MATH 505B HOMEWORK SOLUTIONS

STEVEN HEILMAN

Contents

1. Homework 1

Exercise 1.2 (Continuity of a Probability Law). Let $\mathbf{P}$ be a probability law on a sample space $\mathcal{C}$. Let $A_1, A_2, \ldots$ be sets in $\mathcal{C}$ which are increasing, so that $A_1 \subseteq A_2 \subseteq \cdots$. Then

$$\lim_{n\to\infty} \mathbf{P}(A_n) = \mathbf{P}(\cup_{n=1}^{\infty} A_n).$$

In particular, the limit on the left exists. Similarly, let $A_1, A_2, \ldots$ be sets in $\mathcal{C}$ which are decreasing, so that $A_1 \supseteq A_2 \supseteq \cdots$. Then

$$\lim_{n\to\infty} \mathbf{P}(A_n) = \mathbf{P}(\cap_{n=1}^{\infty} A_n).$$

Solution. First, recall that $A \smallsetminus B := A \cap B^c$ where $A, B \subseteq \mathcal{C}$. Now, let $B_1 := A_1$, let $B_2 := A_2 \smallsetminus A_1$, and for any $n \geq 1$, inductively define $B_n := A_n \smallsetminus A_{n-1}$. We claim that $B_1, B_2, \ldots$ are disjoint, and $\cup_{n=1}^{k} A_n = \cup_{n=1}^{k} B_n$ for any $1 \leq k \leq \infty$.

To see the first statement, let $i, j \geq 1$ with $i > j$. Since $i-1 \geq j$, $A_j \subseteq A_{i-1}$, so $A_{i-1}^c \cap A_j = \emptyset$. So

$$B_i \cap B_j = (A_i \smallsetminus A_{i-1}) \cap (A_j \smallsetminus A_{j-1}) = A_i \cap A_{i-1}^c \cap A_j \cap A_{j-1}^c = \emptyset.$$

Date: May 1, 2021

To see the second statement, let $x\in\cup_{n=1}^{k}A_n$. Let $m\geq 1$ such that $m=\min\{1\leq n\leq k\colon x\in A_n\}$. If $m=1$, then $x\in B_1=A_1$. If $m>1$, then $x\notin A_{m-1}$ so $x\in B_m=A_m\smallsetminus A_{m-1}$. So, in any case, $x\in\cup_{n=1}^{k}B_n$. For the reverse inclusion, let $x\in\cup_{n=1}^{k}B_n$. Then $x\in B_n$ for some $n\geq 1$. So $x\in A_n$ since $B_n\subseteq A_n$. So, $x\in\cup_{n=1}^{k}A_n$. The claim is proven.

Now, using our claim, we have by the second axiom for probability laws,

$$\begin{aligned}\mathbf{P}(\cup_{n=1}^{\infty}A_n)&=\mathbf{P}(\cup_{n=1}^{\infty}B_n)=\sum_{n=1}^{\infty}\mathbf{P}(B_n)=\lim_{k\to\infty}\sum_{n=1}^{k}\mathbf{P}(B_n)\\&=\lim_{k\to\infty}\mathbf{P}(\cup_{n=1}^{k}B_n)=\lim_{k\to\infty}\mathbf{P}(\cup_{n=1}^{k}A_n)=\lim_{k\to\infty}\mathbf{P}(A_k).\end{aligned}$$

The last line used $A_k\supseteq A_{k-1}\supseteq\cdots\supseteq A_1$.

Applying the above result to A_n^c for any $n\geq 1$, and then apply De Morgan's law:

$$\lim_{n\to\infty}\mathbf{P}(A_n)=1-\lim_{n\to\infty}\mathbf{P}(A_n^c)=1-\mathbf{P}(\cup_{n=1}^{\infty}A_n^c)=\mathbf{P}(\cap_{n=1}^{\infty}A_n).$$

□

Exercise 1.3. Let $0<p\leq\infty$. Show that, if $Y_1,Y_2,\ldots\colon\mathcal{C}\to\mathbb{R}$ converge to $Y\colon\mathcal{C}\to\mathbb{R}$ in L_p, then $Y_1,Y_2,\ldots$ converges to Y in probability.

Then, show that the converse is false.

Solution. The first part follows from Markov's inequality since $\mathbf{P}(|Y_n-Y|>\varepsilon)\leq\varepsilon^{-p}\mathbf{E}\,|Y_n-Y|^p$ for all $\varepsilon>0$, $\forall\ n\geq 1$. For the second part, fix $0<p<\infty$. Consider $\mathcal{C}=[0,1]$, $\mathbf{P}$ uniform on $\mathcal{C}$ and $Y_n(t):=n^{1+1/p}1_{(0,1/n]}(t)$ for all $n\geq 1$ $\forall\ t\in[0,1]$. Then $Y_1,Y_2,\ldots$ converges in probability to 0, since $\mathbf{P}(Y_n\neq 0)=1/n$ for all $n\geq 1$, but Y_1,Y_2 does not converge in L_p since $\mathbf{E}\,|Y_n|^p=n^p\to\infty$ as $n\to\infty$. So, if $Y\in L_p$ then $\|Y_n-Y\|_p\geq\|Y_n\|_p-\|Y\|_p\to\infty$ as $n\to\infty$. □

Exercise 1.4. Suppose random variables $Y_1,Y_2,\ldots\colon\mathcal{C}\to\mathbb{R}$ converge in probability to a random variable $Y\colon\mathcal{C}\to\mathbb{R}$. Prove that $Y_1,Y_2,\ldots$ converge in distribution to Y.

Then, show that the converse is false.

Exercise 1.5. Prove the following statement. Almost sure convergence does not imply convergence in L_2, and convergence in L_2 does not imply almost sure convergence. That is, find random variables that converge in L_2 but not almost surely. Then, find random variables that converge almost surely but not in L_2.

Solution. We first find random variables that converge in L_2 but not almost surely. We can do this by a "traveling bump" construction with "decreasing width." Let $\mathcal{C}=[0,1]$ with $\mathbf{P}$ uniform on $\mathcal{C}$. For any integer $n\geq 1$, write $n=2^j+b$ where b is a positive integer with $b<2^j$ and $j\geq 0$ is a nonnegative integer. Let $X_n\colon[0,1]\to\{0,1\}$ be the function

$$X_n:=1_{[2^{-j}b,2^{-j}(b+1)]}.$$

As $n\to\infty$, $j\to\infty$, so $\mathbf{E}X_n^2\to 0$ as $n\to\infty$, so that $X_1,X_2,\ldots$ converges to 0 in L_2. However, $X_1,X_2,\ldots$ does not converge almost surely since, for any $t\in[0,1]$, the sequence $X_1(t),X_2(t),\ldots$ has infinitely many zero and one values.

To find random variables that converge almost surely but not in L_2, we re-use a previous construction. Consider $\mathcal{C} = [0,1]$, $\mathbf{P}$ uniform on $\mathcal{C}$ and $Y_n(t) := n1_{(0,1/n]}(t)$ for all $n \geq 1$ $\forall$ $t \in [0,1]$. Then $Y_1, Y_2, \ldots$ converges almost surely to 0, $\lim_{n\to\infty} Y_n(t) = 0$ for all $t \in [0,1]$, but Y_1, Y_2 does not converge in L_2 since $\mathbf{E}\,|Y_n|^2 = n \to \infty$ as $n \to \infty$. So, if $Y \in L_2$ then $||Y_n - Y||_2 \geq ||Y_n||_2 - ||Y||_2 \to \infty$ as $n \to \infty$. □

2. Homework 2

Exercise 2.1. Let P, Q be stochastic matrices of the same size. Show that PQ is a stochastic matrix. Conclude that, if r is a positive integer, then P^r is a stochastic matrix.

Solution. Since each entry of P and Q is nonnegative, each entry of PQ is also nonnegative. Also, if P, Q are $n \times n$ matrices, then for any $1 \leq i \leq n$,

$$\sum_{j=1}^{n} (PQ)_{ij} = \sum_{j=1}^{n} \sum_{k=1}^{n} P_{ik} Q_{kj}.$$

Since Q is stochastic, $\sum_{j=1}^{n} Q_{kj} = 1$. So, by switching the order of summation, we have

$$\sum_{j=1}^{n} (PQ)_{ij} = \sum_{k=1}^{n} P_{ik} \sum_{j=1}^{n} Q_{kj} = \sum_{k=1}^{n} P_{ik} = 1.$$

In the last line, we used that P is stochastic. We can now conclude that P^r is stochastic by induction on r, since $P^r = P^{r-1}P$. By assumption $P^1 = P$ is stochastic (verifying the base case), and the inductive hypothesis assumes P^{r-1} is stochastic, so that P^r is stochastic, since it is the product of two stochastic matrices. □

Exercise 2.2. Let A, B be events in a sample space. Let $C_1, \ldots, C_n$ be events such that $C_i \cap C_j = \emptyset$ for any $i, j \in \{1, \ldots, n\}$ with $i \neq j$, and such that $\cup_{i=1}^{n} C_i = \mathcal{C}$. Show:

$$\mathbf{P}(A|B) = \sum_{i=1}^{n} \mathbf{P}(A|B,\, C_i)\mathbf{P}(C_i|B).$$

(Hint: consider using the Total Probability Theorem and that $\mathbf{P}(\cdot|B)$ is a probability law.)

Solution. From the Total Probability Theorem applied to $\mathbf{P}(\cdot|B)$, and then using the definition of conditional probability,

$$\begin{aligned}\mathbf{P}(A|B) &= \sum_{i=1}^{n} \mathbf{P}(A \cap C_i|B) = \sum_{i=1}^{n} \frac{\mathbf{P}(A \cap B \cap C_i)}{\mathbf{P}(B)} \\ &= \sum_{i=1}^{n} \frac{\mathbf{P}(A \cap B \cap C_i)}{\mathbf{P}(B \cap C_i)} \frac{\mathbf{P}(B \cap C_i)}{\mathbf{P}(B)} = \sum_{i=1}^{n} \mathbf{P}(A|B,\, C_i)\mathbf{P}(C_i|B).\end{aligned}$$

□

Exercise 2.3. Let $0 < p, q < 1$. Let $P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$. Find the (left) eigenvectors of P, and find the eigenvalues of P. By writing any row vector $x \in \mathbb{R}^2$ as a sum of eigenvectors of

P (whenever possible), find an expression for xP^n for any $n \geq 1$. What is $\lim_{n\to\infty} xP^n$? Is it related to the vector $\pi = (q/(p+q), p/(p+q))$?

Solution. The matrix P has trace $2-p-q$ and determinant $(1-p)(1-q)-pq = 1-p-q$. So, the eigenvalues a, b of P satisfy $ab = 1-p-q$ and $a+b = 2-p-q$. So, $a = 1$ and $b = 1-p-q$. The corresponding (left) eigenvectors are (q,p) and $(1,-1)$. If $p \neq q$, then these vectors form a basis of $\mathbb{R}^2$. So, any vector $x \in \mathbb{R}^2$ can be written as $x = \alpha(q,p) + \beta(1,-1)$, for some $\alpha, \beta \in \mathbb{R}$. Then by definition of (left) eigenvector,

$$xP^n = \alpha(q,p) + \beta(1-p-q)^n(1,-1).$$

So, $\lim_{n\to\infty} xP^n = \alpha(q,p)$. That is, xP^n becomes proportional to the probability distribution $\pi = (q/(p+q), p/(p+q))$ □

Exercise 2.4. Let $G = (V, E)$ be a graph. Let $|E|$ denote the number of elements in the set E, i.e. $|E|$ is the number of edges of the graph. Prove: $\sum_{x\in V} \deg(x) = 2|E|$.

Solution. Let $x \in V$. Then $\deg(x)$ is the number of edges emanating from x. Fix an edge $e \in E$. Then $e = \{x, y\}$ where $x, y \in V$, $x \neq y$. As we sum over all $x \in V$ in $\sum_{x\in V} \deg(x)$, any fixed edge $e \in E$ is counted exactly twice (once for x, once for y, and never again). So, $\frac{1}{2}\sum_{x\in V} \deg(x) = |E|$, as desired. □

Exercise 2.5. Let M, N be stopping times for a Markov chain $X_0, X_1, \ldots$. Show that $\max(M, N)$ and $\min(M, N)$ are stopping times. In particular, if $n \geq 0$ is fixed, then $\max(M, n)$ and $\min(M, n)$ are stopping times

Exercise 2.6. Let A, B be events such that $B \subseteq \{X_0 = x_0\}$. Then $\mathbf{P}(A|B) = \mathbf{P}_{x_0}(A|B)$.

More generally, if A, B are events, then $\mathbf{P}_{x_0}(A|B) = \mathbf{P}(A|B, X_0 = x_0)$.

Solution. Using the definition of conditional probability, and then the definition of $\mathbf{P}_{x_0}$,

$$\begin{aligned}\mathbf{P}_{x_0}(A \mid B) = \mathbf{P}_{x_0}(A \cap B)/\mathbf{P}_{x_0}(B) &= \frac{\mathbf{P}(A \cap B \mid X_0 = x_0)}{\mathbf{P}(B \mid X_0 = x_0)} \\ &= \frac{\mathbf{P}(A \cap B \cap \{X_0 = x_0\})}{\mathbf{P}(B \cap \{X_0 = x_0\})} = \mathbf{P}(A \mid B, X_0 = x_0).\end{aligned}$$

Now, since $B \subseteq \{X_0 = x_0\}$, we get

$$\mathbf{P}_{x_0}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \mathbf{P}(A \mid B).$$

□

Exercise 2.7. Suppose we have a Markov Chain with state space Ω. Let $n \geq 0$, $\ell \geq 1$, let $x_0, \ldots, x_n \in \Omega$ and let $A \subseteq \Omega^\ell$. Using the (usual) Markov property, show that

$$\begin{aligned}&\mathbf{P}((X_{n+1}, \ldots, X_{n+\ell}) \in A \mid (X_0, \ldots, X_n) = (x_0, \ldots, x_n)) \\ &= \mathbf{P}((X_{n+1}, \ldots, X_{n+\ell}) \in A \mid X_n = x_n).\end{aligned}$$

Then, show that

$$\mathbf{P}((X_{n+1}, \ldots, X_{n+\ell}) \in A \mid X_n = x_n) = \mathbf{P}((X_1, \ldots, X_\ell) \in A \mid X_0 = x_n).$$

(Hint: it may be helpful to use the Multiplication Rule.)

Solution. Let $x_0, \ldots, x_{n+\ell} \in \Omega$. Then by the Multiplication rule, (Proposition 2.7 in the notes),

$$\begin{aligned}
&\mathbf{P}(X_{n+1} = x_{n+1}, \ldots, X_{n+\ell} = x_{n+\ell} \,|\, (X_0, \ldots, X_n) = (x_0, \ldots, x_n)) \\
&= \mathbf{P}(X_0 = x_0, \ldots, X_{n+\ell} = x_{n+\ell})/\mathbf{P}(X_0 = x_0, \ldots, X_n = x_n) \\
&= \mathbf{P}(X_{n+\ell} = x_{n+\ell} \,|\, X_0 = x_0, \ldots, X_{n+\ell-1} = x_{n+\ell-1}) \\
&\quad \cdot \mathbf{P}(X_{n+\ell-1} = x_{n+\ell-1} \,|\, X_0 = x_0, \ldots, X_{n+\ell-2} = x_{n+\ell-2}) \\
&\quad\quad \cdots \mathbf{P}(X_{n+1} = x_{n+1} \,|\, X_0 = x_0, \ldots, X_n = x_n).
\end{aligned}$$

From the Markov property, we get

$$\begin{aligned}
&\mathbf{P}(X_{n+1} = x_{n+1}, \ldots, X_{n+\ell} = x_{n+\ell} \,|\, (X_0, \ldots, X_n) = (x_0, \ldots, x_n)) \\
&= \mathbf{P}(X_{n+\ell} = x_{n+\ell} \,|\, X_{n+\ell-1} = x_{n+\ell-1}) \\
&\quad \cdot \mathbf{P}(X_{n+\ell-1} = x_{n+\ell-1} \,|\, X_{n+\ell-2} = x_{n+\ell-2}) \cdots \mathbf{P}(X_{n+1} = x_{n+1} \,|\, X_n = x_n).
\end{aligned}$$

Using the Multiplication rule and the Markov property again,

$$\begin{aligned}
&\mathbf{P}(X_{n+1} = x_{n+1}, \ldots, X_{n+\ell} = x_{n+\ell} \,|\, X_n = x_n)) \\
&= \mathbf{P}(X_n = x_n, \ldots, X_{n+\ell} = x_{n+\ell})/\mathbf{P}(X_n = x_n) \\
&= \mathbf{P}(X_{n+\ell} = x_{n+\ell} \,|\, X_n = x_n, \ldots, X_{n+\ell-1} = x_{n+\ell-1}) \\
&\quad \cdot \mathbf{P}(X_{n+\ell-1} = x_{n+\ell-1} \,|\, X_n = x_n, \ldots, X_{n+\ell-2} = x_{n+\ell-2}) \cdots \mathbf{P}(X_{n+1} = x_{n+1} \,|\, X_n = x_n) \\
&= \mathbf{P}(X_{n+\ell} = x_{n+\ell} \,|\, X_{n+\ell-1} = x_{n+\ell-1}) \\
&\quad \cdot \mathbf{P}(X_{n+\ell-1} = x_{n+\ell-1} \,|\, X_{n+\ell-2} = x_{n+\ell-2}) \cdots \mathbf{P}(X_{n+1} = x_{n+1} \,|\, X_n = x_n).
\end{aligned}$$

Combining the above, we get

$$\begin{aligned}
&\mathbf{P}(X_{n+1} = x_{n+1}, \ldots, X_{n+\ell} = x_{n+\ell} \,|\, (X_0, \ldots, X_n) = (x_0, \ldots, x_n)) \\
&\quad \mathbf{P}(X_{n+1} = x_{n+1}, \ldots, X_{n+\ell} = x_{n+\ell} \,|\, X_n = x_n)).
\end{aligned}$$

Summing over all disjoint points $(x_{n+1}, \ldots, x_{n+\ell}) \in A$ then proves the first assertion.

$$\begin{aligned}
&\mathbf{P}((X_{n+1}, \ldots, X_{n+\ell}) \in A \,|\, (X_0, \ldots, X_n) = (x_0, \ldots, x_n)) \\
&\quad = \mathbf{P}((X_{n+1}, \ldots, X_{n+\ell}) \in A \,|\, X_n = x_n).
\end{aligned}$$

For the second assertion, note that we showed above

$$\begin{aligned}
&\mathbf{P}(X_{n+1} = x_{n+1}, \ldots, X_{n+\ell} = x_{n+\ell} \,|\, X_n = x_n)) \\
&= \mathbf{P}(X_{n+\ell} = x_{n+\ell} \,|\, X_{n+\ell-1} = x_{n+\ell-1}) \\
&\quad \cdot \mathbf{P}(X_{n+\ell-1} = x_{n+\ell-1} \,|\, X_{n+\ell-2} = x_{n+\ell-2}) \cdots \mathbf{P}(X_{n+1} = x_{n+1} \,|\, X_n = x_n). \qquad (*)
\end{aligned}$$

Using the definition of the transition matrix P,

$$\begin{aligned}&\mathbf{P}(X_{n+\ell}=x_{n+\ell}\,|\,X_{n+\ell-1}=x_{n+\ell-1})\\&\qquad\cdot\mathbf{P}(X_{n+\ell-1}=x_{n+\ell-1}\,|\,X_{n+\ell-2}=x_{n+\ell-2})\cdots\mathbf{P}(X_{n+1}=x_{n+1}\,|\,X_n=x_n)\\&=P(x_{n+\ell-1},x_{n+\ell})\cdots P(x_n,x_{n+1})\\&=\mathbf{P}(X_\ell=x_\ell\,|\,X_{\ell-1}=x_{n+\ell-1})\\&\qquad\cdot\mathbf{P}(X_{\ell-1}=x_{n+\ell-1}\,|\,X_{\ell-2}=x_{n+\ell-2})\cdots\mathbf{P}(X_1=x_{n+1}\,|\,X_0=x_n)\end{aligned}$$

Then, using equation $(*)$ for $n=0$,

$$\begin{aligned}&\mathbf{P}(X_\ell=x_\ell\,|\,X_{\ell-1}=x_{n+\ell-1})\\&\qquad\cdot\mathbf{P}(X_{\ell-1}=x_{n+\ell-1}\,|\,X_{\ell-2}=x_{n+\ell-2})\cdots\mathbf{P}(X_1=x_{n+1}\,|\,X_0=x_n)\\&=\mathbf{P}(X_1=x_{n+1},\ldots,X_\ell=x_{n+\ell}\,|\,X_0=x_n))\end{aligned}$$

Combining these three equalities,

$$\mathbf{P}(X_{n+1}=x_{n+1},\ldots,X_{n+\ell}=x_{n+\ell}\,|\,X_n=x_n))=\mathbf{P}(X_1=x_{n+1},\ldots,X_\ell=x_{n+\ell}\,|\,X_0=x_n)).$$

Summing over all disjoint points $(x_{n+1},\ldots,x_{n+\ell})\in A$ then completes the proof. □

Exercise 2.8. Suppose we have a Markov chain $X_0,X_1,\ldots$ with finite state space Ω. Let $y\in\Omega$. Define $L_y:=\max\{n\geq 0\colon X_n=y\}$. Is L_y a stopping time? Prove your assertion.

Solution. No, L_y is not a stopping time. We argue by contradiction. Let $\Omega:=\{1,2\}$. If L_1 were a stopping time, then there exists $B\subseteq\Omega^2$ such that $\{L_1=1\}=\{(X_0,X_1)\in B\}$. But $\{L_1=1\}=\{X_1=1,\,2=X_2=X_3=X_4=\cdots\}$. That is, the B as defined before does not exist. □

Exercise 2.9. Let x,y be points in the state space of a finite Markov Chain $(X_0,X_1,\ldots)$. Let $T_y=\min\{n\geq 1\colon X_n=y\}$ be the first arrival time of y. Let j,k be positive integers. Show that

$$\mathbf{P}_x(T_y>kj\,|\,T_y>(k-1)j)\leq\max_{z\in\Omega}\mathbf{P}_z(T_y>j).$$

(Hint: use Exercise 2.7)

Solution. We first suppose that $x\neq y$. From Exercise 2.7, if $A\subseteq\Omega^\ell$ and if $x_0,\ldots,x_n\in\Omega$,

$$\mathbf{P}((X_{n+1},\ldots,X_{n+\ell})\in A\,|\,(X_0,\ldots,X_n)=(x_0,\ldots,x_n))=\mathbf{P}((X_1,\ldots,X_\ell)\in A\,|\,X_0=x_n).$$

Multiplying both sides by $\mathbf{P}((X_0,\ldots,X_n)=(x_0,\ldots,x_n))$, we get

$$\begin{aligned}&\mathbf{P}((X_{n+1},\ldots,X_{n+\ell})\in A,(X_0,\ldots,X_n)=(x_0,\ldots,x_n))\\&\quad=\mathbf{P}((X_1,\ldots,X_\ell)\in A\,|\,X_0=x_n)\mathbf{P}((X_0,\ldots,X_n)=(x_0,\ldots,x_n))\\&\quad\leq\max_{z\in\Omega}\mathbf{P}((X_1,\ldots,X_\ell)\in A\,|\,X_0=z)\mathbf{P}((X_0,\ldots,X_n)=(x_0,\ldots,x_n))\end{aligned}$$

Summing over all $x_1,\ldots,x_n\neq y$, we get

$$\begin{aligned}&\mathbf{P}((X_{n+1},\ldots,X_{n+\ell})\in A,X_0=x_0,X_1\neq y,\ldots,X_{n-1}\neq y,X_n\neq y)\\&\quad\leq\max_{z\in\Omega}\mathbf{P}((X_1,\ldots,X_\ell)\in A\,|\,X_0=z)\mathbf{P}(X_0=x_0,X_1\neq y,\ldots,X_{n-1}\neq y,X_n\neq y).\end{aligned}$$

Dividing both sides by $\mathbf{P}(X_0 = x_0, X_1 \neq y, \ldots, X_n \neq y)$,

$$\mathbf{P}((X_{n+1}, \ldots, X_{n+\ell}) \in A \,|\, X_0 = x_0, X_1 \neq y, \ldots, X_{n-1} \neq y, X_n \neq y)$$
$$= \max_{z\in\Omega} \mathbf{P}((X_1, \ldots, X_\ell) \in A \,|\, X_0 = z).$$

Using Exercise 2.6, we can rewrite this as

$$\mathbf{P}_{x_0}((X_{n+1}, \ldots, X_{n+\ell}) \in A \,|\, X_1 \neq y, \ldots, X_{n-1} \neq y, X_n \neq y)$$
$$= \max_{z\in\Omega} \mathbf{P}_z((X_1, \ldots, X_\ell) \in A).$$

Or, written another way, and using $A = \{y\}^c \times \cdots \times \{y\}^c$,

$$\mathbf{P}_{x_0}((T_y > n + \ell \,|\, T_y > n) \leq \max_{z\in\Omega} \mathbf{P}_z(T_y > \ell).$$

□

Exercise 2.10. Let x, y be points in the state space of a finite Markov Chain $(X_0, X_1, \ldots)$ with transition matrix P. Let $T_y = \min\{n \geq 1 \colon X_n = y\}$ be the first arrival time of y. Let j be a positive integer. Show that

$$P^j(x, y) \leq \mathbf{P}_x(T_y \leq j)).$$

(Hint: can you induct on j?)

Solution. From the Chapman-Kolmogorov equation Proposition 3.24, $P^j(x, y) = \mathbf{P}(X_j = y \,|\, X_0 = x) = \mathbf{P}_x(X_j = y)$. Since $\{X_j = y\} \subseteq \{T_y \leq j\}$, we have $\mathbf{P}_x(X_j = y) \leq \mathbf{P}_x(T_y \leq j)$, as desired. □

Solution. In the case $j = 1$, we get equality. We now complete the inductive step. Assume the inequality is true for $j-1$, and consider the case of j. Define $T_y' = \min\{n \geq 2 \colon X_n = y\}$. Then $T_y' \geq T_y$, so $\{T_y' \leq j\} \subseteq \{T_y \leq j\}$ for any $j \geq 1$.

$$\begin{aligned}
P^j(x, y) &= (PP^{j-1})(x, y) = \sum_{z\in\Omega} P(x, z) P^{j-1}(z, y) \\
&\leq \sum_{z\in\Omega} P(x, z) \mathbf{P}_z(T_y \leq j - 1) \qquad \text{, by the inductive hypothesis} \\
&= \sum_{z\in\Omega} P(x, z) \mathbf{P}(T_y \leq j - 1 \,|\, X_0 = z) \qquad \text{, by definition of } \mathbf{P}_z \\
&= \sum_{z\in\Omega} P(x, z) \mathbf{P}(T_y' \leq j \,|\, X_0 = x, X_1 = z) \qquad \text{, by Exercise 2.6} \\
&= \sum_{z\in\Omega} \mathbf{P}(X_0 = x, X_1 = z) \mathbf{P}(T_y' \leq j \,|\, X_0 = x, X_1 = z), \\
&= \mathbf{P}_x(T_y' \leq j) \leq \mathbf{P}_x(T_y \leq j).
\end{aligned}$$

The last equality used the Total Probability Theorem. Also, in our use of Exercise 2.6, we let $A \subseteq \Omega^{j-1}$ be the set of points $(x_1, \ldots, x_{j-1})$ such that $x_i = y$ for some $1 \leq i \leq j - 1$. Then

$$\mathbf{P}(T_y \leq j - 1 \,|\, X_0 = z) = \mathbf{P}((X_1, \ldots, X_{j-1}) \in A \,|\, X_0 = z)$$
$$= \mathbf{P}((X_2, \ldots, X_j) \in A \,|\, X_0 = x, X_1 = z) = \mathbf{P}(T_y' \leq j \,|\, X_0 = x, X_1 = z).$$

□

Exercise 2.11. Let x, y be any states in a finite irreducible Markov chain. Show that $\mathbf{E}_x T_y < \infty$. In particular, $\mathbf{P}_y(T_y < \infty) = 1$, so all states are recurrent.

Solution. From Lemma 3.27 in the notes, there exists $0 < \alpha < 1$ and $j > 0$ such that, for any $x, y \in \Omega$ and for any $k > 0$, $\mathbf{P}_x(T_y > kj) \leq \alpha^k$. So, $\mathbf{P}_x(T_y > kj) \leq \alpha^k$. So, using Remark 2.23 in the notes,

$$\mathbf{E}_x T_y = \sum_{i=1}^{\infty} \mathbf{P}(T_y \geq i) = \sum_{j=1}^{\infty} \sum_{k(j-1) < i \leq jk} \mathbf{P}(T_y \geq i)$$
$$\leq \sum_{j=1}^{\infty} k\mathbf{P}(T_y \geq k(j-1)) \leq k \sum_{j=1}^{\infty} \alpha^{j-1} = k/(1-\alpha) < \infty.$$

□

Exercise 2.12 (Simplified Monopoly)**.** Let $\Omega = \{1, 2, \ldots, 10\}$. We consider Ω to be the ten spaces of a circular game board. You move from one space to the next by rolling a fair six-sided die. So, for example $P(1, k) = 1/6$ for every $2 \leq k \leq 7$. More generally, for every $j \in \Omega$ with $j \neq 5$, $P(j, k) = 1/6$ if $k = (j + i) \bmod 10$ for some $1 \leq i \leq 6$. Finally, the space 5 forces you to return to 1, so that $P(5, 1) = 1$. (Note that $\bmod 10$ denotes arithmetic modulo 10, so e.g. $7 + 5 = 2 \bmod 10$.)

Using a computer, find the unique stationary distribution of this Markov chain. Which point has the highest stationary probability? The lowest?

Compare this stationary distribution to the stationary distribution that arises from the doubly stochastic matrix: for all $j \in \Omega$, $P(j, k) = 1/6$ if $k = (j + i) \bmod 10$ for some $1 \leq i \leq 6$.

Solution.

$$P = \begin{pmatrix} 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 \\ 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 \\ 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 \\ 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 1/6 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

By taking several powers of P on the computer, we know that the resulting matrix will converge exponentially fast to a rank one matrix with identical rows. And any such row will approximate the stationary distribution of the Markov chain. The computer outputs the following distribution:

$$(0.1747, 0.0991, 0.1003, 0.1013, 0.1047, 0.0916, 0.0945, 0.0811, 0.0782, 0.0745).$$

By comparison, when P is doubly stochastic, the stationary distribution has all entries .1. $\square$

3. Homework 3

Exercise 3.1 (Knight Moves)**.** Consider a standard 8×8 chess board. Let V be a set of vertices corresponding to each square on the board (so V has 64 elements). Any two vertices $x, y \in V$ are connected by an edge if and only if a knight can move from x to y. (The knight chess piece moves in an L-shape, so that a single move constitutes two spaces moved along the horizontal axis followed by one move along the vertical axis (or two spaces moved along the vertical axis, followed by one move along the horizontal axis.) Consider the simple random walk on this graph. This Markov chain then represents a knight randomly moving around a chess board. For every space x on the chessboard, compute the expected return time $\mathbf{E}_x T_x$ for that space. (It might be convenient to just draw the expected values on the chessboard itself.)

Solution. By inspection, the Markov chain is irreducible. By Corollary 3.37, if π is the unique solution to $\pi = \pi P$, then $\mathbf{E}_x T_x = 1/\pi(x)$. So, it suffices to find $\pi(x)$ for any $x \in \Omega$. From Example 3.50 in the notes, $\pi(x) = \deg(x)/(2|E|)$. (From a previous exercise, we know that $\sum_{x \in V} \deg(x) = 2|E|$.) The following table depicts the degrees of each entry in the chess board

$$\begin{pmatrix} 2 & 3 & 4 & 4 & 4 & 4 & 3 & 2 \\ 3 & 4 & 6 & 6 & 6 & 6 & 4 & 3 \\ 4 & 6 & 8 & 8 & 8 & 8 & 6 & 4 \\ 4 & 6 & 8 & 8 & 8 & 8 & 6 & 4 \\ 4 & 6 & 8 & 8 & 8 & 8 & 6 & 4 \\ 4 & 6 & 8 & 8 & 8 & 8 & 6 & 4 \\ 3 & 4 & 6 & 6 & 6 & 6 & 4 & 3 \\ 2 & 3 & 4 & 4 & 4 & 4 & 3 & 2 \end{pmatrix}.$$

So, $336 = \sum_{x \in V} \deg(x) = 2|E|$, so $\mathbf{E}_x T_x = 1/\pi(x) = (2|E|)/\deg(x) = 336/\deg(x)$. So, the following table depicts $\mathbf{E}_x T_x$ at each point on the chessboard

$$\begin{pmatrix} 168 & 112 & 84 & 84 & 84 & 84 & 112 & 168 \\ 112 & 84 & 56 & 56 & 56 & 56 & 84 & 112 \\ 84 & 56 & 42 & 42 & 42 & 42 & 56 & 84 \\ 84 & 56 & 42 & 42 & 42 & 42 & 56 & 84 \\ 84 & 56 & 42 & 42 & 42 & 42 & 56 & 84 \\ 84 & 56 & 42 & 42 & 42 & 42 & 56 & 84 \\ 112 & 84 & 56 & 56 & 56 & 56 & 84 & 112 \\ 168 & 112 & 84 & 84 & 84 & 84 & 112 & 168 \end{pmatrix}.$$

$\square$

Exercise 3.2. Give an example of a Markov chain where there are at least two different stationary distributions.

Solution. Let $P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Then $(1,0)$ and $(0,1)$ are both stationary distributions for this transition matrix. □

Exercise 3.3. Is there a finite Markov chain where no stationary distribution exists? Either find one, or prove that no such finite Markov chain exists.

(If you want to show that no such finite Markov chain exists, you are allowed to just prove the weaker assertion that: for every stochastic matrix P, there always exists a nonzero vector π with $\pi = \pi P$.)

Solution. Let P be an $n \times n$ stochastic matrix. Let x be the column vector with all entries equal to 1. Since $Px = x$, we know that x is in the null space of $P - I$ with $x \neq 0$. So, the column rank and row rank of $P - I$ are both less than n. So, there exists a linear combination of the rows of $P - I$ that is equal to the zero vector. That is, there must exist a row vector π such that $\pi(P - I) = 0$, so that $\pi P = P$.

To get the stronger result that there is a probability distribution π with $\pi = \pi P$, we can use Brouwer's Fixed point Theorem (which is something not covered in this course). If P is an $n \times n$ stochastic matrix, and if ν is any probability distribution, then νP is also a probability distribution. So, if $\Delta_n := \{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon x_i \geq 0 \, \forall \, 1 \leq i \leq n, \, \sum_{i=1}^n x_i = 1\}$, then Δ_n is a closed, bounded, and convex set, and the function $T(x) := xP$, $x \in \Delta_n$ is a continuous function $T \colon \Delta_n \to \Delta_n$. So, Brouwer's Fixed Point Theorem says there exists $\pi \in \Delta_n$ such that $T\pi = \pi$, i.e. $\pi P = \pi$. □

Exercise 3.4. Let P be the transition matrix for a finite Markov chain with state space Ω. We say that the matrix P is **doubly stochastic** if the columns of P each sum to 1. (Since P is a stochastic matrix, each of its rows already sum to 1.) Let π such that $\pi(x) = 1/|\Omega|$ for all $x \in \Omega$. That is, π is uniform on Ω. Show that $\pi = \pi P$.

Solution.
$$(\pi P)(x) = \sum_{y \in \Omega} \pi(y) P(y,x) = \frac{1}{|\Omega|} \sum_{y \in \Omega} P(y,x) = \frac{1}{|\Omega|} = \pi(x).$$
The penultimate equality used our assumption. □

Exercise 3.5. Give an example of a random walk on a graph that is not reversible.

Solution. Let P be any doubly stochastic matrix that is not symmetric, and such that the Markov chain is irreducible. Then the Markov chain will not be reversible. By Exercise 3.4, the (unique) stationary distribution is uniform, so reversibility reduces to $P(x,y) = P(y,x)$ for all $x, y \in \Omega$. And this equality will not hold when P is not symmetric.

For example, let
$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

The (left) eigenvector of P corresponding to the eigenvalue 1 is $(1,1,1)$. So the unique stationary distribution satisfies $\pi(1)=\pi(2)=\pi(3)=1/3$. But $\pi(1)P(1,2)=(1/3)(1)=1/3$ whereas $\pi(2)P(2,1)=(1/3)(0)=0$. □

Exercise 3.6. Let P be the transition matrix of a finite, irreducible, reversible Markov chain with state space Ω and stationary distribution π. Let $f,g\in\mathbb{R}^{|\Omega|}$ be column vectors. Consider the following bilinear function on f,g, which is referred to as an inner product (or dot product):

$$\langle f,g\rangle := \sum_{x\in\Omega} f(x)g(x)\pi(x).$$

Show that P is self-adjoint (i.e. symmetric) in the sense that

$$\langle f,Pg\rangle = \langle Pf,g\rangle.$$

In particular (for those that have taken 115A), the spectral theorem implies that all eigenvalues of P are real.

Finally, find a transition matrix P such that at least one eigenvalue of P is not real.

Solution. Applying the definitions of the inner product and matrix multiplication, then switching the order of summation and using reversibility,

$$\begin{aligned}\langle f,Pg\rangle &= \sum_{x\in\Omega} f(x)(Pg)(x)\pi(x) = \sum_{x\in\Omega} f(x)\sum_{y\in\Omega}\pi(x)P(x,y)g(y)\\ &= \sum_{y\in\Omega} f(x)\sum_{x\in\Omega}\pi(x)P(x,y)g(y) = \sum_{y\in\Omega} g(y)\sum_{x\in\Omega}\pi(y)P(y,x)f(x)\\ &= \sum_{y\in\Omega} g(y)(Pf)(y)\pi(y) = \langle Pf,g\rangle.\end{aligned}$$

Finally, consider again the doubly stochastic matrix

$$P = \begin{pmatrix} 0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}.$$

The eigenvalues λ of this matrix satisfy $\lambda^3=1$. So, two of the eigenvalues are not real, namely $e^{2\pi\sqrt{-1}/3}$ and $e^{4\pi\sqrt{-1}/3}$. □

Exercise 3.7 (Ehrenfest Urn Model). Suppose we have two urns and n spheres. Each sphere is in either of the first or the second urn. At each step of the Markov chain, one of the spheres is chosen uniformly and random and moved from its current urn to the other urn. Let X_n be the number of spheres in the first urn at time n. Then the transition matrix defining the Markov chain is

$$P(j,k) = \begin{cases} \frac{n-j}{n} & \text{, if } k=j+1\\ \frac{j}{n} & \text{, if } k=j-1\\ 0 & \text{, otherwise.}\end{cases}$$

Show that the unique stationary distribution for this Markov chain is a binomial with parameters n and $1/2$.

Solution. Since any sphere is chosen uniformly at random, it is always possible to move from one configuration of spheres to another configuration. That is, any desired sequence of moves of spheres can occur, with some probability. So, this Markov chain is finite and irreducible, and therefore the stationary distribution is unique, by Theorem 3.36 in the notes. So, it suffices to show that the binomial is stationary. From Proposition 3.46 in the notes, it suffices to show that this distribution is reversible. Indeed, for any $0 \leq j < n$, we have

$$\begin{aligned}\pi(j)P(j,j+1) &= \binom{n}{j}2^{-n}\frac{n-j}{j} = \frac{n!}{(n-j)!j!}2^{-n}\frac{n-j}{n} = \frac{n!}{(n-j-1)!j!}2^{-n}\frac{1}{n}\\ &= \frac{n!}{(n-j-1)!(j+1)!}2^{-n}\frac{j+1}{n} = \binom{n}{j+1}2^{-n}\frac{j+1}{n} = \pi(j+1)P(j+1,j).\end{aligned}$$

Since $P(j,j+1)$ and $P(j+1,j)$ with $0 \leq j < n$ are the only nonzero entries of P, we have verified that P is reversible, as desired. □

Exercise 3.8. Let $V = \{0,1\}^n$ be a set of vertices. We construct a graph from V as follows. Let $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \{0,1\}^n$. Then x and y are connected by an edge in the graph if and only if $\sum_{i=1}^n |x_i - y_i| = 1$. That is, x and y are connected if and only if they differ by a single coordinate.

For any $x \in V$, define $f(x) = \sum_{i=1}^n x_i$, $f\colon V \to \{0,1,\ldots,n\}$. Given $x \in V$, we identify x with the state in the Ehrenfest urn model where the first urn has exactly $f(x)$ spheres. Show that the Ehrenfest urn model is a **projection** of the simple random walk on V in the following sense. The probability that $x \in V$ transitions to any state $z \in V$ such that $y = f(z)$ is equal to: the probability that Ehrenfest model with state $f(x)$ transitions to state y.

Moreover, the unique stationary distribution for the simple random walk on V can be projected to give the unique stationary distribution in the Ehrenfest model. That is, if π is the unique stationary distribution for the simple random walk on V, and if for any $A \subseteq \{0,1,\ldots,n\}$, we define $\mu(A) = \pi(f^{-1}(A))$, then μ is Binomial with parameters n and $1/2$. (Here $f^{-1}(A) = \{x \in V : f(x) \in A\}$.)

Solution. Let $x \in V$. The probability that $x \in V$ transitions to any state $z \in V$ such that $y = f(z)$ is equal to zero, unless $y = f(x)+1$ or $y = f(x)-1$. If $f(x) < n$, we consider the case $y = f(x)+1$. In this case, the number of $z \in V$ such that $f(z) = y$ is the number of n-tuples $(z_1, \ldots, z_n) \in V$ that are obtained by changing a zero entry of x to a one. This number is equal to $n - f(x)$. Also, the degree of every point in V is n. So, the probability that $x \in V$ transitions to any state $z \in V$ such that $y = f(z)$ is $\frac{n-f(x)}{n}$.

Similarly, the probability that Ehrenfest model with state $f(x)$ transitions to state y is $\frac{n-f(x)}{n}$, by the definition of the Ehrenfest model, using $y = f(x)+1$.

A similar argument applies in the case $y = f(x)-1$.

Now, the simple random walk on V has stationary distribution $\pi(x) = \deg(x)/(2|E|) = n/(n2^n) = 2^{-n}$, for any $x \in V$, by Example 3.50 in the notes. If $t \in \{0,1,\ldots,n\}$, then $\mu(f^{-1}(t))$ is, by definition of π, 2^{-n} times the number of n-tupes $(z_1, \ldots, z_n) \in V$ such that

$z_1 + \cdots + z_n = t$. That is, $\mu(f^{-1}(t)) = 2^{-n}\binom{n}{t}$. That is, $\mu(f^{-1}(\cdot))$ is a binomial with parameters n and $1/2$. $\square$

Exercise 3.9 (Birth-and-Death Chains). A birth-and-death chain can model the size of some population of organisms. Fix a positive integer k. Consider the state space $\Omega = \{0, 1, 2, \ldots, k\}$. The current state is the current size of the population, and at each step the size can increase or decrease by at most 1. We define $\{(p_n, r_n, q_n)\}_{n=0}^{k}$ such that $p_n + r_n + q_n = 1$ for each n, and

- $P(n, n+1) = p_n > 0$ for every $0 \leq n < k$.
- $P(n, n-1) = q_n > 0$ for every $0 < n \leq k$.
- $P(n, n) = r_n \geq 0$ for every $0 \leq n \leq k$.
- $q_0 = p_k = 0$.

Show that the birth-and-death chain is reversible.

Solution. We first try to define $\nu(n)$ for any $0 \leq n < k$ so that $\nu(n)P(n, n+1) = \nu(n+1)P(n+1, n)$. Suppose we have defined $\nu(0) := 1$. We then define $\nu(1) := \nu(0)P(0, 1)/P(1, 0) = \nu(0)p_0/q_1$. We then inductively define $\nu(n+1) := \nu(n)P(n, n+1)/P(n+1, n) = \nu(n)p_n/q_{n+1}$, for any $0 \leq n < k$. In this case, $\nu(n)$ is well-defined for any $0 \leq n \leq k$. By construction, ν satisfies the reversibility condition. Since ν may not be a probability distribution, we therefore define $\pi(n) := \nu(n)/\sum_{0 \leq j \leq n} \nu(j)$. Then π satisfies the reversibility condition, and π is a probability distribution. $\square$

Exercise 3.10. Give an explicit example of a Markov chain where every state has period 100.

Solution. Let $\Omega = \{1, \ldots, 100\}$. Define P so that $P(n, n+1) = 1$ for every $1 \leq n < 100$, and $P(100, 1) = 1$. Then $P^{100k}(n, n) = 1$ for every $n \in \Omega$, and for every $k \geq 1$, while $P^j(n, n) = 0$ for any positive integer j that is not a multiple of 100. So, every $n \in \mathbb{N}$ satisfies $\mathcal{N}(n) = \{100, 200, 300, \ldots\}$, so that every state in the Markov chain has period 100. $\square$

4. Homework 4

Exercise 4.1. Let μ, ν be probability distributions on a finite state space Ω. Then

$$||\mu - \nu||_{\mathrm{TV}} = \frac{1}{2}\sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

(Hint: consider the set $A = \{x \in \Omega \colon \mu(x) \geq \nu(x)\}$.)

Solution. Let $A = \{x \in \Omega \colon \mu(x) \geq \nu(x)\}$. Then by definition of the total variation distance,

$$\begin{aligned}||\mu - \nu||_{\mathrm{TV}} &\geq |\mu(A) - \nu(A)| = \mu(A) - \nu(A)\\ &= \sum_{x \in \Omega \colon \mu(x) \geq \nu(x)} (\mu(x) - \nu(x)) = \sum_{x \in \Omega \colon \mu(x) \geq \nu(x)} |\mu(x) - \nu(x)|.\end{aligned}$$

Similarly, if $A' = \{x \in \Omega \colon \mu(x) \leq \nu(x)\}$, then

$$||\mu - \nu||_{\mathrm{TV}} \geq \sum_{x\in\Omega\colon \mu(x)\leq\nu(x)} |\mu(x) - \nu(x)|.$$

Adding the two inequalities,

$$\begin{aligned} 2\,||\mu - \nu||_{\mathrm{TV}} &\geq \sum_{x\in\Omega\colon \mu(x)\geq\nu(x)} |\mu(x) - \nu(x)| + \sum_{x\in\Omega\colon \mu(x)\leq\nu(x)} |\mu(x) - \nu(x)| \\ &= \sum_{x\in\Omega} |\mu(x) - \nu(x)|. \end{aligned}$$

On the other hand, if $B \subseteq \Omega$, then without loss of generality, $\mu(B) \geq \nu(B)$. Let $C := \{b \in B \colon \mu(b) > \nu(b)\}$. Then

$$|\mu(B) - \nu(B)| = \mu(B) - \nu(B) \leq \mu(C) - \nu(C) = |\mu(C) - \nu(C)|.$$

Since μ, ν are probability distributions,

$$|\mu(C^c) - \nu(C^c)| = |1 - \mu(C) - (1 - \nu(C))| = |\mu(C) - \nu(C)|.$$

That is,

$$|\mu(B) - \nu(B)| \leq |\mu(C^c) - \nu(C^c)|.$$

Adding the two inequalities, we get

$$2\,|\mu(B) - \nu(B)| \leq |\mu(C) - \nu(C)| + |\mu(C^c) - \nu(C^c)| = \sum_{x\in\Omega} |\mu(x) - \nu(x)|.$$

□

Exercise 4.2. Let $(X_0, X_1, \ldots)$ be the simple random walk on $\mathbb{Z}$. Show that $\mathbf{P}_0(X_n = 0)$ decays like $1/\sqrt{n}$ as $n \to \infty$. That is, show

$$\lim_{n\to\infty} \sqrt{2n}\,\mathbf{P}_0(X_{2n} = 0) = \sqrt{\frac{2}{\pi}}.$$

Also, show the upper bound

$$\mathbf{P}_0(X_n = k) \leq \frac{10}{\sqrt{n}}, \qquad \forall\, n \geq 0,\, k \in \mathbb{Z}.$$

(Hint 1: first consider the case $n = 2r$ for $r \in \mathbb{Z}$. It may be helpful to show that $\binom{2r}{r+j}$ is maximized when $j = 0$. To eventually deal with k odd, just condition on the first step of the walk.)

(Hint 2: you can freely use **Stirling's formula**:

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}(n/e)^n} = 1.$$

Or, there is a more precise estimate: for any $n \geq 3$, there exists $1/(12n+1) \leq \varepsilon_n \leq 1/(12n)$ such that

$$n! = \sqrt{2\pi} e^{-n} n^{n+1/2} e^{\varepsilon_n}.)$$

Solution. The event $X_{2n} = 0$ occurs when $2n$ fair coin flips result in n heads and n tails. That is,

$$\mathbf{P}_0(X_{2n} = 0) = 2^{-2n}\binom{2n}{n} = 2^{-2n}\frac{(2n)!}{(n!)^2}.$$

Using the precise version of Stirling's formula,

$$\mathbf{P}_0(X_{2n} = 0) = 2^{-2n}\binom{2n}{n} = 2^{-2n}\frac{\sqrt{2\pi}e^{-2n}(2n)^{2n+1/2}e^{\varepsilon_{2n}}}{2\pi e^{-2n}n^{2n+1}e^{2\varepsilon_n}} = \frac{\sqrt{2}}{\sqrt{2\pi}\sqrt{n}}\frac{e^{\varepsilon_{2n}}}{e^{2\varepsilon_n}} = \frac{1}{\sqrt{\pi}\sqrt{n}}\frac{e^{\varepsilon_{2n}}}{e^{2\varepsilon_n}}.$$

Letting $n \to \infty$, we then get

$$\lim_{n\to\infty}\sqrt{2n}\mathbf{P}_0(X_{2n} = 0) = \sqrt{2/\pi}.$$

We now show the upper bound. We claim that $\binom{2r}{r+j}$ is maximized for $0 \leq j \leq r$ such that $j = 0$. To see this, note that $\binom{2r}{r+j} = \frac{(2r)!}{(r-j)!(r+j)!}$, and $\frac{(r-j)!(r+j)!}{(r-j-1)!(r+j+1)!} = \frac{r-j}{r+j+1} < 1$. That is, $\frac{1}{(r-j)!(r+j)!} > \frac{1}{(r-j)-1!(r+j+1)!}$. That is, $\frac{(2r)!}{(r-j)!(r+j)!}$ increases as $j \geq 0$ decreases. That is, $\binom{2r}{r+j}$ is maximized for $0 \leq j \leq r$ when $j = 0$. So, if $n = 2r$ is even, and if $0 \leq k \leq n$ with k even,

$$\mathbf{P}_0(X_{2r} = k) = 2^{-2r}\binom{2r}{k} \leq 2^{-2r}\binom{2r}{r} = 2^{-2r}\binom{2r}{r} = \frac{1}{\sqrt{\pi}\sqrt{r}}\frac{e^{\varepsilon_{2r}}}{e^{2\varepsilon_r}} \leq \frac{5}{\sqrt{r}}.$$

The last equality was shown above. If k is odd, then this probability is zero.

It remains to check X_{2r+1} with k odd. In this case, by conditioning on X_0,

$$\mathbf{P}_0(X_{2r+1} = k) = \frac{1}{2}\mathbf{P}_1(X_{2r} = k) + \frac{1}{2}\mathbf{P}_{-1}(X_{2r} = k).$$

Then, using the Markov property,

$$\mathbf{P}_0(X_{2r+1} = k) = \frac{1}{2}\mathbf{P}_0(X_{2r} = k-1) + \frac{1}{2}\mathbf{P}_0(X_{2r} = k+1). \leq \frac{1}{2}\frac{5}{\sqrt{r}} + \leq \frac{1}{2}\frac{5}{\sqrt{r}} \leq \frac{10}{\sqrt{r}}.$$

□

Exercise 4.3. Show that every state in the simple random walk on $\mathbb{Z}$ is recurrent. (You should show this statement for any starting location of the Markov chain.)

Then, find a nearest-neighbor random walk on $\mathbb{Z}$ such that every state is transient.

Solution. Let $k, r > 0$. Then $\mathbf{P}_k(T_0 > r) \leq \frac{20k}{\sqrt{r}}$ from Theorem 3.66 in the notes. Since $\{T_0 = \infty\} \subseteq \{T_0 > r\}$ for any $r > 0$, we have

$$\mathbf{P}_k(T_0 = \infty) \leq \frac{20k}{\sqrt{r}}, \qquad \forall r > 0.$$

Letting $r \to \infty$, we conclude that $\mathbf{P}_k(T_0 = \infty) = 0$. That is, 0 is recurrent as long as $k > 0$. Since these random walks are symmetric with respect to reflection across 0, we have $\mathbf{P}_k(T_0 > r) = \mathbf{P}_{-k}(T_0 > r)$ for any $k.r > 0$. So, we similarly conclude that $\mathbf{P}_k(T_0 = \infty) = 0$

for any $k < 0$. Finally, $\mathbf{P}_0(T_0 = \infty) = 0$, since, if we condition on the first step of the simple random walk and use Exercise 2.7, then

$$\begin{aligned}\mathbf{P}_0(T_0 = \infty) &= \mathbf{P}_0(T_0 = \infty \,|\, X_1 = 1)\mathbf{P}(X_1 = 1) + P_0(T_0 = \infty \,|\, X_1 = -1)\mathbf{P}(X_1 = -1)\\ &= \frac{1}{2}\mathbf{P}_1(T_0 = \infty) + \frac{1}{2}\mathbf{P}_{-1}(T_0 = \infty) = 0.\end{aligned}$$

So, the state 0 is recurrent, no matter where the random walk starts. Let $j, k \in \mathbb{Z}$. Using translation invariance of the Markov chain,

$$\begin{aligned}\mathbf{P}_j(T_k = \infty) &= \mathbf{P}_j(X_1 \neq k, X_2 \neq k, \ldots)\\ &= \mathbf{P}_{j-k}(X_1 \neq 0, X_2 \neq 0, \ldots = \mathbf{P}_{j-k}(T_0 = \infty) = 0.\end{aligned}$$

So, all states are recurrent.

Finally, define a random walk on $\mathbb{Z}$ such that $P(n, n+1) = 1$ for all $n \in \mathbb{Z}$. Then for any $r > 0$, $P^r(n, n+r) = 1$. That is, $P^r(n, n) = 0$ for all $r \geq 1$ and for all $n \in \mathbb{Z}$. So, $\mathbf{P}_n(T_n < \infty) = 0$ for every $n \in \mathbb{Z}$. That is, every state is transient □

Exercise 4.4. For the simple random walk on $\mathbb{Z}$, show that $\mathbf{E}_0 T_0 = \infty$. Conclude that, for any $x, y \in \mathbb{Z}$, $\mathbf{E}_x T_y = \infty$.

Solution. From Lemma 3.69 in the notes, $\mathbf{P}_1(T_0 > r) = \mathbf{P}_0(-1 < X_r \leq 1) \geq \mathbf{P}_0(X_r = 0)$. From Exercise 4.2, $\lim_{n\to\infty} \sqrt{2n}\,\mathbf{P}_0(X_{2n} = 0) = \sqrt{\frac{2}{\pi}}$. That is, there exists $m \geq 0$ and a constant $c > 0$ such that, for all $n \geq m$, $\mathbf{P}_0(X_{2n} = 0) \geq cn^{-1/2}$. Combining our inequalities,

$$\mathbf{P}_1(T_0 > 2n) \geq cn^{-1/2}.$$

So,

$$\mathbf{E}_1 T_0 = \sum_{n=0}^{\infty} \mathbf{P}_1(T_0 > n) \geq \sum_{n=0}^{\infty} \mathbf{P}_1(T_0 > 2n) \geq \sum_{n=0}^{\infty} cn^{-1/2} = \infty.$$

Let $j \in \mathbb{Z}$. From the Total Expectation Theorem,

$$\mathbf{E}_j T_0 = \mathbf{E}_j(T_0 \,|\, X_1 = j+1)\mathbf{P}(X_1 = j+1) + \mathbf{E}_0(T_0 \,|\, X_1 = j-1)\mathbf{P}(X_1 = j-1).$$

And from Exercise 2.7

$$\begin{aligned}\mathbf{E}_j(T_0 \,|\, X_1 = j+1) &= \sum_{n=0}^{\infty} \mathbf{P}(T_0 > n \,|\, X_0 = j, X_1 = j+1)\\ &= \sum_{n=0}^{\infty} \mathbf{P}(T_0 > n \,|\, X_1 = j+1) = \mathbf{E}_{j+1}(T_0).\end{aligned}$$

Therefore, $\mathbf{E}_j T_0 \geq \frac{1}{2}\mathbf{E}_{j+1} T_0$ and $\mathbf{E}_j T_0 \geq \frac{1}{2}\mathbf{E}_{j-1} T_0$. Iteratively applying these inequalities $|j|$ times, we get

$$\mathbf{E}_j T_0 \geq 2^{-j}\mathbf{E}_1 T_0 = \infty.$$

Let $j, k \in \mathbb{Z}$, $r > 0$. Using translation invariance of the Markov chain,

$$\begin{aligned}\mathbf{P}_j(T_k > r) &= \mathbf{P}_j(X_1 \neq k, \ldots, X_r \neq k)\\ &= \mathbf{P}_{j-k}(X_1 \neq 0, \ldots, X_r \neq 0) = \mathbf{P}_{j-k}(T_0 > r).\end{aligned}$$

Therefore, $\mathbf{E}_j T_k = \mathbf{E}_{j-k} T_0 = \infty$, as desired. □

Exercise 4.5. Let $(X_0, X_1, \ldots)$ be the "corner walk" on $\mathbb{Z}^2$. The transitions are described as follows. From any point $(x, y) \in \mathbb{Z}^2$, the Markov chain adds any of the following four vector to (x, y) each with probability $1/4$: $\{(1, 1,), (1, -1), (-1, 1), (-1, -1)\}$. Using that the coordinates of this walk are each independent simple random walks on $\mathbb{Z}$, conclude that there exists $c > 0$ such that

$$\lim_{n\to\infty} n\mathbf{P}_{(0,0)}(X_{2n} = (0, 0)) = c.$$

That is, $\mathbf{P}_{(0,0)}(X_{2n} = (0, 0))$ is about c/n, when n is large.

Now, note that the usual nearest-neighbor simple random walk on $\mathbb{Z}^2$ is a rotation of the corner walk by an angle of $\pi/4$. So, the above limiting statement also holds for the simple random walk on $\mathbb{Z}^2$.

Solution. Let $W_0, W_1, \ldots$ and let $Y_0, Y_1, \ldots$ be independent simple random walks on $\mathbb{Z}$ such that $W_0 = Y_0 = 0$. By construction, the stochastic process $(W_0, Y_0), (W_1, Y_1), \ldots$ is the corner walk. So, using independence,

$$\mathbf{P}_{(0,0)}(X_{2n} = (0, 0)) = \mathbf{P}_{(0,0)(W_{2n},Y_{2n})=(0,0)} = \mathbf{P}_0(W_{2n} = 0)\mathbf{P}_0(Y_{2n} = 0)$$

From Exercise 4.2, $\lim_{n\to\infty} \sqrt{2n}\, \mathbf{P}_0(W_{2n} = 0) = \sqrt{\frac{2}{\pi}}$. Therefore,

$$\lim_{n\to\infty} n\mathbf{P}_{(0,0)}(X_{2n} = (0, 0)) = \Big(\lim_{n\to\infty} \sqrt{n}\mathbf{P}_0(W_{2n} = 0)\Big)\Big(\lim_{n\to\infty} \sqrt{n}\mathbf{P}_0(W_{2n} = 0)\Big) = \frac{1}{\pi}.$$

Finally, the simple random walk on $\mathbb{Z}^2$ can be identified with the corner walk by the identification $(x, y) \mapsto (x - y, x + y)$, $x, y \in \mathbb{Z}$. That is, if $(A_1, B_1), (A_2, B_2), \ldots$ is the simple random walk on $\mathbb{Z}^2$, then $(A_1 - B_1, A_1 + B_1), (A_2 - B_2, A_2 + B_2), \ldots$ is the corner walk on $\mathbb{Z}^2$. (The latter stochastic process is translation invariant, and each step of the walk moves with probability $1/4$ by any increment: $\{(1, 1,), (1, -1), (-1, 1), (-1, -1)\}$.) □

Exercise 4.6. Let $S_0, S_1, \ldots$ be a random walk with $S_0 = 0$. Let Y be the number of times the random walk takes the value 0. Let $T_0 := \min\{n \geq 1 \colon S_n = 0\}$.

- Y is a geometric random variable with success probability $\mathbf{P}(T_0 = \infty)$.
- $\mathbf{E}Y = \frac{1}{\mathbf{P}(T_0=\infty)}$. (Here we interpret $1/0$ as ∞.)

(Hint: $\{Y = k\} = \{T_0^{(k-1)} < \infty, T_0^{(k)} = \infty\} = \{T_0^{(k-1)} < \infty, T_0^{(k)} - T_0^{(k-1)} = \infty\}$.)

Exercise 4.7. Let $(X_0, X_1, \ldots)$ be a finite, irreducible Markov chain with transition matrix P and state space Ω. For any $x, y \in \Omega$, define

$$G(x, y) := \mathbf{E}_x \sum_{n=0}^{\infty} 1_{\{X_n=y\}} = \sum_{n=0}^{\infty} \mathbf{P}^n(x, y)$$

to be the expected number of visits to y starting from x. Show that the following are equivalent:

(i) $G(x, x) = \infty$ for some $x \in \Omega$.

(ii) $G(x,y)=\infty$ for all $x,y\in\Omega$.
(iii) $\mathbf{P}_x(T_x<\infty)$ for some $x\in\Omega$.
(iii) $\mathbf{P}_x(T_y<\infty)$ for all $x,y\in\Omega$.

So, in an irreducible finite Markov chain, a single state is recurrent if and only if all states are recurrent.

Exercise 4.8. Show that if the Simple Random Walk on $\mathbb{Z}^d$ is recurrent, then this random walk takes every value in $\mathbb{Z}^d$ infinitely many times (with probability 1). And if the Simple Random Walk on $\mathbb{Z}^d$ is transient, then this random walk takes any fixed value in $\mathbb{Z}^d$ only finitely many times (with probability 1).

Solution. We claim that any neighbor of a recurrent state is a recurrent state for the simple random walk on $\mathbb{Z}^d$. To see this, suppose w is a recurrent state, so that the random walk visits w infinitely many times with probability 1. Each time the random walk takes the value w, it moves to any neighbor of w with equal probability. So, the random walk visits any neighbor of w infinitely many times with probability 1. It follows that either all states are recurrent, or none of them are.

The exercise follows from this statement. If the Simple Random Walk on $\mathbb{Z}^d$ is recurrent, then the random walk visits 0 infinitely many times with probability 1. So, the random walk visits all states in $\mathbb{Z}^d$ infinitely many times with probability 1.

If the Simple Random Walk on $\mathbb{Z}^d$ is transient, then all states of $\mathbb{Z}^d$ are transient. So, the random walk takes any fixed value in $\mathbb{Z}^d$ only finitely many times with probability 1. □

Exercise 4.9. Let $0<p<1$. Consider the random walk on $\mathbb{Z}$ such that $\mathbf{P}(X_1=1)=p$ and $\mathbf{P}(X_1=-1)=1-p$. Show that the corresponding random walk $S_0,S_1,\ldots$ is transient when $p\neq 1/2$.

Solution. Since $(S_n+n)/2$ has binomial distribution with parameters n and p, $(S_{2n}/2)+n$ has binomial distribution with parameters $2n$ and p, and $\{S_{2n}=0\}=\{S_{2n}/2+n=n\}$, so

$$\mathbf{P}(S_{2n}=0)=\binom{2n}{n}p^n(1-p)^n.$$

Using Stirling's formula in the form $n!\sim\sqrt{2\pi n}(n/e)^n$, we have

$$\mathbf{P}(S_{2n}=0)=\frac{(2n)!}{(n!)^2}p^n(1-p)^n\sim\frac{1}{\sqrt{2\pi}}\sqrt{\frac{2}{n}}\frac{(2n)^{2n}}{n^{2n}}p^n(1-p)^n=\frac{1}{\sqrt{2\pi n}}(4p(1-p))^n.$$

Since $p\neq 1/2$, $4p(1-p)<1$, so this quantity decays exponentially in n, i.e.

$$\sum_{n=0}^{\infty}\mathbf{P}(S_{2n}=0)<\infty.$$

So, the random walk is transient by Theorem 3.78. □

Exercise 4.10. Let $S_0,S_1,\ldots$ and $S_0',S_1',\ldots$ be independent simple random walks on $\mathbb{Z}^d$. Let $N:=\sum_{n,m\geq 0}1_{S_n=S_m'}$ be the number of pairs of intersections of these two random walks. For any $y\in\mathbb{R}^d$, let $\phi(y):=\mathbf{E}e^{i\langle y,X_1\rangle}$.

- Show $\mathbf{E}N = \lim_{s\to 1^-}\int_{[-\pi,\pi]^d}\frac{1}{|1-s\phi(y)|^2}\frac{dy}{(2\pi)^d}$. (Hint: consider $\mathbf{E}e^{i\langle y,(S_n-S_m')\rangle}$.)
- For what $d\geq 1$ is $\mathbf{E}N<\infty$?
- Let $C:=\{S_n\colon n\geq 0\}\cap\{S_n'\colon n\geq 0\}$ be the intersection set of the two independent random walks. Let $|C|$ denote the cardinality of C. Show that if the simple random walk on $\mathbb{Z}^d$ is transient, then $\mathbf{P}(N=\infty)=1$ if and only if $\mathbf{P}(|C|=\infty)=1$. (Hint: $N=\sum_{x\in C}N_xN_x'$ where $N_x:=\sum_{n\geq 0}1_{S_n=x}$ is the number of visits of the first random walk to x.) In the recurrent case $d=1,2$, Exercise 4.8 implies that $\mathbf{P}(|C|=\infty)=1$. For any $d\geq 1$, note that $N<\infty$ implies $|C|<\infty$. It can also be shown that $\mathbf{P}(N<\infty)\in\{0,1\}$, $\mathbf{P}(|C|=\infty)\in\{0,1\}$, and that $\mathbf{P}(N<\infty)=1$ if and only if $\mathbf{E}N<\infty$ (you don't have to show these things). In summary, $\mathbf{P}(|C|=\infty)=1$ if and only if $\mathbf{E}N=\infty$.
- Hypothesize what happens to $\mathbf{E}N$ when we instead consider the tuples of intersections of $k>2$ independent simple random walks in $\mathbb{R}^d$. (You don't have to prove your hypothesis.)

5. Homework 5

Exercise 5.1. Using the Optional Stopping Theorem, prove Wald's equations:

Let $X_1,X_2,\ldots\colon\mathcal{C}\to\mathbb{R}$ be i.i.d. Let N be a stopping time. Let $S_0,S_1,\ldots$ be the corresponding random walk with $S_0:=0$.

- If $\mathbf{E}N<\infty$, and $\mathbf{E}|X_1|<\infty$, then $\mathbf{E}S_N=\mathbf{E}X_1\mathbf{E}N$.
- If $\mathbf{E}X_1=0,\mathbf{E}X_1^2<\infty$ and $\mathbf{E}N<\infty$, then $\mathbf{E}S_N^2=\mathbf{E}X_1^2\mathbf{E}N$.

Exercise 5.2. Let $1/2<p<1$. Consider the random walk on $\mathbb{Z}$ such that $\mathbf{P}(X_1=1)=p$ and $\mathbf{P}(X_1=-1)=1-p$. Let $S_0,S_1,\ldots$ be the corresponding random walk with $S_0:=0$. Let $N:=\min\{n\geq 1\colon S_n>0\}$. Using Wald's equation for $\min(N,n)$ and then letting $n\to\infty$, show that $\mathbf{E}N=1/\mathbf{E}X_1=1/(2p-1)$.

Solution. Since N is a stopping time, if $n>1$ is fixed, $\min(N,n)$ is a stopping time with $\mathbf{E}\min(N,n)\leq n<\infty$. Also, $\mathbf{E}|X_1|\leq 1<\infty$. Note that $\mathbf{E}X_1=p-(1-p)=2p-1$. So, from Wald's equation,

$$\mathbf{E}S_{\min(N,n)}=\mathbf{E}X_1\mathbf{E}\min(N,n)=(2p-1)\mathbf{E}\min(N,n).\qquad (*)$$

Recall that $\mathbf{E}X_1=p-(1-p)=2p-1>0$ since $p>1/2$. Using Markov's inequality for the fourth moment, $\exists\ c=c(p)>0$ such that, for all $n\geq 1$,

$$\mathbf{P}(S_n\leq 0)\leq\mathbf{P}\left(\left|\frac{S_n-n\mathbf{E}X_1}{n}\right|>\frac{\mathbf{E}X_1}{2}\right)\leq\frac{c}{n^2}.\qquad (**)$$

In particular, $\mathbf{P}(N=\infty)=0$, so the right side of $(*)$ converges monotonically to $(2p-1)\mathbf{E}N$ as $n\to\infty$, by the Monotone Convergence Theorem. Meanwhile, by the definition of N, the left side of $(*)$ can be estimated as

$$\begin{aligned}1 &\geq \mathbf{E}S_{\min(N,n)} = \mathbf{E}S_{\min(N,n)}(1_{N\leq n}+1_{N>n}) = \mathbf{E}S_N 1_{N\leq n} + \mathbf{E}S_n 1_{N>n}\\ &= \mathbf{E}1_{N\leq n} + \mathbf{E}S_n 1_{N>n} = \mathbf{P}(N\leq n) + \mathbf{E}S_n 1_{N>n} \geq \mathbf{P}(N\leq n) - n\mathbf{P}(N>n)\\ &= 1-(n+1)\mathbf{P}(N>n).\end{aligned}$$

(We always have $S_n \geq -n$.) By definition of N, we have $\mathbf{P}(N>n)\leq \mathbf{P}(S_n\leq 0)$. So, from $(**)$ and the Squeeze Theorem, we conclude that $\lim_{n\to\infty}\mathbf{E}S_{\min(N,n)} = 1$, so that $(*)$ says $1=(2p-1)\mathbf{E}N$, as desired. □

Exercise 5.3. Let $X_0 = 0$. Let $(X_0, X_1, \ldots)$ such that $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = 1/2$ for all $i\geq 1$. For any $n\geq 0$, let $Y_n = X_0+\cdots+X_n$. So, $(Y_0, Y_1,\ldots)$ is a symmetric simple random walk on $\mathbb{Z}$. Show that $Y_n^2 - n$ is a martingale (with respect to $(X_0, X_1, \ldots)$).

Solution. Let $m_0, x_0, \ldots, x_n \in \mathbb{Z}$. Then

$$\begin{aligned}&\mathbf{E}(Y_{n+1}^2 - (n+1) - [Y_n^2 - n] \,|\, X_n = x_n, \ldots, X_0 = x_0, Y_0^2 = m_0)\\ &= \mathbf{E}((X_{n+1}+x_n+\cdots+x_0)^2 - (x_n+\cdots+x_0)^2 - 1)\\ &= \mathbf{E}(X_{n+1}^2 - 1) + \mathbf{E}(X_{n+1})(x_n+\cdots+x_0) = 0+0 = 0.\end{aligned}$$

□

Exercise 5.4. Let $1/2 < p < 1$. Let $(X_0, X_1, \ldots)$ such that $\mathbf{P}(X_i = 1) = p$ and $\mathbf{P}(X_i = -1) = 1-p$ for all $i\geq 1$. For any $n\geq 0$, let $Y_n = X_0+\cdots+X_n$. Let $T_0 = \min\{n\geq 1\colon Y_n = 0\}$. Prove that $\mathbf{P}_1(T_0=\infty)>0$. Then, deduce that $\mathbf{P}_0(T_0 = \infty) > 0$. That is, there is a positive probability that the biased random walk never returns to 0, even though it started at 0.

Solution. Let $X_0 := 1$. From Example 4.12 in the notes, if $q := 1-p$, then $(q/p)^{Y_n}$ is a martingale. Since $1/2 < p < 1$, $q/p < 1$, and by definition of T_0, $Y_{n\wedge T_0} \geq 0$ for all $n\geq 0$, so that $0 \leq (q/p)^{Y_{n\wedge T_0}} \leq 1$ for all $n\geq 0$. If $\mathbf{P}_1(T_0=\infty) = 0$, then the Optional Stopping Theorem, Version 2, implies that $\mathbf{E}(q/p)^{Y_0} = \mathbf{E}(q/p)^{Y_{T_0}}$. But $\mathbf{E}(q/p)^{Y_0} = q/p$ and $\mathbf{E}(q/p)^{Y_T} = 1$. Therefore $\mathbf{P}_1(T_0=\infty)>0$. Finally $\mathbf{P}_0(T_0=\infty) = \mathbf{P}_0(T_0=\infty\,|\,X_1=1)\mathbf{P}(X_1=1) + \mathbf{P}_0(T_0=\infty\,|\,X_1=-1)\mathbf{P}(X_1=-1) \geq p\mathbf{P}_0(T_0=\infty\,|\,X_1=1) = p\mathbf{P}_1(T_0=\infty)>0$. □

Solution. Let $m\geq 1$. From the multiplication rule, (Proposition 2.8 in the notes),

$$\mathbf{P}_1(T_{2^m} < T_0) = \prod_{i=1}^m \mathbf{P}_1(T_{2^i} < T_{2^{i-1}} \,|\, T_{2^{i-1}} < T_0). \qquad (*)$$

For any fixed $i\geq 1$, let $T := \min\{n\geq 1\colon X_n = 0 \text{ or } X_n = 2^{i-1}\}$. From the Strong Markov Property (Theorem 3.16 in the notes),

$$\begin{aligned}&\mathbf{P}_1(X_{T+1}>0,\ldots,X_{T+k}>0, X_{t+k+1} = 2^i \,|\, T=j,\, (X_0,\ldots,X_j) = (x_0,\ldots,x_{j-1},2^{i-1}))\\ &\quad= \mathbf{P}_{2^{i-1}}(X_{j+1}>0,\ldots,X_{j+k}>0, X_{j+k+1} = 2^i).\end{aligned}$$

Summing over all possibilities for $x_1,\ldots,x_{j-1}\in\mathbb{Z}$ and $j,k\geq 1$, and using the Total probability Theorem,

$$\mathbf{P}_1(T_{2^i} < T_0 \,|\, T_{2^{i-1}} < T_0) = \mathbf{P}_{2^{i-1}}(T_{2^i} < T_0)\mathbf{P}_1(T<\infty).$$

As shown in class (Example 4.28 for the Gambler's Ruin problem),

$$\mathbf{P}_{2^{i-1}}(T_{2^i} < T_0) = \frac{3^{2^{i-1}} - 3^{2^i}}{1 - 3^{2^i}} \geq \frac{3^{2^i} - 3^{2^{i-1}}}{3^{2^i}} = 1 - 3^{-2^{i-1}}.$$

For any $0 < t < 1/2$, $1-t \geq e^{-3t}$. So, $\mathbf{P}_{2^{i-1}}(T_{2^i} < T_0) \geq e^{-3^{1-2^{i-1}}}$. Substituting into $(*)$, and using $\mathbf{P}_1(T < \infty) = 1$ (which follows since the Markov chain restricted to $\{0, 1, 2, \ldots, 2^{i-1}\}$ is finite and irreducible, so all states are recurrent.),

$$\begin{aligned}\mathbf{P}_1(T_{2^m} < T_0) &\geq \prod_{i=1}^{m} e^{-3^{-2^{i-1}}} = e^{-\sum_{i=1}^{m} 3^{1-2^{i-1}}} \\ &\geq e^{-\sum_{i=1}^{\infty} 3^{1-2^{i-1}}} \geq e^{-\sum_{i=1}^{\infty} 3^{1-i}} = e^{-3/2} > 0.\end{aligned}$$

Note that $T_{2^m} \geq 2^m$ for every $m \geq 1$, so $\{T_0 = \infty\} = \cap_{m=1}^{\infty}\{T_0 \geq T_{2^m}\}$. And the sets $\{T_0 \geq T_{2^m}\}, \{T_0 \geq T_{2^{m+1}}\}, \ldots$ are decreasing. So, using continuity of the probability law,

$$\mathbf{P}_1(T_0 = \infty) = \lim_{m\to\infty} \mathbf{P}_1(T_0 \geq T_{2^m}) \geq e^{-3/2}.$$

That is, $\mathbf{P}_1(T_0 = \infty) \geq e^{-3/2} > 0$, as desired. □

Exercise 5.5 (**Ballot Theorem**)**.** Let a, b be positive integers. Suppose there are c votes cast by c people in an election. Candidate 1 gets a votes and candidate 2 gets b votes. (So $c = a + b$.) Assume $a > b$. The votes are counted one by one. The votes are counted in a uniformly random ordering, and we would like to keep a running tally of who is currently winning. (News agencies seem to enjoy reporting about this number.) Suppose the first candidate eventually wins the election. We ask: with what probability will candidate 1 always be ahead in the running tally of who is currently winning the election? As we will see, the answer is $\frac{a-b}{a+b}$.

To prove this, for any positive integer k, let S_k be the number of votes for candidate 1, minus the number of votes for candidate 2, after k votes have been counted. Then, define $X_k := S_{c-k}/(c-k)$. Show that $X_0, X_1, \ldots$ is a martingale. Then, let T such that $T = \min\{0 \leq k \leq c \colon X_k = 0\}$, or $T = c - 1$ if no such k exists. Apply the Optional Stopping theorem to X_T to deduce the result.

Solution.

$$\begin{aligned}&\mathbf{E}(X_{k+1} - X_k \,|\, S_{c-k} = s_{c-k}, \ldots, S_c = s_c, X_0 = x_0) \\ &\quad = \mathbf{E}\left(\frac{S_{c-k-1}}{c-k-1} - \frac{s_{c-k}}{c-k} \,|\, S_{c-k} = s_{c-k}, \ldots, S_c = s_c, X_0 = x_0\right) \\ &\quad = \mathbf{E}\left(\frac{S_{c-k-1} - S_{c-k} + s_{c-k}}{c-k-1} - \frac{s_{c-k}}{c-k} \,|\, S_{c-k} = s_{c-k}, \ldots, S_c = s_c, X_0 = x_0\right)\end{aligned}$$

Given that $S_{c-k} = s_{c-k}$, $c-k$ votes have been counted, and there are s_{c-k} more votes for candidate 1 than candidate 2, among the first $c-k$ counted votes. So, if there are x votes for candidate 1 after $c-k$ votes have been counted, then there are $c-k-x$ votes for candidate 2. And we know $x = s_{c-k} + (c-k-x)$ by definition of s_{c-k}, so $2x = s_{c-k} + c - k$, and $x = (1/2)(s_{c-k} + c - k)$.

Given that $S_{c-k} = s_{c-k}$, the expected value of $S_{c-k-1} - S_{c-k}$ is the change in the vote tally, with all $c-k$ votes equally likely to be chosen. (That is, we can think of counting the ballots "in reverse." Given the value of S_{c-k}, we can think of $c-k$ votes as sitting in a pile of "counted" votes. Then $S_{c-k-1} - S_{c-k}$ can be found by choosing any of these $c-k$ votes uniformly at random, and placing this vote into the pile of "uncounted" votes.) That is, this (conditional) expected value of $S_{c-k-1} - S_{c-k}$ is

$$(-1) \cdot \frac{x}{c-k} + (1)\frac{c-k-x}{c-k} = \frac{-2x+c-k}{c-k} = -\frac{s_{c-k}}{c-k}.$$

Therefore,

$$\begin{aligned}
&\mathbf{E}(X_{k+1} - X_k \,|\, S_{c-k} = s_{c-k}, \ldots, S_c = s_c, X_0 = x_0) \\
&= \mathbf{E}\left(\frac{-\frac{s_{c-k}}{c-k} + s_{c-k}}{c-k-1} - \frac{s_{c-k}}{c-k} \,|\, S_{c-k} = s_{c-k}, \ldots, S_c = s_c, X_0 = x_0\right) \\
&= s_{c-k}\left(\frac{-\frac{1}{c-k} + 1}{c-k-1} - \frac{1}{c-k}\right) = s_{c-k}\left(\frac{\left(\frac{c-k-1}{c-k}\right)}{c-k-1} - \frac{1}{c-k}\right) = 0.
\end{aligned}$$

We conclude that $X_1, X_2, \ldots$ is a martingale. Then

$$\mathbf{P}(\text{candidate 1 always leads the vote tally}) = \mathbf{E}X_T = \mathbf{E}X_0 = \mathbf{E}S_c/c = \frac{a-b}{a+b}.$$

(Since $a > b$, if the first vote is counted for candidate 2, then X_t will be zero for some t. So, $X_T = 1$ if and only if $S_k > 0$ for all $1 \leq k \leq c$. And $X_T = 0$ otherwise. So, $\mathbf{E}X_T = \mathbf{P}(S_k > 0)$ for all $1 \leq k \leq c$. That is, $\mathbf{E}X_T$ is the probability that candidate 1 always leads the vote tally.) □

Exercise 5.6. Let $X_1, X_2, \ldots$ be i.i.d. random variables with $\mathbf{E}X_i = 0$ for every $i \geq 1$. Suppose there exists $\sigma > 0$ such that $\mathrm{Var}(X_i) = \sigma^2$ for all $i \geq 1$. For any $n \geq 1$, let $S_n = X_1 + \cdots + X_n$. Show that $S_n^2 - n\sigma^2$ is a martingale with respect to $X_1, X_2, \ldots$. (We let $X_0 = 0$.)

Let $a > 0$. Let $T = \min\{n \geq 1 \colon |S_n| \geq a\}$. Using the Optional Stopping Theorem, show that $\mathbf{E}T \geq a^2/\sigma^2$. Observe that a simple random walk on $\mathbb{Z}$ has $\sigma^2 = 1$ and $\mathbf{E}T = a^2$ when $a \in \mathbb{Z}$.

(When applying the Optional Stopping Theorem, you do not have to show that the martingale is bounded.)

Solution.

$$\begin{aligned}
&\mathbf{E}(S_{n+1} - S_n \,|\, X_n = x_n, \ldots, X_0 = x_0, S_0 = s_0) \\
&= \mathbf{E}((X_{n+1} + x_1 + \cdots + x_n)^2 - (n+1)\sigma^2 - (x_1 + \cdots + x_n)^2 + n\sigma^2 \,|\, X_n = x_n) \\
&= \mathbf{E}X_{n+1}^2 - 2(x_1 + \cdots + x_n)\mathbf{E}X_{n+1} - \sigma^2\mathbf{E}X_{n+1}^2 - \sigma^2 = 0.
\end{aligned}$$

From the Optional Stopping Theorem,

$$0 = \mathbf{E}S_0 - 0 = \mathbf{E}S_T^2 - \mathbf{E}T\sigma^2.$$

(Note that $\mathbf{P}(T < \infty) = 1$ by the Central Limit Theorem) Since $S_T^2 \geq a^2$, we have

$$\mathbf{E}T = \sigma^{-2}\mathbf{E}S_T^2 \geq \sigma^{-2}a^2.$$

So, if $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = 1/2$ for all $i \geq 1$, we have $\sigma^2 = 1$, and we recover the simple random walk on $\mathbb{Z}$. And when $a \in \mathbb{Z}$, we have $S_T^2 = a^2$, so the above argument shows $\mathbf{E}T = a^2$. Alternatively, we showed in Example 4.31 in the notes that $\mathbf{E}T = a^2$ when $a \in \mathbb{Z}$.
□

Exercise 5.7 (Azuma's Inequality). In this exercise, we prove a generalization of the Hoeffding inequality to martingales. Let $c_1, c_2, \ldots > 0$. Let $(X_0, X_1, \ldots)$ be a martingale. Assume that $|X_n - X_{n-1}| \leq c_n$ for all $n \geq 1$. Then for any $t > 0$,

$$\mathbf{P}(|X_n - X_0| > t) \leq 2e^{-\frac{t^2}{2\sum_{i=1}^n c_i^2}}.$$

Prove this inequality using the following steps.

- Let $\alpha > 0$. Show that $\mathbf{E}e^{\alpha(X_n - X_0)} = \mathbf{E}[e^{\alpha(X_{n-1}-X_0)}\mathbf{E}(e^{\alpha(X_n - X_{n-1})}|\mathcal{F}_{n-1})]$. (When Y is a random variable, we denote $\mathbf{E}(Y|\mathcal{F}_n) := g(X_0, \ldots, X_n)$ where $g(x_0, \ldots, x_n) := \mathbf{E}(Y|X_0 = x_0, \ldots, X_n = x_n)$ for any $x_0, \ldots, x_n \in \mathbb{R}$.)
- For any $y \in [-1, 1]$, show that $e^{\alpha c_n y} \leq \frac{1+y}{2}e^{\alpha c_n} + \frac{1-y}{2}e^{-\alpha c_n}$.
- Take the conditional expectation of this inequality when $y = (X_n - X_{n-1})/c_n$.
- Now argue as in Hoeffding's inequality.

Using Azuma's inequality, deduce **McDiarmid's Inequality**. Let $X_1, \ldots, X_n$ be independent real-valued random variables. Let $c_1, c_2, \ldots > 0$. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a measurable function such that, for any $1 \leq m \leq n$,

$$\sup_{x_1,\ldots,x_{m-1},x_m,x_m',x_{m+1},\ldots,x_n \in \mathbb{R}} |f(x_1, \ldots, x_n) - f(x_1, \ldots, x_{m-1}, x_m', x_{m+1}, \ldots, x_n)| \leq c_m.$$

Then, for any $t > 0$,

$$\mathbf{P}(|f(X_1, \ldots, X_n) - \mathbf{E}f(X_1, \ldots, X_n)| > t) \leq 2e^{-\frac{t^2}{2\sum_{i=1}^n c_i^2}}.$$

(Note that a linear function f recovers Hoeffding's inequality.)

Solution. We write $e^{\alpha(X_n - X_0)} = e^{\alpha(X_n - X_{n-1})}e^{\alpha(X_{n-1}-X_0)}$. Conditioning on $X_0, \ldots, X_{n-1}$ gives

$$\begin{aligned}
&\mathbf{E}[e^{\alpha(X_n - X_0)} \,|\, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0] \\
&= \mathbf{E}[e^{\alpha(X_n - X_{n-1})}e^{\alpha(X_{n-1}-X_0)} \,|\, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0] \\
&= e^{\alpha(X_{n-1}-X_0)}\mathbf{E}[e^{\alpha(X_n - X_{n-1})} \,|\, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0].
\end{aligned}$$

Taking the expected value of both sides,

$$\mathbf{E}[e^{\alpha(X_n - X_0)} = \sum_{x_0,\ldots,x_{n-1}\in\mathbb{R}} \mathbf{E}[e^{\alpha(X_n - X_0)} \,|\, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0]\mathbf{P}(X_{n-1} = x_{n-1}, \ldots, X_0 = x_0)$$

$$= \sum_{x_0,\ldots,x_{n-1}\in\mathbb{R}} e^{\alpha(X_n - X_{n-1})}\mathbf{E}[e^{\alpha(X_{n-1}-X_0)} \,|\, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0]\mathbf{P}(X_{n-1} = x_{n-1}, \ldots, X_0 = x_0)$$

$$= \mathbf{E}e^{\alpha(X_{n-1}-X_0)}\mathbf{E}[e^{\alpha(X_n - X_{n-1})} \,|\, \mathcal{F}_{n-1}].$$

Since $y \mapsto e^{\alpha c_n y}$ is convex, we have

$$e^{\alpha c_n y} = e^{\alpha c_n \left(\frac{1+y}{2}(1)+\frac{1-y}{2}(-1)\right)} \leq \frac{1+y}{2}e^{\alpha c_n} + \frac{1-y}{2}e^{-\alpha c_n}.$$

Taking the conditional expectation of both sides with $y = (X_n - X_{n-1})/c_n$,

$$\begin{aligned}\mathbf{E}[e^{\alpha(X_n - X_{n-1})} \,|\, \mathcal{F}_{n-1}] &= \mathbf{E}[e^{\alpha c_n y} \,|\, \mathcal{F}_{n-1}] \\ &\leq \mathbf{E}[\frac{1+y}{2}e^{\alpha c_n} + \frac{1-y}{2}e^{-\alpha c_n} \,|\, \mathcal{F}_{n-1}] \\ &= \frac{1}{2}[e^{\alpha c_n} + e^{-\alpha c_n}] = \cosh(\alpha c_n) \leq e^{\alpha^2 c_n^2/2}.\end{aligned}$$

That is,

$$\mathbf{E}e^{\alpha(X_n - X_0)} \leq e^{\alpha^2 c_n^2/2}\mathbf{E}e^{\alpha(X_{n-1}-X_0)}.$$

Now iterate in n and argue as in Hoeffding's inequality. □

6. Homework 6

Exercise 6.1. Let $\lambda > 0$. Let $\tau_1, \tau_2, \ldots$ be independent exponential random variables with parameter λ. For any $n \geq 1$, let $T_n = \tau_1 + \cdots + \tau_n$. Fix positive integers $n_k > \cdots > n_1$ and positive real numbers $t_k > \cdots > t_1$. Then

$$f_{T_{n_k},\ldots,T_{n_1}}(t_k, \ldots, t_1) = f_{T_{(n_k - n_{k-1})}}(t_k - t_{k-1}) \cdots f_{T_{(n_2-n_1)}}(t_2 - t_1) f_{T_{n_1}}(t_1).$$

(Hint: just try to case $k = 2$ first, and use a conditional density function.)

Solution. In the case $k = 2$, if we are given $T_{n_1} = t_1$, then we write $T_{n_2} = (T_{n_2} - T_{n_1}) + T_{n_1} = (T_{n_2} - T_{n_1}) + t_1$. And T_{n_1} is independent of $T_{n_2} - T_{n_1}$. So, if $T_{n_2} = t_2$, and if we condition on $T_{n_1} = t_1$, then $T_{n_2} - T_{n_1} = t_2 - t_1$, and this event is independent of T_{n_1}. So, $f_{T_{n_2}|T_{n_1}}(t_2|t_1) = f_{T_{n_2} - T_{n_1}}(t_2 - t_1)$. Then

$$\begin{aligned}\mathbf{P}(T_{n_2} \leq t_2, T_{n_1} \leq t_1) &= \int_{-\infty}^{t_2}\int_{-\infty}^{t_1} f_{T_{n_2},T_{n_1}}(t_2, t_1) dt_2 dt_1 \\ &= \int_{-\infty}^{t_1}\int_{-\infty}^{t_2} f_{T_{n_2}|T_{n_1}}(t_2|t_1) dt_2 f_{T_{n_1}}(t_1) dt_1 \\ &= \int_{-\infty}^{t_1}\int_{-\infty}^{t_2} f_{T_{n_2}-T_{n_1}}(t_2 - t_1) dt_2 f_{T_{n_1}}(t_1) dt_1 \\ &= \int_{-\infty}^{t_1}\int_{-\infty}^{t_2} f_{T_{n_2-n_1}}(t_2 - t_1) dt_2 f_{T_{n_1}}(t_1) dt_1.\end{aligned}$$

In the last line, we used that $T_{n_2} - T_{n_1} = \tau_{n_2} + \tau_{n_2-1} + \cdots + \tau_{n_1}$ has the same distribution function as $T_{n_2-n_1}$. The case when $k = 2$ is therefore complete. We now consider the case when k is larger.

As before, $f_{T_{n_k}|T_{n_{k-1}},\ldots,T_{n_1}}(t_k|t_{k-1},\ldots,t_1) = f_{T_{n_k}-T_{n_{k-1}}}(t_k - t_{k-1}) = f_{T_{n_k-n_{k-1}}}(t_k - t_{k-1})$, so

$$\begin{aligned}
&\mathbf{P}(T_{n_k} \leq t_k, \ldots, T_{n_1} \leq t_1) \\
&= \int_{-\infty}^{t_k} \cdots \int_{-\infty}^{t_1} f_{T_{n_k},\ldots,T_{n_1}}(t_k,\ldots,t_1) dt_k \cdots dt_1 \\
&= \int_{-\infty}^{t_k} \cdots \int_{-\infty}^{t_1} f_{T_{n_k}|T_{n_{k-1}},\ldots,T_{n_1}}(t_k,\ldots,t_1) f_{T_{n_{k-1}},\ldots,T_{n_1}}(t_{k-1},\ldots,t_1) dt_k \cdots dt_1 \\
&= \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} f_{T_{n_k-n_{k-1}}}(t_k - t_{k-1}) f_{T_{n_{k-1}},\ldots,T_{n_1}}(t_{k-1},\ldots,t_1) dt_1.
\end{aligned}$$

That is, for any $t_1,\ldots,t_k \in \mathbb{R}$,

$$f_{T_{n_k},\ldots,T_{n_1}}(t_k,\ldots,t_1) = f_{T_{n_k-n_{k-1}}}(t_k - t_{k-1}) f_{T_{n_{k-1}},\ldots,T_{n_1}}(t_{k-1},\ldots,t_1).$$

Iterating this equality $k-1$ more times proves the assertion. □

Exercise 6.2. Let $s, t > 0$ and let m, n be nonnegative integers. Let $0 < t_m < t_{m+1} < t_{m+n} < t_{m+n+1}$, and define (using the notation of Exercise 6.1),

$$g(t_m, t_{m+1}, t_{m+n}, t_{m+n+1}) := f_{T_1}(t_{m+n+1} - t_{m+n}) f_{T_{n-1}}(t_{m+n} - t_{m+1}) f_{T_1}(t_{m+1} - t_m) f_{T_m}(t_m).$$

Let $\{N(s)\}_{s\geq 0}$ be a Poisson Process with parameter $\lambda > 0$. Show that

$$\begin{aligned}
&\mathbf{P}(N(s+t) = m+n,\, N(s) = m) \\
&= \int_0^s \left(\int_s^{s+t} \left(\int_{t_{m+1}}^{s+t} \left(\int_{s+t}^{\infty} g(t_m, t_{m+1}, t_{m+n}, t_{m+n+1}) dt_{m+n+1} \right) dt_{m+n} \right) dt_{m+1} \right) dt_m.
\end{aligned}$$

(Hint: use the joint density, and then use Exercise 6.1.)

Solution. By definition of the Poisson Process, $N(s) := \max\{n \geq 0 \colon T_n \leq s\}$ for any $s \geq 0$. Using this definition and Exercise 6.1 with $k = 4$ and $n_4 = m+n+1$, $n_3 = m+n$, $n_2 = m+1$, $n_1 = m$, we get

$$\begin{aligned}
&\mathbf{P}(N(s+t) = m+n,\, N(s) = m) \\
&= \mathbf{P}(T_{m+n+1} > s+t,\, T_{m+n} \leq s+t,\, T_{m+1} > s,\, T_m \leq s,) \\
&= \int_0^s \int_s^{s+t} \int_{t_{m+1}}^{s+t} \int_{s+t}^{\infty} f_{T_1}(t_{m+n+1} - t_{m+n}) f_{T_{n-1}}(t_{m+n} - t_{m+1}) f_{T_1}(t_{m+1} - t_m) \\
&\qquad\qquad \cdot f_{T_m}(t_m) dt_{m+n+1} dt_{m+n} dt_{m+1} dt_m \\
&\int_0^s \left(\int_s^{s+t} \left(\int_{t_{m+1}}^{s+t} \left(\int_{s+t}^{\infty} g(t_m, t_{m+1}, t_{m+n}, t_{m+n+1}) dt_{m+n+1} \right) dt_{m+n} \right) dt_{m+1} \right) dt_m.
\end{aligned}$$

□

Exercise 6.3. Suppose you are running a (busy) car wash. The number of red cars that come to the car wash between time 0 and time $s > 0$ is a Poisson poisson with rate 2. The number of blue cars that come to car wash between time 0 and time $s > 0$ is a Poisson poisson with rate 3. Both Poisson processes are independent of each other. All cars are either red or blue. With what probability will five blue cars arrive, before three red cars have arrived?

Solution. By Theorem 5.19 in the notes, the sum of the two poisson processes is a Poisson process with rate 5. We label this process as $\{N(s)\}_{s\geq 0}$ (so $N(s)$ is the total number of cars that have arrived by time s.) Let $Y_1, Y_2, \ldots$ be independent identically distributed random variables with $\mathbf{P}(Y_i = 1) = 2/5$ and $\mathbf{P}(Y_i = 2) = 3/5$ for all $i \geq 1$. Let $N_1(s)$ be the number of integers $j \leq s$ such that $Y_1 = j$, and let $N_2(s)$ be the number of integers $j \leq s$ such that $Y_2 = j$. Then $\{N_1(s)\}_{s\geq 0}$ and $\{N_2(s)\}_{s\geq 0}$ are independent Poisson processes with rates 2 and 3, respectively.

We are asked to find the probability that there are at least 5 blue cars among the first 7 that arrive. As shown in the proof of Theorem 5.17, if $7 = n_1 + n_2$, then

$$\mathbf{P}(N_1(s) = n_1,\, N_2(s) = n_2 \,|\, N(s) = 7) = \frac{7!}{n_1! n_2!}(2/5)^{n_1}(3/5)^{n_2}.$$

We therefore sum this probability over $(n_1, n_2) \in \{(2,5),(1,6),(0,7)\}$. That is, our desired probability is

$$\sum_{j=5}^{7} \mathbf{P}(N_1(s) = 7-j,\, N_2(s) = j \,|\, N(s) = 7) = \sum_{j=5}^{7} \frac{7!}{(7-j)!j!}(2/5)^{7-j}(3/5)^{j}$$

$$= \frac{7!}{5^7}\sum_{j=5}^{7} \frac{1}{(7-j)!j!}(2)^{7-j}(3)^{j} = \frac{1}{5^7}(3^5 2^2(7)(3) + 3^6(2)(7) + 3^7) = \frac{32805}{78125} = \frac{962}{2291}.$$

□

Exercise 6.4. Let m be a positive integer and let P be an $m \times m$ real matrix.

- Show that the sum
$$\sum_{k=0}^{\infty} \frac{P^k}{k!}$$
converges. That is, e^P is well-defined.
- Show that
$$e^{P+I} = e^P e^I.$$
- Find $m \times m$ matrices P, Q such that $e^{P+Q} \neq e^P e^Q$.

Solution. Let $||P||_\infty$ denote the largest element of the entries of P in absolute value. It follows by induction that $\left|\left|P^k\right|\right|_\infty \leq m^k \,||P||_\infty^k$ for any integer $k \geq 1$. So, for any $j \geq 1$, we have

$$\left|\left|\sum_{k=j}^{\infty} \frac{P^k}{k!}\right|\right|_\infty \leq \sum_{k=j}^{\infty} \frac{\left|\left|P^k\right|\right|_\infty}{k!} \leq \sum_{k=j}^{\infty} \frac{[m\,||P||_\infty]^k}{k!}.$$

The last quantity goes to zero as $j \to \infty$, since $e^{m||P||_\infty}$ is well-defined.

$$\sum_{k=0}^{\infty}\frac{(P+I)^k}{k!}=\sum_{k=0}^{\infty}\frac{\sum_{j=0}^{k}\binom{k}{j}P^j}{k!}=\sum_{j=0}^{\infty}P^j\sum_{k=j}^{\infty}\binom{k}{j}\frac{1}{k!}=\sum_{j=0}^{\infty}P^j\sum_{k=j}^{\infty}\frac{1}{j!(k-j)!}$$
$$=\sum_{j=0}^{\infty}\frac{P^j}{j!}\sum_{k=j}^{\infty}\frac{1}{(k-j)!}=\sum_{j=0}^{\infty}\frac{P^j}{j!}\sum_{\ell=0}^{\infty}\frac{1}{\ell!}=e^Pe^I.$$

Finally, we choose P,Q to be real 2×2 matrices that do not commute. Let $Q=\begin{pmatrix}1&1\\0&0\end{pmatrix}$, $P=\begin{pmatrix}0&1\\0&1\end{pmatrix}$. Then $PQ=0$, so

$$e^{P+Q}=\sum_{k=0}^{\infty}\frac{(P+Q)^k}{k!}=\sum_{k=0}^{\infty}\frac{P^k+Q^k}{k!}=e^P+e^Q.$$

Meanwhile, $Q^k=Q$ and $P^k=P$ for all $k\geq 1$. So,

$$e^P=\sum_{k=0}^{\infty}\frac{P^k}{k!}=P\sum_{k=0}^{\infty}\frac{1}{k!}=e\cdot P$$

Similarly, $e^Q=e\cdot Q$. So,

$$e^{P+Q}=e^P+e^Q=e\cdot(P+Q)=e\cdot\begin{pmatrix}1&2\\0&1\end{pmatrix}\neq e^2\begin{pmatrix}0&0\\0&0\end{pmatrix}=e^2\cdot PQ=e^Pe^Q.$$

□

Exercise 6.5. Let m be a positive integer and let P be an $m\times m$ real matrix. Denote $H_t:=e^{t(P-I)}$ for all $t\geq 0$. Let $f\in\mathbb{R}^m$ be a column vector. Then H_tf denotes multiplying the matrix H_t against the vector f. Show the following:

- $H_0=I$.
- $H_{s+t}=H_sH_t$ for all $s,t\geq 0$. (This identity is an analogue of the Chapman-Kolmogorov equation.)
- $H_t1=1$ for all $t\geq 0$. (Moreover, H_t is a stochastic matrix, for all $t\geq 0$. Here 1 denotes the vector of all ones.)
- $\frac{d}{dt}|_{t=0}H_t=\lim_{t\to 0^+}\frac{H_t-H_0}{t}=(P-I)$.
- For any $f\in\mathbb{R}^m$, we have

$$\frac{d}{dt}H_tf=(P-I)H_tf,\qquad\forall\, t\geq 0.$$

Exercise 6.6 (Markov Property, Continuous-Time). Show that a (finite) continuous-time Markov chain satisfies the following Markov property: for all $x,y\in\Omega$, for any $n\geq 1$, $t>0$ and for any $s>s_{n-1}>\cdots>s_0>0$ and for all events H_{n-1} of the form $H_{n-1}=\cap_{k=0}^{n-1}\{X_{s_k}=x_k\}$, where $x_k\in\Omega$ for all $0\leq k\leq n-1$, such that $\mathbf{P}(H_{n-1}\cap\{X_s=x\})>0$, we have

$$\mathbf{P}(X_{t+s}=y\,|\,H_{n-1}\cap\{X_s=x\})=\mathbf{P}(X_t=y\,|\,X_0=x).$$

Exercise 6.7. Prove the following discrete-time version the above spectral gap inequality from class.

Let P be the transition matrix of a finite, irreducible, reversible Markov chain, with state space Ω and with (unique) stationary distribution π. Let

$$\gamma_* := 1 - \max\{|\lambda| : \lambda \text{ is an eigenvalue of } P \text{ with } \lambda \neq 1\}$$

be the absolute spectral gap of P. Then, for any $f \in \mathbb{R}^\Omega$ and for any integer $k \geq 1$,

$$\mathrm{Var}_\pi(P^k f) \leq (1-\gamma_*)^{2k}\mathrm{Var}_\pi f.$$

Exercise 6.8 (Scaling Invariance). Let $a > 0$. Let $\{B(t)\}_{t\geq 0}$ be a standard Brownian motion. For any $t > 0$, define $X(t) := \frac{1}{\sqrt{a}}B(at)$. Then $\{X(t)\}_{t\geq 0}$ is also a standard Brownian motion.

Solution. It suffices to check properties (i),(ii) and (iii) of standard Brownian motion. Properties (i) and (iii) follow immediately. To verify property (ii), let $0 < s < t$. Then $X(t) - X(s) = a^{-1/2}(B(at) - B(as))$. Since $\{B(t)\}_{t\geq 0}$ is a standard Brownian motion, $B(at) - B(as)$ is a Gaussian random variable with mean zero and variance $a(t-s)$. So, $a^{-1/2}(B(at) - B(as))$ is a Gaussian random variable with mean zero and variance $(a^{-1/2})^2 a(t-s) = t-s$, as desired. □

Exercise 6.9. Let $x_1, \ldots, x_n \in \mathbb{R}$, and if $t_n > \cdots > t_1 > 0$. Using the independent increment property, show that the event

$$\{B(t_1) = x_1, \ldots, B(t_n) = x_n\}$$

has a multivariate normal distribution. That is, the joint density of $(B(t_1), \ldots, B(t_n))$ is

$$f(x_1, \ldots, x_n) = f_{t_1}(x_1)f_{t_2-t_1}(x_2 - x_1)\cdots f_{t_n-t_{n-1}}(x_n - x_{n-1})$$

where

$$f_t(x) = \frac{1}{\sqrt{2\pi t}}e^{-x^2/(2t)}, \qquad \forall x \in \mathbb{R},\, t > 0.$$

Solution. In the case $n = 2$, if we are given $B(t_1) = x_1$, then we write $B(t_2) = (B(t_2) - B(t_1)) + B(t_1) = (B(t_2) - B(t_1)) + x_1$. If $B(t_2) = x_2$, then $B(t_2) - B(t_1) = x_2 - x_1$. And $B(t_1)$ is independent of $B(t_2) - B(t_1)$, so

$$\begin{aligned} f_{B(t_2)|B(t_1)}(x_2|x_1) &= f_{[B(t_2)-B(t_1)+B(t_1)]\,|B(t_1)}(x_2|x_1) = f_{[B(t_2)-B(t_1)+x_1]\,|B(t_1)}(x_2|x_1) \\ &= f_{[B(t_2)-B(t_1)]\,|B(t_1)}(x_2 - x_1|x_1) = f_{B(t_2)-B(t_1)}(x_2 - x_1). \end{aligned}$$

By property (ii) of Brownian motion (in Definition 7.1 in the notes), $B(t_2 - t_1)$ is a Gaussian random variable with mean zero and variance $t_2 - t_1$. So, $f_{B(t_2)-B(t_1)}(x) = f_{t_2-t_1}(x)$ for any $x \in \mathbb{R}$. So, using the definition of conditional density, for any $x_1, x_2 \in \mathbb{R}$,

$$f_{B(t_1),B(t_2)}(x_1, x_2) = f_{B(t_2)|B(t_1)}(x_2|x_1)f_{B(t_1)}(x_1) = f_{t_2-t_1}(x_2 - x_1)f_{t_1}(x_1).$$

We now consider more general $n > 2$. As before, for any $x_1, \ldots, x_n \in \mathbb{R}$,

$$f_{B(t_n)|B(t_{n-1}),\ldots,B(t_1)}(x_n|x_1, \ldots, x_{n-1}) = f_{B(t_n)-B(t_{n-1})}(x_n - x_{n-1}) = f_{t_n-t_{n-1}}(x_n - x_{n-1}).$$

So,

$$
\begin{aligned}
f_{B(t_1),\ldots,B(t_n)}(x_1,\ldots,x_n) &= f_{B(t_n)|B(t_1),\ldots,B(t_{n-1})}(x_n|x_1,\ldots,x_{n-1}) f_{B(t_1),\ldots,B(t_{n-1})}(x_1,\ldots,x_{n-1}) \\
&= f_{t_n-t_{n-1}}(x_n - x_{n-1}) f_{B(t_1),\ldots,B(t_{n-1})}(x_1,\ldots,x_{n-1}).
\end{aligned}
$$

Iterating this equality $n-1$ more times proves the assertion. □

7. Homework 7

Exercise 7.1. Let $\{B(t)\}_{t\geq 0}$ be a standard Brownian motion. Then $\{(B(t))^2 - t\}_{t\geq 0}$ is a (continuous-time) martingale in the following sense: it $t > s > 0$, and if $s > s_n > \cdots > s_1 > 0$, and $x_1,\ldots,x_n \in \mathbb{R}$, then

$$\mathbf{E}((B(t))^2 - t - ((B(s))^2 - s) \mid B(s_n) = x_n, \ldots, B(s_1) = x_1) = 0.$$

More generally, for any $\alpha \in \mathbb{R}$, let $Y(t) := e^{\alpha B(t) - \alpha^2 t/2}$. Show that $\{Y(t)\}_{t\geq 0}$ is a martingale.

Then, using the power series expansion of the exponential function, we have $Y(t) = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} M_n(t)$ for some random variables $M_1(t), M_2(t), \ldots$, for any $\alpha \in \mathbb{R}$. It follows that $\{M_1(t)\}_{t\geq 0}$ is a martingale, $\{M_2(t)\}_{t\geq 0}$ is a martingale, etc. (Starting with the following sentence, you do not have to prove anything.) It turns out that

$$M_n(t) = t^{n/2} p_n(B(t)/\sqrt{t}), \quad \forall\, t \in \mathbb{R}, \quad \forall\, n \geq 1,$$

where p_n is the n^{th} Hermite polynomial, so that

$$p_n(x) = e^{x^2/2}(-1)^n \frac{d^n}{dx^n} e^{-x^2/2}, \quad \forall\, x \in \mathbb{R}, \quad \forall\, n \geq 1.$$

For example, using $n = 3$, we know that $\{(B(t))^3 - 3B(t)\}_{t\geq 0}$ is a martingale.

Solution. The martingale property is shown in Exercise 7.4 below with $\sigma = 1$ and $\mu = 0$. That is, for any $t > s > s_n > \cdots + s_1 > 0$ and for any $x_n, \ldots, x_1 \in \mathbb{R}$,

$$\mathbf{E}(Y(t) - Y(s) \mid B(s_n) = x_n, \ldots, B(s_1) = x_1) = 0.$$

Suppose we then expand Y as a power series in α. Then there exists $M_0(t), M_1(t), \ldots$ such that

$$e^{B(t)\alpha - \alpha^2 t/2} = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} M_n(t).$$

Then

$$
\begin{aligned}
0 &= \mathbf{E}(Y(t) - Y(s) \mid B(s_n) = x_n, \ldots, B(s_1) = x_1) \\
&= \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} \mathbf{E}(M_n(t) - M_n(s) \mid B(s_n) = x_n, \ldots, B(s_1) = x_1).
\end{aligned}
$$

Since this expression is zero for any $\alpha \in \mathbb{R}$, we conclude that all of the coefficients in the power series on the right are zero. That is, for any $n \geq 0$,

$$\mathbf{E}(M_n(t) - M_n(s) \mid B(s_n) = x_n, \ldots, B(s_1) = x_1) = 0.$$

That is, each of the coefficients $M_0(t), M_1(t), \ldots$ is a martingale.

To determine the coefficients $M_n(t)$, note that $M_0 = 1$,

$$M_1(t) = \frac{\partial}{\partial\alpha}|_{\alpha=0} e^{B(t)\alpha - \alpha^2 t/2} = B(t).$$

$$M_2(t) = \frac{\partial^2}{\partial\alpha^2}|_{\alpha=0} e^{B(t)\alpha - \alpha^2 t/2} = (B(t))^2 - t.$$

$$M_3(t) = \frac{\partial^3}{\partial\alpha^3}|_{\alpha=0} e^{B(t)\alpha - \alpha^2 t/2} = ((B(t))^2 - t)B(t) - 2tB(t) = (B(t))^3 - 3tB(t).$$

The more general statement about Hermite polynomials follows by the generating function definition of Hermite polynomials. We have

$$e^{xs - s^2/2} = \sum_{n=0}^{\infty} \frac{s^n}{n!} p_n(x), \qquad \forall\, x, s \in \mathbb{R}.$$

So, using $x = B(t)/\sqrt{t}$, and $s = \sqrt{t}\alpha$, we get

$$e^{B(t)\alpha - \alpha^2 t/2} = e^{\frac{B(t)}{\sqrt{t}}(\sqrt{t}\alpha) - (\sqrt{t}\alpha)^2/2} = \sum_{n=0}^{\infty} \frac{(\sqrt{t}\alpha)^n}{n!} p_n\left(\frac{B(t)}{\sqrt{t}}\right) = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} t^{n/2} p_n\left(\frac{B(t)}{\sqrt{t}}\right).$$

So, by the definition of M_n, we have $M_n(t) = t^{n/2} p_n(B(t)/\sqrt{t})$ for all $t \geq 0$, $n \geq 1$. □

Exercise 7.2. Let $\{B(t)\}_{t\geq 0}$ be a standard Brownian motion.

- Given that $B(1) = 10$, what is the expected length of time after $t = 1$ until $B(t)$ hits either 8 or 12?
- Now, let $\sigma = 2$, and $\mu = -5$. Suppose a commodity has price $X(t) = \sigma B(t) + \mu t$ for any time $t \geq 0$. Given that the price of the commodity is 4 at time $t = 8$, what is the probability that the price is below 1 at time $t = 9$?
- Suppose a stock has a price $S(t) = 4e^{B(t)}$ for any $t \geq 0$. That is, the stock moves according to Geometric Brownian Motion. What is the probability that the stock reaches a price of 7 before it reaches a price of 2?

Solution. Let $T := \inf\{t \geq 1\colon B(t) = 8 \text{ or } B(t) = 12\}$. From the independent increment property, $\{B(t+1) - B(1)\}_{t\geq 0}$ is a standard Brownian motion that is independent of $B(1)$. So, given that $B(1) = 10$, $T := \inf\{t \geq 0\colon B(t+1) - B(1) = -2 \text{ or } B(t+1) - B(1) = 2\}$. So, if $\{Z(t)\}_{t\geq 0}$ is a standard Brownian motion and if $S := \inf\{t \geq 0\colon Z(t) = -2 \text{ or } Z(t) = 2\}$, we have

$$\mathbf{E}(T | B(1) = 10) = \mathbf{E}S.$$

From Proposition 7.13 in the notes, $\mathbf{E}S = 4$. So, $\mathbf{E}(T|B(1) = 10) = 4$.

We now answer the second question. It is given that $X(8) = 4$. That is, $\sigma B(8) + 8\mu = 4$, so $B(8) = (4 - 8\mu)/\sigma$. We want to find the probability that $X(9) < 1$, i.e. $\sigma B(9) + 9\mu < 1$, i.e. $B(9) < (1 - 9\mu)/\sigma$. That is, we want to compute the probability that $B(9) - B(8) + B(8) < (1 - 9\mu)/\sigma$. By the independent increment property, $B(9) - B(8)$ is a standard Gaussian random variable which is independent of $B(8)$. So, we need to compute the probability that

$B(9)-B(8)<(1-9\mu)/\sigma+(8\mu-4)/\sigma$. So, if Y is a standard Gaussian random variable, we need to compute

$$\mathbf{P}\left(Y<\frac{-3-\mu}{\sigma}\right)=\int_{-\infty}^{(-3-\mu)/\sigma}e^{-t^2/2}\frac{dt}{\sqrt{2\pi}}.$$

We now answer the third question. If $S(t)=7$ then $B(t)=\log(7/4)$, and if $S(t)=2$, then $B(t)=\log(1/2)$. So, for any $a\in\mathbb{R}$ let $T_a:=\inf\{t\geq 0\colon B(t)=a\}$. We are asked to compute $\mathbf{P}(T_{\log(7/4)}<T_{\log(1/2)})$. From Proposition 7.11, $\mathbf{P}(T_{\log(7/4)}<T_{\log(1/2)})=\frac{\log 2}{\log(7/4)+\log 2}$. □

Exercise 7.3. Fix $x>0$

- Show the bound $\mathbf{P}(-x<B(t)<x)\geq\frac{x}{20\sqrt{t}}$ holds for all $t>x^2$.
- Show that $\mathbf{E}T_x=\infty$. (Recall we observed something similar for the simple random walk on $\mathbb{Z}$.)

Solution.

Let $x>0$ and let $t>0$. Since $B(t)$ is a Brownian motion, $B(t)$ has density $e^{-y^2/(2t)}\frac{1}{\sqrt{2\pi t}}$. If $t>x^2$, and if $y\in[-x,x]$, then $t>y^2$, $y^2/t<1$ and $-y^2/(2t)>-1/2$. So,

$$\mathbf{P}(-x<B(t)<x)=\int_{-x}^{x}e^{-y^2/(2t)}\frac{dy}{\sqrt{2\pi t}}\geq e^{-1/2}\int_{-x}^{x}dy\frac{1}{\sqrt{2\pi t}}=2xe^{-1/2}(2\pi t)^{-1/2}\geq\frac{x}{20\sqrt{t}}.$$

Now, from the Reflection principle, Proposition 7.15 in the notes,

$$\mathbf{P}(T_x>t)=\mathbf{P}(-x<B(t)<x)\geq\frac{x}{20\sqrt{t}}.$$

So, $\mathbf{E}T_x=\int_0^\infty\mathbf{P}(T_x>t)dt\geq\frac{x}{20}\int_{x^2}^\infty t^{-1/2}dt=\infty$. □

Exercise 7.4. Let $\{X(t)\}_{t\geq 0}=\{\sigma B(t)+\mu t\}_{t\geq 0}$ be a standard Brownian motion with variance $\sigma^2>0$ and drift $\mu\in\mathbb{R}$. Fix $\lambda\in\mathbb{R}$. Then $\{Y(t)\}_{t\geq 0}=\{e^{\lambda X(t)-(\lambda\mu+\lambda^2\sigma^2/2)t}\}_{t\geq 0}$ is a (continuous-time) martingale in the following sense: it $t>s>0$, and if $s>s_n>\cdots>s_1>0$, and $x_1,\ldots,x_n\in\mathbb{R}$, then

$$\mathbf{E}(Y(t)-Y(s)\,|\,B(s_n)=x_n,\ldots,B(s_1)=x_1)=0.$$

Solution. As a preliminary calculation, we let $\alpha>0$ and we compute $\mathbf{E}e^Z$ when Z is a Gaussian random variable with mean zero and variance $\beta^2>0$ with $\beta>0$.

$$\begin{aligned}\mathbf{E}e^Z&=\int_{-\infty}^{\infty}e^ze^{-\frac{z^2}{2\beta^2}}\frac{dz}{\sqrt{2\pi}\beta}=\int_{-\infty}^{\infty}e^{\beta z}e^{-\frac{z^2}{2}}\frac{dz}{\sqrt{2\pi}}\qquad\text{, changing variables}\\&=e^{\beta^2/2}\int_{-\infty}^{\infty}e^{-\frac{(z-\beta)^2}{2}}\frac{dz}{\sqrt{2\pi}}=e^{\beta^2/2}.\end{aligned}$$

Now, using the independent increment property and the stationary property of Brownian motion,

$$\begin{aligned}
&\mathbf{E}(Y(t)-Y(s)\,|\,B(s_n)=x_n,\ldots,B(s_1)=x_1)\\
&=\mathbf{E}(e^{\lambda X(t)-(\lambda\mu+\lambda^2\sigma^2/2)t}-e^{\lambda X(s)-(\lambda\mu+\lambda^2\sigma^2/2)s})\,|\,B(s_n)=x_n,\ldots,B(s_1)=x_1)\\
&=\mathbf{E}(e^{\lambda\sigma(B(t)-B(s)+B(s)-B(s_n)+B(s_n))+\lambda\mu t-(\lambda\mu+\lambda^2\sigma^2/2)t}-e^{\lambda\sigma(B(s)-B(s_n)+B(s_n))+\lambda\mu s-(\lambda\mu+\lambda^2\sigma^2/2)s})\\
&\qquad\qquad\qquad\qquad\qquad\qquad|\,B(s_n)=x_n,\ldots,B(s_1)=x_1)\\
&=e^{\lambda\mu t-(\lambda\mu+\lambda^2\sigma^2/2)t}\mathbf{E}(e^{\lambda\sigma(B(t)-B(s)+B(s)-B(s_n)+x_n)}\,|\,B(s_n)=x_n)\\
&\qquad-e^{\lambda\mu s-(\lambda\mu+\lambda^2\sigma^2/2)s}\mathbf{E}(e^{\lambda\sigma(B(s)-B(s_n)+B(s_n))+\lambda\mu s}\,|\,B(s_n)=x_n))\\
&=e^{\lambda\mu t-(\lambda\mu+\lambda^2\sigma^2/2)t}\mathbf{E}(e^{\lambda\sigma(B(t)-B(s))})\mathbf{E}(e^{\lambda\sigma(B(s)-x_n)})e^{\lambda\sigma x_n}\\
&\qquad\qquad-e^{\lambda\mu s-(\lambda\mu+\lambda^2\sigma^2/2)s}\mathbf{E}(e^{\lambda\sigma(B(s)-x_n)})e^{\lambda\sigma x_n}
\end{aligned}$$

Now, $\lambda\sigma(B(t)-B(s))$ is a Gaussian random variable with mean zero and variance $\lambda^2\sigma^2(t-s)$. So

$$\begin{aligned}
&\mathbf{E}(Y(t)-Y(s)\,|\,B(s_n)=x_n,\ldots,B(s_1)=x_1)\\
&=e^{\lambda\mu t-(\lambda\mu+\lambda^2\sigma^2/2)t}e^{\lambda^2\sigma^2(t-s)/2}\mathbf{E}(e^{\lambda\sigma(B(s)-x_n)})e^{\lambda\sigma x_n}\\
&\qquad-e^{\lambda\mu s-(\lambda\mu+\lambda^2\sigma^2/2)s}\mathbf{E}(e^{\lambda\sigma(B(s)-x_n)})e^{\lambda\sigma x_n}=0.
\end{aligned}$$

□

Exercise 7.5. Let $\{X(t)\}_{t\geq 0}=\{\sigma B(t)+\mu t\}_{t\geq 0}$ be a standard Brownian motion with variance $\sigma^2>0$ and negative drift $\mu<0$. Let $a<0<b$. Let $T:=\inf\{t\geq 0\colon X(t)\in\{a,b\}\}$. Let $\alpha:=2\,|\mu|\,/\sigma^2$. Show that

$$\mathbf{E}T=\frac{1}{\mu}\cdot\frac{b(1-e^{\alpha a})+a(e^{\alpha b}-1)}{e^{\alpha b}-e^{\alpha a}}.$$

Solution. Since $X(t)-\mu t$ is a martingale, $\mathbf{E}(X(T)-\mu T)=0$, so $\mathbf{E}T=\frac{1}{\mu}\mathbf{E}X(T)$ by the Optional Stopping Theorem. Let p be the probability that $X(t)$ takes the value b before taking the value a. From Proposition 7.22 in the notes, $p=\frac{1-e^{\alpha b}}{e^{\alpha b}-e^{\alpha a}}$. So,

$$\mathbf{E}T=\frac{1}{\mu}\mathbf{E}X(T)=\frac{1}{\mu}(bp+a(1-p))=\frac{1}{\mu}\left(b\frac{1-e^{\alpha b}}{e^{\alpha b}-e^{\alpha a}}+a\frac{1-e^{\alpha a}}{e^{\alpha b}-e^{\alpha a}}\right).$$

(Unfortunately $\{X(t\wedge T)-\mu t\wedge T\}_{t\geq 0}$ is not bounded. So, strictly speaking, we cannot apply an analogue of the Optional Stopping Theorem Version 2. But, we can still verify that $\mathbf{P}(T<\infty)=1$. Note that $\mathbf{P}(T<\infty)\geq\mathbf{P}(T_a<\infty)$, and if $T_a<\infty$, then $a=X(T_a)=\sigma B(T_a)+\mu T_a\leq\sigma B(T_a)$. So, if we define $T_a':=\inf\{t\geq 0\colon B(t)=a/\sigma\}$, then $T_a<\infty$ implies $T_a'<\infty$, by property (i) of Brownian motion. So, $\mathbf{P}(T<\infty)\geq\mathbf{P}(T_a<\infty)\geq\mathbf{P}(T_a'<\infty)$, and $\mathbf{P}(T_a'<\infty)=1$ by the Reflection Principle, since $\mathbf{P}(T_a'<\infty)=1-\lim_{s\to\infty}\int_{-a/\sigma}^{a/\sigma}e^{-\frac{y^2}{2s}}\frac{ds}{\sqrt{2\pi s}}=1.$) □

Exercise 7.6. Let $\{X(t)\}_{t\geq 0} = \{\sigma B(t) + \mu t\}_{t\geq 0}$ be a standard Brownian motion with variance $\sigma^2 > 0$ and negative drift $\mu < 0$. Let $a < 0$. Let $T_a := \inf\{t \geq 0\colon X(t) = a\}$. Let $\alpha := 2\,|\mu|\,/\sigma^2$. Show that

$$\mathbf{E}T_a = \frac{a}{\mu}.$$

Solution. Since $X(t) - \mu t$ is a martingale, $\mathbf{E}(X(T_a) - \mu T_a) = 0$, so $\mathbf{E}T_a = \frac{1}{\mu}\mathbf{E}X(T_a) = \frac{a}{\mu}$ by the Optional Stopping Theorem.

(Unfortunately $\{X(t\wedge T) - \mu t\wedge T\}_{t\geq 0}$ is not bounded. So, strictly speaking, we cannot apply an analogue of the Optional Stopping Theorem, Version 2. But we have verified above that $\mathbf{P}(T_a < \infty) = 1$.) □

Exercise 7.7 (Optional)**.** Write a computer program to simulate standard Brownian motion. More specifically, the program should simulate a random walk on $\mathbb{Z}$ with some small step size such as .002. (That is, simulate $B_k(t)$ when $k = 500^2$ and, say, $0 \leq t \leq 1$.)

Solution. Here is a program that plots sample paths of $B_k(t)$ with $0 \leq t \leq 1$, which allows several paths to occur on the same plot, each colored with a randomly chosen color.

```
k=500^2;
length=1;
numpts=length*k;
t=linspace(0,length,numpts);
jumps=2*ceil(2*rand(1,numpts)-1)-1;
y=zeros(1,numpts);
for i=2:numpts
    y(i)=y(i-1)+jumps(i);
end
plot(t,y/sqrt(k),'Color',rand(1,3));
hold on;
```

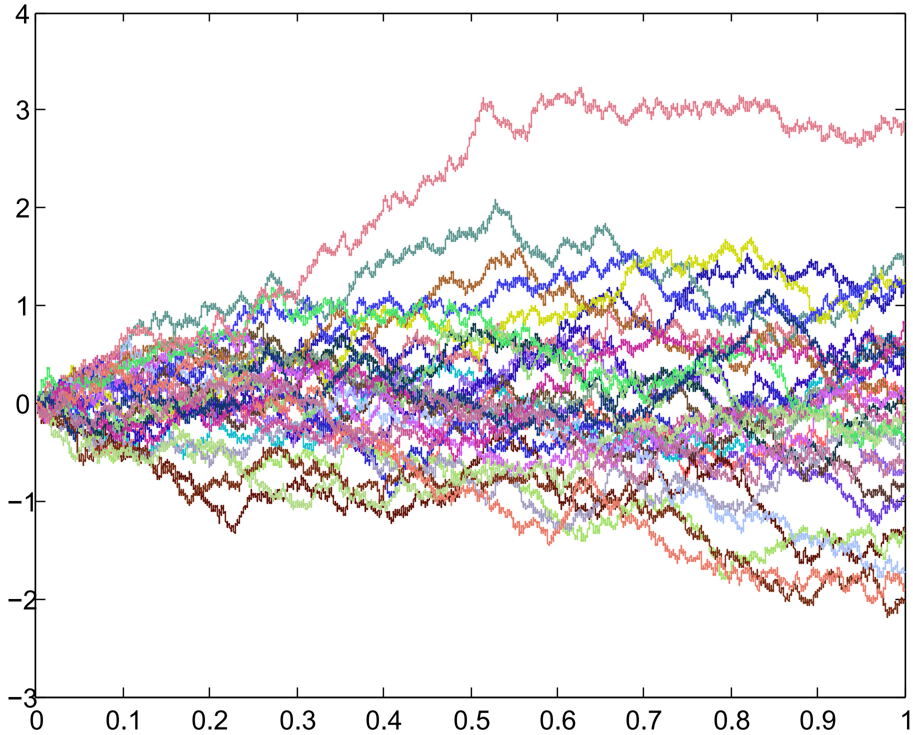
□

Exercise 7.8 (Optional)**.** The following exercise assumes familiarity with Matlab and is derived from Cleve Moler's book, Numerical Computing with Matlab.

The file `brownian.m` plots the evolution of a cloud of particles that starts at the origin and diffuses in a two-dimensional random walk, modeling the Brownian motion of gas molecules.

(a) Modify `brownian.m` to keep track of both the average and the maximum particle distance from the origin. Using loglog axes, plot both sets of distances as functions of n, the number of steps. You should observe that, on the log-log scale, both plots are nearly linear. Fit both sets of distances with functions of the form $cn^{1/2}$. Plot the observed distances and the fits, using linear axes.

(b) Modify `brownian.m` to model a random walk in three dimensions. Do the distances behave like $n^{1/2}$?



The program below answers part (a). We find experimentally that $(1/4)n^{1/2}$ matches the average displacement, and $(3/5)n^{1/2}$ matches the maximum displacement.

```
delta = .002;
x = zeros(100,2);
axis([-1 1 -1 1])
axis square
numpts=10000;
for i=1:numpts
   x = x + delta*randn(size(x));
   maxdist(i)=max(sqrt((x(:,1)).^2 +(x(:,2)).^2));
   avedist(i)=mean(sqrt((x(:,1)).^2 +(x(:,2)).^2));
end
t=linspace(0,1,numpts);
loglog(t,avedist,t,(1/4)*sqrt(t)); title('Average Displacement');
figure; loglog(t,maxdist,t,(3/5)*sqrt(t)); title('Maximum Displacement');
```

The program below answers part (b). We find experimentally that $(1/3)n^{1/2}$ matches the average displacement, and $(7/10)n^{1/2}$ matches the maximum displacement.

```
delta = .002;
x = zeros(100,3);
axis([-1 1 -1 1])
axis square
```

```
numpts=10000;
for i=1:numpts
   x = x + delta*randn(size(x));
   maxdist(i)=max(sqrt((x(:,1)).^2 +(x(:,2)).^2 +(x(:,3)).^2));
   avedist(i)=mean(sqrt((x(:,1)).^2 +(x(:,2)).^2 +(x(:,3)).^2));
end
t=linspace(0,1,numpts);
loglog(t,avedist,t,(1/3)*sqrt(t)); title('Average Displacement');
loglog(t,maxdist,t,((7/10)*sqrt(t))); title('Maximum Displacement');
```

Exercise 7.9 (**Binomial Option Pricing Model**). Let $u,d>0$. Let $0<p<1$. Let $(X_1,X_2,\ldots)$ be independent random variables such that $\mathbf{P}(X_n=\log u)=:p$ and $\mathbf{P}(X_n=\log d)=1-p\ \forall\ n\geq 1$. Let X_0 be a fixed constant. Let $Y_n:=X_0+\cdots+X_n$, and let $S_n:=e^{Y_n}\ \forall\ n\geq 1$. In general, $S_0,S_1,\ldots$ will not be a martingale, but we can still compute $\mathbf{E}S_n$, by modifying $S_0,S_1,\ldots$ to be a martingale.

First, note that if $n\geq 1$, then Y_n has a binomial distribution, in the sense that

$$\mathbf{P}(Y_n=X_0+i\log u+(n-i)\log d)=\binom{n}{i}p^i(1-p)^{n-i},\qquad \forall\, 0\leq i\leq n.$$

Now define

$$r:=p(u-d)-1+d.$$

Here we chose r so that $p=\frac{1+r-d}{u-d}$. For any $n\geq 1$, define

$$M_n:=(1+r)^{-n}S_n.$$

Show that $M_0,M_1,\ldots$ is a martingale with respect to $X_0,X_1,\ldots$. Consequently,

$$(1+r)^{-n}\mathbf{E}S_n=\mathbf{E}S_0,\qquad \forall\, n\geq 0.$$

(This presentation might be a bit backwards from the financial perspective. Typically, r is a fixed interest rate, and then you choose p such that $p=\frac{1+r-d}{u-d}$. That is, you adjust how the random variables behave in order to get a martingale.)

Solution.

$$\begin{aligned}
&\mathbf{E}(M_{n+1}-M_n\,|\,M_0=m_0,X_0=x_0,\ldots,X_n=x_n)\\
&=(1+r)^{-n}\mathbf{E}((1+r)^{-1}S_{n+1}-S_n\,|\,M_0=m_0,X_0=x_0,\ldots,X_n=x_n)\\
&=(1+r)^{-n}\mathbf{E}((1+r)^{-1}e^{X_{n+1}+x_0+\cdots+x_n}-e^{x_0+\cdots+x_n}\,|\,M_0=m_0,X_0=x_0,\ldots,X_n=x_n)\\
&=(1+r)^{-n}e^{x_0+\cdots+x_n}\mathbf{E}((1+r)^{-1}e^{X_{n+1}}-1)\\
&=(1+r)^{-n}e^{x_0+\cdots+x_n}[(1+r)^{-1}(pu+(1-p)d)-1]\\
&=(1+r)^{-n}e^{x_0+\cdots+x_n}[(1+r)^{-1}(1+r)-1]=0.
\end{aligned}$$

In the last line, we used the definition of r. □

Exercise 7.10. There are many ways to try to value an American Put Option. One way is to emulate the formula for a European Put Option which is exercised at time $0\leq t\leq t_0$:

$$e^{-(\mu+\sigma^2/2)t}\mathbf{E}\max(k-S(t),0)$$

We would like to simply take the maximum of the above quantity over all $t \in [0, t_0]$. However, this would be equivalent to knowing the future price of the stock at all times, which is unrealistic. So, we instead consider replacing the variable t by a stopping time. Suppose T is a stopping time. That is, $T(t) \geq 0$ is only allowed to depend on values of $S(t')$ where $t' < t$. Then we could try to maximize the quantity

$$\mathbf{E}e^{-(\mu+\sigma^2/2)T} \max(k - S(T), 0)$$

over all stopping times T where $0 \leq T \leq t_0$. To approximate that quantity, let $0 \leq t_1 \leq t_0$ and just consider stopping times T of the form $T = \inf\{t_1 \leq t \leq t_0 \colon S(t) < S(t') \,\forall\, 0 \leq t' \leq (3/4)t_1\}$, or $T = t_0$ if the set of t inside the infimum is empty. Then, using a computer, compute the maximum over all $0 \leq t_1 \leq t_0$ of

$$\mathbf{E}e^{-(\mu+\sigma^2/2)T} \max(k - S(T), 0)$$

when $\mu = 0, = \sigma = t_0 = S_0 = 1$ and $k = 2$.

This procedure is analogous to the solution of the Secretary Problem.

In order to compute the expected value, use a Monte Carlo simulation of Brownian motion, and take the average value over many runs of the simulation.

The following Matlab problem estimated the price to be .17

```
%Using a bunch of inefficient for-loops, we compute the value of the option
k=500;  % spacing between points is 1/k
numpts=k;
t1index=k/2;  %this is the index of t_1
mu=0;
sigma=1;
Szero=1;
K=2;

for t1index=1:numpts

    for j=1:1000   % average over this many runs
        x=zeros(1,numpts);  % path of the brownian motion
        for i=1:(numpts-1)   %iteratively create the path
            x(i+1) = x(i) + (k^(-1/2))*(1+2*floor(rand - 1/2));
        end
        minval=min(x(1:t1index));  %the minimum of x before t_1
        Stprime= x((t1index+1):numpts);   % values of x after t_1
        T=min(find(Stprime<minval))+t1index;
        if isempty(T)
            T=k;
        end

        output(j) =(exp((-(mu+ (sigma^2)/2))*(T/numpts)))*max(K-(Szero)*exp(x(T)),0);
    end
```

```
    expectedvalue(t1index)=mean(output);  %average of many runs of output
end

finalanswer=max(expectedvalue)   %maximum of the expected value over t_1
```

Exercise 7.11. In each of the following examples, choose a few parameters (e.g. use $\mu = 0$, $\sigma = S_0 = t = 1$ and $k = 2$.), and value the option using several runs of a Monte Carlo simulation of Brownian motion. In each case, we multiply by an exponential term in order to emulate the Black-Scholes formula.

(i) (**Asian Call Option**) The value of an Asian option with strike price $k > 0$ at time $t > 0$ is computed using the average value of the stock from time 0 to time t. That is, if the option is exercised at time $t > 0$, then its value is
$$e^{-(\mu+\sigma^2/2)t}\mathbf{E}\max\left((\frac{1}{t}\int_0^t S(r)dr) - k, 0\right).$$
(ii) (**Lookback Call Option**) The value of a lookback call option with strike price $k > 0$ at time $t > 0$ is computed using the maximum value of the stock between time 0 and time t. That is, if the option is exercised at time $t > 0$, then its value is
$$e^{-(\mu+\sigma^2/2)t}\mathbf{E}\max\left(\max_{0\leq r\leq t} S(r) - k, 0\right).$$
In other words, you can "look back" over the past behavior of the stock, and choose the best price possible over the past.

(iii) (**Lookback Put Option**) The value of a lookback put option with strike price $k > 0$ at time $t > 0$ is computed using the minimum value of the stock between time 0 and time t. That is, if the option is exercised at time $t > 0$, then its value is
$$e^{-(\mu+\sigma^2/2)t}\mathbf{E}\max\left(k - \min_{0\leq r\leq t} S(r), 0\right).$$

Finally, using a Corollary from the notes (which gives the CDF of the maximum of Brownian motion), give an exact formula for the value of the Lookback Call Option. (And check that this formula agrees with the results of your simulation.)

Can you also give an explicit formula for the value of the Lookback Put Option?

The following program estimates the value of the Asian call option with $\mu = \sigma = S_0 = t = 1$ and $k = 2$ to be about .14.

```
k=5000;  % spacing between points is 1/k
numpts=k;
t1index=k/2;  %this is the index of t_1
mu=0;
sigma=1;
Szero=1;
K=2;
```

```
for j=1:10000   % average over this many runs
    x=zeros(1,numpts);  % path of the brownian motion
    for i=1:(numpts-1)   %iteratively create the path
        x(i+1) = x(i) + (k^(-1/2))*(1+2*floor(rand - 1/2));
    end
    output(j) =(exp((-(mu+ (sigma^2)/2))))*max(sum(Szero*exp(x))/k -K,0);
end
expectedvalue=mean(output)  %average of many runs of output
```

To value the lookback call option, we only change the corresponding line to

```
output(j) =(exp((-(mu+ (sigma^2)/2))))*max(max(Szero*exp(x)) -K,0);
```

The value is estimated at about .667

To value the lookback put option, we only change the corresponding line to

```
output(j) =(exp((-(mu+ (sigma^2)/2))))*max(K-min(Szero*exp(x)) ,0);
```

The value is estimated at about .89

From an exercise from the notes, $\mathbf{P}(\max_{0\leq s\leq t} B(s)\geq x)=1-\int_{-x}^{x}e^{-\frac{y^2}{2t}}\frac{dy}{\sqrt{2\pi t}}$. So, $Z:=\max_{0\leq s\leq t} B(s)$ has density

$$f_Z(x)=2e^{-x^2/(2t)}\frac{1}{\sqrt{2\pi t}},\qquad x>0$$

(This next part of the question was unintentionally difficult. Below, for simplicity, we assume that $\mu=0$. More generally, if $\mu\neq 0$ and if $Y:=\max_{0\leq s\leq t}(\sigma B(s)+\mu s)$, then for any $y\geq 0$,

$$\mathbf{P}(Y\leq y)=\Phi\Big(\frac{y-\mu t}{\sigma\sqrt{t}}\Big)-e^{2\mu y/\sigma^2}\Phi\Big(\frac{-y-\mu t}{\sigma\sqrt{t}}\Big).$$

This can be used to price the lookback call option in full generality, but we will not do so below.)

Since $\max_{0\leq s\leq t} e^{B(s)}=e^{\max_{0\leq s\leq t} B(s)}$, the lookback call option (with $\mu=0$) is valued at

$$\begin{aligned}&e^{-\sigma^2t/2}\mathbf{E}\max\left(\max_{0\leq r\leq t}S(r)-k,0\right)\\&=2e^{-\sigma^2t/2}\int_0^\infty\max(S_0e^{\sigma x}-k,0)e^{-x^2/(2t)}\frac{1}{\sqrt{2\pi t}}dx\end{aligned}$$

Choosing $k=2$, $\mu=0$, $\sigma=t=S_0=1$, we get

$$2e^{-1/2}\int_0^\infty\max(e^x-2,0)e^{-x^2/2}\frac{1}{\sqrt{2\pi}}dx$$

This evaluates numerically to about .6488. In fact, we can get a formula similar to the Black-Scholes formula as follows.

$$
\begin{aligned}
&\int_0^\infty \max(S_0 e^{\sigma x} - k, 0) e^{-x^2/(2t)} \frac{1}{\sqrt{2\pi t}} dx \\
&= \frac{1}{\sqrt{2\pi t}} \int_{\max(\frac{\log(k/S_0)}{\sigma},0)}^\infty (S_0 e^{\sigma x} - k) e^{-x^2/(2t)} dx \\
&= \frac{1}{\sqrt{2\pi t}} \int_{\max(\frac{\log(k/S_0)}{\sigma},0)}^\infty S_0 e^{\sigma x} e^{-x^2/(2t)} dx - \frac{k}{\sqrt{2\pi t}} \int_{\max(\frac{\log(k/S_0)}{\sigma},0)}^\infty e^{-x^2/(2t)} dx \\
&= \frac{S_0}{\sqrt{2\pi}} \int_{\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0)}^\infty e^{y\sigma\sqrt{t}} e^{-y^2/2} dy - \frac{k}{\sqrt{2\pi}} \int_{\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0)}^\infty e^{-y^2/2} dy \\
&= \frac{S_0 e^{t\sigma^2/2}}{\sqrt{2\pi}} \int_{\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0)}^\infty e^{-(y-\sigma\sqrt{t})^2/2} dz - k\left(1 - \Phi\left(\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0)\right)\right) \\
&= \frac{S_0 e^{\sigma^2 t/2}}{\sqrt{2\pi}} \int_{\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0)-\sigma\sqrt{t}}^\infty e^{-z^2/2} dy - k\Phi\left(-\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0)\right) \\
&= S_0 e^{\sigma^2 t/2}\left(1 - \Phi\left(\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0) - \sigma\sqrt{t}\right)\right) - k\Phi\left(-\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0)\right) \\
&= S_0 e^{\sigma^2 t/2}\Phi\left(-\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0) + \sigma\sqrt{t}\right) - k\Phi\left(-\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0)\right).
\end{aligned}
$$

So, the price of the option is

$$2S_0\Phi\left(-\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0) + \sigma\sqrt{t}\right) - 2ke^{-\sigma^2 t/2}\Phi\left(-\sqrt{t}\max(\frac{\log(k/S_0)}{\sigma},0)\right).$$

Choosing $k = 2$, $\mu = 0$, $\sigma = t = S_0 = 1$, we get

$$2\Phi(-\log(2) + 1) - 4e^{-1/2}\Phi(-\log(2)) \approx .6488.$$

We now consider the lookback put option. Since $\mathbf{P}(\max_{0\leq s\leq t} B(s) \geq x) = 1 - \int_{-x}^x e^{-\frac{y^2}{2t}} \frac{dy}{\sqrt{2\pi t}}$, multiplying the inequality by -1 and using that $-B(s)$ has the same distribution as $B(s)$,

$$\mathbf{P}(\min_{0\leq s\leq t} B(s) \leq -x) = 1 - \int_{-x}^x e^{-\frac{y^2}{2t}} \frac{dy}{\sqrt{2\pi t}}, \qquad x > 0.$$

So, $Z := \min_{0\leq s\leq t} B(s)$ has density

$$f_Z(x) = 2e^{-x^2/(2t)} \frac{1}{\sqrt{2\pi t}}, \qquad x < 0$$

Since $\min_{0\leq s\leq t} e^{B(s)} = e^{\min_{0\leq s\leq t} B(s)}$, the lookback put option (with $\mu = 0$) is valued at

$$
\begin{aligned}
&e^{-\sigma^2 t/2}\mathbf{E}\max\left(k - \min_{0\leq r\leq t} S(r), 0\right) \\
&= 2e^{-\sigma^2 t/2} \int_{-\infty}^0 \max(k - S_0 e^{\sigma x}, 0) e^{-x^2/(2t)} \frac{1}{\sqrt{2\pi t}} dx.
\end{aligned}
$$

And a formula similar to the Black-Scholes formula can once again be derived.

$$\begin{aligned}
&\int_{-\infty}^{0}\max(k-S_0e^{\sigma x},0)e^{-x^2/(2t)}\frac{1}{\sqrt{2\pi t}}dx\\
&=\frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\min(\frac{\log(k/S_0)}{\sigma},0)}(S_0e^{\sigma x}-k)e^{-x^2/(2t)}dx\\
&=\frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\min(\frac{\log(k/S_0)}{\sigma},0)}S_0e^{\sigma x}e^{-x^2/(2t)}dx-\frac{k}{\sqrt{2\pi t}}\int_{-\infty}^{\min(\frac{\log(k/S_0)}{\sigma},0)}e^{-x^2/(2t)}dx\\
&=\frac{S_0}{\sqrt{2\pi}}\int_{-\infty}^{\sqrt{t}\min(\frac{\log(k/S_0)}{\sigma},0)}e^{y\sigma\sqrt{t}}e^{-y^2/2}dy-\frac{k}{\sqrt{2\pi}}\int_{-\infty}^{\sqrt{t}\min(\frac{\log(k/S_0)}{\sigma},0)}e^{-y^2/2}dy\\
&=\frac{S_0e^{t\sigma^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{\sqrt{t}\min(\frac{\log(k/S_0)}{\sigma},0)}e^{-(y-\sigma\sqrt{t})^2/2}dz-k\Phi\left(\sqrt{t}\min(\frac{\log(k/S_0)}{\sigma},0)\right)\\
&=\frac{S_0e^{t\sigma^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{\sqrt{t}\min(\frac{\log(k/S_0)}{\sigma},0)-\sigma\sqrt{t}}e^{-z^2/2}dy-k\Phi\left(\sqrt{t}\min(\frac{\log(k/S_0)}{\sigma},0)\right)\\
&=S_0e^{t\sigma^2/2}\Phi\left(\sqrt{t}\min(\frac{\log(k/S_0)}{\sigma},0)-\sigma\sqrt{t}\right)-k\Phi\left(\sqrt{t}\min(\frac{\log(k/S_0)}{\sigma},0)\right).
\end{aligned}$$

So, the price of the option is

$$2S_0\Phi\left(\sqrt{t}\min(\frac{\log(k/S_0)}{\sigma},0)-\sigma\sqrt{t}\right)-2ke^{-t\sigma^2/2}\Phi\left(\sqrt{t}\min(\frac{\log(k/S_0)}{\sigma},0)\right).$$

EXERCISES BELOW ARE OPTIONAL

Exercise 7.12. In the discrete binomial model, we can find a price for an American put option using dynamic programming.

Recall this model. Let $u,d>0$. Let $0<p<1$. Let $(X_1,X_2,\ldots)$ be independent random variables such that $\mathbf{P}(X_n=\log u)=:p$ and $\mathbf{P}(X_n=\log d)=1-p\ \forall\ n\geq 1$. Let X_0 be a fixed constant. Let $Y_n:=X_0+\cdots+X_n$, and let $S_n:=e^{Y_n}\ \forall\ n\geq 1$. Let $r:=p(u-d)-1+d$. For any $n\geq 1$, define $M_n:=(1+r)^{-n}S_n$. Recall that $M_0,M_1,\ldots$ is a martingale.

Note that, at time n, the random variable S_n has $n+1$ possible values. Label these values as $S_{n,1}\leq\cdots\leq S_{n,m}$. Let $k>0$. Let $V_{n,m}$ be the value of the American put option at time $n>0$ with strike price k, when S_n has its m^{th} value. Then

$$V_{n,m}=\max\Big(\max(k-S_{n,m},0)\,,\,(1+r)^{-1}(pV_{n+1,m+1}+(1-p)V_{n+1,m})\Big),\qquad\forall\,1\leq m\leq n+1.$$

This recursion formula holds since, at step n, you can either exercise the option at time n, or you can wait and see what happens at time $n+1$. The quantity $\max(k-S_{n,m},0)$ is your revenue from exercising at time n, and the second quantity $(1+r)^{-1}(pV_{n+1,m+1}+(1-p)V_{n+1,m})$ is your expected revenue from waiting until time $n+1$ to exercise the option. So, at time n, you choose the maximum of these two quantities.

Let's solve this recursion in the following example. Suppose $S_0=8$, $p=1/2$, $u=2$, $d=1/2$ (so that $r=1/4$), and $k=10$. And suppose the option expires at time $n=3$ (so that $V_{3,m} = \max(k - S_{3,m}, 0)$ is known for each $1 \leq m \leq 4$.) Then, working backwards, eventually find $V_{0,1}$, the price of the option.

Compare your result in this example with the price of the European put option with the same parameters. (It should be smaller.)

Solution. When $n = 3$, S_n can take four possible values: $S_0u^3, S_0u^2d, S_0u^1d^2$ and S_0d^3. That is, $S_{3,1} = 1$, $S_{3,2} = 4$, $S_{3,3} = 16$ and $S_{3,4} = 64$. The recursion above can be written as

$$V_{n,m} = \max\Big(\max(10 - S_{n,m}, 0)\,,\,(2/5)(V_{n+1,m+1} + V_{n+1,m})\Big), \qquad \forall\, 1 \leq m \leq n+1.$$

We solve this recursion by slowly filling out a tree of values, as follows. Each entry in the matrix is $(S_{n,m}, V_{n,m})$. (Entries are labeled red when the first quantity in the recursion exceeds the second.)

$$\begin{pmatrix} & & & (64,0) \\ & & ? & (16,0) \\ & ? & ? & (4,6) \\ ? & ? & ? & (1,9) \end{pmatrix}, \qquad \begin{pmatrix} & & & (64,0) \\ & & (32,0) & (16,0) \\ & ? & (8,12/5) & (4,6) \\ ? & ? & \color{red}{(2,8)} & (1,9) \end{pmatrix}$$

$$\begin{pmatrix} & & & (64,0) \\ & & (32,0) & (16,0) \\ & (16,24/25) & (8,12/5) & (4,6) \\ (8,348/125) & \color{red}{(4,6)} & \color{red}{(2,8)} & (1,9) \end{pmatrix}$$

So, the value of the put option is 348/125. By contrast, the American call option has price

$$\begin{aligned} &(1+r)^{-3}\mathbf{E}\max(S_3 - 10, 0) \\ &= (1+r)^{-3}\Big(\binom{3}{0}(1-p)^3\max(10 - 8d^3, 0) + \binom{3}{1}p(1-p)^2\max(10 - 8ud^2, 0) \\ &\qquad + \binom{3}{2}p^2(1-p)\max(10 - 8u^2d, 0) + \binom{3}{3}p^3\max(10 - 8u^3, 0)\Big) \\ &= (4/5)^3\Big((1/8)(9) + 3(1/8)(6)\Big) = (4/5)^3(27/8) = 216/125. \end{aligned}$$

□

Exercise 7.13. Let $\{B(t)\}_{t\geq 0}$ be a standard Brownian motion. Let $f\colon \mathbb{R} \to \mathbb{R}$. Assume that $\int_{\mathbb{R}} |f(x)|\,dx < \infty$ and $\int_{\mathbb{R}} f(x)dx = 1$. For any $s > 0$, define

$$X(s) := \frac{1}{\sqrt{s}}\int_0^s f(B(t))dt.$$

Show that $\lim_{s\to\infty}\mathbf{E}X(s) = \sqrt{2/\pi}$. Then, for an optional challenge, show that $\lim_{s\to\infty}\mathbf{E}(X(s))^2 = 1$. (Hint: for the second part, look up the formula for a multivariate normal random variable.)

Solution. First, recall that $B(t)$ has mean zero, variance t, so it has density $(2\pi t)^{-1/2}e^{-x^2/(2t)}$. Then

$$EX(s) = \frac{1}{\sqrt{s}}\int_0^s \int_{\mathbb{R}} f(x)(2\pi t)^{-1/2}e^{-x^2/(2t)}dxdt.$$

Changing variables to $u=t/s$ so that $du=dt/s$, we get

$$EX(s)=(2\pi)^{-1/2}\int_0^1 u^{-1/2}\int_{\mathbb{R}} f(x)e^{-x^2/(2us)}dxdu.$$

Note that $\lim_{s\to\infty} e^{-x^2/(2us)}=1$, so that

$$\lim_{s\to\infty}\mathbf{E}X(s)=(2\pi)^{-1/2}\int_0^1 u^{-1/2}\int_{\mathbb{R}} f(x)dxdu=(2\pi)^{-1/2}\int_0^1 u^{-1/2}du\int_{\mathbb{R}} f(x)dx=\sqrt{\frac{2}{\pi}}.$$

Now, if $t,u>0$, then $(B(t),B(u))$ is a two-dimensional Gaussian random variable such that $\mathbf{E}B(t)B(u)=\min(t,u)$. That is, $(B(t),B(u))$ has the following multivariate normal density

$$\frac{1}{2\pi\sqrt{\left|\det\begin{pmatrix} t & \min(t,u)\\ \min(t,u) & u\end{pmatrix}\right|}}\exp\left(-\frac{1}{2}(x,y)\begin{pmatrix} t & \min(t,u)\\ \min(t,u) & u\end{pmatrix}^{-1}(x,y)^T\right)$$

$$=\frac{1}{2\pi\sqrt{|tu-[\min(t,u)]^2|}}\exp\left(-\frac{(x,y)\begin{pmatrix} u & -\min(t,u)\\ -\min(t,u) & t\end{pmatrix}(x,y)^T}{2(tu-[\min(t,u)]^2)}\right)$$

$$=\frac{1}{2\pi\sqrt{|tu-[\min(t,u)]^2|}}\exp\left(-\frac{(ux^2+ty^2-2xy\min(t,u))}{2(tu-[\min(t,u)]^2)}\right)$$

$$E(X(s))^2=\mathbf{E}\frac{1}{s}\int_0^s f(B(t))dt\int_0^s f(B(u))du$$
$$=\frac{1}{s}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_0^s\int_0^s\frac{f(x)f(y)}{2\pi\sqrt{|tu-[\min(t,u)]^2|}}e^{-\frac{(ux^2+ty^2-2xy\min(t,u))}{2(tu-[\min(t,u)]^2)}}dtdudxdy.$$

Changing variables to $a=t/s$, $b=u/s$, we get

$$E(X(s))^2$$
$$=s\int_{\mathbb{R}}\int_{\mathbb{R}}\int_0^1\int_0^1\frac{f(x)f(y)}{2\pi\sqrt{|abs^2-[\min(as,bs)]^2|}}e^{-\frac{(bx^2s+ay^2s-2xy\min(as,bs))}{2(abs^2-[\min(as,bs)]^2)}}dadbdxdy$$
$$=\int_{\mathbb{R}}\int_{\mathbb{R}}\int_0^1\int_0^1\frac{f(x)f(y)}{2\pi\sqrt{|ab-[\min(a,b)]^2|}}e^{-\frac{1}{s}\frac{(bx^2+ay^2-2xy\min(a,b))}{2(ab-[\min(a,b)]^2)}}dadbdxdy.$$

Letting $s\to\infty$, and using $\int_{\mathbb{R}} f(x)dx=1$, we get

$$\lim_{s\to\infty}\mathbf{E}(X(s))^2=\int_{\mathbb{R}}\int_{\mathbb{R}}\int_0^1\int_0^1\frac{f(x)f(y)}{2\pi\sqrt{|ab-[\min(a,b)]^2|}}dadbdxdy$$
$$=\int_0^1\int_0^1\frac{1}{2\pi\sqrt{|ab-[\min(a,b)]^2|}}dadb=2\int_{b=0}^{b=1}\int_{a=0}^{a=b}\frac{1}{2\pi\sqrt{ab-a^2}}dadb=1.$$

More specifically, using the change of variables $c = (b/2)\sin\theta$, so that $dc = (b/2)\cos\theta$,

$$\int_{a=0}^{a=b} \frac{1}{\sqrt{ab-a^2}}da = \int_{a=0}^{a=b} \frac{1}{\sqrt{-(a-b/2)^2+b^2/4}}da = \int_{c=-b/2}^{c=b/2} \frac{1}{\sqrt{-c^2+b^2/4}}dc$$
$$= \int_{\theta=-\pi/2}^{\theta=\pi/2} \frac{2}{b\sqrt{-\sin^2\theta+1}}(b/2)\cos\theta d\theta = \int_{\theta=-\pi/2}^{\theta=\pi/2} d\theta = \pi.$$

□

Exercise 7.14. Let $t > 0$ and let $\{B(s)\}_{s\geq 0}$ be a standard Brownian motion. Compute the mean and variance of

$$\int_0^t B(s)dB(s).$$

(Hint: start with the Riemann sum, then take a limit.)

Solution. For any $n \geq 1$, consider the Riemann sum on $[0,t]$:

$$X_n := \sum_{i=0}^{n-1} B\left(\frac{ti}{n}\right)\left(B\left(\frac{t(i+1)}{n}\right) - B\left(\frac{ti}{n}\right)\right).$$

From the independent increment property of Brownian motion, each term in the sum is the product of two independent random variables, so

$$\mathbf{E}X_n = \sum_{i=0}^{n-1} \mathbf{E}B\left(\frac{ti}{n}\right) \cdot \mathbf{E}\left(B\left(\frac{t(i+1)}{n}\right) - B\left(\frac{ti}{n}\right)\right) = 0.$$

Each expected value term on the right is zero, since they are each the expected value of a mean zero Gaussian random variable. We now compute the variance.

$$\mathbf{E}X_n^2 = \mathbf{E}\left(\sum_{i=0}^{n-1} B\left(\frac{ti}{n}\right)\left(B\left(\frac{t(i+1)}{n}\right) - B\left(\frac{ti}{n}\right)\right)\right)^2$$
$$= \mathbf{E}\sum_{i,j=0}^{n-1} B\left(\frac{ti}{n}\right)\left(B\left(\frac{t(i+1)}{n}\right) - B\left(\frac{ti}{n}\right)\right)B\left(\frac{tj}{n}\right)\left(B\left(\frac{t(j+1)}{n}\right) - B\left(\frac{tj}{n}\right)\right)$$
$$= \mathbf{E}\sum_{i=0}^{n-1} B\left(\frac{ti}{n}\right)^2\left(B\left(\frac{t(i+1)}{n}\right) - B\left(\frac{ti}{n}\right)\right)^2$$
$$+ 2\mathbf{E}\sum_{i<j} B\left(\frac{ti}{n}\right)\left(B\left(\frac{t(i+1)}{n}\right) - B\left(\frac{ti}{n}\right)\right)B\left(\frac{tj}{n}\right)\left(B\left(\frac{t(j+1)}{n}\right) - B\left(\frac{tj}{n}\right)\right).$$

From the independent increment property, the last term is zero, and

$$\mathbf{E}X_n^2 = \mathbf{E}\sum_{i=0}^{n-1} B\left(\frac{ti}{n}\right)^2\left(B\left(\frac{t(i+1)}{n}\right) - B\left(\frac{ti}{n}\right)\right)^2$$
$$= \sum_{i=0}^{n-1}(ti/n)(t/n) = \frac{t^2}{n^2}\sum_{i=0}^{n-1} i = \frac{t^2}{n^2}\frac{n(n-1)}{2}.$$

So, $\lim_{n\to\infty}\mathbf{E}X_n^2 = \frac{1}{2}t^2$.

Alternatively, by Ito's Formula, $\int_0^t B(s)dB(s) = \frac{1}{2}(B(t))^2 - \frac{1}{2}t$. So,

$$\mathbf{E}\int_0^t B(s)dB(s) = \mathbf{E}\frac{1}{2}(B(t))^2 - \frac{1}{2}t = 0.$$

$$\mathbf{E}\left(\int_0^t B(s)dB(s)\right)^2 = \mathbf{E}\left(\frac{1}{2}(B(t))^2 - \frac{1}{2}t\right)^2 = \frac{1}{4}(\mathbf{E}(B(t))^4 - 2t\mathbf{E}(B(t))^2 + t^2)$$
$$= \frac{1}{4}(3-2+1)t^2 = \frac{1}{2}t^2.$$

□

Exercise 7.15. Let $f\colon\mathbb{R}\to\mathbb{R}$ be a function. Let $t>0$ and let $\{B(s)\}_{s\geq 0}$ be a standard Brownian motion. Find the distribution of

$$\int_0^t f(s)dB(s).$$

That is, find the CDF of $\int_0^t f(s)dB(s)$. (Hint: use Exercise **??**.)

Solution. For any $n\geq 1$, consider the Riemann sum on $[0,t]$:

$$X_n := \sum_{i=0}^{n-1} f\left(\frac{ti}{n}\right)\left(B\left(\frac{t(i+1)}{n}\right) - B\left(\frac{ti}{n}\right)\right).$$

From Exercise **??**, X_n is a sum of independent mean zero Gaussian random variables. So, X_n is a mean zero Gaussian random variable with variance

$$\sum_{i=0}^{n-1}[f(ti/n)]^2(t/n).$$

Letting $n\to\infty$, and realizing this sum as a Riemann sum, we then see that $\int_0^t f(s)dB(s)$ is a mean zero Gaussian random variable with variance

$$\int_0^t (f(s))^2 ds.$$

□

Exercise 7.16. Using Itô's formula, write an expression for $\int_0^1 (B(s))^2 dB(s)$.

Solution. For any $x\in\mathbb{R}$, let $f(x)=x^3$. Then, Itô's formula says, for any $b>0$, we have

$$(B(b))^3 = (B(b))^3 - (B(0))^3 = \int_0^b 3(B(s))^2 dB(s) + 3\int_0^b B(s)ds.$$

That is,

$$\int_0^1 (B(s))^2 dB(s) = \frac{1}{3}(B(1))^3 - \int_0^1 B(s)ds.$$

□

Exercise 7.17. Let $b > 0$. We know from calculus that $\int_0^b e^s ds = e^b - 1$.

Use $f(x) = e^x$, $x \in \mathbb{R}$, in Itô's formula to find a similar expression for $\int_0^b e^{B(s)} dB(s)$. (Note that $e^{B(s)}$ is a Geometric Brownian motion, so now we know how to take the stochastic integral of Geometric Brownian motion.)

Solution. Itô's formula says, for any $b > 0$,

$$e^{B(b)} - 1 = e^{B(b)} - e^{B(0)} = \int_0^b e^{B(s)} dB(s) + \frac{1}{2}\int_0^b e^{B(s)} ds.$$

That is,

$$\int_0^b e^{B(s)} dB(s) = e^{B(b)} - 1 - \frac{1}{2}\int_0^b e^{B(s)} ds.$$

□

Exercise 7.18 (MFE Sample Question, from an old exam)**.** Let $\{Z(t)\}_{t\geq 0}$ be a standard Brownian motion. You are given:

(i) $U(t) := 2Z(t) - 2$, for all $t \geq 0$.
(ii) $V(t) := (Z(t))^2 - t$, for all $t \geq 0$.
(iii) $W(t) := t^2 Z(t) - 2\int_0^t sZ(s)ds$, for all $t \geq 0$.

Which of the processes defined above has/have zero drift? (A stochastic process $\{U(t)\}_{t\geq 0}$ has zero drift if $dU(t) = f(Z(t), t)dZ(t)$ for some function $f\colon \mathbb{R}^2 \to \mathbb{R}$.)

Solution. Let $f(x) = 2x - 2$ for any $x \in \mathbb{R}$. Then Itô's formula says

$$dU(t) = df(Z(t)) = 2dZ(t).$$

So, item (i) has zero drift. Now, let $f(x, y) = y^2 - x$ for any $x, y \in \mathbb{R}$. Then Itô's formula Version 2 says

$$dV(t) = df(Z(t)) = 2Z(t)dZ(t) - dt + dt = 2Z(t)dZ(t).$$

So, item (ii) has zero drift. Finally, let $f(x, y) = x^2 y$ for any $x, y \in \mathbb{R}$. Then Itô's formula Version 2 says

$$df(Z(t)) = t^2 dZ(t) + 2tZ(t)dt.$$

That is, $d[t^2 Z(t) - 2\int_0^t sZ(s)ds] = t^2 dZ(t)$. So, item (iii) also has zero drift. (Note that we have used the usual Fundamental Theorem of Calculus here, to deduce that $d\int_0^t sZ(s)ds = tZ(t)dt$. However, the Fundamental Theorem generally does NOT hold for the stochastic integral. We can rewrite Itô's Lemma as $\int_0^t g'(Z(s))dZ(s) = g(Z(t)) - g(Z(0)) - \frac{1}{2}\int_0^t g''(Z(s))ds$. So, we cannot "differentiate both sides in t" and have both sides be equal, as in the usual Fundamental Theorem of Calculus.) □

Exercise 7.19. Let $f\colon \mathbb{R} \times [0, \infty) \to \mathbb{R}$. We write $f = f(x, t)$, where $(x, t) \in \mathbb{R} \times [0, \infty)$. Let $g\colon \mathbb{R} \to \mathbb{R}$ be a continuous and bounded function with $\int_{\mathbb{R}} |g(x)|\, dx < \infty$. We say that f satisfies the one-dimensional **heat equation** if

$$\frac{\partial f}{\partial t}(x, t) = \frac{\partial^2 f}{\partial x^2}(x, t), \quad \forall\, (x, t) \in \mathbb{R} \times [0, \infty),$$

$$f(x,0)=g(x),\qquad \forall\, x\in\mathbb{R}.$$

Show that f defined by

$$f(x,t):=\frac{1}{\sqrt{4\pi t}}\int_{\mathbb{R}}e^{-\frac{(x-y)^2}{4t}}g(y)dy=\mathbf{E}(g(2B(t)+x)),\qquad \forall\,(x,t)\in\mathbb{R}\times[0,\infty),$$

satisfies the heat equation. (Just check the first condition. You do not have to show that $\lim_{t\to 0^+}f(x,t)=g(x)$ for all $x\in\mathbb{R}$.)

Using a computer, plot the function $f(x,t)$ as a function of x for several different values of $t>0$, using $g=1_{[0,1]}$. Lastly, verify that $\int_{\mathbb{R}}\frac{1}{\sqrt{4\pi t}}e^{-x^2/(4t)}dx=1$ for any $t>0$.

Solution.

$$\frac{\partial}{\partial t}\frac{1}{\sqrt{4\pi t}}\int_{\mathbb{R}}e^{-\frac{(x-y)^2}{4t}}g(y)dy=-\frac{1}{2t\sqrt{4\pi t}}\int_{\mathbb{R}}e^{-\frac{(x-y)^2}{4t}}g(y)dy+\frac{1}{\sqrt{4\pi t}}\int_{\mathbb{R}}\frac{(x-y)^2}{4t^2}e^{-\frac{(x-y)^2}{4t}}g(y)dy.$$

$$\frac{\partial}{\partial x}\frac{1}{\sqrt{4\pi t}}\int_{\mathbb{R}}e^{-\frac{(x-y)^2}{4t}}g(y)dy=\frac{1}{\sqrt{4\pi t}}\int_{\mathbb{R}}\left(-\frac{(x-y)}{2t}\right)e^{-\frac{(x-y)^2}{4t}}g(y)dy.$$

$$\begin{aligned}&\frac{\partial^2}{\partial x^2}\frac{1}{\sqrt{4\pi t}}\int_{\mathbb{R}}e^{-\frac{(x-y)^2}{4t}}g(y)dy\\&=\frac{1}{\sqrt{4\pi t}}\int_{\mathbb{R}}\left(-\frac{1}{2t}\right)e^{-\frac{(x-y)^2}{4t}}g(y)dy+\frac{1}{\sqrt{4\pi t}}\int_{\mathbb{R}}\left(\frac{(x-y)^2}{4t^2}\right)e^{-\frac{(x-y)^2}{4t}}g(y)dy.\end{aligned}$$

So,

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

```
numxpts=100;
for t= [.001,.01,.1,1,10]
    for j=1:numxpts
        x=-5+j*10/numxpts;
    y(j)=quad( @(y) exp(-(x-y).^2 ./(4*t)),0,1);
    end
    figure;
    plot(10*(1:numxpts)/numxpts -5, y);
    title(strcat('t=',num2str(t)));
end
```

□

Exercise 7.20. Let $f\colon\mathbb{R}\times[0,\infty)\to\mathbb{R}$. We write $f=f(x,t)$, where $(x,t)\in\mathbb{R}\times[0,\infty)$. Let $g\colon\mathbb{R}\to\mathbb{R}$ be a continuous and bounded function. We say that f satisfies the one-dimensional **heat equation with forcing term** $h\colon\mathbb{R}\times[0,\infty)\to\mathbb{R}$ if

$$\frac{\partial f}{\partial t}(x,t)=\frac{\partial^2 f}{\partial x^2}(x,t)+h(x,t),\quad \forall\,(x,t)\in\mathbb{R}\times[0,\infty),$$

$$f(x,0)=g(x),\qquad \forall\, x\in\mathbb{R}.$$

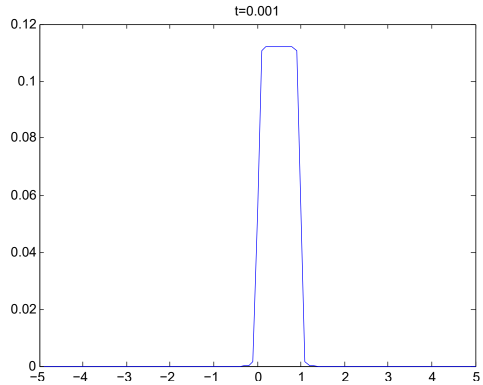


FIGURE 1. $f(x, .001)$ for $x \in [-5, 5]$

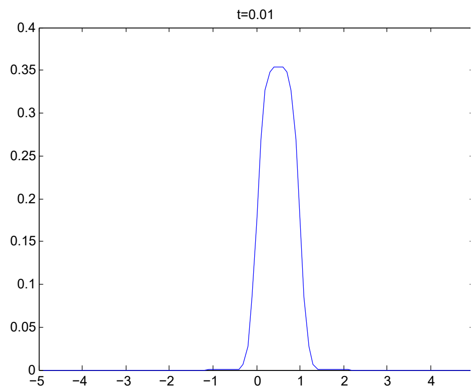


FIGURE 2. $f(x, .01)$ for $x \in [-5, 5]$

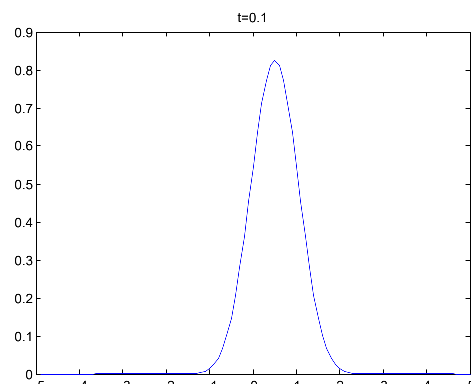


FIGURE 3. $f(x, .1)$ for $x \in [-5, 5]$

For any $(x, t) \in \mathbb{R} \times [0, \infty)$, define $f(x, t)$ so that

$$f(x,t) := \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4t}} g(y)dy + \int_0^t \frac{1}{\sqrt{4\pi(t-s)}} \int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4(t-s)}} h(y,s)dyds.$$

Show that f satisfies the heat equation with forcing term h. (Just check the first condition.)

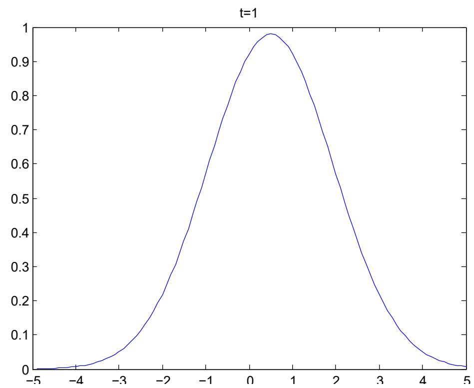


FIGURE 4. $f(x,1)$ for $x \in [-5,5]$

Solution. From the previous exercise, we know that the first term in the definition of $f(x,t)$ satisfies the heat equation (with no forcing term). So, it suffices to show that the second term in the definition of $f(x,t)$ satisfies the heat equation with forcing term.

First, as a Lemma, consider an integral of the form $\int_0^t g(s,t)ds$. To take the derivative with respect to t, we note that

$$\begin{aligned}&\int_0^{t+h} g(s,t+h)ds - \int_0^t g(s,t)ds\\&= \int_0^{t+h} g(s,t+h)ds - \int_0^{t+h} g(s,t)ds + \int_0^{t+h} g(s,t)ds - \int_0^t g(s,t)ds\\&= \int_0^{t+h} [g(s,t+h)-g(s,t)]ds + \int_t^{t+h} g(s,t)ds.\end{aligned}$$

So, dividing this equality by h and letting $h \to 0$, we get

$$\frac{d}{dt}\int_0^t g(s,t)ds = \int_0^t \frac{d}{dt} g(s,t)ds + g(t,t).$$

So,

$$\begin{aligned}&\frac{\partial}{\partial t}\int_0^t \frac{1}{\sqrt{4\pi(t-s)}}\int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4(t-s)}} h(y,s)dyds\\&= \lim_{s\to t}\frac{1}{\sqrt{4\pi(t-s)}}\int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4(t-s)}} h(y,s)dy + \int_0^t\int_{\mathbb{R}} \frac{\partial}{\partial t}\frac{1}{\sqrt{4\pi(t-s)}} e^{-\frac{(x-y)^2}{4(t-s)}} h(y,s)dyds.\end{aligned}$$

From the previous exercise, the first term is $h(x,t)$. Also from the previous exercise, the second term is

$$\int_0^t\int_{\mathbb{R}} \frac{\partial^2}{\partial x^2}\frac{1}{\sqrt{4\pi(t-s)}} e^{-\frac{(x-y)^2}{4(t-s)}} h(y,s)dyds.$$

□

Exercise 7.21. Let $t_0 > 0$. Let $V\colon \mathbb{R}\times[0,t_0]\to\mathbb{R}$. We write $V = V(s,t)$, $s\in\mathbb{R}$, $t\in[0,t_0]$. Let $F\colon \mathbb{R}\to\mathbb{R}$. Let $r\in\mathbb{R}$, let $\sigma>0$. We say that V satisfies the **Black-Scholes** equation if $V(s,t_0)=F(s)$ for all $s\in\mathbb{R}$, and if

$$\frac{\partial V}{\partial t}+rs\frac{\partial V}{\partial s}+\frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{\partial s^2}-rV=0.$$

Show that a solution of this equation is

$$V(s,t):=\frac{e^{-r(t_0-t)}}{\sqrt{2\pi\sigma^2(t_0-t)}}\int_0^\infty \frac{1}{z}e^{-\frac{(\log(s/z)+(r-\sigma^2/2)(t_0-t))^2}{2\sigma^2(t_0-t)}}F(z)dz.$$

(This formula should be nearly identical to the Black-Scholes Option Pricing formula from a remark in the notes, where we take $F(z):=\max(S_0z-k,0)$.) Instead of differentiating V directly, use the following strategy.

First, show that the Black-Scholes equation reduces to the one-dimensional heat equation

$$\frac{\partial U}{\partial \tau}=\frac{\partial^2 U}{\partial x^2},$$

where $V(s,t)=e^{ax+b\tau}U(x,\tau)$, $x=\log s$, $\tau=(\sigma^2/2)(t_0-t)$, $a=(1/2)-r/\sigma^2$, and $b=-(1/2+r/\sigma^2)^2$, and U satisfies the initial condition $U(x,0)=e^{-ax}F(e^x)$ for all $x\in\mathbb{R}$. (Start by differentiating V with respect to s and t, etc.) That is, the Black-Scholes equation is the heat equation, run backwards in time.

Finally, use the formula for U using Exercise 7.20.

Solution. Note that $V(s,t)=e^{a\log s+b(\sigma^2/2)(t_0-t)}U(\log s,(\sigma^2/2)(t_0-t))$. So,

$$\frac{\partial}{\partial t}V(s,t)=-b(\sigma^2/2)V(s,t)+(-\sigma^2/2)e^{ax+b\tau}\frac{\partial U}{\partial \tau}(x,\tau).$$

$$\frac{\partial}{\partial s}V(s,t)=\frac{a}{s}V(s,t)+\frac{1}{s}e^{ax+b\tau}\frac{\partial U}{\partial x}(x,\tau).$$

$$\begin{aligned}\frac{\partial^2}{\partial s^2}V(s,t)=-as^{-2}V(s,t)+\frac{a^2}{s^2}V(s,t)+\frac{a}{s^2}e^{ax+b\tau}\frac{\partial U}{\partial x}(x,\tau)\\ -\frac{1}{s^2}e^{ax+b\tau}\frac{\partial U}{\partial x}(x,\tau)+\frac{a}{s^2}e^{ax+b\tau}\frac{\partial U}{\partial x}(x,\tau)\\ +\frac{1}{s^2}e^{ax+b\tau}\frac{\partial^2 U}{\partial x^2}(x,\tau).\end{aligned}$$

$$\begin{aligned}0 &= \frac{\partial V}{\partial t} + rs\frac{\partial V}{\partial s} + \frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{\partial s^2} - rV\\ &= -b(\sigma^2/2)V(s,t) + (-\sigma^2/2)e^{ax+b\tau}\frac{\partial U}{\partial \tau}(x,\tau)\\ &\quad + arV(s,t) + re^{ax+b\tau}\frac{\partial U}{\partial x}(x,\tau)\\ &\quad - a(\sigma^2/2)V(s,t) + (a^2\sigma^2/2)V(s,t) + (a\sigma^2/2)e^{ax+b\tau}\frac{\partial U}{\partial x}(x,\tau)\\ &\quad - (\sigma^2/2)e^{ax+b\tau}\frac{\partial U}{\partial x}(x,\tau) + (a\sigma^2/2)e^{ax+b\tau}\frac{\partial U}{\partial x}(x,\tau)\\ &\quad + (\sigma^2/2)e^{ax+b\tau}\frac{\partial^2 U}{\partial x^2}(x,\tau) - rV(s,t).\end{aligned}$$

$$\begin{aligned}ar - r + (a^2 - b - a)(\sigma^2/2) &= ar - r + ((1/2) + 2r^2/\sigma^4 - a)(\sigma^2/2)\\ &= ar - r + (2r^2/\sigma^4 + r/\sigma^2)(\sigma^2/2)\\ &= ar - r + (r^2/\sigma^2 + r/2) = r[a + r/\sigma^2 - 1/2] = 0.\end{aligned}$$

So, the sum of the $V(s,t)$ terms is zero.

We now examine the $\partial U/\partial x$ terms.

$$r + a\sigma^2 - \sigma^2/2 = r + \sigma^2/2 - r - \sigma^2/2 = 0.$$

So, the sum of all of the $\partial U/\partial x$ terms is zero.

In summary,

$$0 = (-\sigma^2/2)e^{ax+b\tau}\frac{\partial U}{\partial \tau}(x,\tau) + (\sigma^2/2)e^{ax+b\tau}\frac{\partial^2 U}{\partial x^2}(x,\tau).$$

That is, $-\frac{\partial U}{\partial \tau}(x,\tau) + \frac{\partial^2 U}{\partial x^2}(x,\tau) = 0$. And $F(s) = V(s,t_0) = e^{a\log s}U(\log s, 0) = e^{ax}U(x,0)$. That is, $U(x,0) = e^{-ax}F(e^x)$.

So, using Exercise 7.20, we have

$$U(x,\tau) := \frac{1}{\sqrt{4\pi\tau}}\int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4\tau}} e^{-ay}F(e^y)dy.$$

Going back to the relation between U and V, we have

$$\begin{aligned}
V(s,t) &= e^{a\log s+b(\sigma^2/2)(t_0-t)}U(\log s,(\sigma^2/2)(t_0-t)) \\
&= e^{a\log s+b(\sigma^2/2)(t_0-t)}\frac{1}{\sqrt{4\pi(\sigma^2/2)(t_0-t)}}\int_{\mathbb{R}} e^{-\frac{(\log s-y)^2}{4(\sigma^2/2)(t_0-t)}}e^{-ay}F(e^y)dy \\
&= e^{a\log s+b(\sigma^2/2)(t_0-t)}\frac{1}{\sqrt{2\pi\sigma^2(t_0-t)}}\int_0^\infty \frac{1}{z}e^{-\frac{(\log s-\log z)^2}{4(\sigma^2/2)(t_0-t)}}e^{-a\log z}F(z)dz \\
&= e^{b(\sigma^2/2)(t_0-t)}\frac{1}{\sqrt{2\pi\sigma^2(t_0-t)}}\int_0^\infty \frac{1}{z}e^{-\frac{(\log(s/z))^2}{2\sigma^2(t_0-t)}}e^{a\log(s/z)}F(z)dz \\
&= \frac{e^{-r(t_0-t)}}{\sqrt{2\pi\sigma^2(t_0-t)}}\int_0^\infty \frac{1}{z}e^{-\frac{\left(\log(s/z)+(r-\sigma^2/2)(t_0-t)\right)^2}{2\sigma^2(t_0-t)}}F(z)dz
\end{aligned}$$

Above we used $r+b\sigma^2/2=(r-\sigma^2/2)^2/(2\sigma^2)$

□

Exercise 7.22. Let $a,b,\sigma>0$. Let $f\colon\mathbb{R}\to\mathbb{R}$ satisfying the Vasicek stochastic differential equation for any $t\in\mathbb{R}$.

$$df(t)=a(b-f(t))dt+\sigma dB(t).$$

Show that, for any $t>0$,

$$\mathbf{E}f(t)=b+e^{-at}(f(0)-b),\qquad \mathrm{var}(f(t))=\frac{\sigma^2}{2a}(1-e^{-2at}).$$

More generally, for any $s,t>0$, show that

$$\mathrm{cov}(f(t),f(u))=\mathbf{E}\big((f(t)-\mathbf{E}f(t))(f(u)-\mathbf{E}f(u))\big)=\frac{\sigma^2}{2a}(e^{-a|t-u|}-e^{-a(t+u)}).$$

Conclude that $\lim_{t\to\infty}\mathbf{E}f(t)=b$ and $\lim_{t\to\infty}\mathrm{var}(f(t))=\frac{\sigma^2}{2a}$.

Solution. From the notes, we have

$$f(t)=b+e^{-at}(f(0)-b)+\sigma\int_0^t e^{a(s-t)}dB(s)$$

By Exercise 7.15, the last term is a mean zero Gaussian random variable with variance

$$\sigma^2\int_0^t e^{2a(s-t)}ds=\sigma^2\frac{1}{2a}[e^{2a(s-t)}]_{s=0}^{s=t}=\sigma^2\frac{1}{2a}[1-e^{-2at}].$$

Therefore,

$$\mathbf{E}f(t)=b+e^{-at}(f(0)-b),\qquad \mathrm{var}(f(t))=\frac{\sigma^2}{2a}(1-e^{-2at}).$$

More generally, using the independent increment property of Brownian motion,

$$\begin{aligned}\mathrm{cov}(f(t),f(u)) &= \mathbf{E}f(t)f(u) = \sigma^2\mathbf{E}\int_0^t e^{a(s-t)}dB(s)\int_0^u e^{a(s-u)}dB(s)\\ &= \sigma^2 e^{-a(t+u)}\mathbf{E}\left(\int_0^{\min(t,u)} e^{as}dB(s)\right)^2 = \sigma^2 e^{-a(t+u)}\int_0^{\min(t,u)} e^{2as}ds\\ &= \frac{\sigma^2}{2a}e^{-a(t+u)}(-1+e^{2a\min(t,u)}) = \frac{\sigma^2}{2a}(e^{-a|t-u|}-e^{-a(t+u)}).\end{aligned}$$

□

Exercise 7.23. Using a Monte Carlo simulation, plot several sample paths of the Vasicek stochastic differential equation, with $a=b=\sigma=f(0)=1$.

```
figure;
hold on;
for j=1:5
k=500^2;
length=4;
numpts=length*k;
t=linspace(0,length,numpts);
jumps=(1/sqrt(k))*randn(numpts,1);
y=zeros(1,numpts);
 y(1)=1;
for i=2:numpts
    y(i)=y(i-1)+(1-y(i-1))*(1/k)+sqrt(y(i-1))*jumps(i);
end
plot(t,y,'Color',rand(1,3));
end
```

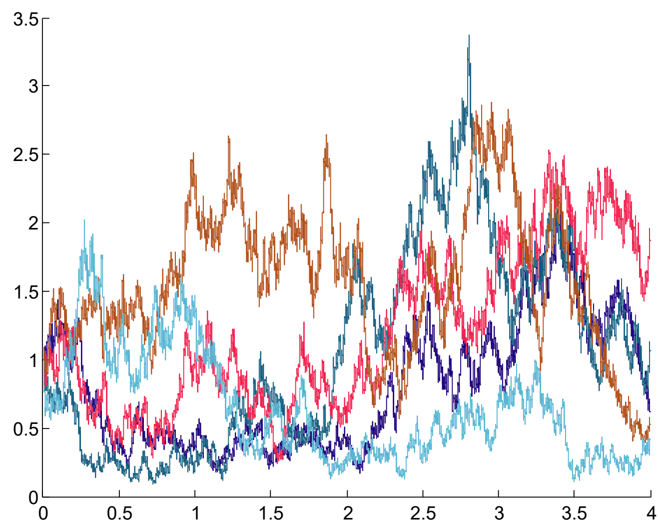


FIGURE 5. Sample Paths of the CIR model with $a=b=\sigma=f(0)=1$. The horizontal axis is the t-axis.

Exercise 7.24 (**Cox-Ingersoll-Ross (CIR) model**). Let $a, b, \sigma > 0$. Let $\{B(t)\}_{t\geq 0}$ be a standard Brownian motion. The **Cox-Ingersoll-Ross model** models an interest rate as a (random) function $f\colon \mathbb{R} \to \mathbb{R}$ satisfying the following stochastic differential equation for any $t > 0$:

$$df(t) = a(b - f(t))dt + \sqrt{f(t)}\sigma dB(t).$$

(Since f is a random function, f is also a function of the sample space, but we omit this dependence from our notation here and below.)

A priori, this stochastic differential equation is not rigorously defined, since $\sqrt{f(t)}$ will not be a real number when $f(t) < 0$. In this exercise, we ignore this issue. (In actuality, if $f(0) > 0$, then $f(t) < 0$ occurs with probability 0.)

Unlike the Vasicek model, we might not be able to get a closed form solution of this equation. Nevertheless, we can still run a Monte Carlo simulation of this stochastic differential equation as follows. Let $f(0) = 1$. Let $i, n > 0$ be integers. Suppose we have inductively determined $f(i/n)$ using a Monte Carlo simulation, and we would like the determine $f((i+1)/n)$. The stochastic differential equation then suggests that

$$f((i+1)/n) \approx f(i/n) + a(b - f(i/n))(i/n) + \sqrt{f(i/n)}\sigma(B((i+1)/n) - B(i/n)).$$

This approximation is known as a **finite difference scheme**.

Using this approximation, plot several sample paths of the CIR model with $a = b = f(0) = \sigma = 1$.

What would be the corresponding finite difference scheme for the Vasicek model?

```
figure;
hold on;
for j=1:5
k=500^2;
length=4;
numpts=length*k;
t=linspace(0,length,numpts);
jumps=(1/sqrt(k))*randn(numpts,1);
y=zeros(1,numpts);
 y(1)=1;
for i=2:numpts
    y(i)=y(i-1)+(1-y(i-1))*(1/k)+jumps(i);
end
plot(t,y,'Color',rand(1,3));
end
```

Exercise 7.25. Let $\{Z(x,t)\}_{x\in\mathbb{R},t\geq 0}$ be a set of independent, standard Gaussian random variables. Suppose $f\colon \mathbb{R}\times[0,\infty) \to \mathbb{R}$ satisfies the stochastic heat equation.

$$\frac{\partial f}{\partial t}(x,t) = \frac{\partial^2 f}{\partial x^2}(x,t) + h(x,t), \quad \forall\,(x,t) \in \mathbb{R}\times[0,\infty),$$
$$f(x,0) = 0, \qquad \forall\, x \in \mathbb{R}.$$

We can explicitly solve this equation by its analogy with Exercise 7.20. That is,

$$f(x,t) := \int_0^t \frac{1}{\sqrt{4\pi(t-s)}} \int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4(t-s)}} Z(y,s)dyds, \qquad \forall\,(x,t)\in\mathbb{R}\times[0,\infty),$$

satisfies the stochastic heat equation. Show that f has the following covariance for any $s,t>0$:

$$\mathbf{E}[f(0,s)f(0,t)] = \frac{1}{2\sqrt{\pi}}(|s+t|^{1/2} - |s-t|^{1/2}).$$

Solution.

$$\begin{aligned}
&\mathbf{E}[f(0,s)f(0,t)]\\
&= \mathbf{E}\int_0^s \frac{1}{\sqrt{4\pi(s-v)}}\int_{\mathbb{R}} e^{-\frac{y^2}{4(s-v)}} Z(y,s)dydv \int_0^t \frac{1}{\sqrt{4\pi(t-u)}}\int_{\mathbb{R}} e^{-\frac{z^2}{4(t-u)}} Z(z,u)dzdu\\
&= \int_0^s \frac{1}{\sqrt{4\pi(s-v)}}\int_{\mathbb{R}} e^{-\frac{y^2}{4(s-v)}} dydv \int_0^t \frac{1}{\sqrt{4\pi(t-u)}}\int_{\mathbb{R}} e^{-\frac{z^2}{4(t-u)}} 1_{\{v=u\}}1_{\{y=z\}}dzdu\\
&= \int_0^{\min(s,t)} \frac{1}{4\pi\sqrt{s-v}\sqrt{t-v}}\int_{\mathbb{R}} e^{-\frac{y^2}{4(s-v)}-\frac{y^2}{4(t-v)}} dydv\\
&= \int_0^{\min(s,t)} \frac{1}{4\pi\sqrt{s-v}\sqrt{t-v}}\int_{\mathbb{R}} e^{-y^2\frac{(t+s-2v)}{4(s-v)(t-v)}} dydv\\
&= \int_0^{\min(s,t)} \frac{1}{2\sqrt{\pi}\sqrt{s-v}\sqrt{t-v}}\frac{\sqrt{(s-v)(t-v)}}{\sqrt{t+s-2v}}dv = \int_0^{\min(s,t)} \frac{1}{2\sqrt{\pi}}\frac{1}{\sqrt{t+s-2v}}dv\\
&= \frac{1}{2\sqrt{\pi}}[-(t+s-2v)^{1/2}]_{v=0}^{v=\min(s,t)} = \frac{1}{2\sqrt{\pi}}(\sqrt{t+s} - \sqrt{t+s-2\min(s,t)})\\
&= \frac{1}{2\sqrt{\pi}}(|t+s|^{1/2} - |t-s|^{1/2}).
\end{aligned}$$

□

USC Mathematics, Los Angeles, CA

E-mail address: stevenmheilman@gmail.com