use random numbers, use NumPy's `default_rng` and report the seed.

Exercise 3.1 *(Reading rates from parity)* [⋆] Assume (P). By Theorem 3.1(i), $K \mapsto C(K) - P(K)$ is affine, with slope $-e^{-rT}$ and intercept the prepaid forward.

(a) Suppose that the stock pays no yield but a known cash dividend D at a date $t_D \in (0, T)$, and that a bond paying 1 at t_D costs e^{-rt_D}. Find the price of the prepaid forward with a static portfolio, and the affine function $C(K) - P(K)$. What does its intercept measure now?

(b) A *box spread* is long a call spread (K_1, K_2) and long a put spread, short the put at K_1 and long the put at K_2, with $K_1 < K_2$. Find its payoff and its price.

(c) With $S_0 = 100$ and $T = 0.5$, the mid quotes of $C(K) - P(K)$ at $K = 90, 100, 110$ are 11.03, 1.23 and -8.57; the stock pays a yield y. Fit an affine function by least squares and deduce r and y, in percent with two decimals. By how much can rounding the quotes to the cent move your r?

(d) The bond with maturity T trades at 0.9802, and a dealer offers to buy or sell the 90/110 box at 19.70. Is there an arbitrage? If so, give the trade and its profit today.

Exercise 3.2 *(Finding the arbitrage in quotes)* [⋆] Let $S_0 = 100$, $T = 1$, $r = y = 0$, and let calls be quoted at

K	80	90	100	110	120
$C(K)$	19.80	12.50	7.20	1.50	0.40

(a) Compute the slopes $s_1, \ldots, s_5$ of Equation (3.4) and say which conditions fail.

(b) For each failure, write down a static portfolio that is an arbitrage: its price today and its payoff as a function of S_T.

(c) Compute the implied volatility of each quote, or explain why it does not exist. How does the second failure show up in the implied volatilities?

Exercise 3.3 *(Default and the left wing)* [⋆⋆] Under (B), let $S_T = 0$ with probability $p \in (0,1)$ and $S_T = \frac{F}{1-p}\exp\left(\sigma\sqrt{T}Z - \frac{\sigma^2 T}{2}\right)$ otherwise, with $Z \sim \mathcal{N}(0,1)$ independent of the default event and $\sigma > 0$.

(a) Check that $\mathbb{E}^{\mathbb{Q}}[S_T] = F$, and express $C(K)$ with the Black–Scholes formula.

(b) Show that the digital put paying $\mathbf{1}_{\{S_T<K\}}$ and the ratio $P(K)/K$ both tend to $e^{-rT}p$ as $K \downarrow 0$, so that the left end of the put curve prices a claim paying 1 at T on default. Does the same hold for every law of S_T, with $p = \mathbb{Q}(S_T = 0)$?

(c) Show that $P(K) \geq pKe^{-rT}$ for every $K \geq 0$, and that for every fixed $v > 0$ the Black–Scholes put price at volatility v is $o(K)$ as $K \downarrow 0$. Deduce that the implied volatility $\sigma(K)$ tends to $+\infty$ as $K \downarrow 0$.

(d) For $S_0 = 100$, $r = 1\%$, $y = 2\%$, $T = 1$, $\sigma = 20\%$ and $p = 2\%$, compute $\sigma(K)$ for $K = 10, 20, 40, 60, 80, 100, 120$ and comment on the shape.

Exercise 3.4 *(Calendar spreads at a fixed strike)* [⋆⋆]

(a) In the Black–Scholes model with $S_0 = 100$, $r = 1\%$, $y = 5\%$ and $\sigma = 20\%$, compute $C(60, T)$ for $T = 1, 2, 5$. Is a call with a longer maturity always worth more?

(b) Compute $e^{yT}C(60e^{(r-y)T}, T)$ for the same maturities and compare with Theorem 3.4.

(c) Under the hypotheses of Theorem 3.4, show that if $y = 0$ and $r \geq 0$, then $T \mapsto C(K,T)$ is nondecreasing for every fixed $K \geq 0$. Which step fails when $y > 0$, and what is the financial reason?

Exercise 3.5 *(Pricing and bounding a digital under the course's smile)* [⋆⋆] This is a computational exercise. Use the setting of Theorem 3.9: $S_0 = 3900$, $T = 1$, $r = 3\%$, $y = 1.3\%$, $\sigma(K) = 25\% + 28\% \log(S_0/K)$, and the strike $K = 4000$.

(a) Price the cash-or-nothing and the asset-or-nothing calls with a flat volatility of 25%.

(b) Price them under the smile.

(c) Calls trade at the strikes $25\mathbb{N}$. Find the cheapest super-replicating and the dearest sub-replicating portfolios of the digital and their prices. Repeat for the strikes $5\mathbb{N}$.

(d) Compare (a), (b) and (c). Is quoting the digital at the flat-volatility price an arbitrage? If so, write the trade.

(e) Which strike spacing would bring the gap between the bounds of (c) below 0.005?

Exercise 3.6 *(A higher-order stencil)* [⋆⋆] Let $D_h f(K) = \big(-f(K-2h) + 16f(K-h) - 30f(K) + 16f(K+h) - f(K+2h)\big)/(12h^2)$.

(a) Let f be six times continuously differentiable on $[K-2h, K+2h]$ with $|f^{(6)}| \leq M_6$

there. Show that $|D_h f(K) - f''(K)| \le M_6 h^4/54$. *Hint:* write the numerator through $f(K+jh) + f(K-jh) - 2f(K)$, $j = 1, 2$, and Taylor's formula with integral remainder.

(b) If the values of f carry errors bounded by ε, bound the error of D_h as in Theorem 3.6, find the best h, and show that the best bound is of order $\varepsilon^{2/3}$ (against $\varepsilon^{1/2}$ for the three-point stencil).

(c) Take the SSVI prices of Figure 3.3 at $T = 1$ (`smile.ssvi_smile` with its defaults), rounded to the cent. On the grids $\{60, 60+h, \dots, 140\}$, $h \in \{1, 2, 3, 4, 5, 6, 8, 10\}$, compute the largest error of the three-point and of the five-point estimates of C''. Which h is best for each, and how do the errors compare with your bounds? (Estimate M_6 numerically on $[40, 160]$.)

Exercise 3.7 *(Durrleman's function)* [⋆⋆] Let σ be a twice differentiable smile, $k = \log(K/F)$ and $w(k) = \sigma(K)^2 T$.

(a) Express $K\sqrt{T}\sigma'(K)$, $K^2 T\sigma'(K)^2$ and $K^2 T\sigma(K)\sigma''(K)$ through w, w' and w'', and show that the bracket of Equation (3.9) equals the function $g(k)$ of Section 3.3.

(b) Show that $g \equiv 1$ for a flat smile. Write $g(0)$ in terms of $w(0)$, $w'(0)$, $w''(0)$, and deduce that a smile with $w''(0) \le 0$ and no butterfly arbitrage at the forward must have $|w'(0)| \le 2/\sqrt{1/w(0) + 1/4}$. Compute $w(0)$, $w'(0)$, $w''(0)$ and $g(0)$ for the course's smile of Theorem 3.9.

(c) For the SSVI slice Equation (3.10), compute $w(0)$, $w'(0)$, $w''(0)$ and $g(0)$ in terms of θ, $\phi = \phi(\theta)$ and ρ.

Exercise 3.8 *(How steep can a smile be?)* [⋆⋆⋆] Assume (B), and let $C(K) = C^{\mathrm{BS}}(K, \sigma(K))$ on an open interval of strikes, with σ differentiable there. Let $M(x) = (1 - N(x))/N'(x)$ be Mills' ratio; you may use that M is positive and strictly decreasing on $\mathbb{R}$, with $M(0) = \sqrt{\pi/2}$.

(a) Use $C'(K) \le 0$ to show that $K\sqrt{T}\,\sigma'(K) \le M(-d_-)$.

(b) Use the convexity of C on $[0, K]$ and $C(0) = S_0 e^{-yT}$ to show that $K\sqrt{T}\,\sigma'(K) \ge -M(d_+)$. Compare with the bound that $C'(K) \ge -e^{-rT}$ gives. (Both bounds appear in M. Roper, *Arbitrage free implied volatility surfaces*, preprint, University of Sydney, 2010, Corollary 2.14.)

(c) At the forward, $K = F$, show that $|\partial\sigma/\partial \log K| \le M(\sigma\sqrt{T}/2)/\sqrt{T} < \sqrt{\pi/2}/\sqrt{T}$: the at-the-money skew can grow at most like $T^{-1/2}$ as $T \downarrow 0$. Check the SSVI surface of Section 3.3 against this bound.

(d) For the course's smile, compute both bounds at $K = S_0$. Find numerically the strikes at which the smile violates (b), and explain why the answer is not the set where its density is negative.

Chapter 4

Static replication and the Kolmogorov equations

Throughout, the market is that of Section 4.1: a pricing measure $\mathbb{Q}$, the bond paying 1 at T costs e^{-rT}, the prepaid forward costs $S_0 e^{-yT}$, $F = S_0 e^{(r-y)T}$, and C, P, D are call, put and digital-call prices for maturity T. "The standard model" is that of Theorem 1.1. You may use the results of the notes, quoting them, and the closed forms and the implied-volatility solver of `bs.py`. In computational parts report your numbers to the precision asked; if you use random numbers, use NumPy's `default_rng` and report the seed.

Exercise 4.1 *(A note with a digital)* [⋆] A five-year note on an index pays, per 100 of notional and with $R = S_T/S_0$,

$$\Pi(R) = \begin{cases} 100R, & R < 0.6, \\ 100, & 0.6 \le R < 0.9, \\ 100 + 400\,(R - 0.9), & 0.9 \le R < 1, \\ 140, & R \ge 1. \end{cases}$$

(a) Sketch Π and write it as a portfolio of shares, calls and one digital, with strikes in units of S_0.

(b) In the standard model with $S_0 = 100$, $T = 5$, $r = 3\%$, $y = 3.5\%$ and $\sigma = 22\%$, price each component and the note, to four decimals. The issuer sells the note at 100: what is its margin?

(c) Does it matter whether the digital pays at $R = 0.6$ exactly? Would it matter if S_T had an atom at $0.6\,S_0$?

(d) A dealer who cannot trade digitals holds instead 40 call spreads, long the call struck at 59 and short the call struck at 60. Show that this position pays at least the digital's payoff in every state, and compute its extra cost.

Exercise 4.2 *(The log put)* [⋆⋆] The *log put* with strike $K > 0$ pays $K\big[\log(K/S_T)\big]^+$.

(a) Using Theorem 4.1 with $\kappa = K$, show that it is replicated by one put struck at K and a strip of puts with strikes $k < K$ and density K/k^2, and write its price in terms of put prices.

(b) Show that, in any model with a pricing measure and $\mathbb{Q}(S_T = 0) = 0$, the log put has a finite price if and only if $\int_0^K P(k)\,k^{-2}\,\mathrm{d}k < \infty$. Give a law of S_T with mean F, without atoms, for which this integral is infinite. Which feature of the law matters?

(c) In the standard model, show that the log put costs $e^{-rT}K\sigma\sqrt{T}\,[N'(d_-) - d_-N(-d_-)]$, and find the price of the *log call*, which pays $K[\log(S_T/K)]^+$, by a parity argument.

(d) For the test case ($S_0 = K = 100$, $T = 1$, $r = 1\%$, $y = 2\%$, $\sigma = 20\%$), compute the log put from (c), and the put and the strip of (a) separately by numerical integration, to five decimals.

Exercise 4.3 *(Two power payoffs from the course)* [⋆⋆] Work in the standard model and use Theorem 4.4.

(a) Show that $2\big(S_T - \sqrt{KS_T}\big)^+ = 2S_T\mathbf{1}_{\{S_T>K\}} - 2\sqrt{K}\,S_T^{1/2}\mathbf{1}_{\{S_T>K\}}$, and deduce that it costs

$$2\Big[S_0e^{-yT}N(d_+) - \sqrt{KS_0}\,e^{-(\frac{1}{2}(r+y)+\frac{1}{8}\sigma^2)T}N(x_0)\Big], \qquad x_0 = \frac{\log(S_0/K) + (r-y)T}{\sigma\sqrt{T}}.$$

Evaluate it in the test case to five decimals.

(b) The *asymptotic call* pays $(K/S_T)(S_T - K)^+$. Price it as a leveraged payoff with $p = -1$: show that the inputs are $r'_{-1} = 2r - y - \sigma^2$ and $y'_{-1} = r$, and evaluate the price for $S_0 = K = 100$, $T = 1$, $r = 0.75\%$, $y = 2\%$, $\sigma = 20\%$ to four decimals.

(c) Replicate the asymptotic call with calls only, and check (b) by integrating over the strip numerically.

(d) Both payoffs have right derivative 1 at K, like the call struck at K. Without computing, which of the two is worth more than the call struck at K with the same inputs (those of (a), respectively (b)), and which less? Check with the numbers.

Exercise 4.4 *(Homogeneity without scale invariance)* [⋆⋆] In the normal (Bachelier) model the forward moves as $\mathrm{d}F_t = \sigma_N\,\mathrm{d}W_t$ under $\mathbb{Q}$, with $F_0 = S_0e^{(r-y)T}$ and a normal volatility $\sigma_N > 0$ measured in money. The call costs

$$C = e^{-rT}\big[(F_0 - K)N(d) + \sigma_N\sqrt{T}\,N'(d)\big], \qquad d = \frac{F_0 - K}{\sigma_N\sqrt{T}}.$$

(a) Show that $C(\lambda S_0, \lambda K, \lambda\sigma_N) = \lambda C(S_0, K, \sigma_N)$ for $\lambda > 0$, and deduce $S_0\partial_{S_0}C + K\partial_K C + \sigma_N\partial_{\sigma_N}C = C$.

(b) Let $r = y = 0$. Show that $C(S_0{+}c, K{+}c) = C(S_0, K)$ for every real c, deduce $\partial_{S_0}C + \partial_K C = 0$ and $\partial_{S_0S_0}C = -\partial_{S_0K}C = \partial_{KK}C$, and compare with Theorem 4.7(iv). What replaces the factor $(S/K)^2$ between the spot gamma and the risk-neutral density?

(c) With $r = y = 0$, $S_0 = K = 100$, $T = 1$ and $\sigma_N = 20$, compute C and $S_0\partial_{S_0}C + K\partial_K C$. How badly does Euler's identity in (S_0, K) fail?

(d) In the constant-elasticity model $\mathrm{d}S_t = \sigma S_t^{\beta}\,\mathrm{d}W_t$ ($r = y = 0$, $0 < \beta < 1$, absorbed at 0), show that λS follows the same model with σ replaced by $\lambda^{1-\beta}\sigma$, and deduce $S_0\partial_{S_0}C + K\partial_K C + (1-\beta)\sigma\partial_\sigma C = C$, assuming the price is differentiable. You may assume that the stochastic differential equation has a unique solution in law.

Exercise 4.5 *(Stationary laws from the forward equation)* [⋆⋆] Let $\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + a(X_t)\,\mathrm{d}W_t$ on an open interval I, with b continuously differentiable, a^2 twice continuously differentiable and $a > 0$ on I, and no discounting. A *stationary density* is a probability density p on I that does not change under the forward equation of Theorem 4.9(iii).

(a) Show that a twice differentiable stationary density satisfies $\frac{1}{2}(a^2p)' - bp = -\phi$ for a constant ϕ, the *probability flux*. Solve this equation when $\phi = 0$.

(b) For the Ornstein–Uhlenbeck process $\mathrm{d}X = \kappa(\theta - X)\,\mathrm{d}t + \gamma\,\mathrm{d}W$ on $I = \mathbb{R}$ ($\kappa, \gamma > 0$), find the stationary law, and show that no stationary density has $\phi \neq 0$.

(c) For the square-root process $\mathrm{d}V = \kappa(\theta - V)\,\mathrm{d}t + \gamma\sqrt{V}\,\mathrm{d}W$ on $I = (0,\infty)$ ($\kappa, \theta, \gamma > 0$), the variance of the Heston model, find the stationary density with zero flux. Show that it is a gamma density, give its mean and variance, and show that it is bounded near 0 if and only if $2\kappa\theta \geq \gamma^2$ (Feller's condition).

(d) Which hypothesis of Theorem 4.9 fails in (c)? Explain in a few lines why the zero-flux condition is the natural boundary condition at 0 here.

Exercise 4.6 *(Dupire's equation)* [⋆⋆] Part (c) is the hardest. In a local-volatility model, $\mathrm{d}S_t = (r-y)S_t\,\mathrm{d}t + \sigma(t,S_t)S_t\,\mathrm{d}W_t$ under $\mathbb{Q}$, with r, y constant. Let $G(T,K)$ be the discounted density of S_T at $K > 0$ (the Arrow–Debreu price density, for a fixed spot S_0), and $C(T,K) = \int_K^\infty (\xi - K)\,G(T,\xi)\,\mathrm{d}\xi$ the call price. Assume that G is $C^{1,2}$ on $(0,\infty)^2$ and satisfies the forward equation

$$\partial_T G = \tfrac{1}{2}\partial_{KK}(\sigma^2K^2G) - \partial_K\big((r-y)KG\big) - rG,$$

which is the forward equation of Theorem 4.9 written in the price variable, and that G and its derivatives decay fast enough at 0 and ∞ for every boundary term below to vanish and for differentiation under the integral sign.

(a) Show that $\partial_K C = -\int_K^\infty G\,\mathrm{d}\xi$ and $\partial_{KK}C = G$.

(b) By integrating the forward equation against $(\xi - K)^+$, show *Dupire's equation*

$$\partial_T C = \tfrac{1}{2}\sigma(T,K)^2K^2\,\partial_{KK}C - (r-y)K\,\partial_K C - yC.$$

(c) Deduce a formula for $\sigma(T,K)^2$ in terms of C and its derivatives. Which no-arbitrage relations of Chapter 3 make the right side well defined and nonnegative?

(d) Check that the Black–Scholes price with constant σ satisfies Dupire's equation. You may use the Greeks of Theorem 2.7.

(e) Compare the role of (T,K) in Dupire's equation with the role of (t,S) in the pricing equation (Equation (1.15)). Which one would you solve to price a single option for many spots, and which to price many strikes for the current spot?

Exercise 4.7 *(A static hedge for a barrier option)* [⋆⋆⋆] In the standard model with $r = y$, a down-and-in call with barrier H and strike K, $0 < H \le K$ and $H < S_0$, pays $(S_T - K)^+$ at T if $\min_{t \le T} S_t \le H$, and nothing otherwise.

(a) Use Theorem 4.6 to show that at any time $t < T$ at which $S_t = H$, the call struck at K and K/H puts struck at H^2/K, both with maturity T, have the same price.

(b) Consider the strategy: at time 0 buy K/H puts struck at H^2/K; if the stock reaches H before T, sell them and buy one call struck at K. Show that it costs nothing after time 0 and pays exactly the down-and-in call. Deduce the price of the down-and-in call, and explain why this is an arbitrage argument that needs no knowledge of the law of the hitting time.

(c) With $S_0 = 100$, $K = 100$, $H = 90$, $T = 1$, $r = y = 2\%$ and $\sigma = 20\%$, compute the price to five decimals. Check it by Monte Carlo: simulate $\log S$ on 500 steps and detect crossings of the barrier between the grid dates with the Brownian-bridge probability $\exp\big(-2(x_k - \log H)(x_{k+1} - \log H)/(\sigma^2 h)\big)$ for consecutive log-prices x_k, x_{k+1} above $\log H$ and step h.

(d) Now let $r = 5\%$ and $y = 0$, with the other inputs of (c). Compute the cost of the exchange in (b) if the barrier is hit with 0.25, 0.5 and 1 year left. Is the put portfolio a super- or a sub-hedge? Explain the sign.

Part II

Monte Carlo methods

Chapter 5

Random numbers

Exercise 5.1 *(Low-order bits)* [⋆] Let (x_i) be an LCG with modulus $m = 2^e$, multiplier a and increment c (Theorem 5.1), and let $1 \le j \le e$.

(a) Show that $y_i = x_i \bmod 2^j$ satisfies $y_{i+1} = (ay_i + c) \bmod 2^j$, and hence that its period is at most 2^j. What does the last bit $x_i \bmod 2$ do when a and c are odd?

(b) Write down $x_i \bmod 2$ and $x_i \bmod 4$ for the course's generator $m = 16$, $a = 9$, $c = 5$, $x_0 = 0$.

(c) The uniforms $u_i = x_i/m$ are rarely affected by this. Why? A program picks one of $k = 2^j$ scenarios at step i by the index $x_i \bmod k$. What goes wrong, and what index should it use instead?

Exercise 5.2 *(The Hull–Dobell conditions are necessary)* [⋆⋆] Consider an LCG with modulus $m \ge 2$, multiplier a and increment c. Show that it does not have full period in each of the following cases, by exhibiting a seed and a residue that the sequence never visits.

(a) $d = \gcd(c, m) > 1$.

(b) Some prime p divides m but not $a - 1$. *Hint:* look at the sequence modulo p and find a fixed point.

(c) 4 divides m, a is odd and 4 does not divide $a - 1$. *Hint:* modulo 4.

(d) Which of $(m, a, c) = (16, 9, 5)$, $(16, 5, 4)$, $(16, 3, 1)$, $(12, 7, 5)$, $(12, 1, 5)$ have full period? (You may use Theorem 5.3.) Comment on the last one.

Exercise 5.3 *(Inversion with atoms and flat stretches)* [⋆] Use Theorem 5.8.

(a) *Pareto.* $F(x) = 1 - (x_0/x)^\alpha$ for $x \ge x_0 > 0$, with $\alpha > 0$. Give a sampler by inversion. For which α does the law have a finite mean, and a finite variance?

(b) *Geometric.* $\mathbb{P}(X = k) = (1 - p)^{k-1}p$ for $k = 1, 2, \dots$, with $p \in (0, 1)$. Show that $F^{-1}(u) = \lceil \log(1 - u)/\log(1 - p) \rceil$ and that $\lceil \log U/\log(1 - p) \rceil$ is also geometric with parameter p.

(c) *An atom.* A loss is 0 with probability 0.3 and otherwise exponential with mean 1. Compute its distribution function F and F^{-1}, and check that F^{-1} has a flat stretch where F jumps.

Exercise 5.4 *(Normal variates by rejection from the Laplace law)* [⋆⋆] Let $g_\lambda(x) = \frac{\lambda}{2}e^{-\lambda|x|}$ ($\lambda > 0$) be the Laplace density, and use Theorem 5.10 with target N'.

(a) Compute $c(\lambda) = \sup_x N'(x)/g_\lambda(x)$, find the λ that minimises it, and give the resulting acceptance probability.

(b) With the optimal λ, show that the acceptance test reduces to $U \le \exp\big(-\frac{1}{2}(|Y|-1)^2\big)$.

(c) Sample Y by inversion with one uniform. How many uniforms does the method use per normal on average? Compare with Box–Muller and with the polar method.

(d) Can one sample the Laplace law by rejection from a normal proposal? Why?

Exercise 5.5 *(Ratio of uniforms)* [⋆ ⋆ ⋆] Let $h \ge 0$ be an integrable function on $\mathbb{R}$ with $\int h = Z > 0$, and $C_h = \{(s,v) \in \mathbb{R}^2 :\ 0 < s \le \sqrt{h(v/s)}\}$.

(a) Show that C_h has area $Z/2$. *Hint:* change variables $(s,v) \mapsto (s,x)$ with $x = v/s$.

(b) Show that if (S,V) is uniform on C_h, then $X = V/S$ has density h/Z.

(c) For $h(x) = e^{-x^2/2}$, show that C_h is contained in the rectangle $(0,1] \times [-\sqrt{2/e}, \sqrt{2/e}]$, and write the acceptance condition for a point of the rectangle without exponentials.

(d) Deduce a rejection algorithm for $\mathcal{N}(0,1)$ that uses two uniforms per trial, and compute its acceptance probability.

Exercise 5.6 *(Square roots of covariance matrices)* [⋆] Use Theorem 5.14.

(a) Let $0 < t_1 < \cdots < t_d$ and $\Sigma_{ij} = \min(t_i, t_j)$, the covariance of a Brownian motion at these times. Show that the Cholesky factor is $L_{ij} = \sqrt{t_j - t_{j-1}}$ for $j \le i$ and 0 otherwise ($t_0 = 0$), and interpret LZ.

(b) For the 2×2 correlation matrix with correlation $\rho \in (-1,1)$, compute the Cholesky and the principal-components factors. What happens to each as $\rho \to \pm 1$?

(c) A desk quotes correlations $\rho_{12} = \rho_{13} = 0.9$ and $\rho_{23} = -0.9$. Run the recursion of Theorem 5.14(ii) by hand. Find a linear combination of the three assets that would have negative variance, and explain the inconsistency in words.

Exercise 5.7 *(Testing moments)* [⋆⋆] This is a computational exercise.

(a) Let $x_1, x_2, \dots$ be independent and identically distributed with variance σ^2 and finite fourth central moment μ_4, and let $s_n^2 = \frac{1}{n}\sum_{i \le n}(x_i - \bar{x}_n)^2$. Show that $\sqrt{n}(s_n^2 - \sigma^2) \to \mathcal{N}(0, \mu_4 - \sigma^4)$ in law. You may use the central limit theorem, the strong law of large numbers and Slutsky's lemma. Deduce the approximate standard errors of $\bar{x}_n$ and s_n^2 for $\mathcal{U}(0,1)$ and for $\mathcal{N}(0,1)$ samples.

(b) You may use that the sample skewness and excess kurtosis are asymptotically normal with variances $6/n$ and $24/n$ for $\mathcal{N}(0,1)$ samples, and $\frac{72}{35}/n$ and $\frac{1152}{875}/n$ (around 0 and

$-\frac{6}{5}$) for $\mathcal{U}(0,1)$ samples. Draw 10^6 uniforms with `numpy.random.default_rng(2024)` and Box–Muller normals from them (both members of each pair). For the nested samples of the first $n = 10^3, 10^4, 10^5, 10^6$ values, tabulate the four sample moments and their z-scores. Which deviations are significant at the 5% level? Is that surprising?

(c) Repeat with RANDU (`uniforms(n, "randu", seed=1)` in the course code). Do the moment tests detect its defect? Why?

Exercise 5.8 *(Birthdays: resolution against period)* [⋆⋆] This is a computational exercise.

(a) An ideal generator with resolution 2^{-b} outputs independent values, each uniform on a set of 2^b points. Show that the expected number of pairs $i < j \le n$ with $u_i = u_j$ is $\binom{n}{2}2^{-b}$. Evaluate it for $n = 2^{20}$ and $b = 32$, and for $b = 53$.

(b) Count the colliding pairs among the first 2^{20} outputs of PCG32 (seed `initstate=42`, `initseq=54`) and among the first 2^{20} outputs $x_1, x_2, \ldots$ of the full-period LCG modulo 2^{32} with $a = 1664525$, $c = 1013904223$, $x_0 = 0$.

(c) Explain the LCG's count, and why it is a failure. Assuming the number of collisions is approximately Poisson for an ideal generator, give a p-value.

(d) What does this say about Theorem 5.5, and about simulations that draw 10^7 or more 32-bit uniforms?

Chapter 6

Monte Carlo pricing and its error

Exercise 6.1 *(The exact variance of a call payoff)* [⋆] In the Black–Scholes model with yield, let $F = S_0 e^{(r-y)T}$ and $d_\pm = \big(\log(S_0/K) + (r - y \pm \tfrac{1}{2}\sigma^2)T\big)/(\sigma\sqrt{T})$. You may use the call price $C = e^{-rT}(FN(d_+) - KN(d_-))$ and the Gaussian identity $\mathbb{E}[e^{aZ}\mathbf{1}_{\{Z>b\}}] = e^{a^2/2}N(a-b)$ for $Z \sim \mathcal{N}(0,1)$.

(a) Show that $\mathbb{E}^{\mathbb{Q}}[S_T^2 \mathbf{1}_{\{S_T>K\}}] = F^2 e^{\sigma^2 T} N(d_+ + \sigma\sqrt{T})$.

(b) Deduce a closed form for the variance σ_C^2 of the discounted payoff $e^{-rT}(S_T - K)^+$, and for the variance of S_T.

(c) Evaluate both standard deviations for the test case, and for the parameters of the course's problem sets: $S_0 = K = 3900$, $T = 1$, $r = 3\%$, $y = 1.3\%$, $\sigma = 25\%$.

(d) A run with $n = 10^4$ paths on the test case returns $\hat{C} = 7.52$ with $\hat{s}_n = 12.9$. At the 5% level, is it consistent with the exact price? How many paths would make a discrepancy of 0.16 significant at that level?

Exercise 6.2 *(When no scenario pays)* [⋆] Let $X = \mathbf{1}_A$ with $p = \mathbb{P}(A) \in (0,1)$, for example the payoff of a digital option, and let $\hat{X}_n$, $\hat{s}_n$ and I_n be as in Theorem 6.1.

(a) Compute $\mathbb{P}(\hat{X}_n = 0)$. What are $\hat{s}_n$ and I_n on this event? Evaluate the probability for $p = 1\%$ and $n = 100$, and the probability that I_n contains p on this event.

(b) Suppose a run has no hits. Show that the set of p for which observing no hit has probability at least 5% is $[0, p^*]$ with $p^* = 1 - 0.05^{1/n}$, and that $p^* \approx 3/n$ for large n. Interpret $[0, 3/n]$ as a confidence interval.

(c) A run of $n = 1000$ paths has 3 hits. Compute the interval I_n, and the exact interval $[p_-, p_+]$, where p_- solves $\mathbb{P}_p(\text{at least 3 hits}) = 2.5\%$ and p_+ solves $\mathbb{P}_p(\text{at most 3 hits}) = 2.5\%$ (numerically). Then compute exactly, by summing binomial probabilities, the coverage of I_n and of the exact interval when $p = 0.003$ and $n = 1000$. Comment.

Exercise 6.3 *(The price of batching)* [⋆⋆] Let $X_1, \dots, X_n$ be independent $\mathcal{N}(\mu, \sigma^2)$, $n = mk$, and let $\hat{V}_m$ be the batch variance estimator of Theorem 6.8 with m batches of size k. You may

use that the sample variance S^2 of ℓ independent $\mathcal{N}(\mu', v)$ variables satisfies $(\ell-1)S^2/v \sim \chi^2_{\ell-1}$, and that $\mathbb{E}[\sqrt{W}] = \sqrt{2}\,\Gamma(j/2+\tfrac{1}{2})/\Gamma(j/2)$ for $W \sim \chi^2_j$.

(a) Compute the relative standard deviation $\sqrt{\mathrm{Var}(\hat{V}_m)}/\mathbb{E}[\hat{V}_m]$ and compare it with that of the plain estimator $\hat{s}_n^2/n$.

(b) Show that the expected half-width of the batch interval $\bar{Y} \pm t_{m-1,\delta}\sqrt{\hat{V}_m}$ is $t_{m-1,\delta}\, c_m\, \sigma/\sqrt{n}$ with $c_m = \sqrt{2/(m-1)}\,\Gamma(m/2)/\Gamma((m-1)/2)$.

(c) For $\delta = 5\%$, tabulate the ratio of this expected half-width to $1.96\,\sigma/\sqrt{n}$ for $m = 2$, 5, 10, 32, 1024. What would you recommend?

(d) If $m = 2^{10}$ batches give a batch standard error equal to the plain one up to noise, what does that tell you, and what does it not tell you?

Exercise 6.4 *(Dependent samples)* [⋆⋆] This exercise is partly computational. Let $|\phi| < 1$ and let $X_0 \sim \mathcal{N}(\mu, \sigma^2)$, $X_i = \mu + \phi(X_{i-1} - \mu) + \sigma\sqrt{1-\phi^2}\,\varepsilon_i$ for $i \geq 1$, with ε_i independent standard normals independent of X_0. Such correlated samples arise, for instance, from Markov chain samplers.

(a) Show that every X_i is $\mathcal{N}(\mu, \sigma^2)$ and that $\mathrm{Cov}(X_i, X_j) = \sigma^2\phi^{|i-j|}$.

(b) With $\hat{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, show that $n\,\mathrm{Var}(\hat{X}_n) \to \sigma^2\frac{1+\phi}{1-\phi}$, while $\mathbb{E}[\hat{s}_n^2] \to \sigma^2$. By what factor does the plain standard error misstate the error for $\phi = 0.9$?

(c) Let the batch means Y_j be built from consecutive blocks of k samples. Show that the correlation of Y_j and Y_{j+1} tends to 0 as $k \to \infty$.

(d) Simulate $n = 2^{20}$ samples with $\mu = 0$, $\sigma = 1$, $\phi = 0.9$. Report the plain and the batch ($m = 2^{10}$) standard errors next to the asymptotic standard deviation from (b), and estimate from 1000 independent runs the coverage of the two nominal 95% intervals for μ.

Exercise 6.5 *(The implied volatility of a simulated price)* [⋆⋆] This exercise is partly computational. A desk prices the test-case call by Monte Carlo and quotes the result as a Black–Scholes implied volatility $\hat{\sigma}_n = g(\hat{C}_n)$, where g is the inverse of $\sigma \mapsto C(\sigma)$. You may use that $\partial C/\partial\sigma = \mathcal{V}$ (vega) and $\partial^2 C/\partial\sigma^2 = \mathcal{V}\,d_+d_-/\sigma$, and that $\sigma_C = 12.570$.

(a) Use Theorem 6.12 to find the asymptotic standard deviation of $\hat{\sigma}_n$. How many paths are needed for a 95% half-width of 0.1 volatility point?

(b) Compute $g''(C)$. Applying the formula of Theorem 6.12(ii) formally, what is the leading bias of $\hat{\sigma}_n$, and how does it compare with the standard deviation at $n = 10^4$? (Part (ii) does not apply as stated: why not?)

(c) Check (a) by simulation: 1000 independent runs of $n = 10^4$ paths.

Exercise 6.6 *(Heavy tails)* [⋆⋆⋆] This exercise is partly computational. Let X have $\mathbb{P}(X > x) = x^{-\alpha}$ for $x \geq 1$, with $1 < \alpha < 2$, and let $M_n = \max_{i\leq n} X_i$.

(a) Show that $U^{-1/\alpha}$ has this law for U uniform on $(0,1)$, compute $\mu = \mathbb{E}[X]$, and show that

$\mathbb{E}[X^2] = \infty$ while $\mathbb{E}[X^q] < \infty$ for $q < \alpha$.

(b) Show that $M_n/n^{1/\alpha}$ converges in distribution and identify the limit.

(c) Show that $\mathbb{E}[M_n]/n \to 0$ but $\mathbb{E}[M_n]/\sqrt{n} \to \infty$. Explain why this suggests an error of order $n^{1/\alpha - 1}$ rather than $n^{-1/2}$ for $\hat{X}_n$.

(d) By simulation, estimate the median of $|\hat{X}_n - \mu|$ for $n = 2^6, 2^8, \dots, 2^{16}$ and $\alpha = 1.2$ and 1.8 (use at least 1000 runs for each n), fit the slope on a log-log scale, and compare with (c).

(e) Estimate the coverage of the nominal 95% interval at $n = 2^{12}$ for $\alpha = 1.2$, 1.8 and 2.5. Comment on the case $\alpha = 2.5$, for which $\mathbb{E}[X^2] < \infty$ but $\mathbb{E}|X|^3 = \infty$.

Exercise 6.7 *(A price relative to the forward)* [⋆⋆] This exercise is partly computational. Since $\mathbb{E}^{\mathbb{Q}}[S_T] = F$ is known, one may estimate the call price by the ratio estimator $\hat{C}_n^R = e^{-rT} F\, \bar{Y}_n / \bar{X}_n$, with $X = S_T$ and $Y = (S_T - K)^+$. You may use the results of Exercise 6.1, and you may assume that the bias expansion of Theorem 6.12(iii) holds for this lognormal X (see the remark that follows it).

(a) Compute $\mathrm{Cov}(X, Y)$ in closed form, and the correlation of X and Y for the test case.

(b) Compute the asymptotic variance of $\hat{C}_n^R$ and compare it with that of the plain estimator. Why is the reduction so much smaller than $1 - \rho^2$, where ρ is the correlation of (a)? *Hint:* write $Y - \theta X$ as a control-variate combination and compare θ with $\mathrm{Cov}(X, Y)/\mathrm{Var}(X)$.

(c) Compute the leading bias of $\hat{C}_n^R$ and check it by simulation at $n = 10$, with enough runs to resolve it.

Exercise 6.8 *(Efficiency with overheads)* [⋆] A pricing run of n samples takes time $c_0 + c_1 n$, with a fixed overhead c_0.

(a) Show that the efficiency of a run of n samples, relative to the limit $n \to \infty$, is $n/(n + n_0)$ with $n_0 = c_0/c_1$. For which n is it 90%?

(b) In the course's timing experiment the relative efficiency at $n = 2$ was about 10^{-3}. Estimate n_0, and the n at which the efficiency reaches 90%.

(c) A new method reduces the standard error at equal n to 57% of the plain one and takes 1.6 times as long per sample. Compute its relative efficiency. Measured on runs of $n = 2^{10}$ samples with the overhead of (b), shared by both methods, what relative efficiency would you find instead?

Exercise 6.9 *(Quadrature in one and in many dimensions)* [⋆⋆] Let $f(x) = \frac{\pi}{2}\sin(\pi x)$ on $[0, 1]$, with $\int_0^1 f = 1$, and let $k \geq 1$ be the number of cells.

(a) Show that for any f with $f(0) = f(1)$ the left and the right Riemann sums coincide with the composite trapezoid rule.

(b) Using the identities

$$\sum_{i=0}^{k-1} \sin\big((2i+1)x\big) = \frac{\sin^2(kx)}{\sin x}, \qquad \sum_{i=0}^{k-1} \sin(2ix) = \frac{\sin\big((k-1)x\big)\sin(kx)}{\sin x},$$

show that with $x = \pi/(2k)$ the midpoint and trapezoid rules equal $x/\sin x$ and $x\cot x$. Deduce their errors to leading order in $1/k$.

(c) For $f_d(x) = \prod_{j=1}^{d} f(x_j)$ on $[0,1]^d$, write the product midpoint rule with $N = k^d$ points in closed form, and its error to leading order. Compute the variance of $f_d(U)$ for U uniform on $[0,1]^d$.

(d) Show that for $d = 4$ the two errors decrease at the same rate in N, and compare the constants. For $N \approx 10^9$, for which $d \in \{1, \dots, 8\}$ is Monte Carlo more accurate (in root mean square) than the product midpoint rule?

Chapter 7

Simulating paths

Exercise 7.1 *(Forward variances and calendar arbitrage)* [⋆] The at-the-money implied volatilities of an index are 20% for six months, 18% for one year and 15% for two years. Take $r = y = 0$ and the model $\mathrm{d}S_t/S_t = \sigma(t)\,\mathrm{d}W_t$ with a deterministic σ.

(a) Using Equation (7.2), find the forward volatilities $\bar\sigma(0, \frac{1}{2})$, $\bar\sigma(\frac{1}{2}, 1)$ and $\bar\sigma(1, 2)$ that reproduce the three quotes.

(b) Show that, in this model, $t \mapsto t\,\sigma_{0,t}^2$ is nondecreasing whatever σ is. Below which value of the two-year volatility can no deterministic σ reproduce the quotes?

(c) With $r = y = 0$ the price process is a martingale. Use the conditional Jensen inequality to show that, in *any* arbitrage-free model with $r = y = 0$, a call with fixed strike is worth at least as much at maturity 2 as at maturity 1. If the two-year volatility were quoted at 12%, compare the two Black–Scholes prices for $S_0 = K = 100$ and describe a trade that exploits the quotes.

Exercise 7.2 *(Bridges forget the drift)* [⋆] Let $X_t = \nu t + \sigma W_t$ with constants $\nu \in \mathbb{R}$, $\sigma > 0$, and $0 \le s < u < t$.

(a) Show that, given $(X_r)_{r \notin (s,t)}$, X_u is normal with mean $\frac{(t-u)X_s + (u-s)X_t}{t-s}$ and variance $\sigma^2 \frac{(u-s)(t-u)}{t-s}$; in particular this conditional law does not depend on ν. You may use Theorem 7.7.

(b) Let $S = S_0 e^X$ be a geometric Brownian motion. Compute $\mathbb{E}[S_u \mid S_s, S_t]$ and evaluate it for $s = 0$, $u = \frac{1}{2}$, $t = 1$, $S_0 = 100$, $S_1 = 110$, $\sigma = 20\%$.

(c) A colleague claims that the conditional expectation in (b) must depend on the interest rate, because S grows on average at the rate $r - y$. What is wrong with the claim?

Exercise 7.3 *(Euler for the Ornstein–Uhlenbeck process)* [⋆] Let $\mathrm{d}X_t = \kappa(\theta - X_t)\,\mathrm{d}t + \xi\,\mathrm{d}W_t$ with $\kappa, \xi > 0$, and let $\bar X_{k+1} = \bar X_k + \kappa(\theta - \bar X_k)h + \xi\,\Delta W_k$ be its Euler scheme with step h.

(a) Find the mean and the variance of $\bar X_k$ as functions of k. Show that the mean converges as

$k \to \infty$ for every starting point if and only if $0 < \kappa h < 2$.

(b) For $0 < \kappa h < 2$, find the limit $\bar{v}$ of $\operatorname{Var} \bar{X}_k$ and compare it with the stationary variance $\xi^2/(2\kappa)$ of X. Expand the ratio to first order in h.

(c) Evaluate the ratio for the Vasicek parameters $\kappa = 2$, $\xi = 2\%$ with monthly and with quarterly steps. What would the exact step of Theorem 7.6 give?

Exercise 7.4 *(Simulating backwards in time)* [⋆⋆] Let $0 < t_1 < \cdots < t_m$.

(a) Show that, given $(W_{t_{k+1}}, \dots, W_{t_m})$, the variable W_{t_k} is normal with mean $\frac{t_k}{t_{k+1}} W_{t_{k+1}}$ and variance $\frac{t_k(t_{k+1}-t_k)}{t_{k+1}}$.

(b) Write an algorithm that generates $(W_{t_1}, \dots, W_{t_m})$ starting from W_{t_m} and moving backwards, and show that it is the construction of Theorem 7.7(iii) for a particular order of the points.

(c) In backward induction (for instance for American options), the paths are needed at t_m, then t_{m-1}, and so on. Compare the memory needed by n paths when they are generated forwards and stored, and when they are generated backwards one date at a time. What must be stored to make the backward pass reproducible?

(d) For a GBM with constant parameters, write the backward step for $\log S_{t_k}$ given $\log S_{t_{k+1}}$.

Exercise 7.5 *(Euler and Milstein for GBM: exact moments)* [⋆⋆] Let $\bar{X}$ and $\tilde{X}$ be the Euler and Milstein schemes for $\mathrm{d}X = \mu X\,\mathrm{d}t + \sigma X\,\mathrm{d}W$, $X_0 = S_0$, with n steps of size $h = T/n$.

(a) Show that $\mathbb{E}\bar{X}_T = \mathbb{E}\tilde{X}_T = S_0(1+\mu h)^n$, $\mathbb{E}\bar{X}_T^2 = S_0^2\big((1+\mu h)^2 + \sigma^2 h\big)^n$ and $\mathbb{E}\tilde{X}_T^2 = S_0^2\big((1+\mu h)^2 + \sigma^2 h + \frac{1}{2}\sigma^4 h^2\big)^n$.

(b) Show that the relative errors of these three moments, compared with those of X_T, are $-\frac{1}{2}\mu^2 T h$, $-(\mu^2 + 2\mu\sigma^2 + \frac{1}{2}\sigma^4)Th$ and $-(\mu^2 + 2\mu\sigma^2)Th$, up to $O(h^2)$.

(c) With the test case ($S_0 = 100$, $\mu = r - y = -1\%$, $\sigma = 20\%$, $T = 1$), compute the error of $\mathbb{E}\bar{X}_T^2$ for $n = 12$ and $n = 24$, and the error of the Richardson combination $2\mathbb{E}\bar{X}^2_{T,(24)} - \mathbb{E}\bar{X}^2_{T,(12)}$.

(d) Show that one Euler step has a negative factor $1 + \mu h + \sigma \Delta W_k$ with probability $N\big(-(1+\mu h)/(\sigma\sqrt{h})\big)$. Evaluate it for $\sigma = 20\%$ with $h = \frac{1}{12}$ and $h = 1$, and for $\sigma = 100\%$ with $h = \frac{1}{4}$. What happens to a Monte Carlo estimate of $\mathbb{E}\log\bar{X}_T$?

Exercise 7.6 *(Richardson with the wrong exponent)* [⋆⋆] This exercise is partly computational. Let $M_n = \max_{0\le k\le n} W_{k/n}$, the maximum of a Brownian path on $[0,1]$ inspected at $n+1$ dates, and recall that $\mathbb{E}\max_{0\le t\le 1} W_t = \sqrt{2/\pi}$. You may use Spitzer's identity for random walks, $\mathbb{E}M_n = \sum_{k=1}^n \frac{1}{k}\mathbb{E}[W^+_{k/n}]$, and the expansion $\sum_{k=1}^n k^{-1/2} = 2\sqrt{n} + \zeta(\frac{1}{2}) + \frac{1}{2\sqrt{n}} + O(n^{-3/2})$, where $\zeta(\frac{1}{2}) \approx -1.46035$.

(a) Show that $\mathbb{E}M_n = (2\pi n)^{-1/2}\sum_{k=1}^n k^{-1/2}$, and deduce an expansion of $\mathbb{E}M_n - \sqrt{2/\pi}$ in powers of $h^{1/2}$, $h = 1/n$, up to $O(h^2)$.

(b) Apply Richardson extrapolation Theorem 7.13 to $\mathbb{E}M_{64}$ and $\mathbb{E}M_{32}$ with $p = 1$ and with $p = \frac{1}{2}$. Which exponent is right, and what is the order of the remaining error in each case?

(c) Estimate $\mathbb{E}M_{16}$, $\mathbb{E}M_{32}$ and $\mathbb{E}M_{64}$ by simulation, all three on the same 10^6 paths, and compare with (a). Then extrapolate your estimates with $p = \frac{1}{2}$.

Exercise 7.7 *(Allocating effort with Milstein and extrapolation)* [⋆⋆] This is a computational exercise. Take GBM with the test-case parameters and the payoff $f(x) = 5\log(1 + e^{(x-100)/5})$ of Figure 7.2, whose variance is $v = \operatorname{Var} f(X_T) \approx 154$. Count the cost of a path with m steps as m, and that of a coupled pair for Richardson extrapolation as $1.5\,m$.

(a) Estimate the weak error constant c of the Milstein scheme, $\mathbb{E}f(\tilde{X}^h_T) - \mathbb{E}f(X_T) \approx c\,h$, and the constant c_2 of the coupled Richardson estimator, $\mathbb{E}R^h - \mathbb{E}f(X_T) \approx c_2 h^2$, by simulation on common paths.

(b) Using Theorem 7.14, find the optimal number of steps and paths and the root-mean-square error for Milstein ($\beta = 1$) and for Richardson-extrapolated Milstein ($\beta = 2$), with budgets 10^6 and 10^8. Compare with Euler and with exact simulation, as in the notes.

(c) For the budget 10^6, run 100 independent replications of the extrapolated estimator with the allocation from (b) (rounded to an even number of fine steps), and compare its root-mean-square error with the prediction. Explain any discrepancy.

Exercise 7.8 *(The principal components of Brownian motion)* [⋆⋆⋆] Let $t_i = ih$, $i = 1, \dots, m$, $h = T/m$, and $C_{ij} = \min(t_i, t_j) = h\min(i, j)$.

(a) Show that $C^{-1} = h^{-1}J$, where J is tridiagonal with $J_{ii} = 2$ for $i < m$, $J_{mm} = 1$ and $J_{i,i\pm1} = -1$.

(b) Show that $v_i^{(k)} = \sin\big((2k-1)i\pi/(2m+1)\big)$, $k = 1, \dots, m$, are eigenvectors of J, and deduce the eigenvalues of C: $\lambda_k = \dfrac{h}{4\sin^2\big((2k-1)\pi/(2(2m+1))\big)}$.

(c) Show that the share of the total variance carried by the first principal component, $\lambda_1/\operatorname{tr} C$, tends to $8/\pi^2$ as $m \to \infty$, and compute it for $m = 64$. Compare with the first normal of the Brownian bridge construction.

(d) Let $T = 1$ and fix k. Show that $\lambda_k/m \to 1/\big((k-\frac{1}{2})^2\pi^2\big)$ and that $v_i^{(k)} = \sin\big((k-\frac{1}{2})\pi t_i\big) + O(1/m)$ uniformly in i. Use this to explain, without proving convergence, the expansion $W_t = \sum_{k\ge1} \frac{\sqrt{2}\,\sin((k-\frac{1}{2})\pi t)}{(k-\frac{1}{2})\pi} Z_k$ of a Brownian path on $[0,1]$, and check that its variance at $t = 1$ is 1, using $\sum_{k\ge1}(2k-1)^{-2} = \pi^2/8$.

Exercise 7.9 *(The strong error of Euler for GBM)* [⋆⋆⋆] This exercise is partly computational. Let $\bar{X}$ be the Euler scheme for $\mathrm{d}X = \mu X\,\mathrm{d}t + \sigma X\,\mathrm{d}W$ with n steps of size $h = T/n$, driven by the same Brownian motion as X.

(a) Expand $\log(1 + \mu h + \sigma\Delta W_k)$ to show that $\log \bar{X}_T - \log X_T = -\frac{1}{2}\sigma^2 \sum_k \big((\Delta W_k)^2 - h\big) + R_h$. You may take for granted that $h^{-1/2}R_h \to 0$ in probability.

(b) Show that $\big(W_T, (2Th)^{-1/2}\sum_k((\Delta W_k)^2 - h)\big)$ converges in distribution to (W_T, N) with $N \sim \mathcal{N}(0,1)$ independent of W_T.

(c) Deduce that $h^{-1/2}(\bar{X}_T - X_T)$ converges in distribution to $-\sigma^2\sqrt{T/2}\,X_T N$, and, assuming uniform integrability, that $\mathbb{E}|\bar{X}_T - X_T| = \sigma^2\,\mathbb{E}[X_T]\sqrt{Th/\pi} + o(h^{1/2})$.

(d) Compare the constant with the strong errors measured in Figure 7.2 (test case) by your own simulation. Why does the same argument give strong order 1 for the Milstein scheme?

Chapter 8

Antithetic variates, moment matching and control variates

Exercise 8.1 *(Antithetic digitals)* [⋆] Let $D = \mathbf{1}_{\{S_T > K\}}$ be a digital call in the Black–Scholes model, $p = \mathbb{Q}(S_T > K)$, and ρ the correlation of D computed from Z and from $-Z$. You may use Theorem 8.2 and Theorem 8.4.

(a) For K at or above the median $Se^{\nu T}$ of S_T, show that $\rho = -p/(1-p)$, and give the standard-error ratio of n antithetic pairs to n plain draws and the efficiency at twice the cost.

(b) For K below the median, compute $\mathbb{E}[D(Z)D(-Z)]$ and show that $\rho = -(1-p)/p$. Conclude that $\rho = -\min(p, 1-p)/\max(p, 1-p)$ for every K.

(c) What happens at $K = Se^{\nu T}$? Explain without any formula why the antithetic estimator is then exact.

(d) Evaluate p, ρ and the efficiency for $K = 90$ and $K = 110$ in the test case.

Exercise 8.2 *(When pairing hurts)* [⋆⋆] This is a computational exercise. Consider the discounted straddle $f(Z) = e^{-rT}|S_T - K|$.

(a) Let $K = Se^{\nu T}$, the median of S_T, and $s = \sigma\sqrt{T}$. Show that $e^{rT}f(Z)/(Ks) \to |Z|$ in L^2 as $s \to 0$, and deduce that the antithetic correlation tends to 1. (You may use that correlations are continuous under L^2 convergence when the limits have positive variance.)

(b) With 10^5 pairs, estimate the antithetic correlation $\rho(K)$ for strikes $K \in [70, 130]$ and find the range of strikes on which antithetic sampling is worse than plain sampling with the same number of payoff evaluations. Explain the result with the even–odd decomposition of Theorem 8.2.

(c) At $K = 100$, use instead the control $X = e^{-rT}(S_T - K)^2$. Compute its mean in closed form, and compare the standard errors of: plain Monte Carlo with $2n$ draws; n antithetic pairs; the control variate with estimated coefficient on $2n$ draws; and the control variate applied to the n antithetic pair averages.

Exercise 8.3 *(Moment matching is a control variate)* [⋆⋆] Let $Z_1, \dots, Z_n$ be i.i.d. $\mathcal{N}(0,1)$ and $Z_i' = (Z_i - \bar{Z}_n)/\tilde{s}_n$ the log-matched sample of Theorem 8.6. Take $f(z) = z^3$, so that $\hat{\mu}_n^{\mathrm{mm}} = \frac{1}{n}\sum_i Z_i'^3$ estimates $\mathbb{E}Z^3 = 0$. You may use the strong law, the central limit theorem and Slutsky's lemma.

(a) Show that $\hat{\mu}_n^{\mathrm{mm}}$ is exactly unbiased.

(b) Show that $\sqrt{n}\,\hat{\mu}_n^{\mathrm{mm}} \Rightarrow \mathcal{N}(0,6)$. *Hint:* expand $\frac{1}{n}\sum_i (Z_i - \bar{Z}_n)^3$.

(c) Compare with the plain estimator $\frac{1}{n}\sum_i Z_i^3$ and with the control-variate estimator using the control Z with its optimal coefficient. Why does the second matched moment, that is the control Z^2, not help here?

Exercise 8.4 *(How many batches?)* [⋆⋆] This is a computational exercise. Estimate the test-case call by log moment matching with $n = 4096$ draws split into $m \in \{4, 16, 64, 256, 1024\}$ separately matched batches, and form the nominal 95% confidence interval from the batch means with the Student quantile on $m - 1$ degrees of freedom.

(a) For each m, repeat the experiment 2000 times and report the bias, the standard deviation of the estimator, the average estimated standard error and the coverage.

(b) Explain the results with Theorem 8.7 and Theorem 8.8. Predict, without simulating, the coverage for $m = 128$.

(c) Give a rule for choosing m as a function of n.

Exercise 8.5 *(Comparing a call and a put)* [⋆⋆] Let $C = e^{-rT}(S_T - K)^+$ and $P = e^{-rT}(K - S_T)^+$, both computed from the same Z, and suppose we estimate $\mathbb{E}C - \mathbb{E}P$ by simulation.

(a) Show that $\mathrm{Cov}(C, P) = -\mathbb{E}[C]\,\mathbb{E}[P]$, and deduce that common random numbers *increase* the variance of the estimated difference compared with independent draws.

(b) Using $S_T(-z) = (Se^{\nu T})^2/S_T(z)$, write $P(-Z)$ as a payoff in S_T and explain why it is strongly correlated with C. Evaluate, for the test case, the variance per draw of the estimated difference with independent draws, with common random numbers, and with $C(Z_i)$ paired with $P(-Z_i)$ (use closed forms or quadrature).

(c) What is the best estimator of $\mathbb{E}C - \mathbb{E}P$? What does this suggest about using the put as a control variate for the call?

Exercise 8.6 *(A fixed coefficient)* [⋆] Let $X, Y \in L^2$ with $\mathrm{Var}\,X > 0$, and let $\beta^* = \mathrm{Cov}(X,Y)/\mathrm{Var}\,X$.

(a) Show that $\mathrm{Var}(Y - \beta X) < \mathrm{Var}\,Y$ if and only if β lies strictly between 0 and $2\beta^*$.

(b) Take $\beta = 1$ and $\mathrm{Var}\,X = \mathrm{Var}\,Y$. Show that the variance factor is $2(1-\rho)$, that the control helps only if $\rho > \frac{1}{2}$, and that if the control doubles the cost per sample it improves efficiency only if $\rho > \frac{3}{4}$.

(c) Let $\beta = \beta^*(1+\varepsilon)$, a coefficient with relative error ε. Show that the variance factor $\mathrm{Var}(Y - \beta X)/\mathrm{Var}\,Y$ is $1 - \rho^2 + \varepsilon^2\rho^2$. For $\rho = 0.9$ and $\rho = 0.999$, compare $\varepsilon = 0$ with

$\varepsilon = 0.2$, and comment on how accurately the coefficient must be known when the correlation is high.

Exercise 8.7 *(The price of estimating the coefficient)* [⋆⋆] Let (X_i, Y_i), $i \le n$, be i.i.d. bivariate normal with correlation ρ and $\operatorname{Var} Y = \sigma_Y^2$, let $\mu_X = \mathbb{E}X$ be known, and let $\hat{\mu}_n^{\mathrm{cv}}$ be the control-variate estimator with the least-squares coefficient, as in Theorem 8.14. You may use: $\bar{X}_n$ and $\sum_i (X_i - \bar{X}_n)^2$ are independent, $\sum_i (X_i - \bar{X}_n)^2 / \operatorname{Var} X \sim \chi^2_{n-1}$, and $\mathbb{E}[1/\chi^2_k] = 1/(k-2)$ for $k > 2$.

(a) Show that, for $n \ge 4$,
$$\operatorname{Var} \hat{\mu}_n^{\mathrm{cv}} = \frac{\sigma_Y^2 (1 - \rho^2)}{n} \, \frac{n-2}{n-3}.$$

(b) For $\rho = 0.5$, for which n does the estimator beat the plain sample mean? And for $\rho = 0.1$?

(c) Discuss what this suggests when the control-variate regression is run on many small batches.

Exercise 8.8 *(The asymptotic square-root call)* [⋆⋆] This is a computational exercise. The *asymptotic square-root call* pays $F(S_T) = 2K\big[1 - \sqrt{K/S_T}\,\big]^+$. You may use its Black–Scholes price
$$e^{-rT} 2K \Big[N(d_-) - \sqrt{K/S}\, e^{-\nu T/2 + \sigma^2 T/8}\, N\big(d_- - \tfrac{1}{2}\sigma\sqrt{T}\big) \Big].$$

(a) Show that F vanishes for $S_T \le K$ and that its right derivative at the strike is $F'(K+) = 1$, and explain why the at-the-money call $(S_T - K)^+$ is a natural control. What does the replication of F by calls say about the coefficient?

(b) With 10^5 draws, estimate the price by plain Monte Carlo, and by control variates with the discounted call as control, with $\beta = 1$ and with the estimated coefficient. Test each estimate against the closed form.

(c) Compute the correlation and β^* (by quadrature or with many draws), and the efficiencies of the two control-variate estimators at twice the cost. Why is the estimated coefficient so much better than the replication coefficient here?

Chapter 9

Importance sampling

Throughout, $\mathbb{Q}$ is the target law, $\widetilde{\mathbb{Q}}$ the sampling law, $\widetilde{\mathbb{E}}$ its expectation, and $L = q/\tilde{q}$ the weight, as in Section 9.1. You may use the results of the notes, quoting them.

Exercise 9.1 *(Higher moments of the Gaussian weight)* [⋆] Let the target be $\mathcal{N}(0,1)$ and the sampling law $\mathcal{N}(c,\beta^2)$, with $\beta > 0$, and let $n \geq 2$ be an integer.

(a) Show that $\widetilde{\mathbb{E}}[L^n] < \infty$ if and only if $\beta^2 > 1 - 1/n$, and that then

$$\widetilde{\mathbb{E}}[L^n] = \frac{\beta^n}{\sqrt{\gamma_n}} \exp\Big(\frac{n(n-1)c^2}{2\gamma_n}\Big), \qquad \gamma_n = n\beta^2 - n + 1.$$

Check that $n = 2$ gives Theorem 9.3(i).

(b) Deduce that every moment of the weight is finite when $\beta \geq 1$, and find the range of β for which the sample variance of the weights itself has finite variance.

(c) Fix $c \neq 0$. Find the scale β^* that minimises $\widetilde{\mathbb{E}}[L^2]$, show that $\beta^* > 1$, and compute β^* and $\widetilde{\mathbb{E}}[L^2]$ at β^* and at $\beta = 1$ for $c = 1$. What does this say about pairing a mean shift with a scale?

Exercise 9.2 *(Exponential tilts of common laws)* [⋆] For each law, find the cumulant generating function ψ, the interval where it is finite, and the tilted law (Theorem 9.6):

(a) the exponential law with rate λ, and the gamma law with shape k and rate λ;

(b) the Bernoulli law with parameter p; write the weight of the tilt of m independent Bernoulli variables as a function of their sum, and identify the law of the sum under the tilt;

(c) the lognormal law of $S_T = S\exp(\nu T + \sigma\sqrt{T}Z)$, $Z \sim \mathcal{N}(0,1)$, $\sigma > 0$: show that $\mathbb{E}[e^{\theta S_T}] = \infty$ for every $\theta > 0$. *Hint:* compare $e^{\theta S_T}$ with $(\theta S_T)^j/j!$ and use $\mathbb{E}[S_T^j]$. Explain why the course tilts $\log S_T$ rather than S_T.

Exercise 9.3 *(An exponential tail)* [⋆⋆] Let X be exponential with rate 1 and $p_t = \mathbb{Q}(X > t) = e^{-t}$, $t > 0$. Estimate p_t by sampling X from the exponential law with rate $\lambda > 0$.

(a) Compute the second moment $M_t(\lambda)$ of the estimator, and show that it is infinite for $\lambda \geq 2$. Which sampling laws are "thinner-tailed" than the target?

(b) Find the rate λ_t^* that minimises M_t, and show that $t\lambda_t^* \to 1$: the sampling mean is placed at the threshold.

(c) Show that $M_t(\lambda_t^*)/p_t^2 \sim et/2$ as $t \to \infty$. Is the estimator logarithmically efficient? Does it have bounded relative error (Theorem 9.13)?

(d) Now let $S = X_1 + X_2$ with X_1, X_2 independent exponential with rate 1, so that $\mathbb{Q}(S > t) = (1+t)e^{-t}$. Tilt both summands to rate $\lambda = 2/t$, which puts the tilted mean of S at t, and show that the relative second moment of the estimator of $\mathbb{Q}(S > t)$ is asymptotic to $e^2 t/8$.

Exercise 9.4 *(Self-normalised importance sampling)* [⋆⋆] Let $\tilde{q} > 0$ wherever $q > 0$, let $X_1, X_2, \ldots$ be independent with density $\tilde{q}$, and assume $\widetilde{\mathbb{E}}[L^2] < \infty$ and $\widetilde{\mathbb{E}}[f^2 L^2] < \infty$. Let $\hat{\mu}_n^{\mathrm{sn}} = \sum_i f(X_i)L(X_i) / \sum_i L(X_i)$.

(a) Show that $\hat{\mu}_n^{\mathrm{sn}} \to \mu$ almost surely, and that $\hat{\mu}_n^{\mathrm{sn}}$ is unchanged if L is multiplied by a constant. Why does this matter outside pricing?

(b) Using Theorem 6.12(i), show that $\sqrt{n}(\hat{\mu}_n^{\mathrm{sn}} - \mu) \Rightarrow \mathcal{N}(0, \widetilde{\mathbb{E}}[L^2(f-\mu)^2])$.

(c) Let $f = \mathbf{1}_A$ with $p = \mathbb{Q}(A)$, and write $M = \widetilde{\mathbb{E}}[L^2 \mathbf{1}_A]$ and $W = \widetilde{\mathbb{E}}[L^2]$. Show that the self-normalised estimator has the smaller asymptotic variance if and only if $2M > p(1+W)$. Evaluate both variances for the target $\mathcal{N}(0,1)$, the sampling law $\mathcal{N}(1,1)$ and $A = \{Z > 1\}$, and explain why self-normalising is a bad idea for rare events.

(d) Assuming that the expansion of Theorem 6.12(iii) applies (it requires a weight bounded away from zero, which is not the case here), show that the bias of $\hat{\mu}_n^{\mathrm{sn}}$ is $(\mu W - \widetilde{\mathbb{E}}[fL^2])/n + O(n^{-3/2})$.

Exercise 9.5 *(Deep out-of-the-money calls)* [⋆⋆⋆] Fix $s > 0$ and, for $u > 0$, let $G_u(z) = (e^{s(z-u)} - 1)^+$: the payoff of a call, divided by its strike, when $S_T = Ke^{s(Z-u)}$ and $Z \sim \mathcal{N}(0,1)$. Let $d \to \infty$.

(a) Show that for $k \in \{1, 2\}$, $\mathbb{E}[G_u(Z)^k] = k!\, s^k N'(u)\, u^{-k-1}(1 + o(1))$ as $u \to \infty$. *Hint:* substitute $z = u + w/u$, use $N'(u + w/u) = N'(u)e^{-w - w^2/(2u^2)}$, and justify the limit by dominated convergence for $u \geq 2ks$.

(b) Deduce that plain Monte Carlo for $p_d = \mathbb{E}[G_d(Z)]$ has $\mathbb{E}[G_d^2]/p_d^2 \sim 2\sqrt{2\pi}\, d\, e^{d^2/2}$.

(c) With the mean shift to the strike, $Z \sim \mathcal{N}(d, 1)$ under $\widetilde{\mathbb{Q}}$, use the mirror formula Equation (9.4) to show that $\widetilde{\mathbb{E}}[Y^2]/p_d^2 \sim \sqrt{\pi/8}\, d$, where Y is the weighted payoff. Compare with the digital in Theorem 9.15.

(d) Explain in words why neither limit depends on s.

Exercise 9.6 *(Credit losses by exponential tilting)* [⋆⋆] *Computational.* A portfolio has $m = 100$ independent obligors, each of which defaults over the horizon with probability 0.01 and then loses one unit. Let N be the number of defaults and $p_\ell = \mathbb{Q}(N \geq \ell)$.

(a) Compute p_ℓ exactly for $\ell = 5, 10, 15, 20$, and the relative error per sample of plain Monte Carlo.

(b) Find the tilt θ_ℓ of the default indicators for which the tilted mean of N is ℓ, the default probability under the tilt, and the bound $e^{-\theta_\ell \ell + m\psi(\theta_\ell)}$ on the estimator (Theorem 9.7).

(c) Implement the tilted estimator with 10^5 samples for each ℓ, report estimates with standard errors and relative errors per sample, and compare with (a). State your seed.

(d) How do the relative errors behave as ℓ grows? With m fixed, the asymptotics in m of Theorem 9.7 do not apply as such; instead compare the ratio of the bound in (b) to p_ℓ with the lattice Bahadur–Rao prefactor $\sigma_\ell \sqrt{2\pi m}\,(1 - e^{-\theta_\ell})$, where σ_ℓ^2 is the variance of one tilted indicator.

Exercise 9.7 *(The GHS drift for a geometric Asian call)* [⋆⋆⋆] *Partly computational.* In the test case of the notes, with $m = 16$ fixings $t_j = j/16$, the geometric average is $G = \exp\left(\frac{1}{m}\sum_j \log S_{t_j}\right)$ with $S_{t_j} = S\exp\left(\nu t_j + \sigma\sqrt{\Delta t}\sum_{i\le j} Z_i\right)$, $\Delta t = 1/m$.

(a) Show that $\log G = \kappa + b^\top Z$ with $\kappa = \log S + \nu\bar{t}$, $\bar{t} = \frac{1}{m}\sum_j t_j$, and $b_i = \sigma\sqrt{\Delta t}\,(m - i + 1)/m$, and compute $|b|^2$.

(b) Show that the GHS drift (Theorem 9.10) for the payoff $(G - K)^+$ is $\mu = \lambda b$, where λ is the unique root of $\lambda(e^{\kappa+\lambda|b|^2} - K) = e^{\kappa+\lambda|b|^2}$ with $\kappa + \lambda|b|^2 > \log K$. Why is the problem one-dimensional?

(c) Using Theorem 9.10(iii), give an upper bound on the relative second moment of the importance-sampling estimator, and compute it for $K = 100, 140, 160$.

(d) Implement the estimator with 10^5 paths for these strikes; compare with the exact price of the geometric Asian call (the law of $\log G$ is normal) and compute the exact variance reduction factors with the mirror formula. Compare your drift with the arithmetic drift of Theorem 9.12.

Exercise 9.8 *(A first-passage probability)* [⋆ ⋆ ⋆] Let $\tilde{W}$ be a Brownian motion under $\mathbb{Q}$, $b, T > 0$, $x = b/\sqrt{T}$, and $p = \mathbb{Q}(\max_{t\le T}\tilde{W}_t \ge b)$, which equals $2N(-x)$ by the reflection principle. Under $\widetilde{\mathbb{Q}}$, let $\tilde{W}_t = B_t + \theta t$ with B a $\widetilde{\mathbb{Q}}$-Brownian motion and $\theta \in \mathbb{R}$ a constant, and let $\tau = \inf\{t : \tilde{W}_t = b\}$. You may use Theorem 9.8 and the optional sampling theorem for bounded stopping times.

(a) Show that $p = \widetilde{\mathbb{E}}[\mathbf{1}_{\{\tau\le T\}}\exp(-\theta b + \theta^2\tau/2)]$: the weight can be evaluated at the hitting time instead of at T.

(b) For $\theta = b/T$, show that the estimator is bounded by $e^{-x^2/2}$ surely.

(c) Deduce that its relative second moment is at most $\sqrt{\pi/2}\,(1 + x^2)/x$, and conclude that it is logarithmically efficient as $x \to \infty$.

(d) Show that using the weight at T instead, $\exp(-\theta\tilde{W}_T + \theta^2 T/2)$, gives another unbiased estimator. Which is better, and why? (Think of conditional expectation.)

Chapter 10

Stratified sampling and conditional Monte Carlo

Exercise 10.1 *(Stratifying a digital)* [⋆] Let $D = e^{-rT}\mathbf{1}_{\{S_T>K\}}$ in the Black–Scholes model, $z^* = (\ln(K/S) - \nu T)/(\sigma\sqrt{T})$, $u^* = N(z^*)$ and $p = 1 - u^* = \mathbb{Q}(S_T > K)$. Stratify $U = N(Z)$ into H equiprobable strata with proportional allocation, n a multiple of H. You may use Theorem 10.1 and Theorem 10.3.

(a) Show that at most one stratum has $\sigma_h > 0$, and that $n \operatorname{Var} \hat{\mu}_n^{\mathrm{st}} = e^{-2rT}\vartheta(1-\vartheta)/H$, where $\vartheta \in [0,1]$ is the fraction of that stratum lying above u^*. Deduce $n \operatorname{Var} \hat{\mu}_n^{\mathrm{st}} \le e^{-2rT}/(4H)$ and compare with plain sampling.

(b) With $n \ge H$ and at least one draw in every stratum, what allocation minimises the variance, and what is the minimal variance?

(c) When is the stratified estimator exact? What does this suggest about where to put the edges of the strata for a payoff with a jump?

(d) For $K = 110$, $H = 10$ and $n = 1000$, give the standard-error ratios to plain sampling of the estimators in (a) and (b).

Exercise 10.2 *(Allocation with costs)* [⋆] Suppose that a draw in stratum h costs $c_h > 0$ (for instance, stratum h is sampled by acceptance–rejection with acceptance probability a_h, so that c_h is proportional to $1/a_h$), and that the budget is $C = \sum_h c_h n_h$.

(a) Treating the $n_h > 0$ as real numbers, show that $\sum_h p_h^2 \sigma_h^2 / n_h$ is minimised under the budget constraint by n_h proportional to $p_h\sigma_h/\sqrt{c_h}$, and that the minimal variance is $\big(\sum_h p_h \sigma_h \sqrt{c_h}\big)^2 / C$.

(b) A plain draw is obtained by first drawing the stratum with probabilities p_h and then sampling in it, so it costs $\bar{c} = \sum_h p_h c_h$ on average. Show that the efficiency of the cost-optimal stratified estimator relative to plain sampling is at least 1.

(c) Take two strata with $p = (0.9, 0.1)$, $\sigma_h = (1, 10)$, $c_h = (1, 25)$ and $\mu_1 = \mu_2$. Compare the products of variance and budget of plain sampling, proportional allocation, the Neyman

allocation of Theorem 10.3 (which ignores costs) and the cost-optimal allocation.

Exercise 10.3 *(One point per stratum)* [⋆⋆] Let $g : [0,1] \to \mathbb{R}$ and $\hat{\mu}_n = \frac{1}{n}\sum_{k=1}^{n} g((k-1+U_k)/n)$ with $U_1, \dots, U_n$ i.i.d. uniform: n equiprobable strata and one point in each. You may use Theorem 10.1.

(a) Let g be continuously differentiable on $[0,1]$. Show that $n^3 \operatorname{Var} \hat{\mu}_n \to \frac{1}{12}\int_0^1 g'(u)^2 \, \mathrm{d}u$. *Hint:* compare $g((k-1+U)/n)$ with its first-order expansion around $(k-1)/n$, and use the uniform continuity of g'.

(b) Deduce the rate of the root-mean-square error, and compare it with plain Monte Carlo and with Theorem 10.1(v).

(c) For the discounted test-case call as a function of the uniform, $g(u) = e^{-rT}(S_T(N^{-1}(u)) - K)^+$, show that $\int_0^1 g'(u)^2 \, \mathrm{d}u = \infty$.

(d) This part is computational. For that call, compute $n^2 \operatorname{Var} \hat{\mu}_n$ and $n^3 \operatorname{Var} \hat{\mu}_n$ for $n = 2^4, 2^6, \dots, 2^{16}$ (exactly, from the stratum variances, or by simulation with independent replications), and describe the rate you observe. Which strata dominate the variance?

Exercise 10.4 *(Collapsed strata)* [⋆⋆] With one draw per stratum no within-stratum variance can be estimated (Theorem 10.4). Let $H = 2J$ equiprobable strata of a uniform U receive one draw Y_h each, so that $\hat{\mu}^{\mathrm{st}} = \frac{1}{H}\sum_h Y_h$, and pair the strata as $(1,2), (3,4), \dots$.

(a) Show that $\hat{v} = \frac{1}{H^2}\sum_{j=1}^{J}(Y_{2j-1} - Y_{2j})^2$ satisfies $\mathbb{E}\hat{v} = \operatorname{Var} \hat{\mu}^{\mathrm{st}} + \frac{1}{H^2}\sum_j (\mu_{2j-1} - \mu_{2j})^2$: it is a conservative estimate of the variance.

(b) Let $Y = g(U)$ with g continuously differentiable on $[0,1]$. Using the expansion of Exercise 10.3(a), show that $\mathbb{E}\hat{v} / \operatorname{Var} \hat{\mu}^{\mathrm{st}} \to 7$ as $H \to \infty$, unless g is constant.

(c) The alternative of Theorem 10.4(iii) splits the H draws into R independent designs of H/R strata each. For the same g, compare the variance of the resulting estimator with that of the single design. Which way of obtaining an error estimate do you prefer, and why?

Exercise 10.5 *(Post-stratification is a regression)* [⋆⋆] Let (ξ_i, Y_i), $i \le n$, be i.i.d., with strata $A_1, \dots, A_H$ of known probabilities $p_h > 0$, and let $D_i^h = \mathbf{1}_{A_h}(\xi_i)$ and $N_h = \sum_i D_i^h$. You may use Theorem 8.15, Theorem 8.14 and Theorem 10.6.

(a) Why can the H indicators not all be used as controls together with an intercept? Use the $H-1$ controls $D^1, \dots, D^{H-1}$, with known means $p_1, \dots, p_{H-1}$.

(b) Show that, when every $N_h \ge 1$, the regression estimator of Theorem 8.15(iii) with these controls equals the post-stratified estimator $\sum_h p_h \bar{Y}_h$.

(c) Compute the R^2 of Theorem 8.15 for these controls, and check that $\sigma^2(1 - R^2)$ is the asymptotic variance of Theorem 10.6(iv).

(d) An estimated control coefficient usually causes a bias of order $1/n$ (Theorem 8.14). Why is the bias here exponentially small?

Exercise 10.6 *(Neyman allocation for a deep out-of-the-money call)* [⋆⋆] This is a computa-

tional exercise. Take the course's deep out-of-the-money call $S = 100$, $K = 150$, $T = 1$, $r = 5\%$, $y = 0$, $\sigma = 20\%$, with $n = 10^5$ draws and H equiprobable strata of $U = N(Z)$.

(a) For $H = 100$, how many strata have $\mu_h = \sigma_h = 0$? Explain why they can be left out of the simulation altogether, and write the resulting estimator.

(b) Compute the σ_h of the other strata and the Neyman allocation of the n draws among them. Give the variance factor $n \operatorname{Var}/\sigma^2$, and compare it with proportional allocation.

(c) Repeat for $H = 10$ and $H = 1000$, and explain the case $H = 10$ in one line.

(d) Confirm (b) by simulation, with 100 repetitions as in the course.

(e) Compare with importance sampling by a mean shift: draw $\tilde{Z} = Z + z^*$, with z^* the value of Z at which $S_T = K$, and average $f(\tilde{Z})e^{-z^*\tilde{Z}+(z^*)^2/2}$, where $f(z)$ is the discounted payoff (you may use that this estimator is unbiased). Give its variance factor at equal n, and comment.

Exercise 10.7 *(Latin hypercube sampling can lose)* [⋆⋆⋆] Let $V_1, \dots, V_n$ be a Latin hypercube sample (Theorem 10.9) and $f \in L^2([0,1]^d)$ with variance σ^2.

(a) Show that $\operatorname{Var} \hat{\mu}_n^{\mathrm{LHS}} = \frac{\sigma^2}{n} + (1 - \frac{1}{n}) \operatorname{Cov}\big(f(V_1), f(V_2)\big)$.

(b) Let $d = 2$ and $f(u) = (u_1 - \frac{1}{2})(u_2 - \frac{1}{2})$. Show that $\mu = 0$, $\sigma^2 = \frac{1}{144}$, and $\operatorname{Cov}(f(V_1), f(V_2)) = c_n^2$ with $c_n = \mathbb{E}\big[(V_{11} - \frac{1}{2})(V_{21} - \frac{1}{2})\big]$. Show that $c_n = -\frac{n+1}{12n^2}$. *Hint:* V_{11} and V_{21} lie in two distinct strata of the first coordinate, chosen uniformly at random.

(c) Deduce that $n \operatorname{Var} \hat{\mu}_n^{\mathrm{LHS}}/\sigma^2 = 1 + (n-1)(n+1)^2/n^4 > 1$, so that Latin hypercube sampling is worse than plain sampling for this f, and check the bound $\sigma^2/(n-1)$ of Theorem 10.10(iv). How close to it is this example?

(d) Reconcile (c) with Theorem 10.10(iii).

(e) Check (c) by simulation for $n = 4$ and $n = 32$.

Exercise 10.8 *(Conditioning a spread option)* [⋆⋆] This is a computational exercise. Under $\mathbb{Q}$ let $S_T^i = S^i \exp\big((r - y_i - \frac{1}{2}\sigma_i^2)T + \sigma_i\sqrt{T}Z_i\big)$, $i = 1, 2$, where (Z_1, Z_2) is a standard normal pair with correlation $\rho \in (-1, 1)$, and let $Y = e^{-rT}(S_T^1 - S_T^2 - K)^+$ with $K \geq 0$. You may use the Black–Scholes formulas and Theorem 10.11.

(a) Writing $Z_1 = \rho Z_2 + \sqrt{1-\rho^2}\, Z^\perp$ with $Z^\perp$ standard normal and independent of Z_2, show that $\mathbb{E}[Y \mid Z_2]$ is a Black–Scholes call price, and identify its spot, strike and volatility.

(b) Take $S^1 = S^2 = 100$, $K = 5$, $T = 1$, $r = 1\%$, $y_1 = y_2 = 2\%$, $\sigma_1 = 20\%$, $\sigma_2 = 30\%$, $\rho = 0.5$. With $n = 10^5$ draws, compare the plain and the conditional estimators: estimates, standard errors and the variance ratio.

(c) Condition instead on Z_1: show that $\mathbb{E}[Y \mid Z_1]$ is a put on S_T^2, and compare the two conditional estimators. Explain which one removes more variance.

(d) Stratify the conditioning variable of your better estimator into 100 equiprobable strata. Which results of the chapter predict what you observe?

Exercise 10.9 *(Hull and White's expansion)* [⋆⋆⋆] Let $\rho = 0$ in Theorem 10.12, with the Heston variance $\mathrm{d}V_t = \kappa(\theta - V_t)\,\mathrm{d}t + \xi\sqrt{V_t}\,\mathrm{d}W_t$ started at $V_0 = \theta$, and let $v = Q_T/T$ be the realised variance. Write $P(v)$ for the Black–Scholes put price as a function of the variance $v = \sigma^2$, the other inputs being those of the test case. You may use that $\operatorname{Var} V_s = \frac{\theta\xi^2}{2\kappa}(1 - e^{-2\kappa s})$ and $\operatorname{Cov}(V_s, V_t) = e^{-\kappa(t-s)} \operatorname{Var} V_s$ for $s \le t$.

(a) Show that $\mathbb{E}v = \theta$ and $\operatorname{Var}(Q_T) = \frac{\theta\xi^2}{\kappa^2}\Big(T - \frac{1-e^{-2\kappa T}}{2\kappa} - \frac{(1-e^{-\kappa T})^2}{\kappa}\Big)$.

(b) With $\mathcal{V} = Se^{-yT}N'(d_+)\sqrt{T}$ the Black–Scholes vega at volatility $\sqrt{v}$, show that $P''(v) = \mathcal{V}\,(d_+d_- - 1)/(4v^{3/2})$. Deduce the approximation $\mathbb{E}[P(v)] \approx P(\theta) + \frac{1}{2}P''(\theta)\operatorname{Var} v$, and explain the sign of the correction at the money.

(c) Evaluate the approximation for $\kappa = 2$, $\theta = 0.04$, $\xi = 0.3$, and compare it with the Fourier price 8.1534 and with your own conditional Monte Carlo estimate (exact variance steps as in the chapter, or any scheme you trust).

(d) Repeat (c) with $\xi = 0.6$ and $\xi = 0.9$, using conditional Monte Carlo as the reference. Where does the expansion break down, and why?

Chapter 11

Quasi-Monte Carlo

Exercise 11.1 *(Discrepancy in one dimension)* [⋆] Let $0 \le x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)} < 1$ be n points sorted in increasing order.

(a) Prove the formula $D_n^* = \max_{1\le i\le n} \max\left(\frac{i}{n} - x_{(i)},\, x_{(i)} - \frac{i-1}{n}\right)$, used for the Kolmogorov–Smirnov statistic in Chapter 5, and deduce that $D_n^* = \frac{1}{2n} + \max_{1\le i\le n} \left|x_{(i)} - \frac{2i-1}{2n}\right|$.

(b) Deduce that $D_n^* \ge \frac{1}{2n}$, with equality if and only if $x_{(i)} = \frac{2i-1}{2n}$ for every i. Show that in dimension d, too, every set of n points has $D_n^* \ge \frac{1}{2n}$.

(c) Compute D_n^* for the first $n = 2^m$ points of the van der Corput sequence in base 2 after adding 2^{-m-1} to each of them. Compute D_3^* for the first three points of the unshifted sequence and compare it with the bound of Theorem 11.6(ii).

Exercise 11.2 *(The L^2 discrepancy and Warnock's formula)* [⋆⋆] For points $P = (x_1, \ldots, x_n)$ of $[0,1)^d$, the *L^2 star discrepancy* is $T_n(P) = \left(\int_{[0,1]^d} \delta_P(a)^2 \,\mathrm{d}a\right)^{1/2}$.

(a) Show that $T_n(P) \le D_n^*(P)$.

(b) Prove Warnock's formula

$$T_n(P)^2 = 3^{-d} - \frac{2^{1-d}}{n} \sum_{i=1}^{n} \prod_{j=1}^{d} (1 - x_{ij}^2) + \frac{1}{n^2} \sum_{i,k=1}^{n} \prod_{j=1}^{d} (1 - \max(x_{ij}, x_{kj})).$$

Hint: expand the square and integrate each term coordinate by coordinate.

(c) Let $x_1, \ldots, x_n$ be independent and uniform on $[0,1)^d$. Show that $\mathbb{E}[T_n^2] = (2^{-d} - 3^{-d})/n$.

(d) Implement the formula in (b) (it costs $O(n^2 d)$ operations) and compute T_n for $n = 2^{10}$ points in $d = 2$ and $d = 8$: pseudo-random points (state your seed), the Halton points (`qmc.halton`) and the Sobol' points (`qmc.sobol`). Compare with $\sqrt{(2^{-d} - 3^{-d})/n}$. What do you conclude in $d = 8$?

Exercise 11.3 *(Computing and losing the variation)* [⋆⋆] Use the formula Equation (11.2) for smooth functions.

(a) Let $f(x) = \prod_{j=1}^{d} g_j(x_j)$ with each g_j continuously differentiable on $[0,1]$. Show that $V_{HK}(f) = \prod_{j=1}^{d} \left(|g_j(1)| + \int_0^1 |g_j'|\right) - \prod_{j=1}^{d} |g_j(1)|$.

(b) For the product of sines $f(x) = \prod_{j=1}^{d} \frac{\pi}{2}\sin(\pi x_j)$, show that $V_{HK}(f) = \pi^d$. Using Exercise 11.1(b), show that for $d = 5$ and $n = 2^{16}$ the Koksma–Hlawka bound is at least 2.3×10^{-3} whatever the points, and compare it with the error 1.9×10^{-4} of the Sobol' rule reported in the notes.

(c) For a function f on $[0,1]^2$, the *Vitali variation* is the supremum, over all grids $0 = s_0 < \cdots < s_k = 1$ and $0 = r_0 < \cdots < r_l = 1$, of $\sum_{i,j} \left|f(s_{i+1}, r_{j+1}) - f(s_i, r_{j+1}) - f(s_{i+1}, r_j) + f(s_i, r_j)\right|$; the Hardy–Krause variation of a non-smooth f is at least its Vitali variation. Show that the Vitali variation of $f = \mathbf{1}\{x_1 + x_2 \le 1\}$ is infinite. Hint: use the uniform grid with $k = l$ and look at the cells crossed by the anti-diagonal.

(d) Show that the Vitali variation of $(x_1 + x_2 - 1)^+$ is finite, but that the three-dimensional analogue of the sum in (c), with third differences over the cells of a grid of $[0,1]^3$, is unbounded for $f = (x_1 + x_2 + x_3 - 1)^+$. Hint: on the uniform grid of side $h = 1/k$, the third difference of $x \mapsto g(x_1 + x_2 + x_3)$ over a cell with lower corner x is $g(s + 3h) - 3g(s + 2h) + 3g(s + h) - g(s)$ with $s = x_1 + x_2 + x_3$.

Exercise 11.4 *(An unbounded integrand in one dimension)* [⋆ ⋆ ⋆] Let $F : (0,1) \to [0,\infty)$ be nondecreasing, continuous and integrable, possibly unbounded near 1, and let $Q_n(F) = \frac{1}{n}\sum_{i=1}^{n} F\left(\frac{2i-1}{2n}\right)$ be the rule with the centred points. You may use Koksma's inequality: if G has bounded total variation $V(G)$ on $[0,1]$, then $|Q_P(G) - I(G)| \le V(G)\, D_n^*(P)$.

(a) Let $x_n = 1 - \frac{1}{2n}$ and $G = \min(F, F(x_n))$. Show that $Q_n(F) = Q_n(G)$ and deduce

$$|Q_n(F) - I(F)| \le \frac{F(x_n) - F(0+)}{2n} + \int_{x_n}^{1} F(u)\,\mathrm{d}u.$$

(b) Let $F(u) = e^{-rT}\left(Se^{\nu T + sN^{-1}(u)} - K\right)^+$ with $s = \sigma\sqrt{T} > 0$, the call of Equation (11.1). Show that $z_n = N^{-1}(x_n) \le \sqrt{2\log n}$, using $1 - N(z) \le \frac{1}{2}e^{-z^2/2}$ for $z \ge 0$, and that both terms of (a) are $O\left(n^{-1}e^{sz_n}\right)$. Hint: $\int_z^\infty e^{sx} N'(x)\,\mathrm{d}x = e^{s^2/2}(1 - N(z - s))$, and $1 - N(z - s) \le (1 - N(z))\, e^{sz}\, \frac{z}{z-s}$ for $z > s$. Conclude that the error is $O(n^{-1+\varepsilon})$ for every $\varepsilon > 0$, although $V_{HK}(F) = \infty$.

(c) For the test case ($S = K = 100$, $T = 1$, $r = 1\%$, $y = 2\%$, $\sigma = 20\%$), compute the error and the bound of (a) for $n = 10, 10^2, \ldots, 10^6$. What is the actual rate, and how pessimistic is the bound?

Exercise 11.5 *(Halton points in high dimension)* [⋆] Let $p < q$ be primes, and consider the Halton coordinates in bases p and q.

(a) Show that for $0 \le i < p$ the points are $(i/p, i/q)$, which lie on the line $y = \frac{p}{q}x$.

(b) For $p = 47$, $q = 53$, count the first 47 points in the box $[0, \frac{1}{2})^2$ and deduce a lower bound on D_{47}^*. Compare it with the star discrepancy of 47 independent uniform points, which is about 0.17 on average.

(c) Show that a random shift modulo 1 (Theorem 11.11) maps the points of (a) onto at most three segments with slope p/q, so that it does not remove the problem. What does a random permutation of the base-p digits $\{0, \dots, p-1\}$, applied to the first digit of the first coordinate, do to these points?

Exercise 11.6 *(Scrambled van der Corput points)* [⋆⋆] Let $y_0, y_1, \dots$ be the nested uniform scramble in base 2 (Theorem 11.11, (c)) of the van der Corput sequence in base 2, let $f(x) = x$, and let $\hat{\mu}_n = \frac{1}{n}\sum_{i<n} y_i$.

(a) Let $N = 2^m$. Show that the set $\{y_0, \dots, y_{N-1}\}$ has the law of a stratified sample with N strata: exactly one point in each interval $[\frac{l}{N}, \frac{l+1}{N})$, uniformly distributed there and independent of the points in the other intervals.

(b) Deduce that $\operatorname{Var}\hat{\mu}_N = \frac{1}{12N^3}$, and compare with Monte Carlo.

(c) Show that y_N lies in the same interval of length $\frac{1}{N}$ as y_0, uniformly distributed in the half of it that does not contain y_0, and that

$$\operatorname{Var}\hat{\mu}_{N+1} = \frac{2N^2 + 2N - 3}{24N^2(N+1)^2} \approx \frac{1}{12(N+1)^2}.$$

So one extra point turns a variance of order n^{-3} into one of order n^{-2}.

(d) Check (b) and (c) by simulation for $m = 6$ and $m = 10$, using `qmc.owen_scramble` on the integer points `qmc.sobol_integers(N + 1, 1)` (the first Sobol' coordinate is the van der Corput sequence). Report the number of scrambles and the seeds.

Exercise 11.7 *(Ridge payoffs and Hermite polynomials)* [⋆ ⋆ ⋆] Let He_k be the Hermite polynomials, defined by $e^{tx - t^2/2} = \sum_{k\geq 0} He_k(x)\, t^k/k!$, and $h_k = He_k/\sqrt{k!}$. You may use that $(h_k)_{k\geq 0}$ is an orthonormal basis of L^2 for the standard normal law. Let X, Y be independent standard normal variables, $c \in [0,1]$ and $X' = cX + \sqrt{1-c^2}\,Y$.

(a) Show that $\mathbb{E}[h_k(X') \mid X] = c^k h_k(X)$ for every k.

(b) Let $g = \sum_k a_k h_k$ be in L^2 and not constant. Show that $\operatorname{Cov}(g(X), g(X')) = \sum_{k\geq 1} a_k^2 c^k$, so that the function κ_g of Equation (11.7) is $\kappa_g(c) = \sum_{k\geq 1} a_k^2 c^k / \sum_{k\geq 1} a_k^2$. Deduce that $\kappa_g(c) \leq c$ for $c \in [0,1]$, with equality at some $c \in (0,1)$ if and only if g is affine.

(c) For the digital payoff $g(x) = \mathbf{1}\{x > \theta\}$, show that $a_k = N'(\theta) He_{k-1}(\theta)/\sqrt{k!}$ for $k \geq 1$. Hint: $\frac{\mathrm{d}}{\mathrm{d}x}(He_{k-1}(x)N'(x)) = -He_k(x)N'(x)$.

(d) For $\theta = 0$, use Sheppard's formula $\mathbb{P}(X > 0, X' > 0) = \frac{1}{4} + \frac{\arcsin c}{2\pi}$ to show that $\kappa_g(c) = \frac{2}{\pi}\arcsin c$. A digital option on the geometric average with $m = 64$ dates, struck at the median of the average, is such a ridge payoff. Using the values of w_1^2 in Table 11.1, compute the share of its variance carried by the first normal under the three constructions, and compare with the call.

Exercise 11.8 *(A payoff that sees only the first date)* [⋆] Take the uniform grid $t_i = i/8$, $i = 1, \dots, 8$, and a payoff that depends on the path only through W_{t_1}.

(a) Under each of the three constructions of Theorem 7.7 and the discussion after it (increments; Brownian bridge in the dyadic order; principal components), write $W_{t_1} = \sum_j a_{1j} Z_j$ and compute the shares a_{1j}^2 / t_1 of $\operatorname{Var} W_{t_1}$ carried by each normal. For the bridge, do it by hand; for principal components, use `paths.construction_matrix`.

(b) For a payoff $g(W_{t_1})$, express the share of the variance carried by the first k normals in terms of these numbers and the function κ_g of Equation (11.7). Which construction should be used with QMC, and what is the truncation dimension (with $p = 0.99$) under each?

(c) Principal components carry the largest share of the variance of the whole path in their first normals. Why is there no contradiction with (a)?

Chapter 12

Multilevel Monte Carlo

Exercise 12.1 *(Uncoupled levels)* [⋆] Suppose that the fine and coarse members of every correction are computed from *independent* Brownian paths, so that $V_\ell = \mathrm{Var}(P_\ell) + \mathrm{Var}(P_{\ell-1})$ for $\ell \geq 1$. Take $\mathrm{Var}(P_\ell) = v$ for every ℓ (nearly true for the test-case call, Figure 12.1), $\tau_0 = 1$ and $\tau_\ell = 3 \cdot 2^{\ell-1}$ for $\ell \geq 1$.

(a) Compute $S_L = \sum_{\ell=0}^{L} \sqrt{V_\ell \tau_\ell}$ in closed form, and the cost $2\varepsilon^{-2} S_L^2$ of the multilevel estimator with variance $\varepsilon^2/2$ and the sample sizes of Theorem 12.2 (ignore the integer ceilings).

(b) Plain Monte Carlo on level L reaches the same variance at the cost $2\varepsilon^{-2} v\, 2^L$. Show that the ratio of the two costs increases to $6(1+\sqrt{2})^2$ as $L \to \infty$, and evaluate it for $L = 6$.

(c) Explain with Theorem 8.11 and Theorem 8.12 why the coupled corrections of Section 12.2 behave so differently.

Exercise 12.2 *(Reading a multilevel run)* [⋆] A multilevel run for the discounted at-the-money digital $e^{-rT}\mathbf{1}_{\{S_T > K\}}$ of the test case with the Euler scheme ($\tau_0 = 1$, $\tau_\ell = 3 \cdot 2^{\ell-1}$) produced the following sample means $\hat{m}_\ell$ of Y_ℓ and variance estimates $\hat{V}_\ell$, and $\mathrm{Var}(P_\ell) \approx 0.242$ on every level.

ℓ	0	1	2	3	4	5	6
$\hat{m}_\ell$	0.4758	−0.01957	−0.01011	−0.00502	−0.002202	−0.001442	−0.0004178
$\hat{V}_\ell$	0.2447	0.01899	0.01794	0.01389	0.01064	0.007714	0.005428

(a) Estimate β from the ratios $\hat{V}_\ell / \hat{V}_{\ell+1}$, $\ell = 3, 4, 5$, and α from the ratios $\hat{m}_\ell / \hat{m}_{\ell+1}$, $\ell = 1, 2, 3$.

(b) Let $\varepsilon = 0.005$. Apply the bias test of step 3 of the algorithm of Section 12.4, with $\alpha = 1$, for $L = 3$ and for $L = 4$. What is the smallest admissible finest level?

(c) For that level, compute the sample sizes of Equation (12.3) with $v = \varepsilon^2/2$ and the total cost. Compare with plain Monte Carlo on the same level with variance $\varepsilon^2/2$.

(d) Repeat (b) and (c) for $\varepsilon = 0.0025$ with the same table. By which factors have the two costs grown? Compare with the factors predicted by Theorem 12.6 and Theorem 6.11 for a halving of ε, and conclude.

Exercise 12.3 *(The refinement factor)* [⋆⋆] Let the level-ℓ step be $h_\ell = TM^{-\ell}$ for an integer $M \geq 2$, the coarse increments being sums of M consecutive fine ones, and let a sample of Y_ℓ, $\ell \geq 1$, cost $M^\ell + M^{\ell-1}$ steps.

(a) For the Euler scheme for GBM and $P = DS_T$, adapt the computation of Theorem 12.5 to show that $V_\ell = \frac{1}{2}\sigma^4 T\,\mathbb{E}[P^2]\,(M-1)\,h_\ell + O(h_\ell^2)$. You may use the expansion of the powers made there. *Hint:* over one coarse step, the product of the M fine factors minus the coarse factor is $\sigma^2 \sum_{i<j} \Delta W_i \Delta W_j$ plus terms of higher order.

(b) Suppose more generally that $V_\ell = \kappa(M-1)h_\ell$ for $\ell \geq 1$, that the finest level is the smallest L with $c\,h_L \leq \varepsilon/\sqrt{2}$, and that the variance is $\varepsilon^2/2$. Show that, as $\varepsilon \to 0$, the cost $2\varepsilon^{-2}S_L^2$ is asymptotic to $2\kappa T\varepsilon^{-2}(\log\varepsilon)^2 g(M)$ with $g(M) = (M - M^{-1})/(\log M)^2$.

(c) Tabulate $g(M)$ for $M = 2, \dots, 8$. Which M is best, and what do $M = 2$ and $M = 4$ cost in comparison?

Exercise 12.4 *(Two levels: statistical Romberg)* [⋆⋆] Kebaier's two-level estimator uses the Euler scheme with a fine step T/n and a coarse step T/m, with m dividing n and coupled paths: $\hat\theta = \frac{1}{n_0}\sum_{i=1}^{n_0} P_m^{(i)} + \frac{1}{n_1}\sum_{j=1}^{n_1}(P_n - P_m)^{(j)}$, the two sums independent. Assume $\mathrm{Var}(P_m) = v$, $\mathrm{Var}(P_n - P_m) = v_1 T/m$, $\mathbb{E}[P_n] - \theta = c_b T/n$, and costs m and n steps per sample of the two sums.

(a) For fixed n and m, find the n_0, n_1 (real) that minimise the cost subject to the variance $\varepsilon^2/2$, and the minimal cost.

(b) Minimise over real m: show that $m^* = \sqrt{v_1 T n / v}$ and that the cost is then $8\varepsilon^{-2}\sqrt{v v_1 T n}$.

(c) Choose n so that the bias is $\varepsilon/\sqrt{2}$ and deduce the growth of the cost as $\varepsilon \to 0$. Compare with plain Monte Carlo and with Theorem 12.6 for the Euler scheme.

(d) For the test-case call, take $v = 158$, $v_1 = 4.7$, $T = 1$ and $n = 256$. Compute m^*, the quantity $\sqrt{vm} + \sqrt{v_1 T n/m}$ for $m = 2$ and $m = 4$, the corresponding quantity $\sqrt{vn}$ for plain Monte Carlo, and compare both with the multilevel value $\sum_{\ell=0}^{8}\sqrt{V_\ell \tau_\ell} \approx 31.5$ obtained from the Euler call levels of Figure 12.1. What do these numbers say about costs?

Exercise 12.5 *(Milstein's level variance, exactly)* [⋆⋆⋆] Take the Milstein scheme for GBM and $P = DS_T$, with $D = e^{-rT}$. Over one coarse step of length $2h$, let d_1, d_2 be the two fine increments, $f_i = 1 + \mu h + \sigma d_i + \frac{1}{2}\sigma^2(d_i^2 - h)$ the fine factors, $F = 1 + 2\mu h + \sigma(d_1 + d_2) + \frac{1}{2}\sigma^2((d_1+d_2)^2 - 2h)$ the coarse factor, and $\Delta = f_1 f_2 - F$. You may use the moments of $\mathcal{N}(0,h)$.

(a) Show that $\Delta = \mu^2 h^2 + \mu\sigma h(d_1 + d_2) + \frac{1}{2}\sigma^3[d_1(d_2^2 - h) + d_2(d_1^2 - h)] + \frac{1}{2}\mu\sigma^2 h[(d_1^2 - h) + (d_2^2 - h)] + \frac{1}{4}\sigma^4(d_1^2 - h)(d_2^2 - h)$, and that $e = \mathbb{E}[\Delta F] = (\mu^2 + 2\mu\sigma^2)h^2 + O(h^3)$ and $q = \mathbb{E}[\Delta^2] = (2\mu^2\sigma^2 + \sigma^6)h^3 + O(h^4)$.

(b) With $v = \mathbb{E}[F^2]$ and $n = 2^{\ell-1}$ coarse steps, show that $\mathbb{E}[Y_\ell^2] = D^2 S_0^2[(v + 2e + q)^n - 2(v + e)^n + v^n]$, and deduce

$$V_\ell = h_\ell^2\Big[\mathbb{E}[P^2]\Big(\tfrac{T}{2}(\sigma^6 + 2\mu^2\sigma^2) + \tfrac{T^2}{4}(\mu^2 + 2\mu\sigma^2)^2\Big) - \tfrac{T^2}{4}\mu^4\,\mathbb{E}[P]^2\Big] + O(h_\ell^3).$$

Which term has no counterpart in the Euler computation of Theorem 12.5, and why does it appear here?

(c) Evaluate the constant for the test case and compare with V_ℓ/h_ℓ^2 computed from the exact formula (`multilevel.gbm_level_exact`) at $\ell = 10$.

Exercise 12.6 *(The best randomisation)* [⋆⋆⋆] In the setting of Theorem 12.10, suppose that every $m_j = \mathbb{E}[Y_j^2]$ is positive, $\inf_j \tau_j > 0$ and $\sum_j \sqrt{m_j \tau_j} < \infty$.

(a) Let $m_j = m\,2^{-\beta j}$ and $\tau_j = \tau\,2^{\gamma j}$ exactly, for all $j \geq 0$, with $\beta > \gamma$, and let $p_j = (1-2^{-r})2^{-rj}$ with $\gamma < r < \beta$. Compute $\mathbb{E}[Z^2]$ and $\mathbb{E}[\tau_J]$ in closed form, find the rate r^* that minimises their product, and show that the minimum equals the lower bound of Theorem 12.10(iv): for exactly geometric moments and costs, the best geometric law cannot be beaten.

(b) For the test-case call with Milstein, a run gave the following second moments m_j and variances V_j of the corrections, with $\tau_0 = 1$ and $\tau_j = 3 \cdot 2^{j-1}$:

j	0	1	2	3	4	5
m_j	214.8	0.04692	0.01247	0.003193	0.0008149	0.0002072
V_j	158.9	0.04101	0.01164	0.00305	0.0007869	0.0002007

For $j \geq 6$ take $m_j = m_5 4^{-(j-5)}$ and $V_j = V_5 4^{-(j-5)}$. With $\theta = 7.36429$, compute p_0, $\mathbb{E}[\tau_J]$ and $\mathrm{Var}(Z)\,\mathbb{E}[\tau_J]$ for the law $p_j \propto \sqrt{m_j/\tau_j}$, which attains the bound of Theorem 12.10(iv); compare with the geometric choice $p_j = (1-2^{-1.5})2^{-1.5j}$ and with the work-normalised variance $\left(\sum_j \sqrt{V_j \tau_j}\right)^2$ of multilevel Monte Carlo with all levels. Why is the geometric choice so far from optimal here?

Exercise 12.7 *(Smoothing a digital)* [⋆⋆] This is a computational exercise on the test-case digital $D\mathbf{1}_{\{S_T > K\}}$, $D = e^{-rT}$.

(a) Let the last step of a path with step h be an Euler step, $S_T = S_{T-h}(1 + \mu h + \sigma \Delta W)$. Show that $\mathbb{Q}(S_T > K \mid \mathcal{F}_{T-h}) = N\big((S_{T-h}(1+\mu h) - K)/(\sigma S_{T-h}\sqrt{h})\big)$ on $\{S_{T-h} > 0\}$, and find it on $\{S_{T-h} \leq 0\}$.

(b) At level ℓ, with $h = h_\ell$ and $n = 2^{\ell-1}$ coarse steps, use (a) for the fine path. For the coarse path, whose last step has length $2h$ and increment $\Delta W_{2n-1} + \Delta W_{2n}$ (the last two fine increments), condition on the coarse path up to $T - 2h$ *and* on the fine increment ΔW_{2n-1}. Write the resulting coarse member and show that it has the mean of the level-$(\ell-1)$ smoothed payoff, so that the corrections satisfy Theorem 12.1 although the coarse member no longer has the law of $P_{\ell-1}$.

(c) Implement this with Euler steps throughout, and with Milstein steps except for the last (Euler) step. Estimate V_ℓ for $\ell \leq 8$ and β in both cases, and compare with Figure 12.1. What happens at level 0?

(d) Explain why smoothing lowers V_ℓ but not its rate with Euler, while it raises the rate with Milstein. *Hint:* compare the difference between the fine and coarse conditional means with the conditional standard deviation $\sigma S\sqrt{h}$.

(e) Run the adaptive algorithm at $\varepsilon = 0.001$ for both versions and compare their costs with the unsmoothed Euler run reported in Section 12.4.

Exercise 12.8 *(A discretely monitored barrier)* [⋆⋆] This is a computational exercise. The course priced a down-and-out call on the test case with barrier $B = 95$ monitored at the twelve monthly dates $t_j = j/12$, $j = 1, \dots, 12$: the price is 5.32979. Let level ℓ use $12 \cdot 2^\ell$ uniform steps, so that the monitoring dates are grid points at every level.

(a) Explain why computing the payoff from the scheme values at the monitoring dates gives coupled corrections, and why there is no monitoring bias. What bias order α do you expect?

(b) Implement the Euler and Milstein couplings. With 2×10^5 samples per level, estimate V_ℓ, the mean of Y_ℓ, the kurtosis of Y_ℓ, and the number of samples in which exactly one of the two paths is knocked out, for $\ell = 0, \dots, 5$. Where can you estimate β?

(c) Price the option with the adaptive algorithm and the Euler scheme to $\varepsilon = 0.01$, and compare with 5.32979.

(d) Adapt the argument of Theorem 12.4(ii) to predict β for the two schemes, and relate the kurtosis to Theorem 12.9.

Chapter 13

Monte Carlo Greeks

Throughout, the model is that of the notes: under $\mathbb{Q}$, $S_T = SG$ with $G = \exp(\nu T + \sigma\sqrt{T}Z)$, $\nu = r - y - \sigma^2/2$, $Z \sim \mathcal{N}(0,1)$, and $S, \sigma, T > 0$. The test case is $S = K = 100$, $T = 1$, $r = 1\%$, $y = 2\%$, $\sigma = 20\%$. The pathwise and likelihood-ratio estimators are those of Equation (13.2) and Equation (13.5). You may use the results of the notes, quoting them, and the closed forms of `bs.py`. In computational parts, use NumPy's `default_rng` with a seed of your choice and report it.

Exercise 13.1 *(Identities that hold draw by draw)* [⋆] Consider the call.

(a) The strike does not enter the law of S_T, but the payoff is $K(S_T/K - 1)^+$ and $S_T/K = (S/K)G$. Use the likelihood ratio in $\log(S/K)$ to derive an unbiased estimator ∂_K^{lr} of $\partial C/\partial K$, and show that $S\,\Delta^{\mathrm{lr}} + K\,\partial_K^{\mathrm{lr}} = e^{-rT}(S_T - K)^+$ for every draw. Compare its per-draw standard deviation at the test case with that of the pathwise $\partial_K^{\mathrm{pw}} = -e^{-rT}\mathbf{1}_{\{S_T > K\}}$.

(b) Differentiate ∂_K^{pw} once more in K by likelihood ratio to get an estimator of $\partial^2 C/\partial K^2$, the discounted risk-neutral density of S_T at K (Breeden–Litzenberger). Show that K^2 times it equals S^2 times the LR-PW gamma of Theorem 13.8 for every draw.

(c) The pathwise derivative of the pathwise delta is 0 almost surely, so there is no draw-by-draw relation between $\mathcal{V}^{\mathrm{pw}}$ and a "pathwise gamma". Show, however, that $\mathcal{V}^{\mathrm{pw}} = \sigma T S^2\,\Gamma^{\text{pw-lr}}$ for every draw, where $\Gamma^{\text{pw-lr}}$ is the PW-LR estimator of Theorem 13.8.

Exercise 13.2 *(Extrapolated differences)* [⋆⋆] Let α be C^5 near θ, let $D_h = (Y(\theta + h) - Y(\theta - h))/(2h)$ for one replication, and let $E_h = (4D_h - D_{2h})/3$, computed with common random numbers (four evaluations of Y per replication).

(a) Show that $\mathbb{E}E_h - \alpha'(\theta) = -\frac{1}{30}\alpha^{(5)}(\theta)h^4 + o(h^4)$.

(b) Suppose that $\mathbb{E}[(Y(\theta_1) - Y(\theta_2))^2] \le c|\theta_1 - \theta_2|$ near θ. Show that $\operatorname{Var} E_h \le C/h$ for a constant C, and use Theorem 13.1(iv) to find the orders of the best bump and of the best mean squared error, assuming the variance is exactly of order $1/h$ and $\alpha^{(5)}(\theta) \neq 0$.

(c) For the delta of the digital call in the test case and $n = 10^6$, find the bump that minimises the root mean squared error of E_h (exact bias from the closed form; variance by quadrature

or by simulation with at least 10^6 draws), and compare with the central difference and with the likelihood-ratio estimator (Theorem 13.2).

(d) Why is extrapolation pointless for a Lipschitz payoff with common random numbers?

Exercise 13.3 *(Smoothing by conditioning)* [⋆⋆] This is partly computational. A digital on a two-date average pays $e^{-rT}\mathbf{1}_{\{(S_{t_1}+S_T)/2>K\}}$ with $t_1 = T/2$; take the test case.

(a) Show that $\mathbb{E}\big[\mathbf{1}_{\{(S_{t_1}+S_T)/2>K\}} \,\big|\, S_{t_1}\big] = \varphi(S_{t_1})$, where $\varphi(s) = N(d(s))$ with $d(s) = \big(\log\frac{s}{2K-s} + \nu(T-t_1)\big)/\big(\sigma\sqrt{T-t_1}\big)$ for $s < 2K$, and $\varphi(s) = 1$ for $s \geq 2K$. Show that φ is continuously differentiable on $(0,\infty)$ with a bounded derivative.

(b) Deduce that $e^{-rT}\varphi(S_{t_1})$ is an unbiased estimator of the price with no larger variance, and that its pathwise delta $e^{-rT}\varphi'(S_{t_1})S_{t_1}/S$ is unbiased. What does the pathwise method give for the original payoff?

(c) Write the likelihood-ratio deltas of the original and of the conditioned payoffs, using the score of the first step.

(d) With 10^6 draws, estimate the delta by the three unbiased estimators of (b) and (c), with standard errors, and compare with the exact delta, computed by one-dimensional quadrature. Which would you use?

Exercise 13.4 *(Combining pathwise and likelihood ratio)* [⋆] Let D_1 and D_2 be two unbiased estimators of the same Greek, computed from the same draw, with finite variances v_1, v_2 and covariance c_{12}, and $v_1 + v_2 > 2c_{12}$.

(a) Show that $D(\lambda) = \lambda D_1 + (1-\lambda)D_2$ is unbiased for every λ, and find the λ^* that minimises its variance and the minimal variance. Explain why this is a control variate (Theorem 8.12) and what happens if λ^* is estimated from the same draws (Theorem 8.14).

(b) For the delta of the test-case call, with $D_1 = \Delta^{\mathrm{pw}}$ and $D_2 = \Delta^{\mathrm{lr}}$, compute the correlation, λ^* and the variance of $D(\lambda^*)$ relative to that of the pathwise estimator alone. Repeat for the vega. Is the combination worth it?

Exercise 13.5 *(The likelihood-ratio delta at the forward)* [⋆⋆] Consider the call struck at the forward, $K = F = Se^{(r-y)T}$, and let $\sigma \downarrow 0$ with S, T, r, y fixed. Write $s = \sigma\sqrt{T}$.

(a) Show that the Black–Scholes delta tends to $e^{-yT}/2$.

(b) Show that $\operatorname{Var}\Delta^{\mathrm{pw}} \to e^{-2yT}/4$ and $\operatorname{Var}\Delta^{\mathrm{lr}} \to \frac{5}{4}e^{-2yT}$, justifying the limits. Why does the blow-up of Theorem 13.7 not occur?

(c) The gamma itself grows like $e^{-yT}/(\sqrt{2\pi}\,Ss)$. Show that the relative standard deviation per draw of the LR-LR gamma estimator, $\operatorname{sd}(\Gamma^{\mathrm{lr}})/\Gamma$, has a finite limit, and compute it.

Exercise 13.6 *(Greeks of the asymptotic square-root call)* [⋆⋆] This is a computational exercise, from the course's homework. The asymptotic square-root call pays $F(S_T) = 2K\big[1 - \sqrt{K/S_T}\,\big]^+$ and has Black–Scholes price $e^{-rT}2K\big[N(d_-) - \sqrt{K/S}\,e^{-\nu T/2+\sigma^2 T/8}N(d_- - \frac{1}{2}\sigma\sqrt{T})\big]$.

(a) Write the pathwise delta and ρ_y, check the hypotheses of Theorem 13.4, and show that $\rho_y^{\mathrm{pw}} = -TS\Delta^{\mathrm{pw}}$.

(b) Write the likelihood-ratio delta and the three gamma estimators of Theorem 13.8, after checking its hypotheses.

(c) Compute the delta and gamma of the closed form (analytically or by finite differences of the closed form). With $n = 10^3$ draws reused for every S_0 in $[Ke^{-2\sigma\sqrt{T}}, Ke^{2\sigma\sqrt{T}}]$, plot the bumped central delta for $\varepsilon = 0.1, 0.01, 0.001$ and the pathwise and likelihood-ratio deltas against S_0. Then, at $S_0 = K$ with $n = 10^5$, report the estimates and standard errors of the two deltas and the three gammas. Which gamma estimator would you use?

Exercise 13.7 *(A barrier monitored once)* [⋆⋆⋆] This is partly computational. A call knocked out if the price is at or below B at the single date $t_1 = T/2$ pays $e^{-rT}(S_T - K)^+ \mathbf{1}_{\{S_{t_1} > B\}}$; take the test case and $B = 90$. Let $C(s, \tau)$ be the Black–Scholes call price with spot s and maturity τ, q the density of S_{t_1}, and $b(S) = (\log(B/S) - \nu t_1)/(\sigma\sqrt{t_1})$.

(a) Show that the price is $V(S) = e^{-rt_1}\mathbb{E}[\mathbf{1}_{\{S_{t_1} > B\}} C(S_{t_1}, T - t_1)]$ and that

$$V'(S) = e^{-rt_1}\mathbb{E}\Big[\mathbf{1}_{\{S_{t_1} > B\}} \partial_s C(S_{t_1}, T - t_1) \frac{S_{t_1}}{S}\Big] + e^{-rt_1} C(B, T - t_1)\, \frac{B}{S}\, q(B).$$

Evaluate both terms by quadrature.

(b) Show that the pathwise estimator $e^{-rT}\mathbf{1}_{\{S_T > K\}}\mathbf{1}_{\{S_{t_1} > B\}} S_T/S$ has mean equal to the first term only. Which hypothesis of Theorem 13.4 fails?

(c) Write the likelihood-ratio delta with the score of the first step.

(d) *One-step survival.* Draw U uniform on $(0, 1)$ and $Z_1 = N^{-1}(N(b) + U(1 - N(b)))$, so that Z_1 has the law of Z given $Z > b$; build S_{t_1} from Z_1 and S_T from an independent Z_2, and let $Y^{\mathrm{s}} = (1 - N(b(S)))\, e^{-rT}(S_T - K)^+$. Show that $\mathbb{E}Y^{\mathrm{s}} = V(S)$, that for fixed (U, Z_2) the map $S \mapsto Y^{\mathrm{s}}$ is Lipschitz near S with an integrable constant, and differentiable except where $S_T = K$ (*hint*: Mills' ratio $(1 - N(z))/N'(z)$ is decreasing), and write its pathwise delta.

(e) With $n = 10^5$, compare the estimators of (b), (c) and (d) with the exact delta.

Exercise 13.8 *(The adjoint of the Black–Scholes formula)* [⋆⋆] This is a computational exercise. Write the call price as a sequence of elementary steps: $v_1 = \log(S/K)$, $v_2 = \sigma\sqrt{T}$, $v_3 = (r - y)T$, $d_+ = (v_1 + v_3)/v_2 + v_2/2$, $d_- = d_+ - v_2$, $a = e^{-yT}$, $b = e^{-rT}$, $C = SaN(d_+) - KbN(d_-)$.

(a) Write the reverse sweep that, from $\bar{C} = 1$, returns $\partial C/\partial x$ for the six inputs $x \in \{S, K, T, r, y, \sigma\}$ in one pass, without any library for automatic differentiation.

(b) Implement it and check it against `bs.greeks` and $\partial_K C = -e^{-rT}N(d_-)$ for the test case. Explain why the adjoint $\bar{d}_+$ of d_+, after both of its uses have been accumulated, is exactly zero.

(c) Count the operations of the forward and reverse sweeps (additions and multiplications, and evaluations of exp, log, N and N'), and compare with central bumps of the six inputs.

Chapter 14

American options by simulation

Exercise 14.1 *(Stopping a sequence of uniforms)* [⋆] Let $Z_1, \dots, Z_m$ be independent and uniform on $[0,1]$, $Z_0 = 0$, $\mathcal{F}_k = \sigma(Z_1, \dots, Z_k)$, and no discounting. You may use Theorem 14.2.

(a) Show that every continuation value is a constant, $C_k = v_{k+1}$, where $v_m = \frac{1}{2}$ and $v_k = \frac{1}{2}(1 + v_{k+1}^2)$ for $1 \le k < m$, and that $U_0 = v_1$. Describe the optimal rule and compute U_0 for $m = 2$ and $m = 5$.

(b) Compute the perfect-foresight bound $\mathbb{E}[\max_k Z_k]$ and its gap for $m = 2$ and $m = 5$.

(c) Write down the martingale M^* of Theorem 14.2(v) and check directly, without Theorem 14.10, that $\max_k (Z_k - M_k^*) = U_0$ for every outcome.

(d) Let $e_k = 1 - v_k$. Show that $e_k = e_{k+1} - \frac{1}{2} e_{k+1}^2$ and deduce $1 - U_0 \le 2/(m+3)$. Show that the perfect-foresight gap is at most $1/m$.

Exercise 14.2 *(Bounds, and exercise at t_0)* [⋆] Consider a Bermudan put on the Black–Scholes stock, $Z_k = e^{-r t_k}(K - S_{t_k})^+$.

(a) Let $r \ge 0$ and let $U_0(s)$ be the Bermudan price as a function of $S_0 = s$, the dates being fixed. Show that $\max\left(K - s, P(0,s)\right) \le U_0(s) \le K$, where P is the European put price, and that U_0 is convex and nonincreasing in s.

(b) Let $r > 0$. Show that $C_0 \le e^{-r t_1} K$, so that immediate exercise is optimal whenever $S_0 \le K(1 - e^{-r t_1})$. Evaluate this bound for the put of Table 14.1 with $\sigma = 20\%$ and $T = 1$, and compare it with the critical price at t_0, which you may compute with `bermudan_grid`.

(c) The in-sample LSM estimate is $\max(Z_0, \hat{c}_0)$, where $\hat{c}_0$ is a mean of N cash flows. Suppose that the cash flows follow the optimal rule from t_1 on, so that $\mathbb{E}[\hat{c}_0] = C_0$. Show that $\mathbb{E}[\max(Z_0, \hat{c}_0)] \ge U_0$. If $Z_0 = C_0$ and $\hat{c}_0$ is normal with standard deviation s, compute the bias.

(d) How does the out-of-sample estimator of the notes treat t_0, and why is it not biased upwards there?

Exercise 14.3 *(Two exercise dates in closed form)* [⋆⋆⋆] Let a put be exercisable at $t_0 = 0$,

$t_1 = T/2$ and $t_2 = T$, with $r > 0$, $y \geq 0$, and let $P(\theta, s)$ be the Black–Scholes European put price with time to maturity θ. Write $N_2(a, b; \rho)$ for the distribution function of a standard bivariate normal vector with correlation ρ.

(a) Show that $\varphi(s) = (K - s) - P(T/2, s)$ is strictly decreasing on $(0, \infty)$, with $\varphi(0+) > 0 > \varphi(K)$. Deduce that there is a unique $s^* \in (0, K)$ such that exercise at t_1 is optimal if and only if $S_{t_1} \leq s^*$.

(b) Show that

$$U_0 = \max\Big(Z_0,\ e^{-rT/2}\mathbb{E}\big[(K - S_{T/2})\mathbf{1}_{\{S_{T/2} \leq s^*\}}\big] + e^{-rT}\mathbb{E}\big[(K - S_T)^+\mathbf{1}_{\{S_{T/2} > s^*\}}\big]\Big),$$

and write the two expectations with N and $N_2(\cdot, \cdot; \rho)$, $\rho = \sqrt{1/2}$.

(c) Compute s^* and U_0 for $S_0 = K = 40$, $r = 6\%$, $y = 0$, $\sigma = 20\%$, $T = 1$, and compare with `bermudan_grid` with $m = 2$ and with an out-of-sample LSM estimate.

Exercise 14.4 *(A better control)* [⋆⋆] Let $P(t, s)$ be the Black–Scholes price at t of the European put with maturity T ($P(T, s) = (K - s)^+$) and $X_k = e^{-rt_k}P(t_k, S_{t_k})$.

(a) Show that $\mathbb{E}[X_\tau] = P(0, S_0)$ for every $\tau \in \mathcal{T}_0$.

(b) In the out-of-sample LSM estimator, use $X_{\hat{\tau}}$ as a control variate, with a coefficient estimated on the training paths. Show that the controlled estimator is still conditionally unbiased for the value of the policy.

(c) Implement it and compare its variance with that of the terminal control $e^{-rT}(K - S_T)^+$ of Theorem 14.9, on the put with $S_0 = 36$, $\sigma = 20\%$, $T = 1$ of Table 14.1 and on the test case. Explain the difference.

Exercise 14.5 *(The European price as a basis function)* [⋆⋆] On the test case, Theorem 14.9 found an LSM policy that exercises a third of the paths although early exercise is worth almost nothing.

(a) What is the continuation value c_k when early exercise is never optimal after t_k? Deduce a natural extra basis function at date k.

(b) Add $s \mapsto P(t_k, s)$ (notation of Exercise 14.4) to the four Laguerre functions at each date, and redo the experiment: out-of-sample estimate, value of the policy, fraction of paths exercised before maturity.

(c) Does the extra function help for the put with $S_0 = 36$, $\sigma = 20\%$, $T = 1$ of Table 14.1?

Exercise 14.6 *(Optimising a martingale)* [⋆⋆] Let M^E be the European put martingale of Section 14.3 and $H(\lambda) = \mathbb{E}[\max_k(Z_k - \lambda M^E_k)]$ for $\lambda \in \mathbb{R}$.

(a) Show that H is convex and that $H(\lambda) \geq U_0$ for every λ.

(b) For the put with $S_0 = 36$, $\sigma = 20\%$, $T = 1$ of Table 14.1, estimate H on a grid of $\lambda \in [0, 1.5]$ from 2000 paths and locate its minimum.

(c) Let $\hat{H}_n(\lambda)$ be the sample mean from n paths. Show that $\mathbb{E}[\min_\lambda \hat{H}_n(\lambda)] \le \min_\lambda H(\lambda)$. Propose a procedure that uses the minimising λ and still gives a valid upper confidence bound for U_0.

Exercise 14.7 *(The gap of a value-function martingale)* [⋆⋆] Let $\hat{V}_k$ be functions with $\mathbb{E}[\hat{V}_k(S_{t_k})^2] < \infty$, set $\hat{U}_k = e^{-rt_k}\hat{V}_k(S_{t_k})$, and define $M_0 = 0$ and $M_{k+1} - M_k = \hat{U}_{k+1} - \mathbb{E}[\hat{U}_{k+1} \mid \mathcal{F}_k]$, as Haugh and Kogan (2004) do with an approximate value function. Assume $\mathbb{E}[Z_k^2] < \infty$.

(a) Show that M is a square-integrable martingale.

(b) Show that $\mathbb{E}[\max_k(Z_k - M_k)] - U_0 \le 2\big(\sum_{k=1}^m \mathbb{E}[(\hat{U}_k - U_k)^2]\big)^{1/2}$.

(c) The bound involves the error of $\hat{V}_k$ everywhere, also far from the exercise boundary. Compare with Theorem 14.11(ii). Which errors of $\hat{V}_k$ do not affect M at all?

Exercise 14.8 *(A call on the maximum of two assets)* [⋆ ⋆ ⋆] Let S^1, S^2 be independent Black–Scholes stocks with $S_0^1 = S_0^2 = S_0$, $\sigma = 20\%$ and $y = 10\%$, let $r = 5\%$, $K = 100$, $T = 3$, and consider the Bermudan call on the maximum, with payoff $(\max(S^1, S^2) - K)^+$ and exercise dates $t_k = k/3$, $k = 0, \dots, 9$.

(a) Implement LSM for this option. Choose a basis (for instance polynomials of degree two in S^1/K and S^2/K, with or without the payoff itself) and justify the choice.

(b) Compute out-of-sample lower bounds for $S_0 = 90, 100, 110$ with 10^5 paths.

(c) Compute Andersen–Broadie upper bounds by nested simulation, with sizes your computer allows, and the confidence interval of Theorem 14.11(iv).

(d) Andersen and Broadie (2004) report the 95% intervals $[8.053, 8.082]$, $[13.892, 13.934]$ and $[21.316, 21.359]$ for these three cases. Compute the price itself on a two-dimensional grid (the transition factorises over the two independent assets). Are your intervals consistent with it? Which half of your interval, below or above the price, is wider, and why?

Part III

Path-dependent and multi-asset payoffs

Chapter 15

Asian options

Exercise 15.1 *(Means)* [⋆] Let $a_1, \dots, a_m > 0$ and weights $w_j > 0$ with $\sum_j w_j = 1$, and let $H = \left(\sum_j w_j/a_j\right)^{-1}$, $G = \prod_j a_j^{w_j}$ and $A = \sum_j w_j a_j$ be the harmonic, geometric and arithmetic means.

(a) Show that $\min_j a_j \le H \le G \le A \le \max_j a_j$, and that $H = G$, or $G = A$, holds if and only if all the a_j are equal.

(b) Order the prices of the calls, and of the puts, with strike K on the harmonic, geometric and arithmetic averages of the same fixings.

(c) Let A be the arithmetic average of the fixings S_{t_j}, $t_j \le T$, and suppose that $\int_t^T (r - y) \ge 0$ for every fixing date t (for instance $r \ge y$). Show that the average-price call is worth at most the European call with maturity T and the same strike. In the test case, where $y > r$, compare the two calls for $m = 12$ at $K = 60$ and $K = 100$.

Exercise 15.2 *(Two fixings)* [⋆] Take $m = 2$ fixings at $t_1 = 0$ and $t_2 = T$, with equal weights, so that $A = \frac{1}{2}(S + S_T)$ and $G = \sqrt{S S_T}$. You may use Theorem 15.2, Theorem 15.8 and Theorem 15.10.

(a) Express the arithmetic call through a European call when $2K > S$, and give its price when $2K \le S$. Evaluate both in the test case, for $K = 100$ and $K = 40$.

(b) Compute μ_G and v_G from Theorem 15.2(i), once with the double sum and once with the tail sums, and price the geometric call for $K = 100$.

(c) Show that the three-moment approximation C_3 is exact here, and that the lower bound $L(x^*)$ equals C for every $K > 0$. For which K is the first term $L(\log K)$ exact? Is the two-moment approximation exact?

Exercise 15.3 *(A seasoned option)* [⋆⋆] A monthly Asian call on the test case ($t_j = j/12$, $K = 100$) is valued at $t = \frac{1}{2}$, after six fixings, with $S_t = 105$. Constant parameters are as in the test case. You may use the code of the chapter for numerical values.

(a) The six known fixings have arithmetic mean 98. Write the value at t as a multiple of the

price of an Asian call on the six remaining fixings, valued at t, with a shifted strike, and compute it.

(b) For the geometric call, the six known fixings have geometric mean 98. Show that its value at t is a Black formula, give its parameters, and compute it.

(c) Now value the arithmetic call at $t = \frac{11}{12}$, just before the last fixing, with $S_t = 105$ and eleven known fixings of arithmetic mean 100. Show that it is a multiple of a European call, and give its value and delta. Which approximations of the chapter are exact in this situation?

Exercise 15.4 *(Thresholds)* [⋆⋆] Use the notation of Theorem 15.9 and Theorem 15.10, with $K > 0$ and $v_G > 0$, and let ϕ_G be the density of $\Lambda = \log G$.

(a) Show that L is differentiable with $L'(x) = -D(h(x) - K)\phi_G(x)$, and deduce again that x^* maximises L. Why does an error δ in a finite threshold change L only by $O(\delta^2)$?

(b) A cheap threshold replaces $h(\log g)$ by $h(\log K) + (g - K)$, its first-order expansion around $g = K$ with slope one. Show that the resulting threshold is $\hat{K} = 2K - h(\log K)$, and that $L(\log \hat{K}) \le L(x^*) \le C$ when $\hat{K} > 0$. What lower bound remains when $\hat{K} \le 0$? ($L(\log \hat{K})$ is the formula published by Curran.)

(c) In the monthly test case ($m = 12$, $K = 100$) compute $h(\log K)$, $\hat{K}$, e^{x^*}, $L(\log K)$, $L(\log \hat{K})$ and $L(x^*)$. Repeat for $\sigma = 80\%$ and compare with the reference price.

Exercise 15.5 *(How good is the geometric control?)* [⋆⋆] This is a computational exercise. Take $K = 100$, $T = 1$ and equally spaced fixings.

(a) For $m \in \{4, 12, 52\}$ and $\sigma \in \{10\%, 20\%, 40\%\}$, estimate $1 - \rho^2$ for the geometric control with 2^{16} paths, and compare it with the bound $D^2\mathbb{E}[(A - G)^2]/\operatorname{Var} Y$ of Theorem 15.4, with $\operatorname{Var} Y$ estimated from the same paths.

(b) How does $1 - \rho^2$ depend on σ, and how on m? Give a heuristic reason for the dependence on σ.

(c) Time the simulation of the paths and the two payoffs, and estimate the efficiency of the geometric control relative to plain Monte Carlo for $m = 12$, $\sigma = 20\%$.

Exercise 15.6 *(An Australian option)* [⋆⋆⋆] An Australian variable-purchase option pays $(A/S_T - k)^+$ at T, where A is the average of the monthly fixings $t_j = jT/12$, $j = 1, \dots, 12$ (so $t_{12} = T$); constant parameters.

(a) Let $X_u = S_{T-u}/S_T$ for $0 \le u \le T$. Show that $(X_u)_{0 \le u \le T}$ has the law of a geometric Brownian motion started at 1 with volatility σ and log-drift $-\nu$, where $\nu = r - y - \frac{1}{2}\sigma^2$. *Hint:* $B_u = W_T - W_{T-u}$ is a Brownian motion.

(b) Deduce that the option is worth an average-price call on X with strike k, rate r, yield $y' = 2r - y - \sigma^2$, volatility σ, fixings at $u_j = T - t_j$ (one of them at 0) and payment at T.

(c) For the test case with $k = 1$, compute the geometric lower bound, the strip upper bound, C_2, C_3, $L(x^*)$, the conditional lognormal approximation and a reference price, and check the reference by simulating $(A/S_T - k)^+$ directly.

(d) A published solution of this exercise uses the yield $2r + y - \sigma^2$. Which price does it give, and how would a simulation detect the error?

Exercise 15.7 *(A weekly two-year Asian)* [⋆⋆] This is a computational exercise. Take weekly fixings $t_j = j/52$, $j = 1, \dots, 104$, $T = 2$, $\sigma = 30\%$, and the other parameters of the test case.

(a) Compute a reference price with the recursion Equation (15.4), and check its convergence in the grid size.

(b) Tabulate the errors of C^G, C_2, C_3, the Edgeworth approximation, $L(\log K)$, $L(x^*)$ and the conditional lognormal approximation, and compare with Table 15.1.

(c) With 2^{16} paths, compare plain Monte Carlo, the geometric control, the three controls and the conditional estimator of Theorem 15.6: standard deviations per path and estimates with their standard errors.

Chapter 16

Barriers, lookbacks and the Brownian bridge

Throughout, $S_t = S_0 e^{X_t}$ with $X_t = \nu t + \sigma W_t$ for a Brownian motion W under $\mathbb{Q}$, $\nu = r - y - \frac{1}{2}\sigma^2$, $\lambda = 2\nu/\sigma^2$; $\overline{X}_T$ and $\underline{X}_T$ are the maximum and minimum of X on $[0,T]$, and p_T is the $\mathcal{N}(\nu T, \sigma^2 T)$ density. The test case is $S_0 = K = 100$, $T = 1$, $r = 1\%$, $y = 2\%$, $\sigma = 20\%$, and $h = T/m$ for m equally spaced monitoring dates $t_j = jh$, $j = 1, \dots, m$.

Exercise 16.1 *(Degenerate cases)* [⋆] Use the test case and a down-and-out call with barrier B unless stated otherwise.

(a) Let $S_0 = B = 100$. Explain why the continuously monitored option is worthless while the option monitored at $t_1 = \frac{1}{2}$ and $t_2 = 1$ is not. Write the latter's price with the bivariate normal distribution function $N_2(\cdot, \cdot; \rho)$ and evaluate it.

(b) With $B = 95$, the continuously monitored down-and-out call costs 3.75412 and the monthly one 5.32979. What does a portfolio of the continuously monitored down-and-out call and the monthly down-and-in call cost, and why is it not the vanilla call?

(c) By Theorem 16.2 and Theorem 16.9(i), the down-and-out price is

$$e^{-rT}\mathbb{E}^{\mathbb{Q}}\Big[g(S_T)\mathbf{1}_{\{S_T > B\}}\Big(1 - e^{-2\log(S_0/B)\log(S_T/B)/(\sigma^2 T)}\Big)\Big].$$

Let $\sigma \downarrow 0$ with r, y fixed, and let $F = S_0 e^{(r-y)T} \neq B$. Show that the price tends to the discounted payoff of the deterministic path $S_0 e^{(r-y)t}$, knocked out if that path reaches B. Evaluate the limit for $B = 95$ with $K = 100$ and with $K = 90$.

(d) Show that, for $S_0 > B$, the price tends to $(S_0 - K)^+$ as $T \downarrow 0$.

Exercise 16.2 *(The image density and the forward equation)* [⋆] Let $b < 0$ and $q_T(x) = p_T(x) - c\, p_T(x - 2b)$ for $x > b$, with a constant c.

(a) Show that q solves the forward equation $\partial_T q = \frac{1}{2}\sigma^2 \partial_x^2 q - \nu\, \partial_x q$ for $T > 0$, whatever c.

(b) Show that $q_T(b) = 0$ for all $T > 0$ if and only if $c = e^{\lambda b}$. Why does the weight $c = 1$ of the driftless case fail when $\nu \neq 0$?

(c) Show that $\int_b^\infty \varphi(x) q_T(x)\,\mathrm{d}x \to \varphi(0)$ as $T \downarrow 0$ for every bounded continuous φ. Does the formula depend on whether the barrier is above or below the spot, or on whether x is above or below b?

(d) With $c = e^{\lambda b}$, show that $\frac{1}{2}\sigma^2 \partial_x q_T(b) = \frac{|b|}{T}\, p_T(b)$, and that this is the density at T of the first time X reaches b. (Differentiate $\mathbb{Q}(\underline{X}_T > b) = \int_b^\infty q_T$ in T.)

Exercise 16.3 *(Exponential barriers)* [⋆⋆] A down-and-out call is knocked out if $S_t \le B e^{\gamma t}$ for some $t \le T$, with $B < S_0$ and a constant γ (continuous monitoring).

(a) Show that its price is $\widetilde{G}(S_0) - (B/S_0)^{2(\nu-\gamma)/\sigma^2}\, \widetilde{G}(B^2/S_0)$, where $\widetilde{G}(x)$ is the Black–Scholes price of $(S_T - K)^+ \mathbf{1}_{\{S_T > B e^{\gamma T}\}}$ for the spot x. You may use Theorem 16.1.

(b) Evaluate it for the test case with $B = 95$ and $\gamma = r - y$ (a barrier that moves with the forward), and compare with the constant barrier.

(c) For which γ is the image weight equal to 1? Interpret.

Exercise 16.4 *(Rebates paid at the hit)* [⋆⋆] Let $B < S_0$, $b = \log(B/S_0)$, $\tau = \inf\{t : X_t = b\}$, and assume $\nu^2 + 2r\sigma^2 \ge 0$.

(a) Show that τ has the density $f(t) = \frac{|b|}{\sigma\sqrt{2\pi t^3}} \exp\big(-\frac{(b-\nu t)^2}{2\sigma^2 t}\big)$ on $(0,\infty)$ (Exercise 16.2(d) or Theorem 16.1(iii)).

(b) Show that a claim paying 1 at τ if $\tau \le T$ costs

$$\mathbb{E}^{\mathbb{Q}}[e^{-r\tau}\mathbf{1}_{\{\tau\le T\}}] = \Big(\frac{B}{S_0}\Big)^{\frac{\nu-\hat\mu}{\sigma^2}} N\Big(\frac{b-\hat\mu T}{\sigma\sqrt{T}}\Big) + \Big(\frac{B}{S_0}\Big)^{\frac{\nu+\hat\mu}{\sigma^2}} N\Big(\frac{b+\hat\mu T}{\sigma\sqrt{T}}\Big), \qquad \hat\mu = \sqrt{\nu^2 + 2r\sigma^2}.$$

Hint: complete the square in $e^{-rt} f(t)$ to obtain the first-passage density of a Brownian motion with drift $\pm\hat\mu$.

(c) Check the cases $r = 0$ and $T \to \infty$.

(d) For the test case with $B = 95$, price the down-and-out call with a rebate of 5 paid at the hit, and with the same rebate paid at T.

Exercise 16.5 *(Images or eigenfunctions)* [⋆⋆⋆] Take $\nu = 0$, $\sigma = 1$, barriers $a < 0 < u$, $w = u - a$, and the image series $q(x) = \sum_{k\in\mathbb{Z}} \big[\varphi_T(x - 2kw) - \varphi_T(x - 2u - 2kw)\big]$ of Theorem 16.4, with φ_T the $\mathcal{N}(0,T)$ density. You may use the Poisson summation formula $\sum_{k\in\mathbb{Z}} f(x + kP) = \frac{1}{P}\sum_{n\in\mathbb{Z}} \hat f(2\pi n/P) e^{2\pi i n x/P}$ for $f = \varphi_T$, whose Fourier transform is $\hat\varphi_T(\xi) = e^{-\xi^2 T/2}$.

(a) Show that $q(x) = \frac{2}{w}\sum_{n\ge1} e^{-n^2\pi^2 T/(2w^2)} \sin\big(\frac{n\pi(-a)}{w}\big) \sin\big(\frac{n\pi(x-a)}{w}\big)$ on (a,u).

(b) Estimate the number of terms each series needs for a relative accuracy ε, as a function of T/w^2. For which T are they equal?

(c) For the test case with $L = 80$, $U = 120$ (put back σ and the drift; `mfe230d.barriers` has both series), find the number of terms each needs for a maximum error below 10^{-10} times the maximum of the density, for $T = 0.1$, 1 and 10.

(d) Deduce from (a) the exponential rate at which the double knock-out call price decays as $T \to \infty$, and check it numerically.

Exercise 16.6 *(Where the continuity correction fails)* [⋆⋆] This is a computational exercise. Use the test case, the recursion of Theorem 16.6 for exact discrete prices (`mfe230d.barriers` implements it) and Theorem 16.7.

(a) For the down-and-out call with $m = 12$, 52 and 252 dates, compute the relative error of the corrected price for barriers at a distance $\delta = \log(S_0/B)/(\sigma\sqrt{h})$ of $0.25, 0.5, 1, 1.5, 2, 3$. What do you observe, and what rule of thumb follows?

(b) For the up-and-out call with $B = 115$, estimate the order in h of the error of the corrected price from $m = 12, 52, 252$, and compare with the down-and-out call at $B = 85$. Why is the up-and-out call harder?

(c) Correct the down-and-in call through in-out parity. Is its relative error larger or smaller than the down-and-out call's, and why?

Exercise 16.7 *(Lookbacks by simulation)* [⋆⋆] This is a computational exercise. Simulate the test case on a monthly grid ($m = 12$) with 10^6 paths.

(a) Estimate the continuously monitored floating-strike lookback put and the fixed-strike lookback call with $K = 110$, drawing the maximum of the bridge on every step (Theorem 16.9(iii)). Compare with Theorem 16.5.

(b) On the same paths, estimate the discretely monitored versions (maximum over the 13 dates $t_0, \dots, t_{12}$). Compare the fixed-strike call with the prediction of Theorem 16.7(iii).

(c) Use the continuous payoff as a control variate for the discrete one. What variance reduction do you get, and why is it so large?

Exercise 16.8 *(A shifted control)* [⋆⋆] This is a computational exercise. In Table 16.2 the continuous down-and-out call, weighted by the bridge survival probabilities, controls the monthly down-and-out call ($B = 95$) with a variance reduction factor near 6.

(a) Replace the control by the same bridge-weighted payoff with the barrier moved to $B' = 95e^{-\beta\sigma\sqrt{h}}$, whose mean is the closed form at B'. With 10^6 paths, estimate the price, the correlation, the optimal coefficient and the variance reduction factor, with and without the vanilla call as a second control.

(b) Explain the improvement with Theorem 16.7. Is the estimator unbiased?

Exercise 16.9 *(Bridges in a corridor)* [⋆⋆⋆] Let $\ell < 0 < u$, $w = u - \ell$, and $x, z \in (\ell, u)$.

(a) Show that, given $X_s = x$ and $X_{s+h} = z$, the probability that X stays in (ℓ, u) on $[s, s+h]$ is

$$\sum_{k\in\mathbb{Z}} \Big[\exp\Big(-\frac{2kw\,(kw - (z-x))}{\sigma^2 h} \Big) - \exp\Big(-\frac{2(u - x + kw)(u - x + kw - (z-x))}{\sigma^2 h} \Big) \Big],$$

that it does not depend on ν, and that it reduces to Theorem 16.9(i) as $\ell \to -\infty$. You may use Theorem 16.4.

(b) Write the bridge-corrected estimator of a continuously monitored double knock-out call from exact grid values, and prove that it is unbiased.

(c) For the test case with $L = 80$, $U = 120$, run it with $m = 1$ and $m = 12$ steps (4×10^5 paths each) and compare with Theorem 16.4. Which m gives the smaller variance, and why? Compare also with checking the corridor at the 12 dates only.

Chapter 17

FX and multi-asset options

Exercise 17.1 *(Forwards seen from both sides)* [⋆] In the Garman–Kohlhagen model Equation (17.2), let $F = X_0 e^{(r_d - r_f)T}$. You may use Theorem 17.3.

(a) Show that $\mathbb{E}^{\mathbb{Q}}[X_T] = F$ and $\mathbb{E}^{\mathbb{Q}_f}[1/X_T] = 1/F$. Why is each identity the natural one for an investor in the corresponding currency?

(b) Compute $\mathbb{E}^{\mathbb{Q}}[1/X_T]$ and compare it with $1/F$.

(c) Suppose that under the real-world measure $\mathbb{P}$, $\mathrm{d}X_t/X_t = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t^{\mathbb{P}}$. For which μ is F an unbiased forecast of X_T? Can F and $1/F$ be unbiased forecasts of X_T and $1/X_T$ at the same time? (This is Siegel's paradox.)

Exercise 17.2 *(Working under the share measure)* [⋆] Work in the standard model with dividend yield y, and in the two-asset model of Theorem 17.6. You may use Theorem 17.3 and Theorem 17.6.

(a) Theorem 17.5 estimates the call by averaging $S_0 e^{-yT}(1 - K/S_T)^+$ with S_T drawn under $\mathbb{Q}^S$. Show that the variance of one such draw is at most $\frac{1}{4}(S_0 e^{-yT})^2$ for every $K > 0$ and $\sigma > 0$, while the variance of the plain draw $e^{-rT}(S_T - K)^+$ under $\mathbb{Q}$ tends to infinity as $\sigma \to \infty$ with the other inputs fixed. Is the share measure always the better choice?

(b) Let N and N' be two numéraires as in Theorem 17.3. Show that $\mathrm{d}\mathbb{Q}^{N'}/\mathrm{d}\mathbb{Q}^{N} = (N'_T/N_T)/(N'_0/N_0)$.

(c) In Margrabe's formula, show that $N(d_+)$ is the probability that $S_{1,T} > S_{2,T}$ under the measure with numéraire $e^{y_1 t}S_{1,t}$, and $N(d_-)$ the same probability under the one with numéraire $e^{y_2 t}S_{2,t}$. Explain with (b), without computing, why the first is the larger.

Exercise 17.3 *(The better and the worse of two assets)* [⋆⋆] Two assets follow Equation (17.3) with $d = 2$. Payoffs are paid at T.

(a) Express the prices of $\max(S_{1,T}, S_{2,T})$ and of $\min(S_{1,T}, S_{2,T})$ with Margrabe's formula Theorem 17.6, and check that they add up to the right amount.

(b) Show that the price of the best-of is nonincreasing in ρ. What are its limits as $\rho \to -1$ and

as $\rho \to 1$? What happens at $\rho = 1$ when $\sigma_1 = \sigma_2$, and why is this obvious without any formula?

(c) Compute both prices for $S_{1,0} = S_{2,0} = 100$, $r = 1\%$, $y_1 = y_2 = 2\%$, $T = 1$, $\sigma_1 = 20\%$, $\sigma_2 = 30\%$ and $\rho = 0.5$, and check them by simulation.

Exercise 17.4 *(A currency option on a foreign share)* [⋆⋆] Work in the market of Theorem 17.8. An *equity-linked* currency call pays $S_T(X_T - \bar{X})^+$ units of domestic currency: a call on the exchange rate, struck at $\bar{X}$, whose notional is one foreign share.

(a) Show that $N_t = X_t S_t e^{yt}$ is a numéraire as in Theorem 17.3(iii) and find the dynamics of X under $\mathbb{Q}^N$.

(b) Show that the price is

$$X_0 S_0 e^{-yT} N(e_+) - \bar{X} S_0 e^{-(y + r_d - r_f + \rho\sigma_S\sigma_X)T} N(e_-),$$

with $e_\pm = \big(\log(X_0/\bar{X}) + (r_d - r_f + \rho\sigma_S\sigma_X \pm \tfrac{1}{2}\sigma_X^2)T\big)/(\sigma_X\sqrt{T})$, and write it as a Black–Scholes price with "pre-washed" inputs. Interpret the second term with Theorem 17.8(ii).

(c) Evaluate it with the inputs of Theorem 17.9 and $\bar{X} = X_0 = 1.2$, and check it by simulation.

Exercise 17.5 *(Hedging a quanto)* [⋆⋆⋆] Work in the market of Theorem 17.8 with $|\rho| < 1$, now under the real-world measure $\mathbb{P}$, with $\mathrm{d}X_t/X_t = \mu_X^{\mathbb{P}}\,\mathrm{d}t + \sigma_X\,\mathrm{d}W_{X,t}$ and $\mathrm{d}S_t/S_t = \mu_S^{\mathbb{P}}\,\mathrm{d}t + \sigma_S\,\mathrm{d}W_{S,t}$, arbitrary constant drifts, and $\mathrm{d}\langle W_X, W_S\rangle = \rho\,\mathrm{d}t$. Three assets are traded: the foreign share with dividends reinvested (domestic value $P_t^1 = X_t S_t e^{yt}$), the foreign money market ($P_t^2 = X_t e^{r_f t}$) and the domestic money market. Let $V(t,s)$ be smooth, and consider the self-financing portfolio holding ϕ_t units of P^1, ψ_t units of P^2 and the rest of its value in domestic cash.

(a) Show that its value can track $V(t,S_t)$ exactly only if $\phi_t P_t^1 = S_t\,\partial_s V$ and $\psi_t P_t^2 = -S_t\,\partial_s V$: the shares are financed by borrowing foreign currency.

(b) Show that it then tracks $V(t,S_t)$ if and only if

$$\partial_t V + (r_f - y - \rho\sigma_S\sigma_X)\,s\,\partial_s V + \tfrac{1}{2}\sigma_S^2 s^2\,\partial_s^2 V = r_d V,$$

and that the quanto price of Theorem 17.8(ii) solves this equation. Where in your computation does the quanto adjustment come from, and why do the real-world drifts drop out?

(c) A hedger who buys the ϕ_t shares with domestic cash, without the foreign loan, keeps a position in the currency. Write the extra term in the change of the portfolio's value, and say which risk it carries.

Exercise 17.6 *(Static bounds for a cross-currency option)* [⋆⋆] Use the inputs of Theorem 17.7. Consider three at-the-money-forward options with $T = 1$, all valued in USD: the GBP/USD call paying $(X_{1,T} - F_1)^+$; the EUR/USD call and put paying $(X_{2,T} - F_2)^+$ and $(F_2 - X_{2,T})^+$; and the GBP/EUR call paying $(X_{1,T} - F_{12}X_{2,T})^+$, where F_1, F_2 and $F_{12} = F_1/F_2$ are the forwards.

(a) Show that $(a-c)^+ - (b-c)^+ \le (a-b)^+ \le (a-c)^+ + (c-b)^+$ for all real a, b, c. Deduce bounds on the price C_{12} of the GBP/EUR call in terms of the prices C_1, C_2, P_2 of the other three, valid in every arbitrage-free model.

(b) In the Garman–Kohlhagen models for the three rates, with volatilities σ_1, σ_2, σ_{12}, show that C_1, $F_{12}C_2$, $F_{12}P_2$ and C_{12} are all of the form $\kappa\, g(\sigma)$ for one constant κ and $g(\sigma) = 2N(\frac{1}{2}\sigma)-1$. Show that the bounds of (a) admit every $\sigma_{12} \in [|\sigma_1 - \sigma_2|, \sigma_1 + \sigma_2]$.

(c) The GBP/EUR call is quoted at a volatility of 0.5%, and on another day at 21.5%. In each case, give a portfolio of the four options with a nonnegative payoff and a negative price, and that price.

(d) Which quotes outside $[1\%, 21\%]$ admit no arbitrage of the kind in (a)? What does this say about the triangle of Section 17.3?

Exercise 17.7 *(Spread options)* [⋆⋆] This is a computational exercise. Take two assets as in Equation (17.3) with $S_{1,0} = S_{2,0} = 100$, $r = 1\%$, $y_1 = y_2 = 2\%$, $T = 1$, $\sigma_1 = 20\%$, $\sigma_2 = 25\%$, $\rho = 0.6$, and the spread call paying $(S_{1,T} - S_{2,T} - K)^+$.

(a) Price it by plain Monte Carlo for $K = 0, 5, 10, 20$, with 2^{18} draws.

(b) Use the exchange option $(S_{1,T} - S_{2,T})^+$, priced by Theorem 17.6, as a control variate. Report the correlation, the coefficient and the reduction of the standard error for each K, and explain the trend.

(c) *Kirk's approximation* treats $S_{2,T}+K$ as lognormal: with $F_i = \mathbb{E}^{\mathbb{Q}} S_{i,T}$, $u = F_2/(F_2+K)$ and $\sigma_K^2 = \sigma_1^2 - 2\rho\sigma_1\sigma_2 u + \sigma_2^2 u^2$, it prices the spread at $e^{-rT}\,\mathrm{Black}(F_1, F_2 + K, \sigma_K^2 T)$. Compare it with your estimates. Is it exact for $K = 0$?

(d) Price the spread to six decimals by conditioning on $S_{2,T}$ and integrating in one dimension, and use the result to judge (a)–(c). Why can Theorem 17.10(iii) not be used here?

Exercise 17.8 *(Dispersion and the geometric control)* [⋆⋆] This is a computational exercise. Take the basket of Theorem 17.12, with $K = 100$, but with the correlations $\rho_{12} = -0.3$, $\rho_{13} = 0.2$, $\rho_{23} = -0.4$.

(a) Check that this is a correlation matrix, and compute $\sqrt{v_G}$, $\sqrt{v_2}$ and the bound of Theorem 17.10(ii).

(b) Price the call by plain Monte Carlo and with the geometric control, with 2^{18} draws. Compare the correlation and the variance reduction with those of Theorem 17.12.

(c) Compute the two-moment and Gentle approximations and the geometric call, and compare them with a reference price (the method of Theorem 17.11, or a long run with the control).

(d) Explain why lower correlations hurt both the control and the approximations.

Part IV

Beyond Black–Scholes

Chapter 18

Hedging in practice

Exercise 18.1 *(Which claims are hedged exactly?)* [⋆] Let F be Lipschitz and $m \geq 1$. You may use Equation (18.4), the Itô isometry, that V is infinitely differentiable on $[0,T) \times (0,\infty)$ and continuous up to T, and that for $0 \leq s < t$ the conditional law of S_t given $\mathcal{F}_s$ has a positive density on $(0,\infty)$.

(a) Show that $e_m = 0$ for every m if $F(s) = a + bs$.

(b) Conversely, suppose that $e_m = 0$ almost surely for one $m \geq 1$. Show that $\Gamma = 0$ on $[0,T) \times (0,\infty)$, and deduce that F is affine.

(c) Which of the following does the m-date hedge replicate exactly: a forward contract; a long call and a short put with the same strike and maturity; a straddle?

Exercise 18.2 *(When the integrability condition fails)* [⋆⋆] The cash-or-nothing call pays $\mathbf{1}_{\{S_T > K\}}$; its gamma is $\Gamma_d(t,s) = -e^{-r\tau} N'(d_-)\, d_+/(s^2\sigma^2\tau)$. Work under $\mathbb{Q}$ ($\mu = r$).

(a) Show that $\varphi(t) = \mathbb{E}[e^{-2rt} S_t^4 \Gamma_d(t,S_t)^2] \geq c\,(T-t)^{-3/2}$ for t close to T and some constant $c > 0$, so that the condition Equation (18.5) fails. *Hint:* restrict the expectation to the event $\{1 \leq d_+(t,S_t) \leq 2\}$.

(b) For $\varepsilon > 0$ let $F_\varepsilon(s) = \big((s-K)^+ - (s-K-\varepsilon)^+\big)/\varepsilon$, a call spread normalised to pay at most 1. Show that F_ε is Lipschitz and satisfies Equation (18.5), so that Theorem 18.3 applies to it.

(c) (Computational.) Compute $I_\varepsilon = \int_0^T \mathbb{E}[e^{-2rt} S_t^4 \Gamma_\varepsilon(t,S_t)^2]\,\mathrm{d}t$ for $K = 100$ and the widths $\varepsilon = 10$, 1, 0.1 and 0.01, where Γ_ε is the gamma of F_ε. How does I_ε grow as $\varepsilon \to 0$, and what does that say about hedging a tight call spread with the asymptotics of Theorem 18.3? *Hint:* the product of two Gaussian kernels $N'(d_+(K_1))N'(d_+(K_2))$ is a constant times a Gaussian kernel centred at the midpoint of their centres, so the closed form `gamma_moment` in the module `code/src/mfe230d/hedging.py` can be reused.

Exercise 18.3 *(A rule of thumb at the money)* [⋆⋆] Take $r = y = 0$ and $S_0 = K$, and work under $\mathbb{Q}$.

(a) Show that for the call

$$\varphi(t) = \mathbb{E}\big[S_t^4 \Gamma(t, S_t)^2\big] = \frac{K^2 \exp\big(-\sigma^2 T^2/(4(T+t))\big)}{2\pi\sigma^2\sqrt{T^2 - t^2}}, \qquad 0 \le t < T.$$

(b) Let V_0 be the price of the call. Deduce that $\sigma^4 T I/(2V_0^2) \to \pi/4$ as $\sigma^2 T \to 0$.

(c) Conclude, from Theorem 18.3, the traders' rule that for short-dated at-the-money options the standard deviation of the hedging error is about $\sqrt{\pi/(4m)}$ times the premium, or $\sqrt{\pi/(4m)}\,\sigma\mathcal{V}$ with $\mathcal{V}$ the vega. Which two limits have been taken, and in which order?

(d) Evaluate the rule $\sqrt{\pi/4}\,\sigma\mathcal{V}$ for the test case (where $r, y \neq 0$) and compare it with the constant $\sigma^2\sqrt{TI_r/2}$ of the chapter.

Exercise 18.4 *(Hedging the log contract)* [⋆] In the setting of Section 18.2 (bounded μ_t and $0 \le \sigma_t \le \bar{\sigma}$), consider the claim $F(s) = -2\log(s/S_0)$, which is not Lipschitz, so that Theorem 18.5 does not apply as stated.

(a) Show that its Black–Scholes price at volatility σ_h is $V_h(t,s) = -2e^{-r\tau}\big(\log(s/S_0) + (r - y - \tfrac{1}{2}\sigma_h^2)\tau\big)$, and that $s^2\Gamma_h(t,s) = 2e^{-r\tau}$.

(b) Following the proof of Theorem 18.5, show that the dealer who sells the claim for $V_h(0,S_0)$ and holds $\Delta_h(t,S_t)$ shares ends with

$$e^{-rT}\big(X_T - F(S_T)\big) = e^{-rT}\Big(\sigma_h^2 T - \int_0^T \sigma_t^2 \,\mathrm{d}t\Big).$$

Why is the passage to the limit $T' \uparrow T$ easier here than for a call?

(c) Interpret the result as a position in realised variance, and contrast it with a call, for which the same realised variance can give very different P&L.

Exercise 18.5 *(The mean error under the real-world measure)* [⋆⋆] This exercise uses the proof of Theorem 18.3, where $\mathbb{E}[e_m] = (\mu - r)\sum_j \mathbb{E}[D_j]$.

(a) Using the expression for $\mathbb{E}[G_t \tilde{S}_t \mid \mathcal{F}_{t_j}]$ in that proof, argue heuristically that

$$\mathbb{E}[e_m] = -(\mu - r)^2 \,\frac{h}{2} \int_0^T \mathbb{E}\big[e^{-rt} S_t^2 \Gamma(t, S_t)\big]\,\mathrm{d}t + o(h).$$

(b) (Computational.) For the test-case call, evaluate the constant with `gamma_s2_moment` of `code/src/mfe230d/hedging.py` (with $\sigma_h = \sigma_r = \sigma$) for $\mu = 10\%$, 20% and -10%, convert to time-T money, and compare with the means of Table 18.1.

(c) Why is the leading term negative whatever the sign of $\mu - r$? Explain in terms of a hedge ratio that is held fixed while the stock drifts.

Exercise 18.6 *(Leland's adjustment when the gamma changes sign)* [⋆⋆] Let the hedger hold a book whose Black–Scholes gamma $\hat{\Gamma}$ may take both signs, for instance a bull call spread (long the 95 call, short the 105 call).

(a) Explain why, in Theorem 18.8, the leading cost of one rebalance becomes $\frac{1}{2}s^2|\hat{\Gamma}(t,s)|\sigma\sqrt{2h/\pi}$, and why a constant $\hat{\sigma} > \sigma$ then makes the losses add up instead of cancel wherever $\hat{\Gamma} < 0$.

(b) Show that balancing the leading cost and the extra time decay at every (t, s), as in Leland's argument, requires the local hedging variance $\sigma^2(1 + A\operatorname{sgn}\hat{\Gamma})$, with A the Leland number, so that the price should solve

$$\partial_t \hat{V} + \tfrac{1}{2}\sigma^2(1 + A\operatorname{sgn}\partial_{ss}\hat{V})s^2\partial_{ss}\hat{V} + (r - y)s\,\partial_s\hat{V} - r\hat{V} = 0.$$

(c) Why is this equation ill-posed when $A > 1$? For $k = 2\%$ and $\sigma = 20\%$, from how many rebalancing dates per year on is $A > 1$? Is that a problem for the call of Table 18.2?

Exercise 18.7 *(How often to rebalance?)* [⋆] A dealer hedges the test-case call at σ with round-trip cost k and chooses m to minimise $\mathbb{E}[\text{costs}] + a \cdot \text{sd}(\text{error})$, for a price of risk $a > 0$. Use the leading-order approximations of the chapter, in time-T money: expected costs $\alpha\sqrt{m}$ with $\alpha = e^{rT}k\,\mathcal{C}/\sqrt{T}$, $\mathcal{C} = 15.58$, and standard deviation $\beta/\sqrt{m}$ with $\beta = 7.02$.

(a) Find the optimal m^* and the minimal value. How does m^* scale with k and with a?

(b) Evaluate them for $k = 2\%$ and $a = 1$.

(c) (Computational.) Simulate the hedge at the integer nearest to m^* with `hedge_experiment` of `code/src/mfe230d/hedging.py` (50,000 paths) and compare the mean cost and the standard deviation of the error with the approximations. Explain the differences.

Exercise 18.8 *(The exact variance of the discrete hedge)* [⋆⋆⋆] Work under $\mathbb{Q}$ with the call, $Y = e^{-rT}(S_T - K)^+$ and M_m as in Equation (18.3). You may use Girsanov's theorem.

(a) Show that $\mathbb{E}^{\mathbb{Q}}[e_m^2] = \operatorname{Var} Y + \mathbb{E}^{\mathbb{Q}}[M_m^2] - 2\mathbb{E}^{\mathbb{Q}}[Y M_m]$.

(b) Show that $\mathbb{E}^{\mathbb{Q}}[M_m^2] = \sum_j e^{-2yt_j}\mathbb{E}^{\mathbb{Q}}[\delta_j^2\tilde{S}_{t_j}^2](e^{\sigma^2 h} - 1)$.

(c) Show that $\mathbb{E}^{\mathbb{Q}}[(\tilde{S}_{t_{j+1}}/\tilde{S}_{t_j})\,Y \mid \mathcal{F}_{t_j}] = e^{-rt_j}V(t_j, S_{t_j}e^{\sigma^2 h})$, and deduce

$$\mathbb{E}^{\mathbb{Q}}[Y M_m] = \sum_j e^{-(r+y)t_j}\,\mathbb{E}^{\mathbb{Q}}\Big[\delta_j\tilde{S}_{t_j}\big(V(t_j, S_{t_j}e^{\sigma^2 h}) - V(t_j, S_{t_j})\big)\Big].$$

(d) (Computational.) Each expectation is an integral against the law of S_{t_j}. Implement the formula with a quadrature rule of your choice; compute the standard-error ratio $\text{sd}(e_m)/\text{sd}(Y)$ for $m = 1$, 12 and 52, simulated as about 0.48, 0.15 and 0.075 after Theorem 8.16; reproduce the column "exact sd" of Table 18.1, and find from which m on the leading term of Theorem 18.3 is within 1% of the exact standard deviation.

Chapter 19

Stochastic volatility

Exercise 19.1 *(One option is enough)* [⋆] Under $\mathbb{P}$ let $\mathrm{d}S = \mu S\,\mathrm{d}t + S\sqrt{V}\,\mathrm{d}B^S$ and $\mathrm{d}V = \kappa^{\mathbb{P}}(\theta^{\mathbb{P}} - V)\,\mathrm{d}t + \xi\sqrt{V}\,\mathrm{d}B$, with $\mathrm{d}\langle B^S, B\rangle_t = \rho\,\mathrm{d}t$, $|\rho| < 1$, and a dividend yield y. Let $C(t,S,V)$ and $C_1(t,S,V)$ be smooth prices of two options, with $\partial_V C_1 \neq 0$, and let $\mathcal{L}^{\mathbb{P}}$ be the generator of (S,V) under $\mathbb{P}$ (first and second derivatives, no time derivative).

(a) Show that a self-financing portfolio long one C, short Δ_1 units of C_1 and short Δ_S shares (dividends reinvested in cash) is locally riskless for suitable Δ_1 and Δ_S, and deduce that

$$\frac{\partial_t C + \mathcal{L}^{\mathbb{P}} C - (\mu + y - r) S\,\partial_S C - rC}{\xi\sqrt{V}\,\partial_V C}$$

takes the same value $\lambda(t,S,V)$ for every option. (No-arbitrage: a riskless portfolio earns r.)

(b) With $\lambda = \kappa_V\sqrt{V}$ for a constant κ_V, show that every option price solves Equation (19.3) with the parameters $\kappa = \kappa^{\mathbb{P}} + \xi\kappa_V$ and $\theta = \kappa^{\mathbb{P}}\theta^{\mathbb{P}}/\kappa$. Can Feller's condition hold under $\mathbb{P}$ and fail under $\mathbb{Q}$?

(c) Why would hedging with the stock and *two* options, as in the course's notes, be redundant?

Exercise 19.2 *(Degenerate cases)* [⋆] You may use Theorem 19.5 and Theorem 10.12.

(a) Solve Equation (19.4) for $\xi = 0$ and show that the price of any European payoff is its Black–Scholes price with total variance $w = \theta T + (V_0 - \theta)(1 - e^{-\kappa T})/\kappa$. Price the put for $V_0 = 0.09$, $\kappa = 1.5$, $\theta = 0.04$ (and the test case S, K, T, r, y). Does ρ matter?

(b) Let $\rho = -1$. Write S_T as an explicit function of (V_T, Q_T). Use it to estimate the at-the-money put by Monte Carlo from exact variance paths (you may use `stratified.cir_exact` on 64 steps and the trapezoidal rule for Q_T), and compare with the Fourier price.

(c) Let $\rho = 0$. Show that $\varphi(u) = e^{iu(\ln S + (r-y)T)}\,\mathbb{E}\big[e^{-\frac{1}{2}(u^2+iu)Q_T}\big]$, and that for $\rho = 0$ the Riccati equation Equation (19.4) is the equation of the Laplace transform $\mathbb{E}[e^{-\lambda Q_T}]$ with $\lambda = \frac{1}{2}(u^2 + iu)$.

Exercise 19.3 *(The law of the variance)* [⋆⋆] Let $\lambda \geq 0$.

(a) Look for $\mathbb{E}[e^{-\lambda V_t}] = \exp(A(t) + B(t)V_0)$, write the equations for A and B, and solve them:

$$\mathbb{E}[e^{-\lambda V_t}] = (1 + 2\lambda c_t)^{-2\kappa\theta/\xi^2} \exp\Big(- \frac{\lambda e^{-\kappa t} V_0}{1 + 2\lambda c_t} \Big), \qquad c_t = \frac{\xi^2 (1 - e^{-\kappa t})}{4\kappa}.$$

Justify that the guess is the Laplace transform (the local martingale you build is bounded).

(b) Deduce that V_t converges in law, as $t \to \infty$, to the gamma law with shape $2\kappa\theta/\xi^2$ and scale $\xi^2/(2\kappa)$, and give its mean and variance.

(c) How does the stationary density behave near 0? Relate this to Feller's condition. Compute the stationary probability that $V < 0.01$ in the test case and for $\kappa = 0.5$, $\theta = 0.04$, $\xi = 1$.

Exercise 19.4 *(Moment explosions)* [⋆⋆] Use the notation of Theorem 19.5(iv).

(a) For $a(a-1) > 0$, solve the real Riccati equation $\partial_\tau D = \frac{1}{2}\xi^2 D^2 - \kappa_a D + \frac{1}{2}a(a-1)$, $D(0) = 0$, and find its blow-up time in each case of Equation (19.8).

(b) Let $\kappa = 0.5$, $\xi = 1.2$, $\rho = 0.6$. Compute $T^*(2)$ by hand. For $T = 2$, what does this imply for the variance of the plain Monte Carlo estimator of a call, and for the damping in Carr and Madan's formula?

(c) For $\kappa = 0.5$, $\xi = 1$, $\rho = -0.9$, and for the Eurostoxx parameters of the notes, compute $a_+(T) = \sup\{a \geq 1 : \mathbb{E}[S_T^a] < \infty\}$ at $T = 1$ and $T = 10$, the admissible range of the damping α, and Carr and Madan's choice of a quarter of its upper end.

Exercise 19.5 *(Digitals and the skew)* [⋆⋆] A cash-or-nothing call pays $\mathbf{1}_{\{S_T > K\}}$; a gap call with trigger K_1 and strike K_2 pays $(S_T - K_2)\mathbf{1}_{\{S_T > K_1\}}$.

(a) Price the gap call through P_1 and P_2 of Theorem 19.7 at the strike K_1, show that it is a call struck at K_1 plus $K_1 - K_2$ cash-or-nothing calls, and evaluate it for $K_1 = 100$, $K_2 = 95$, $\rho = -0.7$.

(b) Compute the cash-or-nothing call at $K = 100$ for $\rho = -0.7$ and $\rho = 0$, and compare it with the Black–Scholes digital evaluated at the at-the-money implied volatility.

(c) Explain the difference with Theorem 3.8, and check it numerically with a finite difference of implied volatilities.

Exercise 19.6 *(Fixing Euler's scheme)* [⋆⋆] This is a computational exercise. Implement the four fixes of the notes (absorption, reflection, partial and full truncation) with the log-Euler step for S.

(a) Show that, with the log-Euler step, every fix reproduces the forward exactly, and that the conditional estimator of Theorem 19.8(ii) applies to every fix.

(b) Estimate the bias of the at-the-money put for $h = \frac{1}{4}$ and $h = \frac{1}{16}$, in the test case with $\rho = -0.7$ and for $V_0 = \theta = 0.04$, $\kappa = 0.5$, $\xi = 1$, $\rho = -0.9$, against the Fourier prices 7.948022 and 4.576954. Report standard errors and seeds.

(c) Rank the fixes. Can you explain the ranking by comparing the drift that each fix applies when $\tilde{V} < 0$?

Exercise 19.7 *(Calibrating to a synthetic smile)* [⋆⋆⋆] This is a computational exercise. You may use `scipy.optimize.least_squares` and the pricers of the chapter.

(a) Generate implied volatilities of out-of-the-money options from the test case with $\rho = -0.7$ for $K \in \{80, 90, 95, 100, 105, 110, 120\}$ and $T \in \{0.25, 0.5, 1, 2\}$, and add independent normal noise with standard deviation 0.1% (in volatility points), with a fixed seed.

(b) Calibrate $(V_0, \kappa, \theta, \xi, \rho)$ by least squares on implied volatilities from a start far from the truth. Report the estimates, the root-mean-square error, and approximate standard errors from the Jacobian. Which parameter is least well determined?

(c) Now use only the slice $T = 1$, fix $\kappa \in \{0.5, 2, 8\}$ in turn and calibrate the other four. Compare the fits and the parameters, and the Feller ratio of each fit. What do you conclude about identifying the model from one maturity?

Exercise 19.8 *(A formula without damping)* [⋆⋆⋆] Let $k = \ln K$ and φ as in the notes.

(a) Show that $\int_{\mathbb{R}} e^{ivz}(v^2 + \tfrac{1}{4})^{-1}\,\mathrm{d}v = 2\pi e^{-|z|/2}$ for real z, and deduce that $\min(S_T, K) = \frac{\sqrt{S_T K}}{2\pi}\int_{\mathbb{R}} e^{iv(X_T - k)}(v^2 + \tfrac{1}{4})^{-1}\,\mathrm{d}v$.

(b) Deduce Lewis's (2001) formula

$$C = S e^{-yT} - \frac{\sqrt{K}\, e^{-rT}}{\pi}\int_0^\infty \frac{\operatorname{Re}\left(e^{-ivk}\varphi(v - \tfrac{i}{2})\right)}{v^2 + \tfrac{1}{4}}\,\mathrm{d}v,$$

and explain why, unlike Carr and Madan's, it needs no condition on the moments of S_T beyond those that always hold.

(c) Obtain the same formula by moving the line of integration of Carr and Madan's Equation (19.10) from $\operatorname{Im} = -(1 + \alpha)$ to $\operatorname{Im} = -\tfrac{1}{2}$, and identify the residue you cross.

(d) Implement it with the midpoint rule on $[0, 125]$ and check the put 7.948021924422 of the notes ($\rho = -0.7$). How small must the step be, and why is it smaller than for Gil-Pelaez's formula?

Chapter 20

Trees and finite differences

Exercise 20.1 *(When is a tree a model?)* [⋆] You may use Theorem 20.2(i).

(a) Show that the Jarrow–Rudd tree, $u, d = e^{\nu\Delta t \pm \sigma\sqrt{\Delta t}}$, has $q \in (0,1)$ if and only if $\sigma\sqrt{\Delta t} < 2$, whatever $r - y$, and that $q = \frac{1}{2} + \frac{1}{24}\sigma^3\Delta t^{3/2} + O(\Delta t^2)$. Compare with the CRR condition when $\sigma \to 0$ with $r \neq y$ fixed.

(b) Take the CRR tree with $n = 1$, $T = 1$, $S_0 = K = 100$, $r = 5\%$, $y = 0$, $\sigma = 1\%$. Compute q and the "price" Equation (20.2) of the put. Exhibit an arbitrage in the one-period model with these u and d.

(c) Show that two CRR steps of length $\Delta t/2$ make one step of a trinomial tree with moves $\pm h$ and 0. Find h and the stretch κ of Equation (20.4), and show that the three probabilities agree with Equation (20.4) up to $O(\Delta t)$.

Exercise 20.2 *(The binomial formula and the Leisen–Reimer tree)* [⋆⋆] Let a tree as in Theorem 20.1 have $q \in (0,1)$, and let a be the smallest number of up moves with $S_0 u^a d^{n-a} > K$.

(a) Show that the European call in the tree is worth

$$S_0 e^{-yT}\Psi(a; n, q') - Ke^{-rT}\Psi(a; n, q), \qquad q' = que^{-(r-y)\Delta t},$$

where $\Psi(a; n, p)$ is the probability that a binomial variable with parameters n and p is at least a. Interpret q' as the probability of an up move under a change of measure.

(b) For the Leisen–Reimer tree of Theorem 20.2(iv)(b), show that $qu + (1-q)d = e^{(r-y)\Delta t}$, that q is the risk-neutral probability, and that $q' = h_n^{-1}(d_+)$.

(c) (Computational.) On the test case with $K = 110$, compute the Leisen–Reimer call for $n = 25,\ 51,\ 101,\ 201,\ 401$, estimate the order of convergence, and check that K lies between the two middle terminal nodes. How many CRR steps give the accuracy of the Leisen–Reimer tree with $n = 25$, and what does that mean for the cost of an American option priced by backward induction?

Exercise 20.3 *(Digitals on a tree)* [⋆⋆] This is a computational exercise. Let D_n be the CRR price of the digital call paying $\mathbf{1}\{S_T > K\}$, and $D = e^{-rT}N(d_-)$.

(a) Repeat the heuristic of Theorem 20.3 for $F(x) = f(x)\mathbf{1}\{x > \log K\}$, with $\alpha_n \in (0,1]$ the offset, in units of $H = 2\delta$, of the first terminal node strictly above $\log K$. Show that it suggests $\sqrt{n}\,(D_n - D) \approx -2e^{-rT}N'(d_-)(\alpha_n - \frac{1}{2})$.

(b) On the test case with $K = 110$, compute $\sqrt{n}(D_n - D)$ for $n = 1000, \dots, 1200$ and compare with (a).

(c) What happens when K is a node? Explain why the computed price then depends on how the payoff is evaluated at that node, and which value is natural.

(d) Compute $n(D_n^{\mathrm{BBS}} - D)$ for the BBS tree of Theorem 20.2(iv)(a), with $K = 100$ and $K = 110$, for a few n between 100 and 1600. What has the smoothing achieved?

Exercise 20.4 *(Small volatility and upwinding)* [⋆⋆] Consider Equation (20.1) with $\nu > 0$ and small σ, on a grid of mesh h.

(a) Show that central differences give $p_- < 0$ as soon as $h\nu > \sigma^2$, so that Theorem 20.6(iv) does not apply for any time step. What does the condition mean in terms of the distance νk that the drift moves the log price in one step, compared with the spread $\sigma\sqrt{k}$?

(b) The *upwind* scheme replaces $\nu\,\partial_x u$ by $\nu(U_{j+1} - U_j)/h$. Show that its implicit version is positive for all k and h. Why is the forward difference the right one for $\nu > 0$?

(c) Show that the upwind operator is consistent of order one only, and that it is consistent of order two with the equation in which σ^2 is replaced by $\sigma^2 + \nu h$.

(d) Take $S_0 = 100$, $T = 1$, $r = 5\%$, $y = 0$, $\sigma = 1\%$ and $h = 0.01$. Find the largest h for which central differences are positive, the minimal number of CRR steps, and the effective volatility of (c). Price the calls with $K = 100$ and $K = 105$ with the implicit upwind scheme and compare with the Black–Scholes prices at σ and at the effective volatility.

Exercise 20.5 *(Positivity of the θ-scheme)* [⋆⋆] Consider the θ-scheme for $\partial_\tau u = \frac{1}{2}\sigma^2\partial_{xx}u$ on $h\mathbb{Z}$, with $\lambda = \sigma^2 k/(2h^2)$ and $\theta \in (0,1)$. You may use the integral and the Neumann-series argument in the proof of Theorem 20.6(iv).

(a) Show that $C = \theta^{-1}I - \theta^{-1}(1-\theta)B$, and deduce that G is positive if and only if $(B^{-1})_{00} \geq 1 - \theta$.

(b) Show that this holds if and only if $\lambda \leq \frac{2-\theta}{4(1-\theta)^2}$. Check $\theta = \frac{1}{2}$ and the limits $\theta \to 0$ and $\theta \to 1$.

(c) Compare, for $\theta = 0.5, 0.6, 0.75, 0.9$, this condition with the sufficient condition of Theorem 20.6(iv) and with the stability condition of Theorem 20.6(i).

(d) On a finite grid with zero boundary values, B_J is the restriction of B to the nodes $1, \dots, J-1$. Show that $B_J^{-1} \leq (B^{-1})_{\{1,\dots,J-1\}}$ entry by entry, and conclude that positivity on the finite grid can only require a smaller λ.

(e) Let $\theta = \frac{1}{2}$ and consider the half-line $\{1, 2, \dots\}$ with a zero value at node 0. Writing $(B^{-1})_{jl} = \beta_{j-l}$ on $h\mathbb{Z}$, show by the method of images that the inverse on the half-line has entries $\beta_{j-l} - \beta_{j+l}$. Given that $\beta_m = \rho^{|m|}/\sqrt{1+2\lambda}$ with $\rho = (\sqrt{1+2\lambda} - 1)/(\sqrt{1+2\lambda} + 1)$,

show that G is positive on the half-line if and only if $\lambda \le 4 - 2\sqrt{2}$, and that the node next to the boundary is the one that binds.

Exercise 20.6 *(How far is far enough?)* [⋆] This is a computational exercise. Price the test-case call on the grid $\log S_0 \pm m\sigma\sqrt{T}$ with the crude boundary values 0 at both ends.

(a) Use Theorem 20.9 with its Gaussian bound to find the smallest m among $6, 6.25, 6.5, 6.75$ for which the truncation error at S_0 is guaranteed below 10^{-6}. (Compute ε as the largest call price at the two ends over $t \le T$.)

(b) Compute the actual truncation error at S_0 for these m, as in Figure 20.3(b), with $h = 0.003$ and 400 Crank–Nicolson steps with Rannacher's start, and compare it with the bound computed with the exact exit probability. Why is the bound not sharp?

Exercise 20.7 *(The penalty method)* [⋆ ⋆ ⋆] Let B be as in Theorem 20.10: $B_{jl} \le 0$ for $j \ne l$, every row sum at least 1. Let z^* solve Equation (20.8), and let z_ρ, for $\rho > 0$, solve $Bz = q + \rho(g - z)^+$ (you may assume that it exists).

(a) Show that $z_\rho \le z^*$. *Hint:* look at a positive maximum of $z_\rho - z^*$.

(b) Let $\varepsilon_\rho = \max_j (g_j - z_{\rho,j})^+$. Show that $z_\rho + \varepsilon_\rho \mathbf{1}$ is a supersolution, so that $z^* \le z_\rho + \varepsilon_\rho \mathbf{1}$.

(c) Show that $\varepsilon_\rho \le \|(Bg - q)^+\|_\infty / \rho$, and that the penalized equation has at most one solution.

(d) (Computational.) For the first implicit half-step of the test-case put ($J = 800$, $N = 400$, Rannacher's start), compute $\|(Bg - q)^+\|_\infty$ and compare $\max_j (z^*_j - z_{\rho,j})$ with the bound for $\rho = 10^2, 10^4, 10^6$. Where is $(Bg - q)^+$ largest, and why?

Exercise 20.8 *(An American straddle)* [⋆ ⋆ ⋆] This is a computational exercise. Consider the American option paying $|S - K|$ on exercise, with $S_0 = K = 100$, $T = 1$, $r = 3\%$, $y = 6\%$, $\sigma = 25\%$, on the Crank–Nicolson grid with $J = 800$, $N = 400$ and Rannacher's start.

(a) Compute its price with PSOR and with the penalty method, and the European price. Find the exercise region at $t = 0$ and explain why it has two parts.

(b) Run Brennan–Schwartz with the exercise region assumed at the bottom, and then at the top. Report the prices and the largest residuals of Equation (20.8). Point to the step of the proof of Theorem 20.10(iii) that fails.

(c) Suppose that the exercise set of one time step is $\{j \le j_1\} \cup \{j \ge j_2\}$ with $j_1 < j_2$, and that j_2 is known. Show that Brennan–Schwartz with the exercise region at the bottom, applied to the rows $j < j_2$ with the values $z_j = g_j$, $j \ge j_2$, moved to the right side, returns z^* on those rows.

(d) Turn (c) into an algorithm that alternates the two one-sided versions, each with the other's exercise region frozen, and test it on this option.