# Compilation of Probability Problems and Solutions

Hashiam Kadhim

## Contents

# 1 Introduction

This is a compilation of probability problems that I've come across in my early days of learning the subject. The references for this compilation are Durrett's "Probabiliy Theory and Examples" and Morters and Peres: Browniaon Motion. Any numbered theorems/exercises/pages in this compilation refer to Durrett's book unless we are in the Brownian motion section, in which case I am referring to Morters' and Peres' book. Some problems are also taken from these books.

# 2 Warm up Problems - Measure Theory

**Problem 1**
Consider the product of three fair-coin toss probability spaces. How many outcomes and how many events are there on this space? Show that the number of heads in three tosses is a random variable defined on this space.

**Solution:**
The set of outcomes for a fair-coin toss is $\Omega = \{H, T\}$ and the events are
$2^\Omega = \{\{\}, \{H\}, \{T\}, \{H, T\}\}$. The product of three fair-coin toss probability spaces will have the following set of outcomes: $\Omega = \{HHH, THH, HTH, HHT, TTH, THT, HTT, TTT\}$, which contain 8 elements. The set of events in this example contained $2^{|\Omega|} = 2^8 = 256$ events.

The number of heads in three tosses is a function $X$ from $\Omega \to \{0, 1, 2, 3\} \subset \mathbb{R}$. Since the set of events $\mathcal{F}$ is the power set of $\Omega$, $X^{-1}(i) \in \mathcal{F}$ for $i \in \{0, 1, 2, 3\}$ so X is measurable. Therefore the number of heads in three tosses is a random variable.

**Problem 2**
Show that the Borel $\sigma$-field on $\mathbb{R}$ is the smallest $\sigma$-field that makes all continuous functions measurable.

**Solution:**
Let $\mathcal{A}$ be the collection of open sets of $\mathbb{R}$ in the standard topology. Then for any continuous function $f : \mathbb{R} \to \mathbb{R}$, and for any $A \in \mathcal{A}$, we have that $f^{-1}(A) \in \mathcal{A} \subset \sigma(\mathcal{A})$. The Borel $\sigma$-field on $\mathbb{R}$ is the smallest $\sigma$-field on $\mathbb{R}$ containing $\mathcal{A}$. Finally, by theorem 1.3.1, f is measurable. We have shown that the Borel $\sigma$-field makes all continuous functions measurable, and that it is the smallest one to do so.

**Problem 3**
Just because $\mathcal{A}$ generates a $\sigma$-field $\mathcal{F}$, the values of P on $\mathcal{A}$ do not in general determine its values on $\mathcal{F}$. To show this, give an example of as measurable space $(\Omega, \mathcal{F})$, a collection $\mathcal{A}$ and probability measures P , Q so that

(i) P (A) = Q(A) for all A $\in \mathcal{A}$,
(ii) F = $\sigma(\mathcal{A})$,

(iii) P = Q.

Note that this can be done on a space with four outcomes.

**Solution:**

Let $\Omega = \{a, b, c, d\}$ and $\mathcal{A} = \{\{a, b\}, \{b, c\}\}$.
Then

$$\mathcal{F} = \sigma(\mathcal{A}) = 2^{\Omega}$$

Let

$$P = \frac{1}{2}(\delta_a + \delta_c), \quad Q = \frac{1}{2}(\delta_b + \delta_d)$$

so

$$P(\{a, b\}) = Q(\{a, b\}) = P(\{b, c\}) = Q(\{b, c\}) = \frac{1}{2}$$

So far conditions (i) and (ii) have been satisfied. But $P \neq Q$ since $P(\{a\}) = \frac{1}{2} \neq Q(\{a\}) = 0$.

**Problem 4**

Given an arbitrary collection of subsets $\mathcal{A}$ of $\Omega$, prove that there exists a unique smallest $\sigma$-algebra $\sigma(A)$ containing $A$.

**Solution:**

let $\Omega$ be a set and let $\mathcal{A} \subset 2^{\Omega}$. We will prove that there exists a unique smallest $\sigma$-algebra containing $\mathcal{A}$.

*claim:* The intersection of a collection of $\sigma$-fields is a $\sigma$-field.

*Proof.* Let $\Omega$ be a set and let $(\mathcal{F}_i)_{i \in I}$ be a collection of $\sigma$-fields on $\Omega$, where $I \neq \varnothing$ is an arbitrary index set.

***first property***

Since $\Omega \in \mathcal{F}_i \; \forall \, i \in I$ we have that $\Omega \in \cap_{i \in I} \mathcal{F}_i$.

***second property***

Let $(E_j)_{j \in \mathbb{N}} \in \cap_{i \in I} \mathcal{F}_i$. Then:

$$(E_j)_{j \in \mathbb{N}} \in \mathcal{F}_i \; \forall \, i \in I$$

So

$$\bigcup_{j \in \mathbb{N}} E_j \in \mathcal{F}_i \; \forall \, i \in I$$

Hence

$$\bigcup_{j \in \mathbb{N}} E_j \in \cap_{i \in I} \mathcal{F}_i$$

***Third property***

Let $E \in \cap_{i\in I}\mathcal{F}_i$.Then:

$$E \in \mathcal{F}_i \; \forall \; i \in I$$

So

$$E^c \in \mathcal{F}_i \; \forall \; i \in I$$

Hence $E^c \in \cap_{i\in I}\mathcal{F}_i$. □

Using this claim, we can proceed to the proof of this problem.

*Proof.* : We begin by showing existence.

Let $B = \{C : C \text{ is a } \sigma\text{-field on } \Omega \text{ containing } \mathcal{A}\}$.
since $2^\Omega \in B$, $B \neq \varnothing$. By the claim, $\cap_{C\in B}C$ is a $\sigma$-algebra and since $\mathcal{A} \in C \; \forall \; C \in B$, $\mathcal{A} \in \cap_{C\in B}C$ We have shown that $D_1 \subset D_2$. Symmetrically, we can show $D_2 \subset D_1$ and so $D_1 = D_2$.

Next we show uniqueness. Suppose $D_1, D_2$ satisfy the properties of $\sigma(\mathcal{A})$. Let $A \in D_1$. Then $A \in C \; \forall C \in B$, meaning $A \in \cap_{C\in B}C$ and so $A \in D_2$. □

**Problem 5**

Show that in the definition of "probability measure P on a measurable space $(\Omega, \mathcal{F})$", we may replace "countably additive" by "finitely additive, and satisfies:

$$\text{if } A_n \downarrow \emptyset \text{ then } P(A_n) \to 0."$$

**Solution:**

consider a probability space $(\Omega, \mathcal{F}, P)$. First we show that countably additive implies finite additivity and satisfies

$$\text{if } A_n \downarrow \varnothing \text{ then } P(A_n) \to 0$$

*Proof.* Suppose $(A_n)_{n\in\mathbb{N}}$ are disjoint sets in $\mathcal{F}$ such that there exists an integer m such that $A_n = \varnothing \; \forall n > m$. Then clearly

$$P(\bigcup_{i=1}^{m} A_n) = P(\bigcup_{i\in\mathbb{N}} A_n) = \sum_{i\in\mathbb{N}} P(A_n) = \sum_{i=1}^{m} P(A_n)$$

and so P is finitely additive. Next suppose $A_n$ is decreasing to the empty set. That is: $A_n \subset A_{n+1}$ and $A_n \downarrow \varnothing$. Let $A = \cap_{i\in N}A_i = \varnothing$ and $C_n = A_n \backslash A_{n+1}$.

then $A_n$ is the disjoint union $(\cup_{k\geq n}C_n) \cup A = (\cup_{k\geq n}C_n)$. so

$$\lim_{n\to\infty} P(A_n) = \lim_{n\to\infty}(P(A) + P(\bigcup_{k\geq n} C_n)) = \lim_{n\to\infty}(0 + \sum_{k\geq n} P(C_n))$$

by countable additivity and using that $A = \varnothing$. The sum on the right is a tail of a convergent sequence, so it goes to 0 as the limit goes to infinity. therefore: $\lim_{n\to\infty} P(A_n) = 0$.

next we prove the other direction: let $(A_n)_{n\in\mathbb{N}}$ be disjoint sets in $\mathcal{F}$. Let $C_n = \cup_{i\geq n+1}A_i$. Then

$$\bigcup_{n\in\mathbb{N}} A_n = A_1 \bigcup A_2 \bigcup ... \bigcup A_n \bigcup C_n$$

so

$$P(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{i=1}^{n} P(A_i) + P(C_n) \tag{1}$$

by finite additivity. Note that $C_{n+1} \subset C_n$ and since $(A_n)$ are disjoint, $\cap_{i\geq 1} C_n = \varnothing$ so $\lim_{n\to\infty} P(C_n) = 0$ and so countable additivity follows from (1) by taking the limit as n goes to infinity.

□

**Problem 6**

Let $\mathcal{B}$ be the field of finite disjoint unions of intervals $(a, b] \subset \mathbb{R}$. For $B \in \mathcal{B}$ define:

$$P(B) = \begin{cases} 1 & \text{if } (0,\epsilon) \subset B \text{ for some} \epsilon > 0 \\ 0 & \text{otherwise} \end{cases}$$

Show that P is finitely additive but not countably additive on $\mathcal{B}$

**Solution:**

We first show finite additivity. Let $A_1, ..., A_n \in \mathcal{B}$. There are 2 cases.

*Case 1:* $P(\cup_{i=1}^n A_i) = 0$

then $\nexists \epsilon > 0$ s.t. $(0, \epsilon) \subset \bigcup_{i=1}^n A_i$. This means that $\nexists \epsilon > 0$ s.t. $(0, \epsilon) \subset A_i \ \forall i \in \{1, ..., n\}$ since $A_i \subset \bigcup_{i=1}^n A_i \ \forall i \in \{1, ..., n\}$. Therefore $P(A_i) = 0 \ \forall i \in \{1, ..., n\}$ and we have that:

$$P(\cup_{i=1}^n A_i) = 0 = \sum_{i=1}^n P(A_i)$$

.

*Case 2:* $P(\cup_{i=1}^n A_i) = 1$

then $\exists \epsilon > 0$ s.t. $(0, \epsilon) \subset \bigcup_{i=1}^n A_i$.

*claim*: $\exists! j \in \{1, ..., n\}$ such that $\exists \epsilon_0 > 0$ and $(0, \epsilon_0) \subset A_j$

*Proof.* We begin by proving existence. suppose existence is not true. i.e. suppose $\forall \epsilon_0 > 0$ and $\forall j \in \{1, ..., n\}$, $(0, \epsilon_0) \not\subset A_j$. By the assumptions in the problem, each $A_j$ can be written as a finite union of intervals $A_j = (a_j^1, b_j^1] \cup ... \cup (a_j^{n_j}, b_j^{n_j}]$. Since $\forall \epsilon_0 > 0$ and $\forall k \in \{1, ..., n_j\}$, $(0, \epsilon_0) \not\subset (a_j^k, b_j^k]$, it is clear to see that either $a_j^k > 0$ or ($a_j^k \leq 0$ and $b_j^k \leq 0$) $\forall k \in \{1, ..., n_j\}$.

Therefore we can find an $\epsilon_j > 0$ such that $(0, \epsilon_j) \subset (a_j^k, b_j^k]^C \ \forall k \in \{1, ..., n_j\}$ meaning $(0, \epsilon_j) \cap A_j = \varnothing$. Finding such an $\epsilon_j$ for each $A_j$, and letting $\epsilon = \min(\epsilon_1, ..., \epsilon_n)$, we see that

$$(0, \epsilon) \cap (\bigcup_{i=1}^n A_i) = \varnothing$$

which contradicts the assumption that $P(\cup_{i=1}^{n} A_i) = 1$.

So far we have proved existence of an $A_j$ such that $P(A_j) = 1$. All that is left is to show that there is only one such $A_j$. Suppose $\exists A_j \neq A_k$ such that $P(A_j) = P(A_k) = 1$. That means $\exists\ \epsilon_j, \epsilon_k > 0$ such that $(0, \epsilon_j) \subset A_j$ and $(0, \epsilon_k) \subset A_k$. This implies that $(0, \min(\epsilon_j, \epsilon_k)) \subset (A_j \cap A_k)$ contradicting the requirement that $A_j$ and $A_k$ are disjoint.

Therefore there exists only one $A_j$ such that $P(A_j) = 1$. Finally this means that

$$\sum_{i=1}^{n} P(A_i) = 1 = P(\bigcup_{i=1}^{n} A_i)$$

□

Next we prove that P is not countably additive. Consider the set $(-1, 1]$. Clearly $P((-1, 1]) = 1$. Now let $A_0 = (-1, 0]$ and $A_n = (\frac{1}{n+1}, \frac{1}{n}]\ \forall n \in \mathbb{N}$. Now it's clear that $(-1, 1]$ is the disjoint union:

$$\bigcup_{i=0}^{\infty} A_i$$

. But none of the sets $A_i$ contain an open interval of the form $(0, \epsilon)$ for some $\epsilon > 0$. Therefore:

$$\sum_{i=0}^{\infty} P(A_i) = 0 \neq P(\bigcup_{i=0}^{\infty} A_i)$$

.

**Problem 7**

Suppose that $B \in \sigma(\mathcal{A})$ for some collection $\mathcal{A}$ of subsets. Show that there exists a countable subcollection $\mathcal{A}_\omega$ so that $B \in \sigma(\mathcal{A}_\omega)$

**Solution:**

Suppose $\sigma(\mathcal{A})$ is a $\sigma$-field on $\Omega$ for some collection $\mathcal{A}$. let:

$$\mathcal{F} = \{A \in \sigma(\mathcal{A}) : \exists \text{ countable subcollection } \mathcal{A}_\omega \text{ of } \mathcal{A} \text{ such that } A \in \sigma(\mathcal{A}_\omega)\}$$

We will show that $\mathcal{F}$ is a $\sigma$-field. Let $E \in \mathcal{F}$. Then $\exists$ countable subcollection $\mathcal{A}_\omega$ of $\mathcal{A}$ such that $E \in \sigma(\mathcal{A}_\omega)$. But this means $E^c \in \sigma(\mathcal{A}_\omega)$ and so $E^c \in \mathcal{F}$ and $E \cup E^c = \Omega \in \mathcal{F}$.

Next suppose $(E_i)_{i \in \mathbb{N}} \in \mathcal{F}$. Then there exists countable subcollections $\mathcal{A}_{\omega_i}$ such that $E_i \in \sigma(\mathcal{A}_{\omega_i})\ \forall i \in \mathbb{N}$. But a countable union of countable sets is countable, so

$$E_i \in \sigma(\bigcup_{i \in \mathbb{N}} \mathcal{A}_{\omega_i})\ \forall i \in \mathbb{N}$$

which means:

$$\bigcup_{i \in \mathbb{N}} E_i \in \sigma(\bigcup_{i \in \mathbb{N}} \mathcal{A}_{\omega_i})$$

and so $\cup_{i\in\mathbb{N}} E_i \in \mathcal{F}$. We have just shown that $\mathcal{F}$ is a $\sigma$-field. But

$$\mathcal{A} \subset \mathcal{F} \subset \sigma(\mathcal{A})$$

and so $\mathcal{F} = \sigma(\mathcal{A})$. Finally to answer the problem, that means $B \in \sigma(\mathcal{A}) = \mathcal{F}$ and so we are done.

# 3 Tricky Probability Spaces

**Problem 8**
Find a probability space $(\Omega, \mathcal{F}, P)$ and events $A_1, ..., A_5 \in \mathcal{F}$ so that:
(i) any 4 of the events are independent, but all 5 are not.
(ii) any 3 of the events are independent but no 4 events are

**Solution:**

We will need the following lemma.

**Lemma 1.** Let $H^n$ be a non-empty finite $n$-dimensional inner product space over the field $K$. Suppose $W = (w_1, ..., w_n)$ is a uniformly distributed random vector on $H^n$. That is each $w_i$ is uniformly distributed on $H$ $\forall i \in \{1, ..., n\}$. Let $\{V_1, ..., V_k\}$ be a set of vectors in $H^n$. Define $Y_i = W \cdot V_i$ $\forall i \in \{1, ..., k\}$. Here $Y_i \in K$.

Then $V_1, ..., V_k \in H$ are linearly independent iff $Y_1, ..., Y_k$ are independent random variables.

*Proof.* First suppose $V_1, ..., V_k$ are linearly independent. WLOG, we can assume that $V_i = e_i$ where $e_i$ is the standard basis vector. This is because we can always rotate and stretch any axis of our frame to correspond with a vector $V_i$ while keeping track of what this does to $W$.

This means that $Y_i = w_i$ $\forall i \in \{1, ..., k\}$ by our assumption. It is clear that the $Y_i$ are independent because each $Y_i$ is the projection of the ith component of $W$ into $H$, and the components of $w_i$ of $W$ are independent of each other because each component is uniformly distributed on $H$.

for the other direction, we will show the contrapostive. Suppose that $V_1, ..., V_k$ are linearly dependent. Suppose the dimension of their span is $0 < m < k$ (if the dimension of their span was 0, then all the vectors would be 0 vectors and the conclusion would be trivial). Again we can rotate and stretch our frame so that we can assume WLOG that $V_i = e_i$ $\forall i \in \{1, ..., m\}$. After this transformation, the rest of the vectors $V_{m+1}, ..., V_k$ will have zeros in their last $k - m$ components. That is to say, $V_{m+1}, ..., V_k \in span(e_1, ..., e_m)$.

So now, $Y_i = w_i$ $\forall i \in 1, ..., m$, and

$$Y_j = \sum_{i=1}^{m} w_i V_j^i$$

$\forall j \in \{m+1, ..., k\}$ where $V_j^i$ is the ith component of $V_j$. Now it is clear that $Y_1, ..., Y_k$ are not independent because $Y_1, ..., Y_m$ completely determine that value of $Y_{m+1}, ..., Y_k$. That is, given the value of $Y_1, ..., Y_m$, the value of $Y_{m+1}, ..., Y_k$ will not be uniformly distributed but completely determined. For example,

$$P(\{Y_1 = 0\} \cap ... \cap \{Y_m = 0\} \cap \{Y_{m+1} \neq 0\} \cap ... \cap \{Y_k \neq 0\}) = 0$$

but

$$\prod_{i=1}^{m} P(\{Y_i = 0\}) \prod_{j=m+1}^{k} P(\{Y_j \neq 0\}) = \frac{(|H|-1)^{k-m}}{|H|^k}$$

where $|H|$ is the number of elements in $H$.

□

**part i)** We can now proceed with the setup of the solution. Let $V_1, ..., V_5$ be vectors in $\mathbb{F}_5^4$ such that any 4 vectors are linearly independent. Such vectors exists, for example, consider the collection of vectors:

$$\{(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1), (1,1,1,1)\}$$

Let $W = (w_1, w_2, w_3, w_4)$ be a uniformly distributed random vector on $\mathbb{F}_5^4$. That is each $w_i$ is uniformly distributed on $\mathbb{F}_5$.

Let $Y_i$ be the inner product of $W$ and $V_i$ in $\mathbb{F}_5$ and let $A_i = \{Y_i = 0\}$ $\forall i \in \{1,2,3,4,5\}$. Clearly any 4 $A_i$ are independent because any 4 $Y_i$ are independent by our lemma and the condition that any 4 of our 5 vectors are linearly independent.

Also note that all 5 events are not independent because the 5 vectors are linearly dependent because our vectorspace has dimension 4. Using the lemma, we find that the probability of the 5th event occuring will be completely determined by the first 4, and therefore the last event will not be uniformly distributed and therefore the 5 events we choose will not be independent.

**part ii)**

Let $V_1, ..., V_5$ be vectors in $\mathbb{F}_5^3$ such that any 3 vectors are linearly independent. Such vectors exists, for example, consider the collection of vectors:

$$\{(1,0,0), (0,1,0), (0,0,1), (2,3,4), (4,3,2)\}$$

(This collection was not entirely trivial to find). Let $W = (w_1, w_2, w_3)$ be a uniformly distributed random vector on $\mathbb{F}_5^3$. That is each $w_i$ is uniformly distributed on $\mathbb{F}_5$.

Let $Y_i$ be the inner product of $W$ and $V_i$ in $\mathbb{F}_5$ and let $A_i = \{Y_i = 0\}$ $\forall i \in \{1,2,3,4,5\}$. Clearly any 3 $A_i$ are independent because any 3 $Y_i$ are independent by our lemma and the condition that any 3 of our 5 vectors are linearly independent.

Also note that any 4 $A_i$ are not independent events because any 4 out of the 5 vectors are linearly dependent because our vectorspace only has dimension 3. Using the lemma, we find that the probability of the 4th event occuring will be completely determined by the first 3, and therefore the last event will not be uniformly distributed and therefore any 4 events we choose will not be independent.

# 4 Laws of Large Numbers

**Problem 9**
(Monte Carlo Integration). (i) let $f$ be a measurable function on $[0, 1]$ with $\int_0^1 |f(x)|dx < \infty$. Let $U_1, U_2, ...$ be independent and uniformly distributed on $[0, 1]$, and let:

$$I_n = n^{-1}(f(U_1) + ... + f(U_n))$$

Show that $I_n \to I \equiv \int_0^1 f dx in probability$
(ii) suppose $\int_0^1 |f(x)|^2 dx < \infty$. Use Chebyshev's inequality to estimate $P(|I_n - I| > a/n^{1/2})$.

**Solution**
**(i)**
First note that since $f$ is measurable on $[0, 1]$ and $U_1, U_2, ...$ are i.i.d. on $[0, 1]$, $f(U_1), f(U_2), ...$ are i.i.d.
Since $\int_0^1 |f(x)|dx < \infty$, $I_n \xrightarrow{p} E(f) = \int_0^1 |f(x)|dx \equiv I$ directly by theorem 2.2.9.

**(ii)**
suppose that $\int_0^1 |f(x)|^2 dx < \infty$. Then $var(f) < \infty$. Since $f(U_1), f(U_2), ...$ are i.i.d. they are uncorrelated.
so by Chebyshev's inequality:

$$P(|I_n - I| > a/n^{1/2}) \leq \frac{E((I_n - I)^2)}{a^2/n} = \frac{var(I_n)}{a^2/n} = \frac{n \cdot var(f)}{a^2/n} = \frac{var(f)}{a^2}$$

where the second last equality comes from theorem 2.2.1.

**Problem 10**
If $X_n$ is any sequence of random variables, there are constants $c_n \to \infty$ so that $X_n/c_n \to 0$ almost surely.

**Solution** Note that since $X_N$ are random variables, $P(|X_n| > \infty) = 1$. Therefore, we can choose $c_n$ such that $P(|X_n| > 2^{-n}c_n) < 2^{-n}$ $\forall n \in \mathbb{N}$. Then

$$\sum_{n=1}^{\infty} P(|X_n| > 2^{-n}c_n) = \sum_{n=1}^{\infty} P(|X_n/c_n| > 2^{-n}) < \sum_{n=1}^{\infty} 2^{-n} < \infty$$

Therefore, by the Borel-Cantelli lemma, $P(|X_n/c_n| > 2^{-n} \text{ i.o.}) = 0$. Since $2^{-n}$ can be made arbitrarily small, $\forall \epsilon > 0$, $P(|X_n/c_n| > \epsilon \text{ i.o.}) = 0$ and hence $X_n/c_n \overset{a.s.}{\to} 0$ by Durrett (lines

5-6 in page 57 ). Also note that increasing $c_n$ only helps our cause, so we can also choose that $c_n \to \infty$ if it is not so already.

**Problem 11**

Let $X_1, X_2, ...$ be i.i.d. with distribution $F$, let $\lambda_n \uparrow \infty$, and let $A_n = \{\max_{1\le m\le n} X_m > \lambda_n\}$. Show that $P(A_n \text{ i.o.}) = 0$ or 1 according as $\sum_{n\ge1}(1 - F(\lambda_n)) < \infty$ or $= \infty$.

**Solution** since $X_1, X_2, ...$ have the same distribution $F$, $P(X_i \le \lambda_n) = F(\lambda_n)\ \forall i$. Hence

$$P(A_n) = 1 - P(\{\underset{1\le m\le n}{min} X_m \le \lambda_n\}) = 1 - F(\lambda_n)$$

Hence if

$$\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty}(1 - F(\lambda_n)) < \infty$$

then $P(A_n \text{ i.o.}) = 0$ by the Borel-Cantelli lemma.

Else if

$$\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty}(1 - F(\lambda_n)) = \infty$$

then $P(A_n \text{ i.o.}) = 1$ by the second Borel-Cantelli lemma and independence of $X_1, X_2, ...$

**Problem 12**

Suppose that $\sum P(A_k) = \infty$. Show that if

$$\limsup_{n\to\infty} \left(\sum_{k=1}^{n} P(A_k)\right)^2 \Big/ \left(\sum_{1\le j,k\le n} P(A_j \hat{A}_k)\right) = \alpha > 0$$

then $P(A_n\text{i.o.}) \ge \alpha$. The case $\alpha = 1$ contains Theorem 2.3.6 in Durrett.

**Solution** for $n \in \mathbb{N}$, let

$$S_n = \sum_{m=1}^{n} 1_{A_m}$$

Note that for $0 \le \alpha \le 1$,

$$E(S_n) = E(S_n 1_{\{S_n < \alpha E(S_n)\}}) + E(S_n 1_{\{S_n \ge \alpha E(S_n)\}})$$

The first term on the right is bounded above by $\alpha E(S_n)$, and the second term is bounded above by $E(S_n^2)^{1/2} P(S_n > \alpha E(S_n))$ by Cauchy-Schwarz. Hence

$$P(S_n > \alpha E(S_n)) \ge (1-\alpha)^2 \frac{E(S_n)^2}{E(S_n^2)}$$

Also note that

$$\underset{n\to\infty}{lim} E(S_n) = \underset{n\to\infty}{lim} \sum_{m=1}^{n} P(A_m) = \infty$$

by hypothesis. Since $S_n(\omega)$ is the number of events in $\{A_1, ..., A_n\}$ that contain the point $\omega$, and $lim\ sup\ A_n$ are the points $\omega$ that are infinitely many $A_i$, we have that

$$\underset{n\to\infty}{lim\ sup}\ A_n = \{\omega : \lim_{n\to\infty} S_n(\omega) = \infty\}$$

Hence for all $\alpha > 0$,

$$\underset{n\to\infty}{lim\ sup}\{\omega : S_n(\omega) > \alpha E(S_n)\} \subset \underset{n\to\infty}{lim\ sup}\ A_n$$

Therefore for $0 < \alpha \le 1$,

$$\begin{aligned}\underset{n\to\infty}{lim\ sup}\ (1-\alpha)^2 \frac{E(S_n)^2}{E(S_n^2)} &\le \underset{n\to\infty}{lim\ sup}\ P(\{\omega : S_n(\omega) > \alpha E(S_n)\}) \text{ by the beginning of problem 4} \\ &\le P(\underset{n\to\infty}{lim\ sup}\{\omega : S_n(\omega) > \alpha E(S_n)\}) \\ &\le P(\underset{n\to\infty}{lim\ sup}\ A_n) \text{ by the above set containment}\end{aligned}$$

which implies

$$\underset{n\to\infty}{lim\ sup}\ \frac{E(S_n)^2}{E(S_n^2)} \le P(\underset{n\to\infty}{lim\ sup}\ A_n)$$

We are done once we realize that

$$E(S_n^2) = \sum_{k=1}^{n}\sum_{m=1}^{n} E(1_{A_k}1_{A_m}) = \sum_{k=1}^{n}\sum_{m=1}^{n} P(A_k \cap A_m)$$

and

$$E(S_n)^2 = \Big(\sum_{m=1}^{n} P(A_m)\Big)^2$$

**Problem 13**

Give an example with $X_n \in \{0, 1\}, X_n \to 0$ in probability, $N(n) \uparrow \infty$ a.s., and $X_N(n) \to 1$ a.s.

**Solution** Let $\Omega = [0, 1]$ and $\mathcal{F}$ be the $\sigma-$algebra of measurable subsets of $\Omega$. Let $P$ be the Lebesgue measure. let:

$$X_n(\omega) = \begin{cases} 1 & \text{if } \frac{k}{2^m} \le \omega < \frac{k+1}{2^m}, \text{ where } n = 2^m + k,\ 0 \le k \le 2^m - 1 \\ 0 & \text{else} \end{cases}$$

i.e. the type writer sequence. This is a standard example to show that $X_n$ converges to 0 in probability but not almost surely (as $n \to \infty$). Now let:

$$N(n) = min\{i : i \ge 2^n \text{ and } X_i > 0\}$$

and note that $N(n) \overset{a.s.}{\to} \infty$ and $X_{N(n)} = 1$ a.s. $\forall n$.

**Problem 14**
Let $X_0 = (1, 0)$ and define $X_n \in \mathbb{R}^2$ inductively by declaring that $X_{n+1}$ is chosen at random from the ball of radius $|X_n|$ centered at the origin, i.e., $X_{n+1}/|X_n|$ is uniformly distributed on the ball of radius 1 and independent of $X_1, ..., X_n$. Prove that $n^{-1} \log |X_n| \to c$ a.s. and compute c.

**Solution** Let

$$Y_1 = \frac{X_1}{|X_0|}, Y_2 = \frac{X_2}{|X_1|}, .., Y_i = \frac{X_i}{|X_{i-1}|}, ...$$

Then $Y_1, Y_2, ..$ are independent and uniformly distributed on the unit ball in $\mathbb{R}^2$. Hence,

$$\begin{aligned}\frac{1}{n} log|X_n| =& \frac{1}{n} log(\frac{|X_n|}{|X_{n-1}|}\frac{|X_{n-1}|}{|X_{n-2}|}...\frac{|X_1|}{|X_0|}) = \frac{1}{n} log(|Y_n||Y_{n-1}|...|Y_1|) \\ =& \frac{1}{n}\sum_{k=1}^{n} log|Y_k| \xrightarrow{a.s.} E\ log|Y_1| \quad \text{by the SLLN}\end{aligned}$$

To find $E\ log|Y_1|$, first note that for $r \in [0, 1]$ and $i \in \mathbb{N}$,

$$P(|Y_i| \leq r) = \frac{\pi r^2}{\pi 1^2}$$

and hence $|Y_i|$ has probability density $f_i(r) = \frac{d}{dr} r^2 = 2r$ with respect to Lebesgue measure on $[0, 1]$. And so

$$c := E\ log|Y_1| = \int_0^1 log(r) f(r) dr = -\frac{1}{2}$$

and we are done.

**Problem 15**
Let $X_1, X_2, ...$ be *i.i.d.* and let $S_n = X_1 + ... + X_n$. Let $p > 0$. If $S_n/n^{1/p} \to 0 a.s.$ then $E|X_1|^p < \infty$.

**Solution** Suppose

$$\frac{S_n}{n^{1/p}} \xrightarrow{a.s.} 0$$

Then,

$$\frac{X_n}{n^{1/p}} = \frac{S_n - S_{n-1}}{n^{1/p}} = \frac{S_n}{n^{1/p}} - \frac{S_{n-1}}{(n-1)^{1/p}} \frac{(n-1)^{1/p}}{n^{1/p}} \xrightarrow{a.s.} 0$$

This means that

$$P(|X_n| > n^{1/p} \text{i.o.}) = 0$$

Hence

$$\sum_{n=1}^{\infty} P(|X_n|^p > n) = \sum_{n=1}^{\infty} P(|X_n| > n^{1/p}) < \infty$$

by the contrapositive of the second Borel-Cantelli lemma, and the assumption that $X_1, X_2, ...$ are independent.
Finally, by lemma 2.2.8,

$$\begin{aligned} E|X_1|^p &= \int_0^\infty py^{p-1}P(|X_1|>y)dy = \int_0^1 py^{p-1}P(|X_1|>y)dy + \int_1^\infty py^{p-1}P(|X_1|>y)dy \\ &\le \int_0^1 py^{p-1}P(|X_1|>y)dy + p\int_1^\infty P(|X_1|>y)dy \\ &\le \int_0^1 py^{p-1}dy + 2\sum_{n=1}^\infty P(|X_1|>n) \quad \text{since } 1<p<2 \\ &\le \int_0^1 py^{p-1}dy + 2\sum_{n=1}^\infty P(|X_1|>n^{1/p}) < \infty \end{aligned}$$

since both terms are finite.

**Problem 16**
Let $X_1 m X_2, ...$ be independent with $EX_n = 0$, $\text{var}(X_n) = \sigma_n^2$.
(i) Show that if $\sum_n \sigma_n^2/n^2 < \infty$ then $\sum_n X_n/n$ converges a.s. and hence $n^{-1}\sum_{m=1}^n X_m \to 0$ a.s.
(ii) Suppose $\sum \sigma_n^2/n^2 = \infty$ and without loss of generality that $\sigma_n^2 \le n^2$ for all $n$. Show that there are independent random variables $X_n$ with $EX_n = 0$ and $\text{var}(X_n) \le \sigma_n^2$ so that $X_n/n$ and hence $n^{-1}\sum_{m\le n} X_m$ does not converge to 0 a.s.

**Solution** If

$$\sum_{n=1}^\infty \frac{X_n}{n} \quad \text{converges a.s.}$$

then

$$\frac{1}{n}\sum_{m=1}^n X_m \to 0 \ \text{ a.s.}$$

by Kronecker's lemma.

**(ii)**
By assumption, $\sigma_n \le n$. Let

$$P(X_n = n) = P(X_n = -n) = \frac{\sigma_n^2}{2n^2}, \;\; P(X_n = 0) = 1 - \frac{\sigma_n^2}{n^2}$$

Now $X_n$ are independent with $E(X_n) = 0$ and $var(X_n) = \sigma_n^2$, but $\forall 0 < \epsilon < 1$,

$$P(|X_n|/n > \epsilon) = \frac{\sigma_n^2}{n^2}$$

Hence

$$\sum_{n=1}^\infty P(|X_n|/n > \epsilon) = \infty$$

Therefore by the second Borel-Cantelli lemma,

$$P(|X_n|/n > \epsilon \text{ i.o.}) = 1 \text{ hence } P(X_n/n \to 0) = 0$$

So $X_n/n$ does not converge to 0 a.s. and hence $n^{-1}\sum_{m\leq n} X_m$ does not converge to 0 a.s.

**Problem 17**

Let $X_1, X_2, ...$ be independent and let $S_{m,n} = X_{m+1} + ... + X_n$. Then

$$P\left(\max_{m<j\leq n} |S_{m,j}| > 2a\right) \min_{m<k\leq n} P(|S_{k,n}| \leq a) \leq P(|S_{m,n}| > a)$$

**Solution** for $m < k \leq n$, define

$$A_{k,\epsilon} = \left\{\omega : \mathop{max}_{m<j\leq k-1} |S_{m,j}| \leq 2\epsilon, |S_{m,k}| > 2\epsilon\right\}$$

Then $A_{k,\epsilon}$ is the event that the indicated maximum occurs on $S_{m,k}$ and not $S_{m,j}$ for $m < j < k$. Thus these sets are disjoint and

$$\{\omega : \mathop{max}_{m<j\leq n} |S_{m,j}| > 2\epsilon\} = \bigcup_{k=m+1}^{n} A_{k,\epsilon}$$

Now note that for all $m < k \leq n$,

$$A_{k,\epsilon} \bigcap \{\omega : |S_{k,n}| \leq \epsilon\} \subset \{|S_{m,n}| > \epsilon\}$$

Hence

$$\bigcup_{k=m+1}^{n} \left(A_{k,\epsilon} \bigcap \{\omega : |S_{k,n}| \leq \epsilon\}\right) \subset \{|S_{m,n}| > \epsilon\}$$

Taking the probability of both sides and using independence, we obtain:

$$\sum_{k=m+1}^{n} P(A_{k,\epsilon}) P(\{\omega : |S_{k,n}| \leq \epsilon\}) \leq P(\{|S_{m,n}| > \epsilon\}) \tag{2}$$

Finally, note that

$$\begin{aligned}
\sum_{k=m+1}^{n} P(A_{k,\epsilon}) P(\{\omega : |S_{k,n}| \leq \epsilon\}) &\geq \mathop{min}_{m<k\leq n} P(\{\omega : |S_{k,n}| \leq \epsilon\}) \left(\sum_{k=m+1}^{n} P(A_{k,\epsilon})\right) \\
&= \mathop{min}_{m<k\leq n} P(\{\omega : |S_{k,n}| \leq \epsilon\}) P(\{\omega : \mathop{max}_{m<j\leq n} |S_{m,j}| > 2\epsilon\})
\end{aligned}$$

Combining this with (1), we obtain the desired inequality:

$$P(\{\omega : \mathop{max}_{m<j\leq n} |S_{m,j}| > 2\epsilon\}) \mathop{min}_{m<k\leq n} P(\{\omega : |S_{k,n}| \leq \epsilon\}) \leq P(\{|S_{m,n}| > \epsilon\})$$

**Problem 18**
Use the previous problem to prove the following: Let $X_1, X_2, ...$ be independent and let $S_n = X_1 + ... + X_n$. If $\lim_{n\to\infty} S_n$ exists in probability then it also exists a.s.

**Solution** Suppose $S_n$ converges in probability. Then $S_n$ is Cauchy in probability. Therefore we have

$$\min_{m\le k,n} P(|S_n - S_k| \le \epsilon) \to 1 \text{ as } m \to \infty$$

Fix an $m_0$ such that for all $m \ge m_0$,

$$\min_{m\le k,n} P(|S_n - S_k| \le \epsilon) > \frac{1}{2}$$

Now for all $m_0 \le m < n$, we have

$$\begin{aligned}
\frac{1}{2}P\Big(\max_{m<j,k\le n}|S_k - S_j| > 4\epsilon\Big) \le& \frac{1}{2}P\Big(\max_{m<j\le n}|S_m - S_j| > 2\epsilon\Big)\\
\le& P\Big(\max_{m<j\le n}|S_m - S_j| > 2\epsilon\Big)\min_{m<k\le n} P(|S_k - S_n| \le \epsilon)\\
\le& P(|S_n - S_m| > \epsilon) \longrightarrow 0 \ \text{ as } \ m, n \to \infty
\end{aligned}$$

where the last inequality is obtained by Problem 9. This implies that

$$P\Big(\max_{m<j,k\le n}|S_k - S_j| > 4\epsilon\Big) \longrightarrow 0 \ \text{ as } \ m, n \to \infty$$

Note that the max is always defined since we are taking the limit of finite values of m and n. Since this is true for all $\epsilon > 0$, the last line implies that the probability that $S_n$ is not Cauchy in probability is 0, hence the probability that $S_n$ is Cauchy in probability is 1. Therefore $S_n$ converges almost surely.

# 5 Random Walks

**Problem 19**
Let $S_0, S_1, S_2, ...$ be simple random walk on the integers (with $S_i = X_1 + ... + X_i$). Let $\mathcal{F}_i = \sigma(X_1, ..., X_i)$, be the natural filtration. Let $a > 1$ be an integer, and let $T$ be the first time $i$ that $|S_i| = a$.
(a) Show that $T$ is a stopping time with respect to the natural filtration.
(b) Show that $\tau = T - 1$ is not a stopping time with respect to the natural filtration.
(c) Show that nevertheless, $X_{\tau+1}, X_{\tau+2}, ...$ is an i.i.d. sequence.
(d) On the other hand, show that the sequence in $(c)$ is not independent of $\mathcal{F}_\tau$.

**Solution (a)** Let $A^c = (-a, a)$. Note that

$$\{T = n\} = \{S_1 \in A^c, ..., S_{n-1} \in A^c, S_n \in A\} \in \mathcal{F}_n$$

So $T$ is a stopping time with respect to the natural filtration.

**(b)**
Note that

$$\{\tau = n\} = \{T - 1 = n\} = \{T = n + 1\} = \{S_1 \in A^c, ..., S_n \in A^c, S_{n+1} \in A\} \notin \mathcal{F}_n$$

since $X_{n+1} \notin \mathcal{F}_n$ So $\tau$ is not a stopping time with respect to the natural filtration.

**(c)**
First note that $X_T$ has the same distribution as $X_1$ since $P(S_n = a-1) = P(S_n = -(a-1))$ and $P(S_{n+1} = 1) = P(S_{n+1} = -1)$. Next it's clear that $X_{T+1}, X_{T+2}, ...$ are independent since they are not constrained (in the way $X_{T-1}$ is) after existing $(-a, a)$. Hence by theorem 4.1.3, $\{X_{T+n}, n \geq 1\}$ have the same distribution as the original sequence and are independent of $F_T$ and in particular $X_T$. Hence $\{X_T, X_{T+1}, ...\}$ are i.i.d.

**Problem 1 (d)**
Let $B = [0, 1]$ and let $n > 1$ be a positive integer. Noting that $\{\tau = n\} \in \mathcal{F}_\tau$, we have:

$$\begin{aligned} P(\{X_{\tau+1} \in B\} \cap \{\tau = n\}) &= P(\{X_{n+1} \in B\} \cap \{S_1 \in A^c, ..., S_n \in A^c, S_{n+1} \in A\} \\ &= P(\{S_1 \in A^c, ..., S_n \in A^c, S_{n+1} \in A\} | \{X_{n+1} \in B\}) P(\{X_{n+1} \in B\}) \\ &\neq P(\{S_1 \in A^c, ..., S_n \in A^c, S_{n+1} \in A\}) P(\{X_{n+1} \in B\}) \end{aligned}$$

Since whether $S_{n+1}$ steps into $A$ or not depends on $X_{n+1}$. Hence the sequence in question $1(c)$ is not independent of $\mathcal{F}_\tau$.

**Problem 20**
Using the same setup as the previous problem,
(a) show that there exists constants $c > 0, b \in (0, 1)$ depending on $a$ only so that $P(T > t) \leq cb^t$ for all $t$.
(b) Show that $T$ has finite mean and variance

**Solution (a)** By example 4.1.5. in Durrett, $P(T > 2na) \leq (1 - 2^{-2a})^n$ Letting $t = 2na$, we have

$$P(T > t) \leq [(1 - 2^{-2a})^{\frac{1}{2a}}]^t$$

finally, $b = (1 - 2^{-2a})^{\frac{1}{2a}} \in (0, 1)$ since $1 - 2^{-2a} \in (0, 1)$.

**(b)**

Note that since $(E(T))^2 < \infty$, all we have to show is that $E(T^2) < \infty$. By lemma 2.2.8. in Durrett,

$$E(T^2) = \int_0^\infty 2t * P(T > t)dt \leq \sum_1^\infty 2t[(1 - 2^{-2a})^{\frac{1}{2a}}]^t < \infty$$

since $[(1 - 2^{-2a})^{\frac{1}{2a}}]^t$ decreases exponentially since $0 < [(1 - 2^{-2a})^{\frac{1}{2a}}] < 1$.

**Problem 21**

Let $X_1, X_2, ...$ be *i.i.d.* with $P(x_1 = 1) = p > 1/2$ and $P(X_1 = -1) = 1 - p$, and let $S_n = X_1 + ... + X_n$. Let $\alpha = \inf\{m : S_m > 0\}$ and $\beta = \inf\{n : S_n < 0\}$.

(i) Show that $P(\alpha < \infty = 1$ and $P(\beta < \infty) < 1$.

(ii) If $Y = \inf S_n$, then $P(Y \leq -k) = P(\beta < \infty)^k$.

(iii) Apply Wald's euqation to $\alpha \wedge n$ and let $n \to \infty$ to get $E\alpha = 1/EX_1 = 1/(2p - 1)$.

**Solution (i)**

Note $E(X_1) = p - (1 - p) = 2p - 1 > 0$ since $p > 1/2$. Hence by the strong law of large numbers, $S_n \to \infty$. Hence $sup\, S_n = \infty$ and $inf\ S_n > -\infty$. Finally, by exercise 4.1.9, this corresponds to the case of $P(\alpha < \infty) = 1$ and $P(\beta < \infty) < 1$.

**(ii)**

We argue by induction. For $k = 1$, the equality is clear because $inf S_n \leq -1 \iff \beta < \infty$. Assume the equality is true for $k = n - 1$. Note that

$$\{inf\ S_m \leq -n\} = \{M = inf\{m : S_m \leq -(n-1)\} < \infty\} \cap \{\underset{j>M}{inf}(S_j + n - 1) < 0)\} \quad (3)$$

$$= \{inf S_m \leq -(n-1)\} \cap \{\underset{j}{inf}(S_j + n - 1) < 0)\} \quad (4)$$

since $(S_j + n - 1) \geq 0$ for $j \leq M$. Hence:

$$\begin{aligned} P(\{inf\ S_m \leq -n\}) &= P(\{M = inf\{m : S_m \leq -(n-1)\} < \infty\})P(\{\underset{j>M}{inf}(S_j + n - 1) < 0)\}) \\ &= P(\{inf S_m \leq -(n-1)\})P(\{\underset{j>M}{inf}(S_j + n - 1) < 0)\}) \\ &= P(\{inf S_m \leq -(n-1)\})P(inf S_j < 0)) \\ &= P(\beta < \infty)^{n-1}P(\beta < \infty) = P(\beta < \infty)^n \end{aligned}$$

Where we used the independence of the sets on the right hand side of (1) in the first equality above and theorem 4.1.4 for the second last equality above.

**(iii)**

By Wald's equation: $E(\alpha \wedge n)E(X_1) = E(S_{\alpha\wedge n})$. From part (i) (the fact that $P(\alpha < \infty) = 1$), we have that $\alpha \wedge n \uparrow \alpha$. Hence $S_{\alpha\wedge n} \to S_\alpha = 1$. By monotone convergence theorem: $E(\alpha \wedge n)E(X_1) \uparrow E(\alpha)E(X_1)$ as $n \to \infty$. Hence By dominated convergence theorem:

$$1 = E(S_\alpha) \leftarrow E(S_{\alpha\wedge n}) = E(\alpha \wedge n)E(X_1) \to E(\alpha)E(X_1) = E(\alpha)(2p - 1)$$

So the equality follows trivially.

**Problem 22 (optimal stopping)**

Let $X_n, n \geq 1$ be i.i.d. with $EX_1^+ < \infty$ and let

$$Y_n = \max_{1\leq m\leq n} X_m - cn$$

That is, we are looking for a large value of $X$, but we have to pay $c > 0$ for each observation.
(i) Let $T = inf\{n : X_n > a\}$, $p = P(X_n > a)$, and compute $EY_T$.
(ii) Let $\alpha$ (possibly $\alpha < 0$) be the unique solution of $E(X_1 - \alpha)^+ = c$. Show that $EY_T = \alpha$ in the case and use the inequality

$$Y_n \leq \alpha + \sum_{m=1}^{n} ((X_m - \alpha)^+ - c)$$

for $n \geq 1$ to conclude that if $\tau \geq 1$ is a stopping time with $E\tau < \infty$, then $EY_\tau \leq \alpha$. The analysis above assumes that you to play at least once. If the optimal $\alpha < 0$, then you shouldn't play at all.

**Solution (i)** First note that $T$ has a geometric distribution with a chance of success of $p$. Hence $E(T) = 1/p$. Also note that since $X_T$ is the first $X_n$ that is larger than a, it has the same distribution as $X_1$ conditional on $X_1 > a$ and $\underset{1\leq m\leq T}{max} X_m = X_T$. So

$$E(Y_T) = E(X_T) - cE(T) = a + E(X - a)^+/p - c/p$$

**(ii)**

Plugging $\alpha = a$ in the above equation gives $E(Y_T) = \alpha$. Using the given inequality and plugging $\tau$ for $n$, we get:

$$Y_\tau \leq \alpha + \sum_{m=1}^{\tau} (X - \alpha)^+ - c\tau$$

Hence Wald's equation gives:

$$E(Y_\tau) \leq \alpha + E(\tau)E(X - \alpha)^+ - cE(\tau) = \alpha + E(\tau)c - cE(\tau) = \alpha$$

# 6 Martingales

**Problem 23**

Suppose $X \geq 0$ and $EX = \infty$. Show that there is a unique $\mathcal{F}$-measurable $Y$ with $0 \leq Y \leq \infty$ so that

$$\int_A XdP = \int_A YdP \text{ for all A} \in \mathcal{F}$$

**Solution**

Letting $Y_M = E(X_M|\mathcal{F})$ as in the hint where $X_M = X \wedge M$, and letting $M \to \infty$, we get by theorem 5.1.2(c): $Y_M = E(X_M|\mathcal{F}) \uparrow$ some limit $Y$ since $X_M \uparrow X$. By definition, $Y$ is $\mathcal{F}$-measurable. Now for all $A \in \mathcal{F}$, the conditional expectation definition gives:

$$\int_A X \wedge MdP = \int_A Y_M dP$$

Taking the limit as $M \uparrow \infty$, monotone convergence theorem then gives:

$$\int_A XdP = \int_A YdP$$

Hence $Y$ satisfies (i) and (ii) on page 189, hence the short uniqueness proof on page 190 applies and we are done.

**Problem 24**

Let $\text{var}(X|\mathcal{F}) = E(X^2|\mathcal{F}) - E(X|\mathcal{F})^2$. Show that

$$\text{var}(X) = E(\text{var}(X|\mathcal{F})) + \text{var}(E(X|\mathcal{F}))$$

**Solution**

First note that $E(E(X|\mathcal{F})) = E(X)$ and similarly, $E(E(X^2|\mathcal{F})) = E(X^2)$. Hence

$$var(E(X|\mathcal{F})) = E((E(X|\mathcal{F}))^2) - (E(E(X|\mathcal{F})))^2 = E((E(X|\mathcal{F}))^2) - (E(X))^2$$

and

$$E(var(X|\mathcal{F})) = E(E(X^2|\mathcal{F})) - E((E(X|\mathcal{F}))^2) = E(X^2) - E((E(X|\mathcal{F}))^2)$$

so finally:

$$E(var(X|\mathcal{F})) + var(E(X|\mathcal{F})) = E(X^2) - (E(X))^2 = var(X)$$

**Problem 25**

Show that if $X$ and $Y$ are random variables with $E(Y|\mathcal{G}) = X$ and $EY^2 = EX^2 < \infty$, then $X = Y$ a.s.

**Solution**
First we compute:

$$E(XE(Y|\mathcal{G})) = E(E(XE(Y|\mathcal{G})|\mathcal{G})) = E((E(Y|\mathcal{G}))^2)$$

hence:

$$E((X - E(Y|\mathcal{G}))^2) = E(X^2) - 2E(XE(Y|\mathcal{G})) + E((E(Y|\mathcal{G}))^2) = E(X^2) - E((E(Y|\mathcal{G}))^2)$$

which we use to conclude:

$$0 = E(Y^2) - E(X^2) = E(Y^2) - E(E(Y|\mathcal{G})) = E((Y - E(Y|\mathcal{G}))^2) = E((Y - X)^2)$$

hence $X = Y$ a.s. since $(X - Y)^2 = 0$ a.s.

**Problem 26**
The result in the last problem implies that if $EY^2 < \infty$ and $E(Y|\mathcal{G})$ has the same distribution as $Y$, then $E(Y|\mathcal{G}) = Y$ a.s. Prove this under the assumption that $E|Y| < \infty$.

**Solution**
First, Jensen's inequality implies that

$$|E(X|\mathcal{G})| \leq E(|X||\mathcal{G})$$

The conditions in the problem imply that we have an equality in the above. Hence, when $E(X|\mathcal{G}) \geq 0$, we have that $E(X|\mathcal{G}) = E(|X||\mathcal{G})$ a.s. and when $E(X|\mathcal{G}) < 0$, we have that $-E(X|\mathcal{G}) = -E(|X||\mathcal{G})$ a.s. Written more formally:

$$E(|X| - X; E(X|\mathcal{G}) \geq 0) = 0 \quad \text{and} \quad E(|X| - X; E(X|\mathcal{G}) < 0) = 0$$

Hence $sgn(X) = sgn(E(X|\mathcal{G}))$ a.s. Taking $X = Y - c$ for any real c as in the hint and doing the above again, we get that $sgn(Y - c) = sgn(E(Y|\mathcal{G}) - c)$ a.s. for all real $c$. Hence $Y = E(X|\mathcal{G})$ a.s.

**Problem 27**
(a) Let $M$ be a martingale, with $EM_n^2$ finite for all $n$. Show that if $i < j \leq k < l$ then $E[(M_l - M_k)(M_j - M_i)] = 0$
(b) Show directly (without use of martingale convergence theorem) that if $\sup_{n\geq 0} EM_n^2 < \infty$, then there is a random variable $M_\infty$ so that $M_n \to M_\infty$ in $L^2$. Also show that $EM_\infty = EM_0$

**Solution**
Let $i < j \leq k < \ell$. Then by theorem 5.1.5:

$$\begin{aligned} E((M_\ell - M_k)(M_j - M_i)) &= E[E[(M_\ell - M_k)(M_j - M_i)|\mathcal{F}_j]] \\ &= E[(M_j - M_i)E[(M_\ell - M_k)|\mathcal{F}_j]] \\ &= E[(M_j - M_i)(M_j - M_j)] = 0 \end{aligned}$$

Where we used

$$E[M_\ell|\mathcal{F}_j] = E[E[M_\ell|\mathcal{F}_{\ell-1}]|\mathcal{F}_j] = E[M_{\ell-1}|\mathcal{F}_j] = ... = E[M_j|\mathcal{F}_j] = M_j$$

and similarly for $M_k$.

**(b)**
Since $\sup_{n\geq 0} E(M_n^2) < \infty$, there exists $K$ such that $E(M_n^2) \leq K$ for all $n \geq 0$. Let $X_n = M_n - M_{n-1}$ for $n > 0$.

Note trivially that: $|2xy| \leq x^2+y^2$, we have that $|(x+y)^2| \leq x^2+y^2+|2xy| \leq 2(x^2+y^2)$. Hence for all $n \geq 1$, $E[(M_n - M_0)^2] \leq 2(E[M_n^2] + E[M_0^2]) \leq 4K$. Note also that for any $j \neq k$, $E(X_jX_k) = 0$ by part a. Hence

$$E[(M_n - M_0)^2] = E[(X_n + X_{n-1} + ... + X_1)^2] = \sum_{k=1}^{n} E[X_k^2]$$

since the $E(X_jX_k)$ terms vanish. Combining these results, we get:

$$\sum_{k=1}^{n} E[X_k^2] = E[(M_n - M_0)^2] \leq 4K$$

$\sum_{k=1}^{n} E[X_k^2]$ is increasing in $n$ and bounded above, hence it converges. Therefore we have $E[X_k^2] \overset{k\to\infty}{\longrightarrow} 0$. But this means that $M_n$ is Cauchy, and hence converges to some $M_\infty$ in $L^2$. Now note that for all $k > 0$, by theorem 5.1.5:

$$E[M_k] = E[E[M_k|\mathcal{F}_{k-1}]] = E[M_{k-1}] = ... = E[M_0]$$

but convergence in $L^2$ for random variables implies convergence in $L^1$, hence

$$E[M_\infty] = \lim_n E[M_n] = \lim_n E[M_0] = E[M_0]$$

**Problem 28 (Conditional expectation for $L^2$ random variables)**
(a) Let A be a subspace of $L^2(\Omega, \mathcal{F}, P)$ that is closed in $L^2$. show that for any $X \in L^2$, there exists a $Y \in A$ that minimizes $||X - Y||_2$. Show that $X - Y$ is orthogonal to all elements of $A$.
(b) Show that if $\mathcal{G} \subset \mathcal{F}$ is a sub sigma-algebra, then $A = L^2(\Omega, \mathcal{G}, P)$ is a closed subspace of $L^2(\omega, \mathcal{F}, P)$
(c) Show that in a the setup of (a) and (b), $Y = E(X|\mathcal{G})$.

**Solution**
$L^2(\Omega, \mathcal{F}, P)$ is a Hilbert space, hence this entire problem (including showing that $X - Y$ is orthogonal to all elements of A) follows directly from Hilbert's projection theorem. I'm sure you agree that there's no point in copying it here.

**(b)**
Let $f, g \in L^2(\Omega, \mathcal{G}, P)$, $c \in \mathbb{R}$. Then $cf + g \in \mathcal{G}$. Furthermore $E((cf + g)^2) \leq 2E((cf)^2) +$

$2E(g^2) = 2c^2E(f^2) + 2E(g^2) < \infty$. Hence $cf + g \in L^2(\Omega, \mathcal{G}, P)$ and $L^2(\Omega, \mathcal{G}, P)$ is a subspace. To show closedness, let $f_n$ be a sequence in $L^2(\Omega, \mathcal{G}, P)$ that converges to some $f \in L^2(\Omega, \mathcal{F}, P)$. Then by theorem 1.3.5 in Durrett, $f = \lim f_n = \limsup f_n \in \mathcal{G}$ Hence $f \in L^2(\Omega, \mathcal{G}, P)$.

**(c)**
Let $X \in L^2(\Omega, \mathcal{F}, P)$. Let $Y = E[X|\mathcal{G}]$ where $\mathcal{G} \subset \mathcal{F}$. By theorem 5.1.4 in Durrett $E|Y|^2 \leq E|X|^2$. Therefore $Y \in L^2$. The rest follows directly from theorem 5.1.8 in Durrett and we are done.

What follows is an alternative proof of $Y \in L^2$ for "fun" (and not because I already wrote it out before stumbling across the theorem): Suppose $X$ is non-negative on the set $A \in \mathcal{G}$. Let $B := \{x \in A : Y(x) \geq 0\}$. Then $P(A \setminus B) = 0$ or else we would have that $Y$ is negative on a set $C \in \mathcal{G}$ of positive measure but then $C \subset A$ and so $0 \geq \int_C X dP = \int_C Y dP < 0$ which is a contradiction.

Now for $A \in \mathcal{G}$ Let $A^+ := \{x \in A : X(x) \geq 0\}$, $A^- := \{x \in A : X(x) < 0\}$, $A_+ := \{x \in A : Y(x) \geq 0\}$, $A_- := \{x \in A : Y(x) < 0\}$. Note that by the above paragraph: $P(A^+ \setminus A_+) = 0$ and similarly with $A^-$ and $A_-$. So we find that for all $A \in \mathcal{G}$:

$$\begin{aligned} \int_A |X| &= \int_A X^+ - \int_A X^- = \int_{A^+} X - \int_{A^-} X = \int_{A^+} Y - \int_{A^-} Y \\ &= \int_{A_+} Y - \int_{A_-} Y = \int_A Y^+ - \int_A Y^- = \int_A |Y| \end{aligned}$$

Hence for all $A \in \mathcal{G}$ (and in particular $A = \Omega$), we see that $\int_A |X|^2 = \int_A |Y|^2$.

**Problem 29**
Use regular conditional probability to get the condition Holder inequality from the unconditional one, i.e. show that if $p, q \in (1, \infty)$ with $1/p + 1/q = 1$, then

$$E(|XY||\mathcal{G}) \leq E(|X|^p|\mathcal{G})^{1/p} E(|Y|^q|\mathcal{G})^{1/q}$$

**Solution**
To do this problem, first we do problem 5.1.14 (in Durret): if $f = 1_A$ then the problem is true from definition. Now we just use the regular construction (simple functions by linearity, then non-negative functions by monotone convergence, and finally general functions by $f = f^+ - f^-$).

Now we just apply problem 5.1.14 and the usual Holder inequality after fixing $\omega$:

$$\int |X(\omega')Y(\omega')|\mu(\omega, dx) \leq \left(\int |X(\omega')|\mu(\omega, dx)\right)^{1/p} \left(\int |Y(\omega')|\mu(\omega, dx)\right)^{1/q}$$

**Problem 30**
Suppose $f$ is superharmonic on $R^d$. Let $\xi_1, \xi_2, ...$ be i.i.d. uniform on $B(0,1)$, and define $S_n$ by $S_n = S_{n-1} + \xi_n$ for $n \geq 1$ and $S_0 = x$. Show that $X_n = f(S_n)$ is a supermartingale.

**Solution**
Since $f$ is superharmonic on $\mathbb{R}^d$, it is continuous on $\mathbb{R}^d$. Hence, since $S_n \in B(x,n)$, and $f$ is bounded on $B(x,n)$ by continuity, $E(X_n) = E(f(S_n)) < \infty$ for all $n$ so condition $(i)$ is satisfied in the definition.

Letting $\mathcal{F}_n$ be the natural filtration on $S_n$, we see that $X_n = f(S_n)$ is adapted to $\mathcal{F}_n$ by Theorem 1.3.2. and continuity of $f$. Hence condition $(ii)$ is satisfied.

Finally, we compute:

$$E(X_{n+1}|\mathcal{F}_n) = E(f(S_n + \xi_{n+1})|\mathcal{F}_n) = \frac{1}{|B(S_n,1)|}\int_{B(S_n,1)} f(y)dy \leq f(S_n) = X_n$$

and so condition $(iii)$ is satisfied.

**Problem 31**
Give an example of a submartingale $X_n$ so that $X_n^2$ is a supermartingale.

**Solution**
Let $X_n = -1/n$ on $[0,1]$. Then $E(X_n) = -1/n$ is increasing but $E((X_n)^2) = 1/n^2$ is decreasing.

**Problem 32**
Give an example of a martingale $X_n$ with $X_n \to -\infty$ a.s.

**Solution**
Let $P(\xi_n = -1) = 1 - 1/2^n n$, $P(\xi_n = \frac{1-1/2^n}{2^n}) = 1/2^n$. Then $E(\xi_n) = 0$ for all $n$. Hence $X_n = \xi_1 + ... + \xi_n$ is clearly a martingale. Now

$$\sum_{n=1}^{\infty} P(\xi_n \neq -1) = \sum_{n=1}^{\infty} \frac{1}{2^n} = 1 < \infty$$

Hence, by Borel-Cantelli: $P(\xi_n \neq -1 \ i.o.) = 0$. Therefore $X_n/n \to -1$ meaning $X_n \to -\infty$.

**Problem 33**
Let $Y_1, Y_2, ...$ be nonnegative i.i.d random variables with $EY_m = 1$ and $P(Y_m = 1) < 1$
(i) show that $X_n = \prod_{m\leq n} Y_m$ defines a martingale.
(ii) Use theorem 5.2.9 and an arguement by contradiction to show that $X_n \to 0$ a.s.
(iii) Use the strong law of large numbers to conclude that $(1/n)\log X_n \to c < 0$.

**Solution (i)**
To check condition $(i)$, just note that $E(|X_n|) = E(X_n) = \prod_{m\leq n} E(Y_m) = 1$. Condition $(ii)$ is met by letting $\mathcal{F}_n$ to be the natural filtration. To see condition $(iii)$, just note that:

$$E(X_{n+1}|\mathcal{F}_n) = E(X_n Y_{n+1}|\mathcal{F}_n) = X_n E(Y_{n+1}|\mathcal{F}_n) = X_n$$

**(ii)**
Since $Y_n$ are i.i.d and $P(Y_n \neq 1) < 1$, choose $\delta > 0$ such that $P(|Y_n - 1| > \delta) = \epsilon > 0$. Since $X_n$ is non-negative, for all $\alpha > 0$,

$$P(X_n \geq \alpha)P(|Y_n - 1| > \delta) \leq P(X_n|Y_n - 1| \geq \alpha\epsilon) = P(|X_{n+1} - X_n| \geq \alpha\epsilon) \to 0$$

Where the convergence of the term on the right is given by the fact that $X_n \to X$ a.s. by theorem 5.2.9. By assumption $P(|Y_n - 1| > \delta) = \epsilon > 0$ for all $n$, so it must be that $P(X_n \geq \alpha) \to 0$. Since this is true for all $\alpha > 0$, we have that $X_n \to 0$ in probability. Hence there exists a subsequence $X_{n_k} \to 0$ a.s. But we already know that $X_n \to X$ a.s. by theorem 5.2.9. Hence $X_n \to 0$ a.s.

**(iii)**
First by Jensen's inequality: $E(\log(Y_k)) \leq \log(E(Y_k)) = 0$, however since $P(Y_m = 1) < 1$, we have a strict inequality. So now by the SLLN:

$$\frac{1}{n}\log(X_n) = \frac{1}{n}\sum_{k=1}^{n}\log(Y_k) \to E(\log(Y_1)) < 0$$

**Problem 34**
Suppose $y_n > -1$ for all $n$ and $\sum |y_n| < \infty$. Show that $\prod_{m=1}^{\infty}(1 + y_m)$ exists.

**Solution**
This is equivalent to showing $\sum_{m=1}^{\infty}\log(1 + y_m) < \infty$ i.e. $\sum_{m=k}^{\infty}\log(1 + y_m) \overset{k\to\infty}{\longrightarrow} 0$. $\sum_{m=1}^{\infty}|y_m| < \infty$ means that $|y_k| \overset{k\to\infty}{\longrightarrow} 0$ and hence $\sum_{m=1}^{\infty}|y_m|^2 < \infty$. Note that for $|y| < 1/2$:

$$\log(1 + y) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{y^k}{k} \geq y - \frac{y^2}{2}\left(\sum_{k=0}^{\infty}2^{-k}\right) = y - y^2$$

Also note trivially that $\log(1+y) \leq y$. Hence for large enough $N$, and $m \geq N$, $|y_m| < 1/2$, so:

$$\sum_{k=N}^{\infty}(y_k - y_k^2) \leq \sum_{k=N}^{\infty}\log(1 + y_k) \leq \sum_{k=N}^{\infty}|y_k|$$

taking the limit as $N \to \infty$ we see that $\sum_{m=k}^{\infty}\log(1 + y_m) \overset{k\to\infty}{\longrightarrow} 0$ and so we are done.

**Problem 35**
Let $X_n$ and $Y_n$ be positive integrable and adapted to $\mathcal{F}_n$. Suppose

$$E(X_{n+1}|\mathcal{F}_n) \leq (1 + Y_n)X_n$$

with $\sum Y_n < \infty$. a.s. Prove that $X_n$ converges a.s. to a finite limit by finding a closely related supermartingale.

**Solution**

From the previous problem, $\prod_{m=1}^{\infty}(1+Y_m)$ exists. Let

$$Z_n = \frac{X_n}{\prod_{m=1}^{n-1}(1+Y_m)}$$

Note that just as in problem 1,

$$E|X_n| \leq E\left[\left|X_1 \prod_{m=1}^{n-1}(1+Y_m)\right|\right]$$

Hence $E|Z_n| \leq E|X_n| < \infty$. Since $Z_n \in \mathcal{F}_n$:

$$E[Z_{n+1}|\mathcal{F}_n] \leq \frac{E[X_{n+1}|\mathcal{F}_n]}{\prod_{m=1}^{n}(1+Y_m)} \leq \frac{X_n}{\prod_{m=1}^{n-1}(1+Y_m)} = Z_n$$

And since $Y_m, X_n$ are positive, $Z_n$ is a positive submartingale. Hence theorem 5.2.9 implies that $Z_n \overset{n\to\infty}{\longrightarrow} Z_\infty$ a.s. Now since $\prod_{m=1}^{\infty}(1+Y_m)$ exists, we find

$$X_n = Z_n \prod_{m=1}^{n-1}(1+Y_m) \overset{n\to\infty}{\longrightarrow} Z_\infty \prod_{m=1}^{\infty}(1+Y_m) \quad \text{a.s.}$$

**Problem 36**

Use the random walks in problem 30 to conclude that in $d \leq 2$, nonnegative superharmonic functions must be constant. The example $f(x) = |x|^{2-d}$ show that this is false in $d > 2$.

**Solution**

From problem 30 above, we showed that $X_n$ is a submartingale. Since $f$ is non-negative, $X_n = f(S_n) \geq 0$. Hence by theorem 5.2.9, $X_n$ converges to some $X_\infty$ almost surely. Since $f$ is superharmonic, it is continuous. Suppose $f$ is not constant. Then there exists constants $a < b$ such that $A := \{f < a\}, B := \{f > b\}$ are non-empty. Since $S_n$ has finite variance and mean of $x = S_0$, and $d \leq 2$, by theorem 4.2.7 (for $d = 1$) aor theorem 4.2.8 (for $d = 2$), $S_n$ visits $A$ and $B$ infinitely often. Hence $\liminf f(S_n) \leq a < b \leq \limsup f(S_n)$ which contradicts that $f$ is superharmonic i.e. that the second derivative vanishes.

**Problem 37**

Let $X_n$ be a martingale atabped to $\mathcal{F}_n$. Let $\mathcal{F}_\infty$ be the sigma field generated by the union of $\mathcal{F}_n$. Show that the following are equivalent:
1. $\{X_n\}$ is uniformly integrable
2. $X_n$ converges in $L^1$ to $X$
3. $X_n$ converges in $L^1$ and almost surely to $X$ 4. $X_n = E(X|\mathcal{F}_n)$ for some random variable $X \in \mathcal{F}_\infty$.

For the last three, show that the $X$ have to be the same.

**Solution**

($1 \Rightarrow 3$): $\{X_n\}$ uniformly integrable gives $\sup E|X_n| < \infty$ for all $n$, hence by martingale convergence theorem, $X_n \to X$ almost surely. Theorem 5.5.2 implies $L^1$ convergence.
($3 \Rightarrow 2$): trivial
($2 \Rightarrow 4$): This follows immediately from Lemma 5.5.5
($4 \Rightarrow 1$): This follows immediately from theorem 5.5.1

If $X_n$ converges in $L^1$ to $X_\infty$, then there's a subsequence that converges a.s. to $\tilde{X}$, but since every subsequence converges to $X_\infty$, $X_\infty = \tilde{X}$ a.s. and in $L^1$. Finally, from theorem 5.5.7, $X_n = [\tilde{X}_\infty|\mathcal{F}_n] \to [\tilde{X}_\infty|\mathcal{F}_\infty] = \tilde{X}_\infty$ a.s., and in $L^1$. But $X_n \to X_\infty$ a.s. and in $L^1$. Hence $X_\infty, \tilde{X}$, and $\tilde{X}_\infty$ are the same.

**Problem 38**

Let $Z_1, Z_2, ..$ be i.i.d. with $E|Z_i| < \infty$, let $\theta$ be an independent random variable with finite mean, and let $Y_i = Z_i + \theta$. If $Z_i$ is normal$(0,1)$ then in statistical terms, we have a sample from a normal population with variance 1 and unknown mean. The distribution of $\theta$ is called the prior distribution, and $P(\theta \in \cdot|Y_1, ..., Y_n)$ is called the posterior distribution after $n$ observations. Show that $E(\theta|Y_1, ..., Y_n) \to \theta$ a.s.

**Solution**

Let $\mathcal{F}_n = \sigma(Y_1, ..., Y_n)$ and $\mathcal{F}_\infty = \sigma(\cup \mathcal{F}_n)$. Note that by the SLLN,

$$\frac{1}{n}\sum_{k=1}^{n} Y_k = \theta + \frac{1}{n}\sum_{k=1}^{n} Z_k \overset{n\to\infty}{\longrightarrow} \theta + 0$$

Hence $\theta \in \mathcal{F}_\infty$. Then by theorem 5.5.7:

$$E[\theta|\mathcal{F}_n] \to E[\theta|\mathcal{F}_\infty] = \theta$$

**Problem 40**

Let $\Omega = [0,1), I_{k,n} = [k2^{-n}, (k+1)2^{-n})$, and $\mathcal{F}_n = \sigma(I_{k,n} : 0 \le k < 2^n)$. $f$ is said to be *Lipschitzcontinuous* if $|f(t) - f(s)| \le K|t-s|$ for $0 \le s, t < 1$. Show that $X_n = (f((k+1)2^{-n}) - f(k2^{-n}))/2^{-n}$ on $I_{k,n}$ defines a martingale, $X_n \to X_\infty$ a.s. and $L^1$, and

$$f(b) - f(a) = \int_a^b X_\infty(\omega) d\omega$$

**Solution**

Let $f_{k,n} := (f((k+1)2^{-n}) - f(k2^{-n}))/2^{-n}$ so that $X_n = f_{k,n}$ on $I_{k,n}$. The trick is to note that $I_{k,n} = I_{2k,n+1} \cup I_{2k+1,n=1}$. and $P(I_{2k,n+1}) = P(I_{2k+1,n=1}) = P(I_{k,n})/2$. Hence

$$E(X_{n+1}|\mathcal{F}_n) = \frac{f_{2k,n+1} + f_{2k+1,n+1}}{2} = f_{k,n} = X_n$$

And so $X_n$ is a martingale. Next, since $t, s \in [0,1)$, $|t-s| < 1$ and $|f(t)-f(s)| \le K|t-s| < K$. Hence $0 \le |X_n| \le K$ for all $n$. Therefore $X_n$ is uniformly integrable and therefore converges to $X_\infty$ a.s. and in $L^1$. Finally, note that lemma 5.5.5 gives us that $E(X_\infty|\mathcal{F}_n) = X_n$ for all $n$. Hence since $I_{k,n} \in \mathcal{F}_n$, for $a = k2^{-n}$ and $b = (k+1)2^{-n}$:

$$\int_a^b X_\infty = \int_a^b E(X_\infty|\mathcal{F}_n) = \int_a^b X_n = \int_a^b X_n = f(b) - f(a)$$

By combining integrals over these sets, we can obtain this results for $a = k2^{-n}$ and $b = j2^{-n}$ for $0 \le k < j \le 2^n$. Now for any $a, b$ we can take sequences $\{k_i\}$, $\{n_i\}$, $\{j_i\}$ such that $k_i 2^{-n_i} \to a$ and $j_i 2^{-n_i} \to b$ and use the continuity of $f$ (since $f$ is Lipschitz continuous), to obtain the result for all $a$ and $b$.

**Problem 41**

Let $\Omega = [0,1), I_{k,n} = [k2^{-n}, (k+1)2^{-n})$, and $\mathcal{F}_n = \sigma(I_{k,n} : 0 \le k < 2^n)$.Suppose $f$ is integrable on $[0,1)$. $E(f|\mathcal{F}_n)$ is a step function and $\to f$ in $L^1$. From this it follows immediately that if $\epsilon > 0$, there is a step function $g$ on $[0,1]$ with $\int |f-g| dx < \epsilon$.

**Solution**

$E[f|\mathcal{F}_n]$ is a step function by example 5.1.3 since $I_{k,n}$ are disjoint for $o \le k < 2^n$. Since $E[f|\mathcal{F}_n]$ is uniformly integrable, it converges in $L^1$ to $E[f|\mathcal{F}_\infty] = f$ since $f \in \mathcal{F}_\infty$.

**Problem 42 (exercise 5.5.6 in Durret)**

Let $Z_n$ be a branching process with offspring distribution $p_k$ (see end of section 5.3 for definitions). Use exercise 5.5.5 in Durret to show that if $p_0 > 0$ then $P(lim_n Z_n = 0 or \infty) = 1$.

**Solution**

Suppose $p_0 > 0$, then $P(Z_{n+1} = 0|Z_1, ..., Z_n) \ge p_0^k > 0$ on $\{Z_n \le k\}$. Note that if $Z_N = 0$ for some $N$, then $Z_k = 0$ for all $k > N$. Hence $\{Z_n = 0 \text{ for some } n \ge 1\} = \{\lim_n Z_n = 0\}$. Hence

$$P(\{\lim_n Z_n = 0\} \cup \{\lim_n Z_n = \infty\}) = 1$$

By Durrett exercise 5.5.5.

**Problem 43**

Show that if $\mathcal{F}_n \uparrow \mathcal{F}_\infty$ and $Y_n \to Y$ in $L^1$ then $E(Y_n|\mathcal{F}_n) \to E(Y|\mathcal{F}_\infty$ in $L^1$.

**Solution**

First note that

$$E(|E[Y_n|\mathcal{F}_n] - E[Y|\mathcal{F}_n]|) = E(|E[Y_n - Y|\mathcal{F}_n]|) \le E(E(|Y_n - Y| \,|\mathcal{F}_n)) = E(|Y_n - Y|) \to 0$$

By Jensen's and since $Y_n \to Y$ in $L^1$. Also Note that

$$E(|E[Y|\mathcal{F}_n] - E[Y|\mathcal{F}]|) \to 0$$

By theorem 5.5.7. Hence By the triangle inequality:

$$E(|E[Y_n|\mathcal{F}_n] - E[Y|\mathcal{F}_\infty]|) \le E(|E[Y_n|\mathcal{F}_n] - E[Y|\mathcal{F}_n]|) + E(|E[Y|\mathcal{F}_n] - E[Y|\mathcal{F}]|) \to 0$$

**Problem 44**

Let $T$ be a stopping time so that for some $b, \epsilon > 0$ and every $n$ we have $P(T \leq n+b|\mathcal{F}_n) > \epsilon$. Show that $ET < \infty$.

**Solution**

First note that for all $n$, $\epsilon < P(T \leq n+b|\mathcal{F}_n) = E(1_{\{T\leq n+b\}}|\mathcal{F}_n)$. Now since $\{T > n\} = \{T \leq n\}^c \in \mathcal{F}_n$, for all $n$ we have:

$$\begin{aligned} p(n < T \leq n+b) &= \int_{\{n<T\leq n+b\}} 1 dP = \int_{\{n<T\}} 1_{\{T\leq n+b\}} dP \\ &= \int_{\{n<T\}} E(1_{\{T\leq n+b\}}|\mathcal{F}_n) dP \\ &= \int_{\{n<T\}} P(T \leq n+b|\mathcal{F}_n) dP \\ &\geq \epsilon P(T > n) \end{aligned}$$

So

$$P(T > n) \leq \frac{1}{\epsilon} P(n < T \leq n+b)$$

Hence since $T > 0$

$$\begin{aligned} ET = E|T| = \sum_{n=0}^{\infty} P(T > n) \leq \frac{1}{\epsilon} \sum_{n=0}^{\infty} P(n < T \leq n+b) \leq \frac{b}{\epsilon} \sum_{n=0}^{\infty} P(n < T \leq n+1) \\ = \frac{b}{\epsilon} \sum_{n=0}^{\infty} P(T = n+1) \leq \frac{b}{\epsilon} < \infty \end{aligned}$$

**Problem 45**

Prove that if $\{X_i\}_{i\in I}$ are uniformly integrable and so are $\{Y_j\}_{j\in j}$, then so are $\{X_i + Y_j\}_{i\in I, j\in j}$

**Solution**

We want to show that for all $\epsilon > 0$ there exists $M > 0$ such that

$$\sup_{(i,j)\in I\times J} E(|X_i + Y_j|; |X_i + Y_j| > M) < \epsilon$$

Since $\{X_i\}_{i\in I}$ and $\{Y_j\}_{j\in J}$ are uniformly integrable, for any $\epsilon_0 > 0$ there exists $M_0 > 0$ such that

$$\sup_{i\in I} E(|X_i|; |X_i| \geq M_0/2) < \epsilon_0, \quad \sup_{j\in J} E(|Y_j|; |Y_j| \geq M_0/2) < \epsilon_0$$

Note that

$$E(\max(|X_i|, |Y_j|); |X_i| \geq M_0/2) < 2\epsilon_0, \quad E(\max(|X_i|, |Y_j|); |Y_j| \geq M_0/2) < 2\epsilon_0$$

Hence

$$\begin{aligned} E(|X_i + Y_j|; |X_i + Y_j| > M_0) &\leq E(2\max(|X_i|, |Y_j|); |X_i + Y_j| > M_0) \\ &\leq 2E(\max(|X_i|, |Y_j|); \{|X_i| \geq M_0/2\} \cup \{|Y_j| \geq M_0/2\}) \\ &\leq 2E(\max(|X_i|, |Y_j|); |X_i| \geq M_0/2) + 2E(\max(|X_i|, |Y_j|); |Y_j| \geq M_0/2) \\ &\leq 4\epsilon_0 + 4\epsilon_0 = 8\epsilon_0 \end{aligned}$$

Therefore given any $\epsilon > 0$, choose $\epsilon_0 = \epsilon/8$ and follow the above recipe.

**Problem 46**

Show that if $Y$ is a random variable with values in $[-c, c]$ and $EY = 0$, then $Ee^{\lambda Y} \leq \cosh(\lambda c) \leq e^{\lambda^2 c^2/2}$.

**Solution**

Since $Y$ ranges from $-c$ to $c$, so does

$$Y = \left(\frac{Y+c}{2c}\right)c + \left(1 - \frac{Y+c}{2c}\right)(-c)$$

Also note that $\frac{Y+c}{2c}$ ranges from 0 to 1. Hence by convexity of the exponential and the fact that $E(Y) = 0$:

$$\begin{aligned} Ee^{\lambda Y} &\leq E\left[\left(\frac{Y+c}{2c}\right)e^{\lambda c} + \left(1 - \frac{Y+c}{2c}\right)e^{-\lambda c}\right] \\ &= \frac{1}{2}e^{\lambda c} + e^{-\lambda c} - \frac{1}{2}e^{-\lambda c} = cosh(\lambda c) \end{aligned}$$

To show the other inequality, we expand the Taylor series of both sides.

$$\cosh(\lambda c) = \sum_{n=0}^{\infty} \frac{(\lambda c)^{2n}}{(2n)!} \leq \sum_{n=0}^{\infty} \frac{(\lambda c)^{2n}}{n! 2^n} = \sum_{n=0}^{\infty} \frac{(\frac{\lambda^2 c^2}{2})^n}{n!} = e^{\lambda^2 c^2/2}$$

since $(2n)! \geq n! 2^n$ for all $n \geq 0$.

# 7 Markov Chains

**Problem 46**
Show that a finite aperiodic Markov chain is irreducible if and only if there is $n_0$ so that for al $n \geq n_0$ and $x, y \in S$ we have $p^n(x, y) > 0$.

**Solution**
Suppose there exists $n_0$ such that for all $n \geq n_0$, and $x, y \in S$, we have $p^n(x, y) > 0$. Hence for each $x, y \in S$, $\rho_{x,y} = P_x(T_y < \infty) > 0$. Hence the Markov chain is irreducible. Now suppose that the Markov chain is irreducible. Given $x, y \in S$, we have $\rho_{x,y} = P_x(T_y < \infty) > 0$, hence there exists $n_{x,y}$ such that $p^{n_{x,y}}(x, y) > 0$. Note that since the Markov chain is aperiodic, for any $m > 0$, $y \in S$, we have $p^m(y, y) > 0$. Hence for $N = m + n_{x,y} > n_{x,y}$, we have that $p^N(x, y) = p^{m+n_{x,y}}(x, y) \geq p^{n_{x,y}}(x, y)p^m(y, y) > 0$. Therefore $p^N(x, y) > 0$ for all $N > n_{x,y}$. Since the Markov chain is finite, we can let $n_0 = \max_{x,y \in S}\{n_{x,y}\}$.

**Problem 47**
Let $X_n$ be a finite irreducible Markov chain started at $x$. Let $T_0 = 0$ and let $T_i$ be the $i$th positive time visiting $x$.
(a) Show that the finite sequences $S_i = (X_{T_i}, X_{T_i+1}, ..., X_{T_{i+1}-1}$ are i.i.d. as $i$ varies.
(b) Let $N_t^y$ be the number of visits to $y$ by time $t$. Show that for all $y$ we have $N_t^y/t$ converges a.s.
(c) Show that $(N_t^y - m_t)/\sigma_t$ converges to a standard normal random variable in distribution for the right choice of $m_t$ and $\sigma_t$.

**Solution (a)**
Since $X_n$ is finite irreducible, it is recurrent. First note that $S_i = S_1 \circ \theta_{T_{i-1}}$. Since $X(T_{k-1}) = x$ a.s. Hence

$$P_x(S_1 \circ \theta_{T_{k-1}} = S_i | \mathcal{F}_{T_{i-1}}) = P_x(S_1 = S_i)$$

Hence $S_i$ is independent of $\mathcal{F}_{T_{i-1}}$ and hence $S_1, ..., S_{i-1}$. The result follows by induction on $i$.

**(b)**
By part a, $X_n$ is recurrent. Hence the result follows directly from Theorem 6.6.1 in Durrett.

**(c)**
Let $R_k = T_k - T_{k-1}$, $R_0 = 0$. Then

$$\frac{T_k}{N_{T_k}^y} = \frac{T_k}{k} = \frac{R_1 + ... + R_k}{k} \to ER_1$$

By the strong law of large numbers. For $T_{k-1} \leq t < T_k$, we have:

$$\frac{1}{ER_1} \leftarrow \frac{k}{k-1}\frac{k-1}{T_{k-1}} = \frac{k}{T_{k-1}} \geq \frac{N_t^y}{t} \geq \frac{k}{T_k} \to \frac{1}{ER_1}$$

Hence $T_t/t \to ER_1$. So if $T_k \leq t < T_{k+1}$, by CLT:

$$\frac{N_t^y - \frac{R_1+...+R_{N_t^y}}{ER_1}}{\frac{\sigma(R_1)\sqrt{N_t^y}}{ER_1}} = \frac{N_{T_k}^y - \frac{R_1+...+R_k}{ER_1}}{\frac{\sigma(R_1)\sqrt{k}}{ER_1}} = \frac{k - \frac{R_1+...+R_k}{ER_1}}{\frac{\sigma(R_1)\sqrt{k}}{ER_1}} = \frac{kER_1 - R_1... - R_k}{\sigma(R_1)\sqrt{k}} \Rightarrow \mathcal{N}(0,1)$$

**Problem 48**

Let $D$ be a finite set of size $n$ and let $Y_i$ be independent uniform random variables taking values in $D$. Let $N_t$ be the number of different values of $Y_i$ up to time $t$, and let $T_i$ be so that $T_1 + ... + T_k$ is the first time $t$ so that $N_t = k$.

(a) Show that $T_k$ are independent and $T_k$ has Geometric distribution with parameter $1/k$

(b) Let $\alpha \in (0,1]$. Compute the asymptotics of the mean and variance of $T_1 + ... + T_{n-\lfloor n^\alpha \rfloor}$ as $n \to \infty$

(c) Now let $X_n$ be a lazy random walk on the d-dimensional hypercube started at $(0,...,0)$. Show that there exists $\epsilon > 0$ so that the random vector $X_{\lfloor \epsilon d \log d \rfloor}$ has at least $d^{2/3}$ zero entries with probability tending to 1 as $d \to \infty$

(d) Show that the mixing time of $X_n$ is at least $\epsilon d \log d$ for large enough d.

**Solution (a)**

If $N_t = k$ then the there are $n-k$ values left in our set. So our chance of success on the first (and each) attempt is $\frac{n-k+1}{n}$. Since the attempts are independent, we have a geometric distribution and the $T_k$ are independent.

**(b)**

Since $T_k$ has a geometric distribution, $ET_1 = 1$, $ET_2 = \frac{1}{(n-1)/n} = \frac{n}{n-1}$, ... , $ET_{n-\lfloor n^\alpha \rfloor} = \frac{1}{(n-(n-\lfloor n^\alpha \rfloor)+1)/n} = \frac{n}{1+\lfloor n^\alpha \rfloor}$. Hence

$$\begin{aligned}
\sum_{k=1}^{n-\lfloor n^\alpha \rfloor} ET_k &= 1 + \frac{n}{n-1} + ... + \frac{n}{1+\lfloor n^\alpha \rfloor} \\
&= n\left(\frac{1}{n} + \frac{1}{n-1} + ... + \frac{1}{1+\lfloor n^\alpha \rfloor}\right) \\
&= n\left(1 + \frac{1}{2} + ... + \frac{1}{n}\right) - n\left(1 + \frac{1}{2} + ... + \frac{1}{1+\lfloor n^\alpha \rfloor}\right) \\
&\sim n\log(n) - n\log(1+\lfloor n^\alpha \rfloor) \\
&= n\log(\frac{n}{1+\lfloor n^\alpha \rfloor}) \\
&\sim n\log(\frac{n}{\lfloor n^\alpha \rfloor}) \\
&\sim n(1-\alpha)\log(n)
\end{aligned}$$

Similarly, the variance of a geometric distribution with parameter $p$ is $(1-p)/p^2$. Hence

$$\begin{aligned}
Var(T_1 + ... + T_{n-\lfloor n^\alpha \rfloor}) &= Var(T_1) + ... + Var(T_{n-\lfloor n^\alpha \rfloor}) \\
&= 0 + n \sum_{k=2}^{n-\lfloor n^\alpha \rfloor} \frac{(k-1)}{(n-k+1)^2} \\
&\sim n \int_2^{n-\lfloor n^\alpha \rfloor} \frac{x-1}{(n-x+1)^2} dx \sim n \int_2^{n-\lfloor n^\alpha \rfloor} \frac{x}{(n-x)^2} dx \\
&= n \int_{n-2}^{(\lfloor n^\alpha \rfloor} \frac{x-n}{x^2} dx \\
&= n \left[ \ln(\lfloor n^\alpha \rfloor) - \ln(n-2) + n \left( \frac{1}{\lfloor n^\alpha \rfloor} - \frac{1}{n-2} \right) \right] \\
&\sim \frac{n^2}{\lfloor n^\alpha \rfloor} \sim n^{2-\alpha}
\end{aligned}$$

**(c)**

Let $Y_n = T_1 + ... + T_{n-\lfloor n^\alpha \rfloor}$, $\epsilon = 1/6$, $\alpha = 2/3$. Then by part b, for large $d$, $E(Y_d) \geq \frac{d}{6} \log d$ and $Var(Y_d) \leq 2d^{2-2/3}$. Hence for large $d$, we have:

$$\begin{aligned}
P(Y_d \leq \frac{d}{6} \log d) &\leq P(|Y_d - EY_d| \leq \frac{d}{6} \log d - EY_d) \\
&= P(|Y_d - EY_d| \geq (1/3 - 1/6) d \log d) = P(|Y_d - EY_d| \geq \frac{d}{6} \log d)
\end{aligned}$$

Hence by Chebyshev's inequality:

$$\begin{aligned}
P(Y_d \leq \frac{d}{6} \log d) &\leq P(|Y_d - EY_d| \geq \frac{d}{6} \log d) \\
&\leq \frac{Var Y_d}{(\frac{d}{6} \log d)^2} \leq \frac{100}{d^{2/3} (\log d)^2} \to 0
\end{aligned}$$

Hence $P(Y_d \geq \frac{d}{6} \log d) \to 1$ as $d \to \infty$.

**(d)**

Since the stationary distribution is clearly uniform, we expect half of the entries to be 0. But for large $d$, we expect to have at least $d^{2/3}$ zero entries at time $\epsilon d \log d$. Hence the expected total variation is bounded by $\left| \frac{d^{2/3} - d/2}{d} \right| = |d^{-1/3} - 1/2| \to 1/2$ and so the mixing time is at least $\epsilon d \log d$ since the value above does not converge to a time less than half.

**Problem 49**

Let $X_t$ be a biased lazy random walk on the $n$-cycle. Namely, with probability $1/2$ $X_t$ stays where it is, with probability $p/2$ it moves to the right and otherwise it moves to the left. Assume $p \in (1/2, 3/4)$. Show that the mixing time of $X_t$ is at most $100n^2$.

**Solution**

Since I missed Lecture, I will use a similar resource to the lecture notes (Markov Chains and Mixing Times by Levin, Peres, Wilmer). Consider the lazy random walk on $\{1, ..., n\}$. where we have 1/2 chance of staying in the same place. Let $\tau$ be the hitting time of 1 or $n$. Then by proposition 2.1 in Levin,

$$E_k(\tau) \leq k(n-k) \leq \frac{n^2}{4}$$

since the chance to step in any direction is less than half (i.e. that of a simple random walk). Note that by differentiating, the middle part above is maximized when $k = n/2$ hence we get the right hand term. Now we just construct a coupling $(X_t, Y_t)$ of two particles performing the lazy biased random walks on the n-cycle. They must not take steps at the same time so that they don't jump over each other. When they meet, they make identical moves. They way they move could be dictated by the following rule: a fair coin is flipped. If the result is heads, the biased lazy random walk on $X_t$ is iterated by one step. Otherwise the same is done for $Y_t$. If we let $D_t$ be the clockwise distance of the particles and $\tau$ be the hitting time of this random walk to the points 0 or $n$ (i.e. when the points collide), just as in example 5.3.1 in Levin, using the above and corollary 5.3 in Levin, we obtain:

$$d(t) \leq \max_{x,y \in \mathbb{Z}_n} P(\tau > t) \leq \frac{\max_{x,y} E_{|x-y|}(\tau)}{t} \leq \frac{n^2}{4t}$$

The right side is equal to 1/4 for $t = n^2$. Hence $t_{mix} \leq n^2$.

**Problem 50**

Let $\xi_1, \xi_2, ...$ be i.i.d. $\in \{1, 2, .., N\}$ and taking each value with probability $1/N$. Show that $X_n = |\{\xi_1, ..., \xi_n\}|$ is a markov chain and compute its transition probability.

**Solution**

The probability of adding a new value (at time $n+1$ depends on the number of values we have seen at time $n$. Hence $X_n$ is a Markov chain. The transition probabilities are $p(j, j) = j/N$ since there is a $j/N$ chance that $\xi_{N+1} \leq j$, $p(j, j+1) = 1 - j/N$, $p(i, j) = 0$ otherwise.

**Problem 51 (Durrett Exercise 6.3.6)**

Let $h(x) = P_x(\tau_A < \tau_B)$. Suppose $A\hat{B} = \emptyset$, $S - (A \cup B)$ is finite, and $P_x(\tau_{A\cup B} < \infty) > 0$ for all $x \in S - (A \cup B)$.
(i) Show that

$$h(x) = \sum_y p(x, y)h(y) \text{ for } x \notin A \cup B$$

(ii) Show that if $h$ satisfies the equation above, then $h(X(n \wedge \tau_{A\cup B}))$ is a martingale
(iii) use this and exercise 6.3.5 in Durret to conclude that $h(x) = P_x(\tau_A < \tau_B)$ is the only solution of the above equality that is 1 on $A$ and 0 on $B$.

**Solution**

We take the expected value for the case when $x \notin A \cup B$ (and therefore $1_{(\tau_A < \tau_b)} \circ \theta_1 = 1_{(\tau_A < \tau_b)}$), we get:

$$P_x(\tau_A < \tau_B) = E_x(1_{(\tau_A<\tau_b)} \circ \theta_1) = E_x(E_x(1_{(\tau_A<\tau_b)} \circ \theta_1|\mathcal{F}_1)) = E_x h(X_1)$$

**(ii)**
Let $N = \tau_{A\cup B}$. Note that $X_{(n+1)\wedge N} = X_{n\wedge N}$ on $\{N \le n\} \in \mathcal{F}_n$. Hence by exercise 1.1 in ch.5,

$$Eh(X_{n+1\wedge N}|\mathcal{F}_n) = E(h(X_{n\wedge N}|\mathcal{F}_n) = h(X_{n\wedge N})$$

Note that $X_{(n+1)\wedge N} = X_{n+1}$ on $\{N > n\} \in \mathcal{F}_n$. Hence by the Markov property, part (i) and exercise 1.1 in ch.5,

$$\begin{aligned} Eh(X_{n\wedge N}|\mathcal{F}_n) &= E(h(X_{n+1}|\mathcal{F}_n) = E(h(X_1 \circ \theta_n|\mathcal{F}_n) \\ &= E_{X_n} h(X_1) = h(X_n) = h(X_{n\wedge N}) \end{aligned}$$

**(iii)**
By exercise 6.3.5, $N < \infty$ a.s. Also note that $h$ is bounded since $S - (A \cup B)$ is finite. Hence by the bounded convergence theorem and Martingale property:

$$h(x) = E_x h(X_{n\wedge N}) \to E_X h(X_N) = P_x(\tau_A < \tau_B)$$

**Problem 52 (Durrett Exercise 6.3.10)**
Let $\tau_A = \inf\{n \ge 0 : X_n \in A\}$ and $g(x) = E_x\tau_A$. Suppose that $S - A$ is finite and for each $x \in S - A$, $P_x(\tau_A < \infty) > 0$.
(i) Show that

$$g(x) = 1 + \sum_y p(x,y)g(y) \text{ for} x \notin A$$

(ii) Show that if $g$ satisfies the above equality, $g(X(n \wedge \tau_A)) + n \wedge \tau_A$ is a martingale
(iii) Use this to conclude that $g(x) = E_x\tau_A$ is the only solution to the above equality that is 0 on $A$.

**Solution (i)**
Note that $\tau_A \circ \theta_1 = \tau_A - 1$ if $x \notin A$. Just as in the above, we take expected value:

$$g(x) - 1) = E_x(\tau_A - 1) = E_x(\tau_A \circ \theta_1) = E_x E_x(\tau_A \circ \theta_1|\mathcal{F}_1) = E_x g(X_1)$$

**(ii)**
Just as in the previous problem, we consider the two sets. First, on $\{\tau_A \le n\} \in \mathcal{F}_n$, we have that

$$g(X_{n+1\wedge\tau_A}) + (n+1) \wedge \tau_A = g(X_{n\wedge\tau_A}) + n \wedge \tau_A$$

so by exercise 1.1 in ch.5,

$$E_x(g(X_{n+1\wedge\tau_A}) + (n+1) \wedge \tau_A|\mathcal{F}_n) = E_x(g(X_{n\wedge\tau_A}) + n \wedge \tau|\mathcal{F}_n) = g(X_{n\wedge\tau_A}) + n \wedge \tau_A$$

On $\{\tau_A > n\} \in \mathcal{F}_n$, we have that

$$g(X_{n+1\wedge\tau_A}) + (n+1) \wedge \tau_A = g(X_{n+1}) + n + 1$$

so by exercise 1.1 in ch.5 and part (i):

$$\begin{aligned} E_x(g(X_{n+1\wedge\tau_A}) + (n+1) \wedge \tau_A|\mathcal{F}_n) &= E_x(g(X_{n+1}) + n + 1|\mathcal{F}_n) = g(X_n) - 1 + n + 1 \\ &= g(X_n) + n \end{aligned}$$

**(iii)**

By exercise 6.3.5, $P_y(\tau_A > kN) \leq (1-\epsilon)^k \; \forall y \notin A$ so that $E_y\tau_A < \infty$. Just as in the previous question, any solution is bounded since $S - A$ is finite. Hence by monotone, bounded convergence theorems and the martingale property:

$$g(x) = E_x(g(X_{n\wedge\tau_A}) + n \wedge \tau_A) \to E_x\tau_A$$

**Problem 53 (Durrett Exercise 6.4.9)**

$f$ is said to be superharmonic if $f(x) \geq \sum_y p(x,y)f(y)$, or equivalently if $f(X_n$ is a supermartingal. Suppose $p$ is irreducible. Show that if $p$ is recurrent if and only if every nonnegative superharmonic function is constant.

**Solution**

We start by using a contrapositive arguement. Suppose the chain is transient. Hence for some x. $P_y(T_y < \infty) < 1$ and $f(x) = P_x(T_y < \infty)$ is a non constant superharmonic function. Now suppose $f \geq 0$ is superharmonic. Let $f(X_n) = Y_n$. Then $Y_n$ is a supermartingale, so $Y_n$ converges to a limit $Y$ a.s. If $X_n$ is recurrent then for any x, $X_n = x$ infinitely often, hence $f(x) = Y$ and this function is a constant.

# 8 Brownian Motion

**Problem 54**
Show that for a Brownian motion B
(a) for all $t \geq 0$ we have $P(t \text{is a local maximum}) = 0$
(b) almost surely local maxima exist
(c) almost surely, there exist times $t_1, t_2 \in (0,1)$, $D^*B(t) \geq 0$ and $D_*B(t) \leq 0$. Here

$$D^*f(t) = \limsup_{h\downarrow 0} \frac{f(t+h) - f(t)}{h}$$

and $D_8$ is the same with lim inf.

**Solution**
By the discussion in page 13 of Morters and Peres: Browniaon Motion, it suffices to consider a standard Brownian motion $B$. For $t \geq 0$, let $Y_{t_0}(t) = B(t+t_0) - B(t_0)$ and let $X_{t_0}(t) = tY(\frac{1}{t})$. Then by theorem 2.3 (Markov property) and theorem 1.9 (time), $Y_{t_0}, X_{t_0}$ are standard Brownian motions. Now suppose $t_0$ is a local max of $B$. Then for sufficiently small $t > 0$, $Y_{t_0}(t) \leq 0$ and hence for sufficiently large $t > 0$, $X_{t_0}(t) \leq 0$ and so $\limsup_{t\to\infty} X_{t_0}(t) \leq 0$. So $\{t_0 : t_0 \text{ is a local max for } B\} \subset \{\limsup_{t\to\infty} X_{t_0}(t) \leq 0\}$. But by proposition 1.23, $\limsup_{t\to\infty} X_{t_0}(t) = +\infty$ almost surely. Hence

$$P\{t_0 : t_0 \text{ is a local max for } B\}) \leq P(\{\limsup_{t\to\infty} X_{t_0}(t) \leq 0\}) = 0$$

**(b)**
Consider the interval $I = (0,1)$. We claim that there exists a local maximum in $I$. If this were not the case, then either $B$ is monotone increasing on $I$, or there exists a $t_0 \in I$ such that $B$ is monotone decreasing on $(0, t_0)$ and increasing on $(t_0, 1)$. However, since these are open intervals, we can fit non-degenerate closed interval in each one. But by proposition 1.22, almost surely, $B$ is not monotone on any closed interval. Hence the probability of both the cases above is 0, and so almost surely, $B$ has a local max.

**(c)**
By a symmetric argument as in part (b), almost surely, $B$ has a local min on $(0,1)$. Let $t_0, t_1 \in (0,1)$ be a local max and min of $B$ respectively. Then for sufficiently small $h > 0$, $B(t_0 + h) - B(t_0) \leq 0$ and $B(t_1 + h) - B(t_1) \geq 0$ Hence:

$$D^*B(t_0) = \limsup_{h\downarrow 0} \frac{B(t_0+h) - B(t_0)}{h} \leq 0, \quad D_*B(t_1) = \liminf_{h\downarrow 0} \frac{B(t_1+h) - B(t_1)}{h} \geq 0$$

**Problem 55**
Show that, for every point $x \in \mathbb{R}$, there exists a two-sided Brownian motion $(B(t) : t \in \mathbb{R})$ with $B(0) = x$ which has continuous paths, independent increments, and the property that

for all $t \in \mathbb{R}$ and $h > 0$, the increments $B(t+h) - B(t)$ are normally distributed with expectation zero and variance $h$.

**Solution**

Let $B_1, B_2$ be independent standard Brownian motions (so they have the same distribution). Let $\tilde{B_1} = B_1 + x/2, \tilde{B_2} = B_2 + x/2$. Then $\tilde{B_1}, \tilde{B_2}$ are Brownian motions with the same distributions as $B_1$. Define $\tilde{B}_1(t) = \tilde{B}_1(0), \tilde{B}_2(t) = \tilde{B}_2(0)$ for negative $t$. Now let $E(t) = \tilde{B}_2(-t)$. Then $E(-\cdot)$ has the same distribution as $B_1(\cdot)$. Finally, let $B(t) = \tilde{B}_1(t) + E(t)$. Then $B(0) = B_1(0) + x/2 + B_2(0) + x/2 = x$. The continuity of the paths follows from the fact that the sum of 2 continuous paths is continuous. The increments are independent since $\tilde{B}_1$ and $E$ are independent, and they each have independent increments. Finally, note that $h > 0$ so to check the last condition, we have a 3 cases. Case $(i)$ $t \geq 0$. Then

$$\begin{aligned} B(t+h) - B(t) &= \tilde{B}_1(t+h) - \tilde{B}_1(t) + \tilde{B}_2(-t-h) - \tilde{B}_2(-t) \\ &= B_1(t+h) + x/2 - B_1(t) - x/2 + B_2(0) + x/2 - B_2(0) - x/2 \\ &= B_1(t+h) - B_1(t) \end{aligned}$$

which is normally distributed with mean 0 and variance $h$. Next we have case $(ii)$ $t < 0$ and $t + h < 0$. Then

$$\begin{aligned} B(t+h) - B(t) &= \tilde{B}_1(t+h) - \tilde{B}_1(t) + \tilde{B}_2(-t-h) - \tilde{B}_2(-t) \\ &= B_1(0) + x/2 - B_1(0) - x/2 + B_2(t+h) + x/2 - B_2(t) - x/2 \\ &= B_2(t+h) - B_2(t) \end{aligned}$$

which is normally distributed with mean 0 and variance $h$. Next we have case $(iii)$ $t < 0$ and $t + h \geq 0$. Then

$$\begin{aligned} B(t+h) - B(t) &= \tilde{B}_1(t+h) - \tilde{B}_1(t) + \tilde{B}_2(-t-h) - \tilde{B}_2(-t) \\ &= B_1(t+h) + x/2 - B_1(0) - x/2 + B_2(0) + x/2 - B_2(t) - x/2 \\ &= B_1(t+h) - B_2(t) \end{aligned}$$

which is normally distributed with mean 0 and variance $h$ since $B_1$ and $B_2$ have the same distribution and are both Brownian motions.

**Problem 56**

Use the time inversion formula and properties of random walks to show that almost surely $b$ takes on both positive and negative values in every nonempty interval $(0, 1)$

**Solution**

By time inversion $B(t) \stackrel{d}{=} tB(1/t)$. Hence we have $B(t)$ takes positive and negative values for $t \in (0, \epsilon)$ almost surely iff $tB(1/t)$ takes positive and negative values for $t \in (0, \epsilon)$ almost surely. But by corollary 1.11 and Chung-Fuchs theorem, $B(t)$ is recurrent and therefore so is $B(1/t)$ for $t \in (0, \epsilon)$. In particular, $B(1/t)$ visits positive and then negative values infinitely many time for $t \in (0, \epsilon)$. Hence so does $tB(1/t)$ and consequently $B(t)$.

**Problem 57**

Recall that a random vector $X$ has (mean zero) multivariate normal distribution if it can be

written as $MZ$ for $Z$ a vector with independent standard normal entries and $M$ a possibly rectangular deterministic matrix.
(a) Show that the distribution of a standard normal vector is determined by its covariances $EX_iX_j$
(b) Show that if $X, Y$ are independent standard normal then $X-Y$, $X+Y$ are independent normals of variance 2.

**Solution (a)**
Let $X = MZ$ as in the problem where $Z = (Z_1, ..., Z_n)$ with independent standard normal entries (i.e. $Z_i \sim \mathcal{N}(0,1)$). Let $\sigma_i = 1$ be the variance of these entries and $I = \Sigma = E[ZZ^t]$ be the covariance matrix of $Z$ and $\Phi = E[XX^t]$ be the covariance matrix of $X$. Then using our knowledge of the characteristic function:

$$\begin{aligned}\varphi(\lambda) = E[e^{i\lambda X}] = \prod_{k=1}^{n} E[e^{i\lambda_k X_k}] &= e^{\frac{-1}{2}\lambda^t M\Sigma M^t\lambda} \\ = e^{\frac{-1}{2}\lambda^t ME[\Sigma\Sigma^t]M^t\lambda} &= e^{\frac{-1}{2}\lambda^t ME[ZZ^t]M^t\lambda} \\ = e^{\frac{-1}{2}\lambda^t E[MZZ^tM^t]\lambda} &= e^{\frac{-1}{2}\lambda^t\Phi\lambda}\end{aligned}$$

Hence the distribution of $X$ is determined by it's covariance.

**(b)**

$$\begin{aligned}\varphi_{X+Y}(\lambda) = \varphi_X(\lambda)\varphi_Y(\lambda) = e^{-\lambda^2/2}e^{-\lambda^2/2} = e^{-2\lambda^2/2} \\ \varphi_{X-Y}(\lambda) = \varphi_X(\lambda)\varphi_{-Y}(\lambda) = e^{-\lambda^2/2}e^{-(-\lambda)^2/2} = e^{-2\lambda^2/2}\end{aligned}$$

Hence $X+Y$, $X-Y \sim \mathcal{N}(0,2)$. Their independence is a basic lemma in Durrett.

**Problem 58**
(a) Show that for any continuous function $f$ on $(0,\infty)$, if $\lim_{h\downarrow 0} f(h) = 0$ over rational $h$ then the same is true over real $h$
(b) Define $R(t) = tB(1/t)$ for $t > 0$. We have checked that $R(t)$ is continuous on $(0,\infty)$ and has the same distribution as $B$ on this interval. Use part (a) to show that $\lim_{h\to 0} R(h) = 0$

**Solution (a)**
Let $h_n \downarrow 0$ be rational, $k_n \downarrow 0$ be irrational. Suppose $f(h_n) \to 0$ and $f(k_n) \to c \neq 0$. Then for for $\epsilon = c/100$, for any $\delta > 0$, we can choose $N > 0$ s.t. for all $n > 0$, $|h_n - k_n| < \delta$, $|f(h_n) - 0| < c/100$ and $|f(k_n) - c| < c/100$. But now $|f(h_n) - f(k_n)| \geq c - c/50 > \epsilon$ which contradicts continuity of $f$.

**(b)**
Let $h_n \downarrow 0$ be rational just as in part a. Then by page 27 of Morters and Peres, $\lim_{n\to\infty} R(h_n) = 0$ almost surely. And since $R$ is almost surely continuous on $(0,\infty)$, $\lim_{t\to 0} R(t) = 0$ almost surely by part $(a)$.

**Problem 59**
(a) Use the $SMP$ and problem 56 to show that for every $t$ a.s. if $T$ is the first time after $t$

so that $B(T) = 0$, then there are zeros of $B$ in the interval $(T, T + \epsilon)$ for all $\epsilon > 0$.
(b) Let $Z$ be the set of zeros of $BM$. Use part (a) for all rational $t$ to conclude that a.s. every point in $Z$ is a limit point of $Z$. (It follows that $Z$ is uncountable)

**Solution (a)**

By problem 56, $T$ is almost surely finite. Hence by the SMP, $X(t) = \{B(T+t) - B(T) : t \geq 0\}$ is a standard Brownian motion independent of $\mathcal{F}^+(T)$ (defined in Morters and Peres). And so by problem 3 again, there are zeros of $X(t)$ in every interval $t \in (0, \epsilon)$, $\epsilon > 0$. But hence this means There are zeros of $B(T+t) - B(T) = B(T+t) + 0$ in every interval $t \in (0, \epsilon)$, $\epsilon > 0$. So there are zeros of $B(t)$ in every interval $t \in (T, T + \epsilon)$, $\epsilon > 0$.

**(b)**

Suppose $z \in Z$ (so $B(z) = 0$) and suppose $z$ doesn't have a limit from the left (or else we are done already). Then by part $(a)$, there exists $z_n$ such that $B(z_n) = 0$ and $z_n \in (z, z + 1/2^n)$. But now $z_n \to z$ and we are done.