Math 710 Homework

Austin Mohr

June 16, 2012

1. For the following random "experiments", describe the sample space Ω. For each experiment, describe also two subsets (events) that might be of interest, and describe how you might assign probabilities to these events.

   (a) The USC football team will play 12 games this season. The experiment is to observe the Win-Tie-Loss record.

   Solution: Define the sample space Ω to be the set

   $$\{(a_1, \ldots, a_{12}) \mid a_i \in \{\text{"Win", "Tie", "Loss"}\}\},$$

   where each a_i reflects the result of the i^{th} game.

   One interesting event might be the event in which $a_i =$ "Win" for all i, corresponding to an undefeated season. Another interesting event is the set

   $$\{(a_1, \ldots, a_{12}) \mid \exists j \in [12] \text{ such that } a_i = \text{"Loss"} \ \forall i \leq j \text{ and } a_i = \text{"Win"} \ \forall i > j\}.$$

   This event corresponds to all possible seasons in which the Gamecocks lose their first j games (here, j is nonzero), but rally to win the remaining games.

   To assign probabilities to each element of the sample space, we define a probability function p_i for each a_i. This can be accomplished by considering the history of the Gamecocks versus the opposing team in game i and setting

   $$\begin{aligned} p_i(\text{"Win"}) &= \frac{\text{Games won against team } i}{\text{Total games against team } i} \\ p_i(\text{"Tie"}) &= \frac{\text{Games tied against team } i}{\text{Total games against team } i} \\ p_i(\text{"Loss"}) &= \frac{\text{Games lost against team } i}{\text{Total games against team } i}. \end{aligned}$$

   Now, for each elementary event $\omega = (a_1, \ldots, a_{12})$, set

   $$P(\omega) = \prod_{i \in [12]} p_i(a_i).$$

Finally, for any subset A of Ω, define

$$P(A) = \sum_{\omega \in A} P(\omega).$$

(b) Observe the change (in percent) during a trading day of the Dow Jones Industrial Average. Letting X denote the random variable corresponding to this change, we are observing

$$X = 100 \frac{\text{Value at Closing} - \text{Value at Opening}}{\text{Value at Opening}}.$$

Solution: Strictly speaking, X may take on any real value. In the interest of cutting down the sample space somewhat, we may round X to the nearest integer. Thus, $\Omega = \mathbb{Z}$.

One interesting event is $X = 0$, corresponding to no net change for the day. Another interesting event is $X = 100$, corresponding to a doubling in value for the day.

An elementary event corresponds to specifying a single value m for X. A very rough way to define this probability to examine the (rounded) percent change data for all days that the DJIA has been monitored and set

$$P(m) = \frac{\text{Occurrences of } m}{\text{Number of days in data set}}.$$

For an arbitrary subset of $\mathbb{Z}$, we extend linearly, as before.

(c) The DJIA is actually monitored continuously over a trading day. The experiment is to observe the trajectory of values of the DJIA during a trading day.

Solution: Suppose we sample the data every second and compile it into a piecewise linear function f. The trajectory at time t (in seconds after the opening bell) is given by $g(t) = f(t) - f(t-1)$, where we take $g(0) = 0$. As before, g may take on any real value. We may combat this by partitioning the real line into intervals of the form $[x, x+\epsilon)$ for some fixed $\epsilon > 0$. Our elementary events, therefore, are ordered pairs $(t, [x, x+\epsilon))$, corresponding to the trajectory at time t falling into the interval $[x, x+\epsilon)$. The sample space Ω is the collection of all such elementary events.

One interesting (and highly suspicious) event might be $\{(t, [0, \epsilon)) \mid \text{any } t\}$, corresponding to a day in which the DJIA saw nearly no change throughout the day. Anoter interesting event might be

$$\{(t, I) \mid I = \bigcup_{x>0} [x, x+\epsilon), \text{ any } t\},$$

corresponding to the event where the DJIA saw positive trajectory throughout the day.

The probabilities might be assigned as in part b, where we now further divide the data to reflect the value of t. That is, we do not want the probability of seeing a given trajectory, but the probability of seeing a given trajectory at a given time.

(d) Let G be the grid of points $\{(x, y) \mid x, y \in \{-1, 0, 1\}\}$. Consider the experiment where a particle starts at the point $(0, 0)$ and at each timestep the particle either moves (with equal probability) to a point (in G) that is "available" to its right, left, up, or down. The experiment ceases the moment the particle reaches any of the four points $(-1, -1)$, $(-1, 1)$, $(1, -1)$, $(1, 1)$.

Solution: One natural probability to assess is the probability that the experiment ceases after n steps. We note, however, that it is possible (though infinitely-unlikely) that the experiement *never* ceases. Thus, we take the sample space to be $\Omega = \mathbb{Z}^+ \cup \infty$.

One interesting event is that the experiment ceases after exactly 2 steps (the minimum steps required to reach a termination state). Another intesting event is that the experiment takes at least 100 (or any constant number) steps before ceasing.

This problem suggests that an exact solution may be found using Markov chains. Barring that, we might run a computer simulation to gather data. From this data, we can set

$$P(m) = \frac{\text{Number of occurrences of } m}{\text{Total number of trials}},$$

and then extend linearly to more general events.

(e) The experiment is to randomly generate a point on the surface of the unit sphere.

Solution: Given the abstract nature of the problem, we decline to impose any artificial discretization as was done in previous problems. Now, any point in $\mathbb{R}^3$ may be specified by a spherical coordinate (r, θ, ϕ), where r denotes radial distance, θ inclination, and ϕ azimuth. Since we are restricted to the unit sphere, we may discard r and consider ordered pairs (θ, ϕ). Thus,

$$\Omega = \{(0, 0)\} \cup \{(\pi, 0)\} \cup \{(\theta, \phi) \mid \theta \in (0, \pi), \phi \in [0, 2\pi)\}$$

(the restrictions on θ and ϕ are to ensure a unique representation of each point).

One interesting event might be $\{(0, 0)\} \cup \{(\pi, 0)\}$, corresponding to the random point lying at either the north or south pole of the sphere. Another interesting event might be $\{(\frac{\pi}{2}, \phi) \mid \phi \in [0, 2\pi)\}$, corresponding to the random point lying somewhere along the equator.

As a point in the plane has measure zero, we cannot assign probabilities to elementary events and extend. Instead, given a subset A of Ω, we must set $P(A)$ to be the measure of A as a subset of $\mathbb{R}^2$.

2. (Secretary Problem) You have in your possession N balls, each labelled with a distinct symbol. In front of you are N urns that are also labelled with the same symbols as the balls. Your experiment is to place the balls at random into these boxes with each box getting a single ball.

   (a) Write down the sample space Ω of this experiment. How many elements are there in Ω?

   Solution: For simplicity, let the symbols be the first N integers. Thus,

   $$\Omega = \{(a_1, \ldots, a_N) \mid a_i \in [N] \, \forall i\},$$

   where $a_i = j \in [N]$ means that the bucket labelled i received the ball labelled j. Observe that Ω is simply the collection of all permutations of the N distinct objects, so $|\Omega| = N!$.

   (b) What probabilities will you assign to the elementary events in $|\Omega|$?

   Solution: Each elementary event is equally-likely, so $P(\omega) = \frac{1}{N!}$ for all $\omega \in \Omega$.

   (c) Define a match to have occurred in a given box if the ball placed in this box has the same label as the box. Let A_N be the event that there is at least one match. What is the probability of A_N?

   Solution: For each $i \in [N]$, let B_i denote the set of arrangments having a match in bucket i. Thus, $\bigcup_{i\in[N]} B_i$ is the collection of all arrangements having at least one match. By the inclusion-exclusion principle, we have

   $$\begin{aligned}
   \bigcup_{i\in[N]} B_i &= \sum_{i\in[N]} |B_i| - \sum_{\substack{i,j\in[N] \\ i\neq j}} |B_i \cap B_j| + \cdots \\
   &= \binom{N}{1}|B_1| - \binom{N}{2}|B_1 \cap B_2| + \cdots \qquad \text{(since } |B_i| = |B_j| \text{ for all } i, j) \\
   &= \binom{N}{1}(N-1)! - \binom{N}{2}(N-2)! + \cdots \\
   &= \sum_{i\in[N]} (-1)^{i-1} \frac{N!}{i!}.
   \end{aligned}$$

   Therefore,

   $$\begin{aligned}
   P(A_N) &= \frac{1}{N!} \sum_{i\in[N]} (-1)^{i-1} \frac{N!}{i!} \\
   &= \sum_{i\in[N]} (-1)^{i-1} \frac{1}{i!}.
   \end{aligned}$$

   (d) When you let $N \to \infty$, does the sequence of probabilities $P(A_N)$ converge?

Solution: It is well known that

$$\sum_{i\in\mathbb{N}}(-1)^{i-1}\frac{1}{i!}=\frac{1}{e}.$$

(e) Is the answer in (d) surprising to you in the sense that it did not coincide with your initial expectation of what the probability of at least one match is when N is large? Provide some discussion.

Solution: I recall that, when first encountering this problem, I was unable to form a conjecture either way. On the one hand, as N grows, the chance of placing a given ball in the right bucket is approaching 0. On the other hand, the number of chances to get a match (i.e. the number of balls and buckets involved) is growing without bound. Whenever an infinite number of terms are involved, strange things may happen. Regardless, I suspected to find the probability to be 0 or 1. That is converges to something in between is quite astonishing. That it involves e is a nice feature, though not terribly surprising considering the importance of factorials in the problem.

3. A box contains N identically-sized balls with K of them colored red and $N-K$ colored blue. Consider the following two random experiments.

Experiment 1: Draw n balls in succession without replacement, taking into account the order in which the balls are drawn.

Experiment 2: Draw n balls in succession without replacement, disregarding the order in which the balls are drawn.

If you let A_k be the event that there are exactly k red balls in the sample, do you get the same probability with Experiment 1 and Experiment 2? Justify your answer.

Solution: The probability is the same in both experiments. To see this, suppose there are ℓ distinct ways to draw a total of k red balls in which order matters (as in Experiment 1). Associated with each such event is a probability p_i of witnessing the i^{th} ordering. Since these events are elementary (and so disjoint), we have $P(A_k)=\sum_{i\in[\ell]}p_i$ in Experiment 1. In Experiment 2, an elementary event is drawing exactly k red balls in any order. This event may be viewed, however, as the collection of the ℓ equivalent orderings, and so we still compute $P(A_k)=\sum_{i\in[\ell]}p_i$.

4. Prove the following basic results from set theory. Here, $A, B, C, \ldots$ are subsets of some sample space Ω.

(a) $A\cup(B\cup C)=(A\cup B)\cup C$

Solution:

$$\begin{aligned} x \in A \cup (B \cup C) &\Leftrightarrow x \in A \text{ or } x \in B \cup C \\ &\Leftrightarrow x \in A \text{ or } x \in B \text{ or } x \in C \\ &\Leftrightarrow x \in A \cup B \text{ or } x \in C \\ &\Leftrightarrow x \in (A \cup B) \cup C \end{aligned}$$

(b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Solution: Observe that $A \cup (B \cap C) \subset A \cup B$ and $A \cup (B \cap C) \subset A \cup C$. Hence, $A \cup (B \cap C) \subset (A \cup B) \cap (A \cup C)$.

Next, let $x \in (A \cup B) \cap (A \cup C)$. Thus, $x \in A \cup B$ and $x \in A \cup C$. If $x \notin A$, then $x \in B$ and $x \in C$. That is, $x \in B \cap C$. Hence, $x \in A \cup (B \cap C)$.

(c) (DeMorgan's Laws) Let $\{A_\alpha \mid \alpha \in \mathcal{A}\}$ for some index set $\mathcal{A}$ where each A_α is a subset of Ω. Prove that

$$\left(\bigcup_{\alpha \in \mathcal{A}} A_\alpha \right)^c = \bigcap_{\alpha \in \mathcal{A}} A_\alpha^c.$$

Solution:

$$\begin{aligned} x \in \left(\bigcup_{\alpha \in \mathcal{A}} A_\alpha \right)^c &\Leftrightarrow x \notin \bigcup_{\alpha \in \mathcal{A}} A_\alpha \\ &\Leftrightarrow x \notin A_\alpha \text{ for all } \alpha \in \mathcal{A} \\ &\Leftrightarrow x \in A_\alpha^c \text{ for all } \alpha \in \mathcal{A} \\ &\Leftrightarrow x \in \bigcap_{\alpha \in \mathcal{A}} A_\alpha^c \end{aligned}$$

(d) Let $A_1, A_2, \ldots$ be a sequence of subsets of Ω. Define the sequence $B_1, B_2, \ldots$ according to

$$\begin{aligned} B_1 &= A_1 \\ B_2 &= A_1^c \cap A_2 \\ &\vdots \\ B_n &= A_1^c \cap A_2^c \cap \cdots \cap A_{n-1}^c \cap A_n \\ &\vdots \end{aligned}$$

Prove that $B_1, B_2, \ldots$ is a pairwise disjoint sequence and that, for each n,

$$\bigcup_{j \in [n]} A_j = \bigcup_{j \in [n]} B_j$$

so that, in particular,

$$\bigcup_{j\in\mathbb{N}} A_j = \bigcup_{j\in\mathbb{N}} B_j.$$

Solution:

To see that the sequence is pairwise disjoint, let $i, j \in \mathbb{N}$ with $i \neq j$. Without loss of generality, let $i < j$. It follows immediately that $B_i \cap B_j \subset A_i \cap A_i^c = \emptyset$.

For the second claim, observe first that $B_j \subset A_j$ for all $j \in [n]$, so $\bigcup_{j\in[n]} A_j \supset \bigcup_{j\in[n]} B_j$. For the reverse inclusion, let $x \in \bigcup_{j\in[n]} A_j$. This implies that, for some subset S of $[n]$, $x \in A_i$ for all $i \in S$. Let i_0 be the least element of S. Thus, $x \in A_{i_0}$ and $x \notin A_j$ for all $1 \leq j < i_0$. In other words,

$$\begin{aligned} x &\in \left(\bigcap_{j<i_0} A_j^c\right) \cap A_{i_0} \\ &= B_{i_0} \\ &\subset \bigcup_{j\in[n]} B_j. \end{aligned}$$

Since $\mathbb{N}$ is well-ordered under the usual order, we conclude that

$$\bigcup_{j\in\mathbb{N}} A_j = \bigcup_{j\in\mathbb{N}} B_j.$$

5. Let $\Omega = [-1, 1]$ and define, for each $n \in \mathbb{N}$, the subset of Ω given by $A_n = [-1 + \frac{1}{2n}, 1 - \frac{1}{n}]$. Obtain

$$\limsup A_n = \bigcap_{n\in\mathbb{N}} \bigcup_{n\geq k} A_k$$

and

$$\liminf A_n = \bigcup_{n\in\mathbb{N}} \bigcap_{n\geq k} A_k.$$

Do the two sets coincide?

Solution: Observe first that $(-1 + \frac{1}{2n})$ is a bounded, monotone sequence of real numbers, and so converges (in particular to -1). Similarly, $(1 - \frac{1}{n})$ converges to 1. Hence, $\limsup A_n = \liminf A_n = \lim A_n = [-1, 1]$.

Let us now evaluate each righthand side explicitly. Let $x \in [-1, 1]$ and consider, for any $\epsilon > 0$, define the basic open neighborhood $U = (x-\epsilon, x+\epsilon)$. Choose N so that $\frac{1}{n} < \epsilon$ for all $n \geq N$. Thus, $U \cap A_n \neq \emptyset$ for all $n \geq N$. As $A_N \subset \bigcup_{n\geq k} A_n$, for any k, it follows that $B \cap \bigcap_{n\in\mathbb{N}} \bigcup_{n\geq k} A_k \neq \emptyset$. Hence, $\bigcap_{n\in\mathbb{N}} \bigcup_{n\geq k} A_k = [-1, 1]$. Similarly, as $A_N \subset \bigcap_{n\geq N} A_n$, $B \cap \bigcup_{n\in\mathbb{N}} \bigcap_{n\geq k} A_k \neq \emptyset$. Hence, $\bigcup_{n\in\mathbb{N}} \bigcap_{n\geq k} A_k = [-1, 1]$.

1. Let $\Omega = (0, 1]$ and consider the class of subsets of Ω given by

$$\mathfrak{F}_0 = \left\{ A = \cup_{j=1}^{n} I_j : \ I_j = (a_j, b_j] \subset \Omega, n \in \{0, 1, 2, \ldots\} \right\}.$$

That $\Omega \in \mathfrak{F}_0$ and if $A \in \mathfrak{F}_0$ then $A^c \in \mathfrak{F}_0$ are immediate. Show formally that $\mathfrak{F}_0$ is closed under finite unions, that is, if $A_1, A_2, \ldots, A_n$ are in $\mathfrak{F}_0$, then $\cup_{i=1}^{n} A_i \in \mathfrak{F}_0$. [These set of properties establishes that $\mathfrak{F}_0$ is a field.]

Solution: Let $A_1, \ldots, A_n \in \mathfrak{F}_0$. Now, each A_i can be written as $\bigcup_{j=1}^{n_i} (a_{i_j}, b_{i_j}]$. Thus, we have

$$\begin{aligned} \bigcup_{i=1}^{n} A_i &= \bigcup_{i=1}^{n} \bigcup_{j=1}^{n_i} (a_{i_j}, b_{i_j}] \\ &= \bigcup_{k \in I} (a_k, b_k], \end{aligned}$$

where I is the collection of all indices i_j appearing in $\bigcup_{j=1}^{n_i} (a_{i_j}, b_{i_j}]$ as i ranges from 1 to n. As n is finite and n_i is finite for all i, it follows that $|I|$ is finite, and so $\bigcup_{i=1}^{n} A_i \in \mathfrak{F}_0$.

2. For the Ω and $\mathfrak{F}_0$ in Problem 1, define the set function $P : \mathfrak{F}_0 \to \Re$ via: for $A = \cup_{i=1}^{n} (a_j, b_j]$ where $(a_j, b_j] \cap (a_i, b_i] = \emptyset$ for $i \neq j$, we have

$$P(A) = \sum_{i=1}^{n} (b_i - a_i).$$

Establish that P is indeed a function by showing that for two different representations of A, you obtain the same value for $P(A)$ according to the preceding definition.

Solution: Let $A \in \mathfrak{F}_0$ be given with representations $\bigcup_{i=1}^{n} (a_i, b_i]$ and $\bigcup_{i=1}^{m} (c_i, d_i]$. Choose any maximally connected interval $(x, y]$ of A. Thus, after appropriate reordering of the indices, $(x, y]$ is of the form

$$\begin{aligned} (x, y] &= (a_1, a_2] \cup (a_2, a_3] \cup \cdots \cup (a_{n-1}, a_n] \\ &= (x, a_2] \cup (a_2, a_3] \cup \cdots \cup (a_{n-1}, y]. \end{aligned}$$

Using the other representation of A, we have also that

$$\begin{aligned} (x, y] &= (c_1, c_2] \cup (c_2, c_3] \cup \cdots \cup (c_{n-1}, c_m] \\ &= (x, c_2] \cup (c_2, c_3] \cup \cdots \cup (c_{n-1}, y]. \end{aligned}$$

Now, using the first representation of A,

$$\begin{aligned} P(A) &= \sum_{i=1}^{n} (b_i - a_i) \\ &= (a_2 - a_1) + (a_3 - a_2) + \cdots + (a_n - a_{n-1}) \\ &= a_n - a_1 \\ &= y - x. \end{aligned}$$

Using the second representation of A,

$$\begin{aligned}P(A) &= \sum_{i=1}^{n}(d_i - c_i)\\ &= (c_2 - c_1) + (c_3 - c_2) + \cdots + (c_m - c_{m-1})\\ &= c_m - c_1\\ &= y - x.\end{aligned}$$

Now, since the maximally connected intervals in A are disjoint, $P(A)$ is just the sum of its value on these intervals. As we have demonstrated that P agrees on all maximally connected intervals of A under both representations, it follows that P agrees on all of A.

3. Let Ω be an uncountable sample space. A subset $A \subset \Omega$ is said to be co-countable if A^c is countable. Show that the class of subsets

$$\mathfrak{C} = \{A \subset \Omega : \ A \text{ is countable or co-countable}\}$$

is a σ-field of subsets in Ω.

Solution: As $\Omega^c = \emptyset$, we see that $\Omega \in \mathfrak{C}$.

If $A \in \mathfrak{C}$, then A is either countable or co-countable. Hence, A^c is either co-countable or countable, respectively. Thus, $A^c \in \mathfrak{C}$.

Let $\{A_i\}$ be a countable collection of elements of $\mathfrak{C}$. If each A_i is countable, then $\bigcup\{A_i\}$ is countable. If at least one of the elements, say A_1, is co-countable, then we have

$$\begin{aligned}\left(\bigcup\{A_i\}\right)^c &= \bigcap\{A_i^c\}\\ &\subseteq A_1^c,\end{aligned}$$

which is countable. That is, $\bigcup\{A_i\}$ is co-countable.

4. Let Ω be an uncountable set, and let $\mathcal{C} = \{\{\omega\} : \ \omega \in \Omega\}$ be the class of singleton subsets of Ω. Show that the σ-field generated by $\mathcal{C}$ is the σ-field consisting of countable and co-countable sets.

Solution: Let $\mathfrak{C}$ denote the σ-field generated by $\mathcal{C}$, and let $\mathfrak{D}$ denote the σ-field consisting of all countable and co-countable subsets of Ω.

Since any singleton set is countable, it is clear that $\mathcal{C} \subseteq \mathfrak{D}$. As $\mathfrak{C}$ is the intersection of all σ-fields containing $\mathcal{C}$, it follows that $\mathfrak{C} \subseteq \mathfrak{D}$.

Let now $\mathfrak{C}_0$ be any σ-field containing $\mathcal{C}$. As $\mathfrak{C}_0$ contains every singleton subset of Ω and is closed under countable union, it follows that $\mathfrak{C}_0$ contains every countable subset of Ω. Now, since $\mathfrak{C}_0$ contains every countable subset of Ω and is closed under complementation, it follows that $\mathfrak{C}_0$ contains every co-countable subset of Ω. Hence, $\mathfrak{D} \subset \mathfrak{C}_0$. As $\mathfrak{C}_0$ was chosen arbitrarily, we see that $\mathfrak{D}$ is contained in *any* σ-field containing $\mathcal{C}$, and so $\mathfrak{D} \subseteq \mathfrak{C}$.

5. Suppose $\mathcal{C}$ is a non-empty class of subsets of Ω. Let $\mathfrak{A}(\mathcal{C})$ be the minimal field over $\mathcal{C}$ (i.e., the field generated by $\mathcal{C}$). Show that $\mathfrak{A}(\mathcal{C})$ consists of sets of the form

$$\cup_{i=1}^{m} \cap_{j=1}^{n_i} A_{ij}$$

where $A_{ij} \in \mathcal{C}$ or $A_{ij}^c \in \mathcal{C}$, and where the m sets $\cap_{j=1}^{n_i} A_{ij}, i = 1, 2, \ldots, m$ are disjoint.

Solution: In what follows, let

$$\mathcal{F} = \left\{ \bigcap_{j=1}^{m} A_j \mid A_j \in \mathcal{C} \text{ or } A_j^c \in \mathcal{C} \right\}$$

and

$$\mathcal{D} = \left\{ \bigsqcup_{i=1}^{n} F_i \mid F_i \in \mathcal{F}, \{F_i\}_{i \in [n]} \text{ pairwise disjoint} \right\}.$$

With this notation, we must show that $\mathfrak{A}(\mathcal{C}) = \mathcal{D}$.

To see that $\mathcal{D} \subseteq \mathfrak{A}(\mathcal{C})$, notice that $\mathfrak{A}(\mathcal{C})$ is a field containing $\mathcal{C}$, and so is closed under complementation, finite unions, and finite intersections of elements of $\mathcal{C}$. In particular, $\mathfrak{A}(\mathcal{C})$ must contain any element of the form specified by $\mathcal{D}$.

We show next that $\mathfrak{A}(\mathcal{C}) \subseteq \mathcal{D}$. To accomplish this, note first that $\mathcal{C} \subseteq \mathcal{D}$. Thus, if we can show that $\mathcal{D}$ is itself a field, we will have the desired inclusion, as $\mathfrak{A}(\mathcal{C})$ is the *smallest* field containing $\mathcal{C}$.

Before proceeding, we establish a useful lemma.

Lemma 0.1. *If $F \in \mathcal{F}$, then $F^c \in \mathcal{D}$.*

Proof. Let $F \in \mathcal{F}$ be of the form $F = \bigcap_{j \in [m]} A_j$, where A_j or A_j^c belongs to $\mathcal{C}$ for each j and the collection of all A_j is pairwise disjoint. It follows that

$$\begin{aligned} F^c &= \bigcup_{j \in [m]} A_j^c \\ &= A_1^c \cup (A_2^c \cap A_1) \cup \cdots \cup (A_m^c \cap A_1 \cap \cdots \cap A_{m-1}). \end{aligned}$$

Consider a typical term $A_k^c \cap A_1 \cap \cdots \cap A_{k-1}$ in the union. This set is evidently an element of $\mathcal{F}$, as A_i or A_i^c belongs to $\mathcal{C}$ for all i (similarly for A_k^c). Moreover, the collection of all terms in the union is pairwise disjoint via disjointification. Thus, $F^c \in \mathcal{D}$. □

We now return to the task of showing that $\mathcal{D}$ is a field.

Let $A \in \mathcal{C}$. It follows that $\mathcal{F}$ contains $A \cap A^c = \emptyset$. Thus, by the lemma, $\emptyset^c = \Omega$ belongs to $\mathcal{D}$.

Next, we show that $\mathcal{D}$ is closed under finite intersections. To that end, let $D_1, D_2 \in \mathcal{D}$ with $D_1 = \bigsqcup_{i\in[n_1]} F_i$ and $D_2 = \bigsqcup_{j\in[n_2]} F'_j$. Now,

$$\begin{aligned} D_1 \cap D_2 &= \left(\bigsqcup_{i\in[n_1]} F_i\right) \cap \left(\bigsqcup_{j\in[n_2]} F'_j\right) \\ &= \bigcup_{j\in[n_2]} (F_1 \cap F'_j) \cup \cdots \cup \bigcup_{j\in[n_2]} (F_{n_1} \cap F'_j). \end{aligned}$$

Observe that this last line is a union of elements of the form $F_i \cap F'_j$. As $\mathcal{F}$ is closed under finite intersection, each term of the union is a member of $\mathcal{F}$. Moreover, the collection of all these terms is pairwise disjoint. To see this, consider some distinct $F_{i_1} \cap F'_{j_1}$ and $F_{i_2} \cap F'_{j_2}$ in the union. Without loss of generality, let $F_{i_1} \neq F_{i_2}$. By definition of D_1, it must be that F_{i_1} and F_{i_2} are disjoint, and thus $F_{i_1} \cap F'_{j_1}$ and $F_{i_2} \cap F'_{j_2}$ are disjoint. All told, we have that $D_1 \cap D_2 \in \mathcal{D}$. Proceeding by induction, we have that $\mathcal{D}$ is closed under finite intersections.

Finally, we show that $\mathcal{D}$ is closed under complementation. To that end, pick any $D \in \mathcal{D}$ with $D = \bigsqcup_{i\in[n]} F_i$. It follows that $D^c = \bigcap_{i\in[n]} F_i^c$. By the lemma, each F_i^c is an element of $\mathcal{D}$. As $\mathcal{D}$ is closed under finite intersections, $D^c \in \mathcal{D}$.

Taken together, we have verified that $\mathcal{D}$ is indeed a field containing $\mathcal{C}$, and so $\mathfrak{A}(\mathcal{C}) \subseteq \mathcal{D}$. Therefore, $\mathfrak{A}(\mathcal{C}) = \mathcal{D}$, as desired.

6. Let $\mathfrak{C}$ be a class of subsets of Ω. It is said to be a *monotone class* if for $A_1 \subset A_2 \subset \ldots$ in $\mathfrak{C}$, then $\cup_{n=1}^{\infty} A_n = \lim A_n \in \mathfrak{C}$ and for $A_1 \supset A_2 \supset \ldots$ in $\mathfrak{C}$, then $\cap_{n=1}^{\infty} A_n = \lim A_n \in \mathfrak{C}$. Prove that if $\mathfrak{C}$ is a field and a monotone class, then it is a σ-field.

   Solution: Since $\mathfrak{C}$ is a field, we have $\Omega \in \mathfrak{C}$ and closure under complementation.

   Let now $\{A_i\}$ be a countable collection of elements of $\mathfrak{C}$. Define, for all k, $B_k = \bigcup_{i=1}^{k} A_i$. Thus, $\{B_i\}$ is a monotone sequence of subsets of $\mathfrak{C}$, and so $\bigcup\{B_i\} \in \mathfrak{C}$. It is clear, however, that $\bigcup\{A_i\} = \bigcup\{B_i\}$, and so we conclude that $\bigcup\{A_i\} \in \mathfrak{C}$, thus verifying that $\mathfrak{C}$ is a σ-field.

7. Let $\Omega = \Re$ and consider the two classes of subsets given by

$$\mathcal{C}_1 = \{(a, b) : \ a, b \text{ are rationals in } \mathbb{R}\};$$

$$\mathcal{C}_2 = \{[a, b] : \ a, b \text{ are in } \mathbb{R}\}.$$

   Establish that these two classes of sets generate the same σ-field. (Their common generated σ-field is the Borel σ-field in $\Re$.)

   Solution: Let $\mathfrak{C}_1$ denote the σ-field generated by $\mathcal{C}_1$ and $\mathfrak{C}_2$ the σ-field generated by $\mathcal{C}_2$. Thus, we have that

$$\mathfrak{C}_1 = \bigcap\{\mathfrak{D} \mid \mathfrak{D} \text{ is a } \sigma\text{-field and } \mathcal{C}_1 \subset \mathfrak{D}\}$$

and

$$\mathfrak{C}_2 = \bigcap \{\mathfrak{D} \mid \mathfrak{D} \text{ is a } \sigma\text{-field and } \mathcal{C}_2 \subset \mathfrak{D}\}.$$

We proceed by showing that any σ-field containing $\mathcal{C}_1$ contains $\mathcal{C}_2$ and vice versa, thus establishing that $\mathfrak{C}_1 = \mathfrak{C}_2$.

Let $\mathfrak{D}_1$ be any σ-field containing $\mathcal{C}_1$. We need to show that, for any $a, b \in \mathbb{R}$, $[a, b] \in \mathfrak{D}_1$. To that end, choose two sequences $\{a_n\}$ and $\{b_n\}$ of rational numbers with $a_n \nearrow a$ and $b_n \searrow b$. Now, $(a_n, b_n) \in \mathfrak{D}_1$ for all n, and so $\mathfrak{D}_1$ contains $\bigcap_{n \in \mathbb{N}} (a_n, b_n) = [a, b]$.

Let $\mathfrak{D}_2$ be any σ-field containing $\mathcal{C}_2$. We need to show that, for any $a, b \in \mathbb{Q}$, $(a, b) \in \mathfrak{D}_2$. To that end, choose two sequences $\{a_n\}$ and $\{b_n\}$ of real numbers with $a_n \searrow a$ and $b_n \nearrow b$. Now, $[a_n, b_n] \in \mathfrak{D}_2$ for all n, and so $\mathfrak{D}_2$ contains $\bigcup_{n \in \mathbb{N}} [a_n, b_n] = (a, b)$.

1. Let $\mathcal{S}$ be a semi-algebra of subsets of Ω. Denote by $\mathfrak{A}(\mathcal{S})$ the algebra or field generated by $\mathcal{S}$ and by $\sigma(\mathcal{S})$ the σ-algebra generated by $\mathcal{S}$. Prove that

$$\sigma(\mathfrak{A}(\mathcal{S})) = \sigma(\mathcal{S}).$$

Solution: Since $\mathcal{S} \subset \mathfrak{A}(\mathcal{S})$, we have immediately that $\sigma(\mathcal{S}) \subset \sigma(\mathfrak{A}(\mathcal{S}))$.

To demonstrate the reverse inclusion, it suffices to show that $\sigma(\mathcal{S})$ is itself a σ-algebra containing $\mathfrak{A}(\mathcal{S})$. To that end, recall that every element of $\mathfrak{A}(\mathcal{S})$ is of the form

$$\bigsqcup_{i=1}^{m} S_i$$

where $S \in \mathcal{S}$. Now, as $\sigma(\mathcal{S})$ is countable union, such elements belong to $\sigma(\mathcal{S})$. Thus, $\sigma(\mathcal{S})$ is a σ-algebra containing $\mathfrak{A}(\mathcal{S})$, and so $\sigma(\mathfrak{A}(\mathcal{S})) \subset \sigma(\mathcal{S})$.

2. Let $\Omega = \mathcal{C}[0, 1]$, the space of continuous functions on $[0, 1]$. For $t \in [0, 1]$ and $a, b \in \Re$, define the subset of Ω given by

$$A(t; a, b) = \{f \in \Omega : f(t) \in (a, b]\}.$$

Gather these subsets into a collection $\mathcal{C}_0$, that is,

$$\mathcal{C}_0 = \{A(t; a, b) : t \in [0, 1], a \in \Re, b \in \Re\}.$$

(a) Demonstrate that $\mathcal{C}_0$ is not a semi-algebra.
Solution: Let $A(t; a, b) \in \mathcal{C}_0$. Now,

$$\begin{aligned} A(t; a, b)^c &= \{f \in \Omega \mid f(t) \notin (a, b]\} \\ &= \{f \in \Omega \mid f(t) \in (-\infty, a] \cup (b, \infty)\}. \end{aligned}$$

Evidently, $A(t; a, b)^c$ cannot be represented by a finite union of elements of $\mathcal{C}_0$, and so $\mathcal{C}_0$ is not a semi-algebra.
Additionally, $\Omega \notin \mathcal{C}_0$, as there is no finite interval $(a, b]$ such that all continuous functions satisfy, for example, $f(0) \in (a, b]$.

(b) Describe the structure of the typical element of $\mathcal{S}_0 \equiv \mathfrak{S}(\mathcal{C}_0)$, the semi-algebra generated by $\mathcal{C}_0$.

Solution: We claim that a typical element of $\mathcal{S}_0$ has the form $\bigcap_{i=1}^n A_i$ where, for each i, $A_i = A(t_i; a_i, b_i)$ or $A_i = A(t_i; a_i, b_i)^c = A(t_i; -\infty, a_i) \cup A(t_i; b_i, \infty)$ with $t_i \in [0,1], a_i, b_i \in \mathfrak{R}$, where we understand that intervals of the form $(a, \infty]$ should be interpreted as (a, ∞).

By the observations in part a, it suffices to show that $\mathcal{S}_0$ as it is defined above is a σ-algebra (the new elements we've included are required at a minimum to patch up the deficiencies of $\mathcal{C}_0$). Now, it is clear that $\emptyset \in \mathcal{S}_0$ (take the intersection of any disjoint balls) and $\Omega \in \mathcal{S}_0$ (Ω can be represented, for example, by $A(0; 0, 0)^c$). By construction, $\mathcal{S}_0$ is closed under complementation. It is also closed under finite intersection, since intervals of the form $(a, b]$ with $a, b \in \mathfrak{R} \cup \{-\infty, \infty\}$ are closed under finite intersection. Therefore, $\mathcal{S}_0$ is a semi-algebra, and so must be the smallest semi-algebra containing $\mathcal{C}_0$.

(c) Describe the structure of the typical element of $\mathfrak{A}(\mathcal{C}_0)$, the algebra generated by $\mathcal{C}_0$.

Solution: By a previous homework, we know that the elements of $\mathfrak{A}(\mathcal{C}_0)$ are of the form

$$\bigsqcup_{i=1}^{m} \bigcap_{j=1}^{n_j} A_{ij},$$

where, for all i and j, $A_{ij} \in \mathcal{C}_0$ or $A_{ij}^c \in \mathcal{C}_0$ and the m sets $\bigcap_{j=1}^{n_j} A_{ij}$ are pairwise disjoint. In this particular example, we know that $A_{ij} = A(t; a, b)$ and that $A_{ij}^c = A(t; a, b)^c = \{f \in \Omega \mid f(t) \in (-\infty, a] \cup (b, \infty)\}$ for some $t \in [0, 1]$ and $a, b \in \mathfrak{R}$. Adopting the conventions as in part b, we can represent any element of $\mathfrak{A}(\mathcal{C}_0)$ as

$$\bigsqcup_{i=1}^{m} \bigcap_{j=1}^{n_j} A_i.$$

(d) Denoting by $\mathcal{B}_0 = \sigma(\mathcal{C}_0)$, the σ-field generated by $\mathcal{C}_0$, determine if the subset of Ω given by

$$E = \{f \in \Omega : \sup_{t \in [0,1]} |f(t)| \leq B\}$$

is an element of $\mathcal{B}_0$. [By the way, $\mathcal{B}_0$ is called the σ-field generated by the cylinder sets.]

Solution: Let the sequence (r_n) be an enumeration of $\mathfrak{Q} \cap [0, 1]$ and define

$$F = \bigcap_{n=1}^{\infty} A(r_n, -B, B).$$

Since $\mathcal{B}_0$ is closed under countable intersection, we have that $F \in \mathcal{B}_0$. We claim that $E = F$.

Evidently, $E \subseteq F$, as any function f satisfying $\sup_{t\in[0,1]} |f(t)| \leq B$ satisfies $|f(r_n)| \leq B$ for all n.

Suppose now, for the purpose of contradiction, that $F \not\subseteq E$. That is, there exists continuous f and $t_0 \in \mathfrak{R} \setminus \mathfrak{Q}$ such that $|f(t_0)| > B$. Pick any $0 < \epsilon < B - |f(t_0)|$. By the continuity of f, there is $\delta > 0$ such that, whenever $|x - y| < \delta$, $|f(x) - f(y)| < \epsilon$. Now, as $\mathfrak{Q}$ is dense in $\mathfrak{R}$, we can find a rational number r such that $|t_0 - r| < \delta$, yet

$$\begin{aligned} |f(t_0) - f(r)| &> ||f(t_0)| - |f(r)|| \\ &\geq ||f(t_0)| - B| \\ &> \epsilon, \end{aligned}$$

which is contrary to the continuity of f. Hence, $F \subseteq E$, as desired.

3. Same Ω as in Problem 2. Define the metric (distance) function $d : \Omega \times \Omega \to \Re$ via

$$d(f, g) = \sup_{t\in[0,1]} |f(t) - g(t)|.$$

For $\epsilon > 0$ and $f \in \Omega$, define the subset $B(f; \epsilon)$ of Ω according to

$$B(f; \epsilon) = \{g \in \Omega : d(f, g) < \epsilon\}.$$

These are the open balls in Ω. Gather these open balls in the collection $\mathcal{S}_1$, that is,

$$\mathcal{S}_1 = \{B(f; \epsilon) : f \in \Omega, \epsilon > 0\}.$$

(a) Determine if $\mathcal{S}_1$ is a semi-algebra in Ω; and if it is not, find the semi-algebra generated by $\mathcal{S}_1$.

Solution: $\mathcal{S}_1$ does not contain Ω, and so is not a semi-algebra. To see this, observe that, for any $f \in \Omega$ and finite $\epsilon > 0$, there is a continuous function on $[0, 1]$ not contained in $B(f; \epsilon)$ (for example, the constant function $M + 2\epsilon$ with $M = \sup_{t\in[0,1]} |f(t)|$).

For the same reason,

$$B(f; \epsilon)^c = \{g \in \Omega \mid d(f, g) \geq \epsilon\}$$

cannot be represented as a finite union of $B(f_i, \epsilon_i)$ (for example, the constant function $M + 2\epsilon$ with $M = \max_{i\in[n]}\{\sup_{t\in[0,1]} |f_i(t)|\}$ will not be contained in this union).

As before, we represent elements of $\mathfrak{S}(\mathcal{S}_1)$ by $\bigcap_{i=1}^n S_i$ where, for all i, $S_i \in \mathcal{S}_1$ or $\mathcal{S}_i^c \in \mathcal{S}_1$.

(b) Determine the algebra generated by $\mathcal{S}_1$, that is, $\mathfrak{A}(\mathcal{S}_1)$.

Solution: In a previous homework, we have established this result in more generality. In this particular case, we have

$$\mathfrak{A}(\mathcal{S}_1) = \bigsqcup_{i=1}^{m} \bigcap_{j=1}^{n_j} S_{ij}$$

with $S_{ij} \in \mathcal{S}_1$ or $S_{ij} \in \mathcal{S}_2$ for all i, j and the m sets $\bigcap_{j=1}^{n_j} S_{ij}$ are pairwise disjoint.

(c) Denote by $\mathcal{B}_1 = \sigma(\mathcal{S}_1)$, the σ-field generated by $\mathcal{S}_1$, so that by definition this is the Borel σ-field associated with the metric d. Determine if the subset of Ω defined in item (d) in Problem 2 is an element of this Borel σ-field.

Solution: The set

$$E = \left\{ f \in \Omega : \sup_{t \in [0,1]} |f(t)| \leq B \right\}$$

can be represented in $\mathcal{B}_1$ as

$$\bigcap_{n=1}^{\infty} B\left(0; B + \frac{1}{n}\right).$$

Evidently, $E \subset B\left(0; B + \frac{1}{n}\right)$ for each n, and so $E \subset \bigcap_{n=1}^{\infty} B\left(0; B + \frac{1}{n}\right)$. Let now $g \in \bigcap_{n=1}^{\infty} B\left(0; B + \frac{1}{n}\right)$. This implies that $\sup_{t\in[0,1]} |g(t)| < B + \frac{1}{n}$ for all n. Hence, $\sup_{t\in[0,1]} |g(t)| \leq B$, and so $g \in E$. Therefore, $E = \bigcap_{n=1}^{\infty} B\left(0; B + \frac{1}{n}\right)$.

4. Investigate the relationship between $\mathcal{B}_0$ in Problem 2 and $\mathcal{B}_1$ in Problem 3. Are these σ-fields identical; or is one strictly containing the other, and if so, which one is the larger σ-field?

Solution: We claim that $\mathcal{B}_0 = \mathcal{B}_1$. To prove this, we demonstrate that any basic open ball of $\mathcal{B}_0$ can be represented as a countable union of basic open balls of $\mathcal{B}_1$, and vice versa.

Pick some $A(t; a, b) \in \mathcal{B}_0$. It is (relatively) well-known that the collection of continuous piecewise linear functions on $[0, 1]$ is countable and dense in the collection of all continuous functions on $[0, 1]$. Denote by $\mathcal{F}$ the subset of continuous piecewise linear functions on $[0, 1]$ contained in $A(t; a, b)$. Now, for each $f \in \mathcal{F}$, define $\epsilon_f = \min\{f(t) - a, b - f(t)\}$. We claim that

$$A(t; a, b) = \bigcup \{B(f, \epsilon_f) \mid f \in \mathcal{F}\}.$$

By our choice of ϵ_f, each $B(f, \epsilon_f) \subseteq A(t; a, b)$. Conversely, if $g \in A(t; a, b)$, then the density of $\mathcal{F}$ in $A(t; a, b)$ guarantees that there is $f \in \mathcal{F}$ such that $d(f, g) < \epsilon_f$, and so $g \in B(f, \epsilon_f)$. Hence, $\mathcal{B}_0 \subseteq \mathcal{B}_1$.

For the reverse inclusion, pick some $B(f,\epsilon) \in \mathcal{B}_1$ and let (r_n) be an enumeration of $\mathfrak{Q} \cap [0,1]$. We claim that

$$B(f,\epsilon) = \bigcap_{n=1}^{\infty} A\left(r_n; f(r_n) - \frac{\epsilon}{2}, f(r_n) + \frac{\epsilon}{2}\right).$$

By construction, $A\left(r_n; f(r_n) - \frac{\epsilon}{2}, f(r_n) + \frac{\epsilon}{2}\right) \subset B(f,\epsilon)$ for each n. Conversely, if $f \in B(f,\epsilon)$, then the density of the rationals in $[0,1]$ guarantees that $f \in A\left(r_n; f(r_n) - \frac{\epsilon}{2}, f(r_n) + \frac{\epsilon}{2}\right)$ for each n. Hence, $\mathcal{B}_1 \subseteq \mathcal{B}_0$.

5. As you might have already noticed, the typical element of $\mathcal{S}_0$ in Problem 2 is of form $\cap_{i=1}^n A_{t_i}^*$ where the t_is are distinct and A_t^* is either of form $A(t;a,b)$ or $A(t;a,b)^c$. Define the (set)-function $\mathbb{P} : \mathcal{S}_0 \to \Re$ according to

$$\mathbb{P}\left[\cap_{i=1}^n A_{t_i}^*\right] = \prod_{i=1}^n \left[\int_{A_{t_i}^*} \phi(v)dv\right]$$

where $\phi(z) = (2\pi)^{-1/2} \exp\{-v^2/2\}$ is the standard normal density function. *Assuming* that $\mathbb{P}$ is σ-additive on $\mathcal{S}_0$, provide a reason [do not prove anything, just a reason!] why you could conclude that there is a unique probability measure $\mathbb{W}$ on $\mathcal{B}_0$ which extends P, that is, $\mathbb{W}|_{\mathcal{S}_0} = \mathbb{P}$. [This probability measure $\mathbb{W}$ is infact the *Wiener measure* on $(\mathcal{C}[0,1], \mathcal{B}_0)$].

Solution: The desired result is a direct application of the first and second extension theorems. Since $\mathbb{P}$ is σ-additive on the semi-algebra $\mathcal{S}_0$, we can extend it uniquely to a probability measure $\mathbb{P}'$ on $\mathfrak{A}(\mathcal{S}_0)$. Applying the second extension theorem to $\mathbb{P}'$, we can obtain the desired probability measure $\mathbb{W}$ on $\sigma(\mathcal{S}_0) = \mathcal{B}_0$.

Problem 1

Let Ω be some non-empty set, and let A and B be two subsets of Ω which are not disjoint and with $\Omega \neq A \cup B$. Define the semi-algebra

$$\mathfrak{S} = \{\emptyset, \Omega, A, B, AB, A^cB, AB^c, A^cB^c\}.$$

Define the function $P : \mathfrak{S} \to \Re$ according to the following specification:

$P(\emptyset) = 0; P(\Omega) = 1; P(A) = .2; P(B) = .6; P(AB) = .12; P(A^cB) = .48; P(AB^c) = .08; P(A^cB^c) = .32$

Find the σ-field, $\mathfrak{A}$, generated by $\mathfrak{S}$, i.e., enumerate all the elements of this σ-field. [Note: This should coincide with the algebra or field generated by $\mathfrak{S}$.]

Proof. Since $\mathfrak{S}$ is finite, the σ-field generated by $\mathfrak{S}$ coincides with the field generated by $\mathfrak{S}$. Thus, it suffices to consider $\mathfrak{A}(\mathfrak{S})$. Now, we know that $\mathfrak{A}(\mathfrak{S})$ is the collection of all sums of finite families of mutually disjoint subsets of Ω in $\mathfrak{S}$ (Resnick, Lemma 2.4.1). A first calculation gives the following as the elements of $\mathfrak{A}$.

$\emptyset$	$A \cup A^cB$	$A \cup A^cB \cup A^cB^c$	$AB \cup A^cB \cup AB^c \cup A^cB^c$
Ω	$A \cup A^cB^c$	$B \cup AB^c \cup A^cB^c$	
A	$B \cup A^cB$	$AB \cup A^cB \cup AB^c$	
B	$B \cup A^cB^c$	$AB \cup AB^c \cup A^cB^c$	
AB	$AB \cup A^cB$	$A^cB \cup AB^c \cup A^cB^c$	
A^cB	$AB \cup AB^c$		
AB^c	$AB \cup A^cB^c$		
A^cB^c	$A^cB \cup AB^c$		
	$A^cB \cup A^cB^c$		
	$AB^c \cup A^cB^c$		

Many of these elements, however, are redundant. For example,

$$\begin{aligned} A \cup A^cB \cup A^cB^c &= A \cup A^c(B \cup B^c) \\ &= A \cup A^c\Omega \\ &= A \cup A^c \\ &= \Omega. \end{aligned}$$

Removing redundant elements gives the following representation of $\mathfrak{A}$.

$\emptyset$	$A \cup A^cB$
Ω	$A \cup A^cB^c$
A	$B \cup A^cB^c$
B	$AB \cup A^cB^c$
AB	$A^cB \cup AB^c$
A^cB	$A^cB \cup A^cB^c$
AB^c	$AB^c \cup A^cB^c$
A^cB^c	

□

Find the (unique) extension of P to $\mathfrak{A}$. You must enumerate the values of this extension for each possible element of $\mathfrak{A}$.

Proof. Since $\mathfrak{S}$ is a semi-algebra, the unique extension P' of P to $\mathfrak{A}$ is defined by

$$P'\left(\sum_{i \in I} S_i\right) = \sum_{i \in I} P(S_i),$$

(Resnick, Theorem 2.4.1). By direct calculation, we define P' on each element of $\mathfrak{A}$.

$P(\emptyset) = 0$	$P(A \cup A^cB) = .68$
$P(\Omega) = 1$	$P(A \cup A^cB^c) = .52$
$P(A) = .2$	$P(B \cup A^cB^c) = .92$
$P(B) = .6$	$P(AB \cup A^cB^c) = .44$
$P(AB) = .12$	$P(A^cB \cup AB^c) = .56$
$P(A^cB) = .48$	$P(A^cB \cup A^cB^c) = .8$
$P(AB^c) = .08$	$P(AB^c \cup A^cB^c) = .4$
$P(A^cB^c) = .32$	

□

Problem 2

Show that a σ-field cannot be countably infinite, i.e., either it has a finite cardinality or it is at least as large as $\mathbb{R}$.

Proof. Let σ be an infinite σ-field of subsets of Ω and suppose we can find a sequence $\{A_n \mid n \in \mathbb{N}\}$ of pairwise disjoint subsets of σ. Consider the function f from the collection of infinite binary strings into σ defined by

$$f(s) = \bigcup_{i \in I_s} A_i,$$

where I_s is the set of indices on which s is 1. Now, given two binary strings s_0 and s_1,

$$\begin{aligned} f(s_0) = f(s_1) &\Rightarrow \bigcup_{i \in I_{s_0}} A_i = \bigcup_{i \in I_{s_1}} A_i \\ &\Rightarrow I_{s_0} = I_{s_1} \qquad \text{(since the } A_i \text{ are pairwise disjoint)} \\ &\Rightarrow s_0 = s_1. \end{aligned}$$

Hence, f is injective, and so the cardinality of σ is at least $\aleph_1$.

It remains to show that we can indeed produce a sequence of pairwise disjoint subsets of σ. To that end, let $A \in \sigma$. Since σ is closed under complementation, $A^c \in \sigma$. Now, $A \cup A^c = \Omega$, which is infinite (since σ is infinite). Hence, one of A or A^c is infinite. Without loss of generality, let A be infinite and set $A_1 = A^c$.

Consider next the collection of subsets

$$\{A \cap B \mid B \in \sigma\}.$$

Observe that each $A \cap B$ is an element of σ, since σ is closed under countable intersection. Now,

$$\bigcup \{A \cap B \mid B \in \sigma\} = A,$$

which is infinite, so some element C of $\{A \cap B \mid B \in \sigma\}$ is infinite. Set $A_2 = C^c$ and consider next

$$\{C \cap B \mid B \in \sigma\}.$$

Proceeding in this way, we generate a sequence $\{A_n\}$ of elements of σ. The A_i are disjoint since, by construction, $A_k \subset A_j^c$ for all $k > j$. Using this sequence in the above argument yields the desired result. □

Problem 3

Let $\mathfrak{B}$ be a σ-field of subsets of Ω and let $A \subset \Omega$ which is not in $\mathfrak{B}$. Show that the smallest σ-field generated by $\{\mathfrak{B}, A\}$ consists of sets of form

$$AB_1 \cup A^c B_2, \quad B_1, B_2 \in \mathfrak{B}.$$

Proof. Denote by σ the σ-field generated by $\{\mathfrak{B}, A\}$ and by $\mathfrak{C}$ the collection $\{AB_1 \cup A^c B_2 \mid B_1, B_2 \in \mathfrak{B}\}$. We have immediately that $\mathfrak{C} \subseteq \sigma$, since σ is closed under complementation, countable intersection, and countable union, by definition. It remains to show that $\mathfrak{C} \subseteq \sigma$. To accomplish this, we need only show that $\mathfrak{C}$ is itself a σ-field and appeal to the minimality of σ.

Evidently, $\Omega \in \mathfrak{C}$, since $\Omega \in \mathfrak{B}$.

To establish closure under complementation, choose any $C \in \mathfrak{C}$, so that $C = AB_1 \cup A^c B_2$ for $B_1, B_2 \in \mathfrak{B}$. It follows that

$$\begin{aligned}
C^c &= (AB_1 \cup A^c B_2)^c \\
&= (AB_1)^c \cap (A^c B_2)^c \\
&= (A^c \cup B_1^c) \cap (A \cup B_2^c) \\
&= A^c A \cup A^c B_2^c \cup B_1^c A \cup B_1^c B_2^c \\
&= A^c B_2^c \cup B_1^c A \cup B_1^c B_2^c \\
&= A^c B_2^c \cup B_1^c A \cup (AB_1^c B_2^c \cup A^c B_1^c B_2^c) \\
&= A(B_1^c \cup B_1^c B_2^c) \cup A^c(B_2^c \cup B_1^c B_2^c) \\
&= AB_1^c \cup A^c B_2^c.
\end{aligned}$$

Since $\mathfrak{B}$ is closed under complementation, B_1^c and B_2^c belong to $\mathfrak{B}$, and so C^c belongs to $\mathfrak{C}$.

To establish closure under countable unions, choose $C_i \in \mathfrak{C}$ for every natural number, so that each C_i is of the form $AB_i \cup A^c B_i'$ for $B_i, B_i' \in \mathfrak{B}$. Now,

$$\begin{aligned}
\bigcup_{i \in \mathbb{N}} C_i &= \bigcup_{i \in \mathbb{N}} (AB_i \cup A^c B_i') \\
&= \bigcup_{i \in \mathbb{N}} AB_i \cup \bigcup_{i \in \mathbb{N}} A^c B_i') \\
&= A\left(\bigcup_{i \in \mathbb{N}} B_i\right) \cup A^c \left(\bigcup_{i \in \mathbb{N}} B_i'\right).
\end{aligned}$$

Since $\mathfrak{B}$ is closed under countable unions, we have that both $\bigcup_{i \in \mathbb{N}} B_i$ and $\bigcup_{i \in \mathbb{N}} B_i'$ belong to $\mathfrak{B}$, and so $\bigcup_{i \in \mathbb{N}} C_i$ belongs to $\mathfrak{C}$.

Hence, $\mathfrak{C}$ is a σ-field, and so contains σ. Therefore, σ and $\mathfrak{C}$ coincide, as desired. □

Problem 6

If $\mathfrak{S}_1$ and $\mathfrak{S}_2$ are two semialgebras of subsets of Ω, show that

$$\mathfrak{S}_1\mathfrak{S}_2 = \{A_1A_2 : A_1 \in \mathfrak{S}_1, A_2 \in \mathfrak{S}_2\}$$

is again a semialgebra of subsets of Ω.

Proof. For ease of notation, let $\mathfrak{S}$ denote $\mathfrak{S}_1\mathfrak{S}_2$. We show directly that $\mathfrak{S}$ is a semialgebra.

Since both $\mathfrak{S}_1$ and $\mathfrak{S}_2$ are semialgebras, $\Omega \in \mathfrak{S}_1$ and $\Omega \in \mathfrak{S}_2$. Thus, $\Omega = \Omega\Omega \in \mathfrak{S}$.

To see that $\mathfrak{S}$ is closed under finite intersection, pick $S, S' \in \mathfrak{S}$, so that $S = S_1S_2$ and $S' = S'_1S'_2$ with $S_1, S'_1 \in \mathfrak{S}_1$ and $S_2, S'_2 \in \mathfrak{S}_2$. It follows immediately that

$$\begin{aligned} SS' &= (S_1S_2)(S'_1S'_2) \\ &= (S_1S'_1)(S_2S'_2). \end{aligned}$$

Since $\mathfrak{S}_1$ and $\mathfrak{S}_2$ are both semialgebras, $S_1S'_1 \in \mathfrak{S}_1$ and $S_2S'_2 \in \mathfrak{S}_2$. Thus, $SS' \in \mathfrak{S}$. By a simple induction argument, we have that $\mathfrak{S}$ is closed under finite intersection.

It remains to show that, given $S \in \mathfrak{S}$, S^c can be written as the union of a finite collection of pairwise disjoint elements of $\mathfrak{S}$. To that end, let $S = S_1S_2$ with $S_1 \in \mathfrak{S}_1$ and $S_2 \in \mathfrak{S}_2$. It follows that,

$$\begin{aligned} S^c &= (S_1S_2)^c \\ &= S_1^c \cup S_2^c \\ &= \sum_{i=1}^{m} A_i \cup \sum_{i=1}^{n} B_i, \end{aligned}$$

where the A_i are pairwise disjoint elements of $\mathfrak{S}_1$ and the B_i are pairwise disjoint elements of $\mathfrak{S}_2$. Denote $\sum_{i=1}^{m} A_i$ by A. Observe that,

$$\sum_{i=1}^{m} A_i \cup \sum_{i=1}^{n} B_i = \sum_{i=1}^{m} A_i \cup \sum_{i=1}^{n} B_iA^c.$$

Now, the A_i are pairwise disjoint because $\mathfrak{S}_1$ is a semialgebra. Similarly, the B_i are pairwise disjoint because $\mathfrak{S}_2$ is a semialgebra, and so the B_iA^c are pairwise disjoint. Finally, any A_i is disjoint from any B_jA^c, since A_i is disjoint from A^c. Thus, $\{A_i \mid i \in [m]\} \cup \{B_iA^c \mid i \in [n]\}$ is a pairwise disjoint collection of elements from $\mathfrak{S}_1$ and $\mathfrak{S}_2$. We proceed by showing that each A_i and B_iA^c can be represented by a union of disjoint elements of $\mathfrak{S}$.

Evidently, every A_i belongs to $\mathfrak{S}_1$, since each can be represented as $A_i\Omega$. Now, for any $k \in [n]$, we have

$$\begin{aligned} B_k A^c &= B_k \left(\sum_{i=1}^{m} A_i \right)^c \\ &= B_k \prod_{i=1}^{m} A_i^c \\ &= B_k \prod_{i=1}^{m} \sum_{j=1}^{m_i} C_{ij}, \end{aligned}$$

where, for each $i \in [m]$, $\{C_{ij} \mid j \in [m_i]\}$ is a pairwise disjoint collection of elements of $\mathfrak{S}_1$. Let now M denote the collection of m-tuples $\prod_{i=1}^{m}[m_i]$. We may rewrite the above as

$$B_k \prod_{i=1}^{m} \sum_{j=1}^{m_i} C_{ij} = B_k \sum_{x \in M} \prod_{i=1}^{m} C_{ix(i)}.$$

Now, since $\mathfrak{S}_1$ is a semialgebra, we have that $\prod_{i=1}^{m} C_{ix(i)} \in \mathfrak{S}_1$ for each $x \in M$. Moreover, the collection of all such elements is pairwise disjoint. That is, given $x, y \in M$ with $x \neq y$,

$$\prod_{i=1}^{m} C_{ix(i)} \cap \prod_{i=1}^{m} C_{iy(i)} = \emptyset,$$

since, for some $j \in [m]$, $x(j) \neq y(j)$ and $C_{j\ell} \cap C_{j\ell'} = \emptyset$ for any j and $\ell \neq \ell'$.

All told, we have shown that, for each $k \in [n]$,

$$\begin{aligned} B_k A^c &= B_k \sum_{x \in M} \prod_{i=1}^{m} C_{ix(i)} \\ &= \sum_{x \in M} B_k \prod_{i=1}^{m} C_{ix(i)}, \end{aligned}$$

which is a disjoint union of elements of $\mathfrak{S}$. Therefore, S^c can be represented as a disjoint union of elements of $\mathfrak{S}$, thus completing the proof that $\mathfrak{S}$ is a semialgebra. $\square$

Problem 7

Let $\mathfrak{B}$ be a σ-field of subsets of Ω and let $Q : \mathfrak{B} \to \Re$ satisfying the following conditions:

(i) Q is finitely additive on $\mathfrak{B}$.

(ii) $Q(\Omega) = 1$ and $0 \leq Q(A) \leq 1$ for all $A \in \mathfrak{B}$.

(iii) If $A_i \in \mathfrak{B}$ are pairwise disjoint and $\sum_{i=1}^{\infty} A_i = \Omega$, then $\sum_{i=1}^{\infty} Q(A_i) = 1$.

Show that Q is σ-additive, so that it is, in fact, a probability measure on $\mathfrak{B}$.

Proof. Let $\{A_n\}$ be a countable sequence of pairwise disjoint elements of $\mathfrak{B}$ and let A denote $\sum_{n \in \mathbb{N}} A_n$. Since $\mathfrak{B}$ is a σ-field, $A \in \mathfrak{B}$, and so $A^c \in \mathfrak{B}$. Now, $\{A_n \mid n \in \mathbb{N}\} \cup \{A^c\}$ is a pairwise disjoint collection of elements of $\mathfrak{B}$ whose union is Ω, so it follows that

$$\begin{aligned}
\sum_{n \in \mathbb{N}} Q(A_n) + Q(A^c) &= 1 && \text{(by property iii)} \\
&= Q(\Omega) && \text{(by property ii)} \\
&= Q(A \cup A^c) \\
&= Q(A) + Q(A^c) && \text{(by property i)} \\
&= Q\left(\sum_{n \in \mathbb{N}} A_n\right) + Q(A^c).
\end{aligned}$$

Thus,

$$\sum_{n \in \mathbb{N}} Q(A_n) = Q\left(\sum_{n \in \mathbb{N}} A_n\right),$$

as desired. □

Problem 1

Proposition 0.2. *Let $X : (\Omega, \mathfrak{F}) \to (\Re, \mathfrak{B})$ where $\mathfrak{B}$ is the Borel σ-field in $\Re$. Let $S : (\Omega, \mathfrak{F}) \to (\mathcal{S}, \mathfrak{A})$ where $\mathfrak{A}$ is a σ-field in $\mathcal{S}$. Denote by $\mathfrak{F}_S = S^{-1}(\mathfrak{A})$ the sub-σ-field induced by S. The function X is $\mathfrak{F}_S/\mathfrak{B}$-measurable if and only if there exists a measurable $h : (\mathcal{S}, \mathfrak{A}) \to (\Re, \mathfrak{B})$ such that $X(\omega) = h[S(\omega)]$ for every $\omega \in \Omega$.*

Proof. ($\Leftarrow$) It follows immediately from the existence of h that

$$\begin{aligned}
X^{-1}(\mathfrak{B}) &= S^{-1}(h^{-1}(\mathfrak{B})) \\
&\subseteq S^{-1}(\mathfrak{A}) \\
&\subseteq \mathfrak{F}_S,
\end{aligned}$$

and so X is $\mathfrak{F}_S/\mathfrak{B}$-measurable. □

Problem 2

Proposition 0.3. *Let $\{X_n : n = 1, 2, 3, \ldots\}$ be a sequence of random variables defined on $(\Omega, \mathfrak{F})$, and let N be a positive integer-valued random variable defined on $(\Omega, \mathfrak{F})$. The function $Y = X_N$ is a random variable.*

Proof. Let $a \in \mathbb{R}$ and consider $Y^{-1}(-\infty, a]$. It follows from the definition of Y that

$$\begin{aligned}
Y^{-1}(-\infty,a] &= \{\omega \in \Omega \mid Y(\omega) \in (-\infty,a]\} \\
&= \{\omega \in \Omega \mid X_N(\omega) \in (-\infty,a]\} \\
&= \{\omega \in \Omega \mid N(\omega) \in (-\infty,a] \text{ and } X_{N(\omega)}(\omega) \in (-\infty,a]\} \\
&= \{\omega \in \Omega \mid N(\omega) \in (-\infty,a]\} \cap \{\omega \in \Omega \mid X_n(\omega) \in (-\infty,a], n \in (-\infty,a] \cap \mathbb{Z}^+\} \\
&= \{\omega \in \Omega \mid N(\omega) \in (-\infty,a]\} \cap \left(\bigcup_{n \in (-\infty,a] \cap \mathbb{Z}^+} \{\omega \in \Omega \mid X_n(\omega) \in (-\infty,a]\} \right).
\end{aligned}$$

Now, N is a random variable, so $\{\omega \in \Omega \mid N(\omega) \in (-\infty,a]\} \in \mathfrak{F}$. Similarly, X_n is a random variable for each $n \in \mathbb{N}$, so each $\{\omega \in \Omega \mid X_n(\omega) \in (-\infty,a]\} \in \mathfrak{F}$. Finally, as $\mathfrak{F}$ is closed under countable union and countable intersection, we see that $Y^{-1}(-\infty,a] \in \mathfrak{F}$, and so Y is a random variable. □

Problem 3

Proposition 0.4. *If X is a random variable, then $|X|$ is also a random variable.*

Proof. Let $a \in \mathbb{R}$ and consider $|X|^{-1}(-\infty,a]$. We know that

$$\begin{aligned}
|X|^{-1}(-\infty,a] &= \{\omega \in \Omega : |X(\omega)| \in (-\infty,a]\} \\
&= \{\omega \in \Omega : |X(\omega)| \in [0,a]\} \\
&= \{\omega \in \Omega : X(\omega) \in [-a,a]\}.
\end{aligned}$$

Now, since X is measurable and $[-a,a] \in \mathfrak{B}$, $|X|^{-1}(-\infty,a] \in \mathfrak{F}$, and so $|X|$ is a random variable. □

Proposition 0.5. *If $|X|$ is a random variable, X need not be a random variable.*

Proof. Let N denote some nonmeasurable subset of $\mathbb{R}$ and define $X : (\mathbb{R}, \mathfrak{B}) \to (\mathbb{R}, \mathfrak{B})$ via

$$X(y) = \begin{cases} -1 & \text{if } y \in N \\ 1 & \text{if } y \notin N. \end{cases}$$

We see that $|X| \equiv 1$, and so

$$|X|^{-1}(-\infty,a] = \begin{cases} \emptyset & \text{if } a < 1 \\ \mathbb{R} & \text{if } a \geq 1. \end{cases}$$

Hence, $|X|$ is a random variable. At the same time, we have $X^{-1}(-\infty,-1] = N$, and so X is not a random variable. □

Problem 4

Proposition 0.6. *Let* $(\Omega, \mathfrak{B}, P)$ *be* $([0,1], \mathfrak{B}(0,1], \lambda)$ *where* λ *is the Lebesgue measure on* $[0,1]$*. Define the process* $\{X_t : 0 \leq t \leq 1\}$ *according to*

$$X_t(\omega) = I\{t = \omega\}.$$

Each X_t *is a random variable.*

Proof. Observe first that the range of each X_t is $\{0,1\}$, so it suffices to consider preimages of the generators of the σ-algebra $\{\emptyset, \{0\}, \{1\}, \{0,1\}\}$, namely $\{0\}$ and $\{1\}$. Let now $s \in [0,1]$ be arbitrary but fixed. We see that

$$X_s^{-1}\{0\} = (0,1] \setminus \{s\}$$
$$X_s^{-1}\{1\} = \{s\}.$$

As both $(0,1] \setminus \{s\}$ and $\{s\}$ belong to $\mathfrak{B}(0,1]$, we conclude that X_s is a random variable.

By the above, we see the σ-field generated by $\{X_t : 0 \leq t \leq 1\}$ consists of all sets A such that A itself or A^c is a countable union of singletons. □

Problem 5

Proposition 0.7. *If* X *and* Y *are random variables on* $(\Omega, \mathfrak{F}, P)$*, then*

$$\sup_{A \in \mathfrak{B}} |P\{X \in A\} - P\{Y \in A\}| \leq P\{X \neq Y\}.$$

Proof. For any $A \in \mathfrak{B}$,

$$\{X \neq Y\} \supseteq X^{-1}(A) \cup Y^{-1}(A) \setminus (X^{-1}(A) \cap Y^{-1}(A)),$$

and so

$$\begin{aligned} |P(X \neq Y)| &\geq |P(X \in A) + P(Y \in A) - P((X \in A) \cap (Y \in A))| \\ &\geq |P(X \in A) - P(Y \in A)|. \end{aligned}$$

As A was arbitrary, we have that $P(X \neq Y) \geq \sup_{A \in \mathfrak{B}} |P(X \in A) - P(Y \in A)|$, as desired. □

Problem 6

Proposition 0.8. *If* $\{A_n : n = 1, 2, 3, \ldots\}$ *is an independent sequence of events, then*

$$P\left\{\bigcap_{n=1}^{\infty} A_n\right\} = \prod_{n=1}^{\infty} P\{A_n\}.$$

Proof. Define for each $n \in \mathbb{N}$ the set $B_n = \bigcap_{k=1}^{n} A_k$. We see that $\{B_n\}$ is a nonincreasing sequence of sets, and thus $\lim_{n\to\infty} B_n = \bigcap_{n\in\mathbb{N}} B_n$. By definition of the B_n, however, we have also that $\bigcap_{n\in\mathbb{N}} B_n = \bigcap_{n\in\mathbb{N}} A_n$. It now follows that

$$\begin{aligned} P\left\{\bigcap_{n=1}^{\infty} A_n\right\} &= P\left\{\lim_{n\to\infty} B_n\right\} \\ &= \lim_{n\to\infty} P(B_n) \\ &= \lim_{n\to\infty} P\left(\bigcap_{k=1}^{n} A_k\right) \\ &= \lim_{n\to\infty} \prod_{k=1}^{n} P(A_k) \\ &= \prod_{n=1}^{\infty} P(A_n). \end{aligned}$$

□

Problem 7

Proposition 0.9. *If X and Y are independent random variables and f, g are measurable and real-valued, then $f(X)$ and $g(Y)$ are independent.*

Proof. Let $A, B \in \mathfrak{B}$. Since f and g are measurable and real-valued, there exist $A', B' \in \mathfrak{B}$ such that $f^{-1}(A) = A'$ and $g^{-1}(B) = B'$. Since X and Y are independent random variables, we have that $X^{-1}(A')$ and $Y^{-1}(B')$ are independent. Thus, we have shown for any $A, B \in \mathfrak{B}$ that $X^{-1}(f^{-1}(A))$ and $Y^{-1}(g^{-1}(B))$ are independent. In other words, $f(X)$ and $g(Y)$ are independent. □

Problem 8

Proposition 0.10. *A random variable X is independent of itself if and only if there is some constant c such that $P\{X = c\} = 1$.*

Proof. (⇒) Choose some event $\omega \in \Omega$ with nonzero probability and set $c = X(\omega)$. Since X is independent of itself, we have

$$\begin{aligned} 0 &= P(\{\omega\} \cap (\Omega \setminus \{\omega\})) \\ &= P(\{\omega\}) P(\Omega \setminus \{\omega\}). \end{aligned}$$

Since $P(\{\omega\}) > 0$, it must be that $P(\Omega \setminus \{\omega\}) = 0$. Thus, we conclude that $P(\{\omega\}) = 1$, and so $P(X = c) = 1$.

(⇐) Let $A, B \in \mathfrak{B}$. We consider three cases.

If $c \in A$ and $c \in B$, then $P(X \in A) = P(X \in B) = 1$. Moreover, $c \in A \cap B$, so $P([X \in A] \cap [X \in B]) = 1$.

If $c \in A$ and $c \notin B$, then $P(X \in A) = 1$ and $P(X \in B) = 0$. Moreover, $c \notin A \cap B$, so $P([X \in A] \cap [X \in B]) = 0$.

If $c \notin A$ and $c \notin B$, then $P(X \in A) = 0$ and $P(X \in B) = 0$. Moreover, $c \notin A \cap B$, so $P([X \in A] \cap [X \in B]) = 0$.

In any case, we have $P([X \in A] \cap [X \in B]) = P(X \in A)P(X \in B)$, and so A and B are independent. Therefore, X is independent of itself. □

Problem 9

Consider the experiment of tossing a fair coin some number of times, so that each of the possible outcomes in the sample space are equally likely.

Proposition 0.11. *There are three events A, B, and C such that every pair are independent, but $P(ABC) \neq P(A)P(B)P(C)$.*

Proof. The desired events can be constructed using two flips. Set

$$\begin{aligned} A &= \{TT, TH\} \\ B &= \{TH, HT\} \\ C &= \{HT, HH\}. \end{aligned}$$

We have $P(A) = P(B) = P(C) = \frac{1}{2}$ and $P(AB) = P(AC) = P(BC) = \frac{1}{4}$, so $P(AB) = P(A)P(B)$, $P(AC) = P(A)P(C)$, and $P(BC) = P(B)P(C)$. That is, the events are pairwise independent. At the same time, we have $P(ABC) = 0$ (since $ABC = \emptyset$), which is not equal to $P(A)P(B)P(C) = \frac{1}{8}$. □

Proposition 0.12. *There are three events A, B, and C such that $P(ABC) = P(A)P(B)P(C)$, but with at least one possible pair not independent.*

Proof. The desired events can be constructed using three flips. Set

$$\begin{aligned} A &= \{TTT, TTH, THT, THH\} \\ B &= \{TTT, TTH, THT, HTT\} \\ C &= \{TTT, THH, HTT, HTH\}. \end{aligned}$$

We have $P(A) = P(B) = P(C) = \frac{1}{2}$ and $P(ABC) = \frac{1}{8}$, so $P(ABC) = P(A)P(B)P(C)$. At the same time, we have $P(AB) = \frac{3}{8}$, which is not equal to $P(A)P(B) = \frac{1}{4}$. □

Proposition 0.13. *There are four events A, B, C, and D such that these events are independent.*

Proof. The desired events can be constructed using four flips. Set

$$\begin{aligned} A &= \{TTTT, TTTH, TTHT, THTT, THTH, THHT, THHH, HTTT\} \\ B &= \{TTTT, TTTH, TTHT, TTHH, THTT, HTTH, HTHT, HHTH\} \\ C &= \{TTTT, TTTH, TTHH, THTH, THHT, HTTH, HTHH, HHTT\} \\ D &= \{TTTT, TTHT, TTHH, THHH, HTTT, HTHT, HTHH, HTTT\}. \end{aligned}$$

For clarity, we list the possible intersections explicitly.

$$\begin{aligned} AB &= \{TTTT, TTTH, TTHT, THTT\} \\ AC &= \{TTTT, TTTH, THTH, THHT\} \\ AD &= \{TTTT, TTHT, THHH, HTTT\} \\ BC &= \{TTTT, TTTH, TTHH, HTTH\} \\ BD &= \{TTTT, TTHT, TTHH, HTHT\} \\ ABC &= \{TTTT, TTTH\} \\ ABD &= \{TTTT, TTHT\} \\ BCD &= \{TTTT, TTHH\} \\ ABCD &= \{TTTT\} \end{aligned}$$

From the above, it is routine to check (but lengthy to write out) that the events A, B, C, D are indepedent. □

Problem 1

Consider then the experiment where a computer generates successive letters independently from the Roman alphabet randomly.

Proposition 0.14. *The string "MOHR" will appear infinitely often with probability one.*

Proof. Break the infinite string generated by the computer into disjoint blocks of four characters each (i.e. the first four characters comprise the first block, the second four characters comprise the second block, and so on). Let A_i be the event that the i^{th} block is the string "MOHR". Note that the A_i are independent since the blocks are disjoint. Now, for all i, $P(A_i) = \frac{1}{26^4}$. Hence,

$$\sum_{n=1}^{\infty} P(A_n) = \infty.$$

By Borel-Cantelli, A_n occurs infinitely often with probability one. That is, the string "MOHR" appears infinitely often with probability one. □

Problem 3

Proposition 0.15. *If $\{A_n\}$ are independent events satisfying $P(A_n) < 1$ for all n, then*

$$P\left\{\bigcup_{n=1}^{\infty} A_n\right\} = 1 \text{ if and only if } P(A_n \text{ i.o.}) = 1.$$

Proof. ($\Rightarrow$) Suppose that

$$P\left\{\bigcup_{n=1}^{\infty} A_n\right\} = 1.$$

It follows that

$$\begin{aligned} 0 &= P\left\{\left(\bigcup_{n=1}^{\infty} A_n\right)^c\right\} \\ &= P\left\{\bigcap_{n=1}^{\infty} A_n^c\right\} \\ &= \prod_{n=1}^{\infty} P(A_n^c) && \text{(by independence)} \\ &= \prod_{n=1}^{\infty} 1 - P(A_n). \end{aligned}$$

Now, since $P(A_n) < 1$ for all n, the fact that $\prod_{n=1}^{\infty} 1 - P(A_n) = 0$ implies that $P(A_n) > 0$ infinitely often. Thus,

$$\begin{aligned} 0 &= \prod_{n=1}^{\infty} e^{-P(A_n)} \\ &= e^{-\sum_{n=1}^{\infty} P(A_n)}. \end{aligned}$$

Therefore,

$$\sum_{n=1}^{\infty} P(A_n) = \infty,$$

and so $P(A_n \text{ i.o.}) = 1$ by Borel-Cantelli.

($\Leftarrow$) Define B_n to be the event $\bigcup_{k \geq n} A_k$. Observe that $\{B_n\}$ is a non-increasing sequence of events. Now,

$$\begin{aligned} 1 &= P(A_n \text{ i.o.}) \\ &= P\left(\bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k\right) \\ &= P\left(\bigcap_{n=1}^{\infty} B_n\right) \\ &= \lim_{n \to \infty} P(B_n). \end{aligned}$$

As the B_n are non-increasing, the sequence $\{P(B_n)\}$ is non-increasing, and so $P(B_n) = 1$ for all n. In particular,

$$\begin{aligned} 1 &= P(B_1) \\ &= P\left(\bigcup_{n=1}^{\infty} A_n\right). \end{aligned}$$

□

In the above proposition, the condition $P(A_n) < 1$ for all n cannot be dropped. To see this, consider the sequence $\{A_n\}$ where $A_1 = \Omega$ and A_n are sets of measure zero for all $n \geq 2$. We have that

$$\begin{aligned} P\left\{\bigcup_{n=1}^{\infty} A_n\right\} &\geq P(\Omega) \\ &= 1, \end{aligned}$$

yet

$$\begin{aligned} P(A_n \text{ i.o.}) &= P\left(\bigcap_{n=1}^{\infty} \bigcup_{k\geq n} A_k\right) \\ &\leq P\left(\bigcap_{n=2}^{\infty} \bigcup_{k\geq n} A_k\right) \\ &= 0 \qquad \text{(since all } A_k \text{ are sets of measure zero).} \end{aligned}$$

Problem 5

In a sequence of independent Bernoulli random variables $\{X_n \mid n = 1, 2, \cdots\}$ with

$$P(X_n = 1) = p = 1 - P(X_n = 0),$$

let A_n be the event that a run of n consecutive 1's occurs between trials 2^n and 2^{n+1}.

Proposition 0.16. *If $p \geq \frac{1}{2}$, then $P(A_n$ i.o.$) = 1$.*

Proof. Break the trials into $\frac{2^n}{n}$ disjoint blocks of length n. The probability of not getting all 1's in a given block is $1 - p^n$, and so the probability of not getting all 1's among any of the blocks is

$$(1 - p^n)^{\frac{2^n}{n}}.$$

Now,

$$\begin{aligned}(1-p^n)^{\frac{2^n}{n}} &\leq \left(e^{-p^n}\right)^{\frac{2^n}{n}} \\ &= e^{-\frac{(2p)^n}{n}}.\end{aligned}$$

Hence,

$$\begin{aligned}P(A_n) &= 1-(1-p^n)^{\frac{2^n}{n}} \\ &\geq 1-e^{-\frac{(2p)^n}{n}}.\end{aligned}$$

We consider two cases. If $p > \frac{1}{2}$,

$$\begin{aligned}\frac{(2p)^n}{n} &= \frac{(1+\epsilon)^n}{n} \qquad\qquad (\text{some } \epsilon > 0) \\ &\to \infty\end{aligned}$$

Thus,

$$\begin{aligned}\sum_{n=1}^{\infty} P(A_n) &\geq \sum_{n=1}^{\infty} 1-e^{-\frac{(2p)^n}{n}} \\ &= \sum_{n=1}^{\infty} 1 - \sum_{n=1}^{\infty} e^{-\frac{(2p)^n}{n}} \\ &= \infty,\end{aligned}$$

and so $P(A_n \text{ i.o.}) = 1$ by Borel-Cantelli.

If $p = \frac{1}{2}$, then we use

$$e^{-\frac{1}{n}} = \sum_{k=0}^{\infty} \frac{\left(-\frac{1}{n}\right)^k}{k!}.$$

Now,

$$\begin{aligned}P(A_n) &= 1 - \sum_{k=0}^{\infty} \frac{\left(-\frac{1}{n}\right)^k}{k!} \\ &= 1 - 1 + \frac{1}{n} - \sum_{k=2}^{\infty} \frac{\left(-\frac{1}{n}\right)^k}{k!} \\ &= \frac{1}{n} - \sum_{k=2}^{\infty} \frac{\left(-\frac{1}{n}\right)^k}{k!}.\end{aligned}$$

Thus,

$$\begin{aligned}\sum_{n=1}^{\infty} P(A_n) &= \sum_{n=1}^{\infty} \left(\frac{1}{n} - \sum_{k=2}^{\infty} \frac{\left(-\frac{1}{n}\right)^k}{k!} \right) \\ &= \sum_{n=1}^{\infty} \frac{1}{n} - \sum_{n=1}^{\infty}\sum_{k=2}^{\infty} \frac{\left(-\frac{1}{n}\right)^k}{k!} \\ &= \infty,\end{aligned}$$

and so $P(A_n \text{ i.o.}) = 1$ by Borel-Cantelli. ☐

Problem 7

Proposition 0.17. *If the event A is independent of the π-system $\mathcal{P}$ and $A \in \sigma(\mathcal{P})$, then $P(A)$ is either 0 or 1.*

Proof. Since A is independent of $\mathcal{P}$, $\sigma(A)$ is independent of $\sigma(\mathcal{P})$ by the Basic Criterion. As $A \in \sigma(\mathcal{P})$, A is independent of itself. Thus,

$$\begin{aligned}P(A) &= P(A \cap A) \\ &= P(A)P(A) \\ &= P(A)^2,\end{aligned}$$

and so $P(A)$ is either 0 or 1. ☐

Problem 9

Proposition 0.18. *If $\sigma(X_1, \ldots, X_{n-1})$ and $\sigma(X_n)$ are independent for all $n \geq 2$, then $\{X_n \mid n = 1, 2, \ldots\}$ is an independent collection of random variables.*

Proof. Let a finite collection $\{A_{i_j} \mid A_{i_j} \in X_{i_j}^{-1}(\mathfrak{B}), j \in [k]\}$ be given. Without loss of generality, suppose $i_\ell < i_m$ for $\ell < m$. Consider the event $A_{i_1} \cap \cdots \cap A_{i_k}$. Since

$$\begin{aligned}A_{i_1} \cap \cdots \cap A_{i_{k-1}} &\in \sigma(X_{i_1}, \ldots, X_{i_{k-1}}) \\ &\subset \sigma(X_1, X_2, \ldots, X_{i_k - 1}),\end{aligned}$$

which is independent from $\sigma(X_{i_k})$, it follows that

$$P(A_{i_1} \cap \cdots \cap A_{i_{k-1}} \cap A_{i_k}) = P(A_{i_1} \cap \cdots \cap A_{i_{k-1}})P(A_{i_k}).$$

Proceeding inductively, we conclude that

$$P(A_{i_1} \cap \cdots \cap A_{i_{k-1}} \cap A_{i_k}) = \prod_{j=1}^{k} P(A_{i_j}),$$

as desired. ☐

Problem 10

Proposition 0.19. *Given a sequence of events* $\{A_n \mid n = 1, 2, \dots\}$ *with* $P(A_n) \to 1$*, there exists a subsequence* $\{n_k\}$ *tending to infinity such that*

$$P\left(\bigcap_k A_{n_k}\right) > 0.$$

Proof. Since $P(A_n) \to 1$, we have that, for all $\epsilon > 0$ there exists N such that $P(A_n) > 1 - \epsilon$ for all $n \geq N$. Choose a sequence of positive ϵ_k for $k \geq 1$ satisfying

$$\sum_k \epsilon_k < 1.$$

(For example, $\epsilon_k = \frac{1}{4^k}$ suffices.) From the above, we can find corresponding n_k for each k such that $P(A_{n_k}) > 1 - \epsilon_k$. Now,

$$\begin{aligned} P\left(\bigcap_k A_{n_k}\right) &= 1 - P\left(\left(\bigcap_k A_{n_k}\right)^c\right) \\ &= 1 - P\left(\bigcup_k A_{n_k}^c\right) \\ &\geq 1 - \sum_k P(A_{n_k}^c) \\ &> 1 - \sum_k \epsilon_k \\ &> 0. \end{aligned}$$

□

Problem 1

Let $(\Omega, \mathfrak{F}, P)$ be a probability space, and let $\{A_n\}$ and A belong to $\mathfrak{F}$. Let X be a random variable defined on this probability space with $X \in L_1$.

Proposition 0.20.

$$\lim_{n\to\infty} \int_{|X|>n} X dP = 0$$

Proof. Define the sequence of random variables X_n for $n \in \mathbb{N}$ via

$$X_n(\omega) = \begin{cases} X(\omega) & \text{if } |X(\omega)| > n \\ 0 & \text{otherwise.} \end{cases}$$

Observe that, for all n,

$$\int_{|X|>n} X dP = \int_\Omega X_n dP.$$

Since

$$\left|\int_\Omega X_n dP\right| \le \int_\Omega |X_n| dP,$$

it suffices to show that $\int_\Omega |X_n|dP \to 0$.

Now, $|X_n| \to 0$ almost everywhere, since $|X|$ is finite almost everywhere. Moreover, $|X_n| \le |X|$ for all n. Thus, by the Dominated Convergence Theorem,

$$\begin{aligned}\lim_{n\to\infty} \int_\Omega |X_n| dP &= \int_\Omega \lim_{n\to\infty} |X_n| dP \\ &= \int_\Omega 0 dP \\ &= 0.\end{aligned}$$

□

Proposition 0.21. *If*

$$\lim_{n\to\infty} P(A_n) = 0,$$

then

$$\lim_{n\to\infty} \int_{A_n} X dP = 0.$$

Proof. Define the sequence of random variables X_n for $n \in \mathbb{N}$ via $X_n = X \cdot 1_{A_n}$. Observe that, for all n,

$$\int_{A_n} X dP = \int_\Omega X_n dP.$$

Since

$$\left|\int_\Omega X_n dP\right| \le \int_\Omega |X_n| dP,$$

it suffices to show that $\int_\Omega |X_n|dP \to 0$.

Now, $|X_n| \to 0$, since $P(A_n) \to 0$. Moreover, $|X_n| \le |X|$ for all n. Thus, by the Dominated Convergence Theorem,

$$\begin{aligned}\lim_{n\to\infty} \int_\Omega |X_n| dP &= \int_\Omega \lim_{n\to\infty} |X_n| dP \\ &= \int_\Omega 0 dP \\ &= 0.\end{aligned}$$

□

Proposition 0.22.

$$\int_A |X| dP = 0$$

if and only if

$$P\{A \cap [X > 0]\} = 0.$$

Proof. Observe first that

$$\int_A |X|dP = \int_{A\cap[|X|>0]} |X|dP + \int_{A\cap[|X|=0]} |X|dP$$
$$= \int_{A\cap[|X|>0]} |X|dP + 0.$$

Thus, $\int_A |X|dP = 0$ if and only if $\int_{A\cap[|X|>0]} |X|dP = 0$ if and only if $P\{A \cap [|X| > 0]\} = 0$ (since we integrate over those $\omega \in \Omega$ for which $|X(\omega)| > 0$). □

Proposition 0.23. *Let $X \in L_2$. If $V(X) = 0$, then $P[X = E(X)] = 1$.*

Proof. Let $\epsilon > 0$. By Chebychev's Inequality,

$$P[|X - E(X)| \geq \epsilon] \leq \frac{V(X)}{\epsilon^2}$$
$$= 0.$$

Equivalently,

$$P[|X - E(X)| < \epsilon] = 1$$

for all $\epsilon > 0$. Therefore,

$$P[X = E(X)] = 1.$$

□

Problem 2

If X and Y are independent random variables and $E(X)$ exists, then, for all $B \in \mathfrak{B}(\mathfrak{R})$,

$$\int_{[Y\in B]} XdP = E(X)P\{Y \in B\}.$$

Proof. For ease of notation, let $A = [Y \in B]$. Thus,

$$\int_{[Y\in B]} XdP = \int_A XdP$$
$$= \int_\Omega X \cdot 1_A dP$$
$$= E(X \cdot 1_A).$$

Now, write $X = X^+ - X^-$. Since each of X^+ and X^- is measurable, we can find sequences of nonneagtive, simple random variables $\{X_n^+\}$ and $\{X_n^-\}$ with $X_n^+ \uparrow X^+$ and $X_n^- \uparrow X^-$. By the Monotone Convergence Theorem, $E(X_n^+) \uparrow E(X^+)$ and $E(X_n^-) \uparrow E(X^-)$. Thus, by linearity of expectation, $E(X_n^+ - X_n^-) \to E(X)$. Moreover, $E(X_n^+ \cdot 1_A - X_n^- \cdot 1_A) \to E(X \cdot 1_A)$. We show next that $E(X_n^+ \cdot 1_A - X_n^- \cdot 1_A) \to E(X)P(A)$ and so conclude that $E(X \cdot 1_A) = E(X)P(A)$.

For each n, we have

$$X_n^+ = \sum_{i=1}^{n} a_i 1_{A_i},$$

where the a_i are constants and the A_i partition Ω. It follows that

$$\begin{aligned} X_n^+ \cdot 1_A &= \sum_{i=1}^{n} a_i 1_{A_i} 1_A \\ &= \sum_{i=1}^{n} a_i 1_{A_i \cap A}, \end{aligned}$$

and so

$$\begin{aligned} E(X_n^+ \cdot 1_A) &= \sum_{i=1}^{n} a_i P(A_i \cap A) \\ &= \sum_{i=1}^{n} a_i P(A_i) P(A) \qquad \text{(by independence of } X \text{ and } Y) \\ &= P(A) \sum_{i=1}^{n} a_i P(A_i) \\ &= P(A) E(X_n^+) \\ &\to P(A) E(X^+). \end{aligned}$$

Similarly, $E(X_n^- \cdot 1_A) \to P(A)E(X^-)$, and so

$$\begin{aligned} E(X_n^+ \cdot 1_A - X_n^- \cdot 1_A) &= E(X_n^+ \cdot 1_A) - E(X_n^- \cdot 1_A) \\ &\to P(A)E(X^+) - P(A)E(X^-) \\ &= P(A)E(X^+ - X^-) \\ &= P(A)E(X), \end{aligned}$$

as desired. □

Problem 3

Proposition 0.24. *For all $n \geq 1$, let X_n and X be uniformly bounded random variables. If*

$$\lim_{n\to\infty} X_n = X,$$

then

$$\lim_{n\to\infty} E|X_n - X| = 0.$$

Proof. The random variable that is identically K belongs to L_1, since

$$\begin{aligned}\int_\Omega K dP &= K \cdot P(\Omega) \\ &= K.\end{aligned}$$

Thus, the identically K random variable is a dominating random variable for the X_n, and so by the Dominated Convergence Theorem, $E|X_n - X| \to 0$. □

Problem 4

On the Lebesgue interval $(\Omega = [0,1], \mathfrak{B}([0,1]), P = \lambda)$ define the random variables

$$X_n = \frac{n}{\log n} I_{0,\frac{1}{n}}.$$

Proposition 0.25. *For X_n defined as above,*

$$\lim_{n\to\infty} X_n = 0$$

and

$$\lim_{n\to\infty} E(X_n) = 0,$$

yet the X_n are unbounded.

Proof. For any $x \in [0,1]$, we can choose N such that $\frac{1}{N} < x$. Thus, $X_N(x) = 0$, since $x \notin [0, \frac{1}{N}]$. As x was arbitrary, we conclude that $X_n \to 0$.

Next, observe that, for all n,

$$\begin{aligned}E(X_n) &= \frac{n}{\log n} \lambda \left[0, \frac{1}{n}\right] \\ &= \frac{n}{\log n} \cdot \frac{1}{n} \\ &= \frac{1}{\log n},\end{aligned}$$

and so $E(X_n) \to 0$.

Finally, $\frac{n}{\log n} \to \infty$, and so the X_n are unbounded. Hence, $X_n \to 0$ and $E(X_n) \to 0$, yet the condition in the Dominated Convergence Theorem fails. □

Problem 5

Proposition 0.26. *Let $X_n \in L_1$ for all $n \geq 1$ satisfying*

$$\sup_n E(X_n) < \infty.$$

If $X_n \uparrow X$, then $X \in L_1$ and $E(X_n) \to E(X)$.

Proof. Since $X_n \uparrow X$, we see that $X_n^+ \uparrow X^+$ and $X_n^- \downarrow X^-$. Since each of X_n^+ and X_n^- belongs to L_1 for all n, we have by the Monotone Convergence Theorem that $E(X_n^+) \to E(X^+)$ and $E(X_n^-) \to E(X^-)$. By linearity of expectation, this implies that $E(X_n) \to E(X)$. Since $\sup_n E(X_n)$ is finite, so is $E(X)$ by uniqueness of limits.

To show that $X \in L_1$, it remains to rule out the case that $E(X^+) = E(X^-) = \infty$ (if only one of them is infinite, then $E(X) = \pm\infty$, but $\sup_n E(X_n) < \infty$). Observe, however, $X_n^- \downarrow X^-$. Thus, if $E(X^-) = \infty$, $E(X_n^-) = \infty$ for all n, contradicting the fact that $X_n^- \in L_1$ for all n. □

Problem 6

Proposition 0.27. *For any positive random variable X,*

$$E(X) = \int_{[0,\infty)} P(X > t)dt.$$

Proof. We may view the area of integration as a subset A of the product space $\Omega \times [0, \infty)$ where

$$A = \{(\omega, t) \mid X(\omega) > t\}.$$

with product measure

$$P' = P \times \mu.$$

Now, by Fubini's Theorem,

$$\begin{aligned}\int_{\Omega\times[0,\infty)} 1_A dP' &= \int_\Omega \int_{[0,\infty)} 1_A(\omega, t)dtdP \\ &= \int_\Omega X(\omega)dP \\ &= E(X).\end{aligned}$$

On the other hand,

$$\begin{aligned}\int_{\Omega\times[0,\infty)} 1_A dP' &= \int_{[0,\infty)} \int_\Omega 1_A(\omega, t)dPdt \\ &= \int_{[0,\infty)} P\{\omega \mid X(\omega) > t\}dt \\ &= \int_{[0,\infty)} P(X > t)dt.\end{aligned}$$

□

Proposition 0.28. *For any positive random variable X and any constant $\alpha > 0$,*

$$E(X^\alpha) = \alpha \int_{[0,\infty)} t^{\alpha-1} P(X > t)dt.$$

Proof. By direct computation, we have $X^\alpha(\omega) = \int_0^{X(\omega)} \alpha t^{\alpha-1} dt$. It follows that

$$\begin{aligned}
E(X^\alpha) &= \int_\Omega X(\omega) P(d\omega) \\
&= \int_\Omega \int_0^{X(\omega)} \alpha t^{\alpha-1} dt P(d\omega) \\
&= \int_0^\infty \int_{\{P(X>t)\}} \alpha t^{\alpha-1} P(d\omega) dt \qquad \text{(by Fubini's Theorem)} \\
&= \int_0^\infty \alpha t^{\alpha-1} P(X > t) dt.
\end{aligned}$$

□

Problem 7

Proposition 0.29. *Let X be a nonnegative random variable and let $\delta > 0$, $0 < \beta < 1$, and C be constants. If*

$$P\{X > n\delta\} \leq C\beta^n$$

for all $n \geq 1$, then $E(X^\alpha) < \infty$ for all $\alpha > 0$.

Proof. By the previous problem, it is equivalent to show that $\alpha \int_{[0,\infty)} t^{\alpha-1} P(X > t) dt$ is finite.

To begin, pick N such that $t^{\alpha-1}P(X > t)$ is strictly decreasing in t for all $t \geq N$. Such an N exists, as $t^{\alpha-1}$ is a polynomial in t and $P(X > t)$ decays exponentially. It follows that

$$\begin{aligned}
\alpha \int_{[N,\infty)} t^{\alpha-1} P(X > t) dt &\leq \alpha \sum_{n=N}^\infty \delta \cdot (n\delta)^{\alpha-1} P(X > n\delta) \quad \text{(since } t^{\alpha-1}P(X > t) \text{ is strictly decreasing)} \\
&\leq \alpha\delta^\alpha \sum_{n=N}^\infty n^{\alpha-1} C\beta^n \\
&= C\alpha\delta^\alpha \sum_{n=N}^\infty n^{\alpha-1} \beta^n \\
&< \infty \qquad \text{(by the ratio test).}
\end{aligned}$$

Since $\int_{[0,N]} t^{\alpha-1} P(X > t) dt$ is also finite, we conclude that $\int_{[0,\infty]} t^{\alpha-1} P(X > t) dt$ is finite. □

Problem 2

Let $\{X_n \mid n = 1, 2, \dots\}$ be a sequence of random variables with

$$P\{X_n = \pm n^3\} = \frac{1}{2n^2} \text{ and } P\{X_n = 0\} = 1 - \frac{1}{n^2}.$$

Proposition 0.30. *For the sequence described above,*

$$P\left\{\lim_{n\to\infty} X_n = 0\right\} = 1.$$

Proof. Let A_n be the event that X_n is nonzero. Formally,

$$A_n = \{\omega \in \Omega \mid X_n(\omega) \neq 0\}.$$

We see that

$$\begin{aligned} P(A_n) &= 1 - P\{X = 0\} \\ &= 1 - \left(1 - \frac{1}{n^2}\right) \\ &= \frac{1}{n^2}, \end{aligned}$$

and so

$$\begin{aligned} \sum_{n=1}^{\infty} P(A_n) &= \sum_{n=1}^{\infty} \frac{1}{n^2} \\ &< \infty. \end{aligned}$$

Thus, by Borel-Cantelli,

$$P([A_n \text{i.o}]) = P\left(\limsup_{n\to\infty} A_n\right) = 0.$$

By taking complements, we have

$$\begin{aligned} 1 &= P\left(\limsup_{n\to\infty} A_n^c\right) \\ &= P\left(\liminf_{n\to\infty} [X_n = 0]\right), \end{aligned}$$

and so

$$P\left(\lim_{n\to\infty} X_n = 0\right) = 1.$$

□

Proposition 0.31. *For the sequence described above,* $\lim_{n\to\infty} \mathbb{E}(X_n)$ *is either* $\pm\infty$ *or is undefined.*

Proof. For each n, $P\{X_n = \pm n^3\} = \frac{1}{2n^2}$. Hence,

$$P\{X_n^+ = n^3\} \geq \frac{1}{4n^2}$$

or

$$P\{X_n^- = n^3\} \geq \frac{1}{4n^2}$$

(possibly both). Suppose the former is true. It follows that

$$\begin{aligned}\mathbb{E}(X_n^+) &\geq n^3 P\{X_n^+ = n^3\}\\ &\geq n^3 \frac{1}{4n^2}\\ &= \frac{n}{4},\end{aligned}$$

and so

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}(X_n^+) &\geq \lim_{n\to\infty} \frac{n}{4}\\ &= \infty.\end{aligned}$$

Similarly, if $P\{X_n^- = n^3\} \geq \frac{1}{4n^2}$, then $\lim_{n\to\infty} \mathbb{E}(X_n^-) = \infty$. Therefore,

$$\lim_{n\to\infty} \mathbb{E}(X_n) = \begin{cases} \infty & \text{if } \lim_{n\to\infty} \mathbb{E}(X_n^+) = \infty \text{ and } \lim_{n\to\infty} \mathbb{E}(X_n^-) < \infty \\ -\infty & \text{if } \lim_{n\to\infty} \mathbb{E}(X_n^+) < \infty \text{ and } \lim_{n\to\infty} \mathbb{E}(X_n^-) = \infty \\ \text{undefined} & \text{if } \lim_{n\to\infty} \mathbb{E}(X_n^+) = \infty \text{ and } \lim_{n\to\infty} \mathbb{E}(X_n^-) < \infty. \end{cases}$$

□

Problem 3

Let $(\Omega, \mathfrak{F}, P)$ be a probability space.

Definition 0.32. *Two random variables X and Y are said to be independent provided that, for any $A, B \in \mathfrak{B}(\mathfrak{R})$,*

$$P[X^{-1}(A) \cdot Y^{-1}(B)] = P(X^{-1}(A)) \cdot P(Y^{-1}(B)).$$

Proposition 0.33. *Two random variables X and Y are independent if and only if, for every pair f and g of non-negative continuous functions on $(\mathfrak{R}, \mathfrak{B}(\mathfrak{R}))$,*

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)].$$

Proof. ($\Rightarrow$) Let f and g be any non-negative continuous functions on $(\mathfrak{R}, \mathfrak{B}(\mathfrak{R}))$. Since continuous functions between metric spaces are measurable (Resnick, 3.2.3), f and g are measurable. Since the composition of measurable functions is measurable (Resnick, 3.2.2), $f(X)$ and $g(Y)$ are measurable. Now, $\sigma(f(X)) \subseteq \sigma(X)$ (since $f \in \mathfrak{B}(\mathfrak{R})/\mathfrak{B}(\mathfrak{R})$) and $\sigma(g(Y)) \subset \sigma(Y)$ (since $g \in \mathfrak{B}(\mathfrak{R})/\mathfrak{B}(\mathfrak{R})$). Hence, $f(X)$ and $g(Y)$ are independent, since X and Y are independent.

Define now $Z_1 = f(X)$ and $Z_2 = g(Y)$. By the above, Z_1 and Z_2 are independent random variables. Thus, by Fubini's Theorem (as in Resnick, 5.9.2),

$$\mathbb{E}(Z_1 Z_2) = \mathbb{E}(Z_1)\mathbb{E}(Z_2),$$

but this is precisely

$$\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y)).$$

(⇐) (Idea) Let a and b be real numbers and take $f = 1_{(0,a]}$ and $g = 1_{(0,b]}$. The support of $f(X)$ is $\{\omega \mid X(\omega) \leq a\}$ and the support of $g(Y)$ is $\{\omega \mid Y(\omega)\}$. Thus, the measures of the supports are $P(X \leq a)$ and $P(Y \leq b)$, respectively. Using the fact that

$$\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y)),$$

I would like to derive that

$$P(X \leq a, Y \leq b) = P(X \leq a) \cdot P(Y \leq b).$$

Since a and b were arbitrary, we could conclude that X and Y are independent by the Factorization Criterion (Resnick, 4.2.1). Perhaps this may be accomplished by looking at the appropriate approximations of $f(X)$ and $g(Y)$ by simple functions (where the probabilty of the support becomes more evident in the computation). □

For each n, let X_n and Y_n be a pair of independent random variables and define

$$\lim_{n\to\infty} X_n = X \text{ and } \lim_{n\to\infty} Y_n = Y.$$

Proposition 0.34. *The functions X and Y are independent random variables.*

Proof. (Idea) We have, for each n and for all continuous, non-negative f and g,

$$\mathbb{E}(f(X_n)g(Y_n)) = \mathbb{E}(f(X_n))\mathbb{E}(g(Y_n)).$$

If we could switch limits with integrals, we would have

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}(f(X_n)g(Y_n)) &= \lim_{n\to\infty} \mathbb{E}(f(X_n))\mathbb{E}(g(Y_n)) \\
\mathbb{E}\left(\lim_{n\to\infty} f(X_n)g(Y_n)\right) &= \mathbb{E}\left(\lim_{n\to\infty} f(X_n)\right)\mathbb{E}\left(\lim_{n\to\infty} g(Y_n)\right) \\
\mathbb{E}(f(X)g(Y)) &= \mathbb{E}(f(X))\mathbb{E}(g(Y)),
\end{aligned}$$

where the last step makes use of the continuity of f and g. Thus, appealing again to part b, we could conclude that X and Y are independent. I fail to see how to accomplish the interchange, however, as the X_n need not be monotone nor does there appear to be any bounding function. □

Problem 5

Suppose $\{p_k \mid k \geq 0\}$ is a probability mass function on $(\Omega = \{0, 1, 2, \dots\}, \mathfrak{P} = \mathfrak{P}(\Omega))$, where $\mathfrak{P}(\cdot)$ denotes the power set, so that $p_k \geq 0$ and $\sum_k p_k = 1$. Define for all $A \subset \Omega$,

$$P(A) = \sum_{k\in A} p_k.$$

Proposition 0.35. *The function P defined above is a probability measure on $(\Omega, \mathfrak{P})$.*

Proof. Since $p_k \geq 0$ for all k, $P(A) \geq 0$ for all $A \subset \Omega$.

We have, by definition of the probability mass function,

$$\begin{aligned} P(\Omega) &= \sum_{k \in \Omega} p_k \\ &= 1. \end{aligned}$$

Let $\{A_n\}$ be a countable sequence of disjoint events and let $A = \bigcup \{A_n\}$. It follows that

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n\right) &= P(A) \\ &= \sum_{k \in A} p_k \\ &= \sum_{n=1}^{\infty} \sum_{k \in A_n} p_k \qquad \text{(since the } A_n \text{ are disjoint)} \\ &= \sum_{n=1}^{\infty} P(A_n). \end{aligned}$$

□

Define the generating function $\Psi : ([0,1], \mathfrak{B}[0,1]) \to (\mathfrak{R}, \mathfrak{B})$ via

$$\Psi(s) = \sum_{k=0}^{\infty} p_k s^k.$$

Proposition 0.36. *The function Ψ defined above satisfies*

$$\Psi'(s) \equiv \frac{d}{ds}\Psi(s) = \sum_{k=1}^{\infty} k p_k s^{k-1}$$

for $0 \leq s \leq 1$.

Proof. Define the function $X_n = \sum_{k=1}^{n} k p_k s^{k-1}$. Observe that $0 \leq X_n \uparrow X$, and so by the Monotone Convergence Theorem

$$\begin{aligned} \lim_{n\to\infty} \mathbb{E}(X_n) &= \mathbb{E}\left(\lim_{n\to\infty} X_n\right) \\ \lim_{n\to\infty} \mathbb{E}\left(\sum_{k=1}^{n} k p_k s^{k-1}\right) &= \mathbb{E}\left(\lim_{n\to\infty} \sum_{k=1}^{n} k p_k s^{k-1}\right) \\ \lim_{n\to\infty} \sum_{k=0}^{n} p_k s^k &= \mathbb{E}\left(\lim_{n\to\infty} \sum_{k=1}^{n} k p_k s^{k-1}\right) \\ \Psi(s) &= \mathbb{E}\left(\Psi'(s)\right). \end{aligned}$$

□

Proposition 0.37. *If X has probability measure P, then $\mathbb{E}(X) = \lim_{s\uparrow 1} \Psi'(s)$.*

Proof. We have,

$$\begin{aligned}
\mathbb{E}(X) &= \int_{\Omega} X(\omega) dP \\
&= \sum_{k=0}^{\infty} kP(X = k) \\
&= \sum_{k=1}^{\infty} kp_k \\
&= \lim_{s\uparrow 1} \Psi'(s).
\end{aligned}$$

□

Problem 6

Let $X_1, X_2, \ldots, X_n \in L_2(P)$ be random variables defined on a probability space $(\Omega, \mathfrak{F}, P)$. For each $i, j \in \{1, 2, \ldots, n\}$, define the covariances

$$\sigma_{ij} = \mathbb{C}(X_i, X_j) = \mathbb{E}\{[X_i - \mu_i][X_j - \mu_j]\},$$

where

$$\mu_i = \mathbb{E}(X_i) \text{ and } \sigma_i^2 = \sigma_{ii} = \mathbb{V}(X_i) = \mathbb{E}[(X_i - \mu_i)^2].$$

Lemma 0.38. *For any random variable X and real numbers a and b,*

$$\mathbb{V}(aX + b) = a^2 \mathbb{V}(X).$$

Proof. It follows from the linearity of expectation that,

$$\begin{aligned}
\mathbb{V}(aX + b) &= \mathbb{E}[(aX + b - \mathbb{E}(aX + b))^2] \\
&= \mathbb{E}[(aX + b - a\mathbb{E}(X) - b)^2] \\
&= \mathbb{E}[(a(X - \mathbb{E}(X)))^2] \\
&= a^2 \mathbb{E}[(X - \mathbb{E}(X))^2] \\
&= a^2 \mathbb{V}(X).
\end{aligned}$$

□

Lemma 0.39. *For any random variables X and Y,*

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + 2\mathbb{C}(X, Y) + \mathbb{V}(Y).$$

Proof. It follows from the linearity of expectation that,

$$\begin{aligned}\mathbb{V}(X+Y) &= \mathbb{E}[(X+Y-\mathbb{E}(X)-\mathbb{E}(Y))^2] \\ &= \mathbb{E}[((X-\mathbb{E}(X))+(Y-\mathbb{E}(Y)))^2] \\ &= \mathbb{E}[(X-\mathbb{E}(X))^2+2(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))+(Y-\mathbb{E}(Y))^2] \\ &= \mathbb{E}[(X-\mathbb{E}(X))^2]+\mathbb{E}[2(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))]+\mathbb{E}[(Y-\mathbb{E}(Y))^2] \\ &= \mathbb{V}(X)+2\mathbb{C}(X,Y)+\mathbb{V}(Y).\end{aligned}$$

□

Proposition 0.40. *For all i and j,*

$$\sigma_{ij} \le |\sigma_{ij}| \le \sigma_i\sigma_j.$$

Moreover, $|\sigma_{ij}| = \sigma_i\sigma_j$ if and only if, for some α and β, we have $P\{X_j = \alpha + \beta X_i\} = 1$.

Proof. For all real numbers x, we have $x \le |x|$, so certainly $\sigma_{ij} \le |\sigma_{ij}|$.

For the second inequality, let t be a real variable. It follows from the lemmas that

$$\begin{aligned}0 &\le \mathbb{V}[tX_i + X_j] \\ &= \mathbb{V}(tX_i) + 2\mathbb{C}(X_i, X_j) + \mathbb{V}(X_j) \\ &= \sigma_i^2 t^2 + 2\sigma_{ij} t + \sigma_j^2.\end{aligned}$$

Viewing this as a non-negative quadratic in t, we have that

$$0 \ge 4\sigma_{ij}^2 - 4\sigma_i^2\sigma_j^2,$$

and so

$$|\sigma_{ij}| \le \sigma_i\sigma_j.$$

For the remaining claim, observe that

$$\begin{aligned}|\sigma_{ij}| = \sigma_i\sigma_j &\Leftrightarrow \sigma_{ij}^2 = \sigma_i^2\sigma_j^2 \\ &\Leftrightarrow 0 = 4\sigma_{ij}^2 - 4\sigma_i^2\sigma_j^2.\end{aligned}$$

Hence, $\mathbb{V}[tX_i + X_j]$ has a unique real root t_0. Now, the variance of a random variable is equal to 0 if and only if it is constant with probability one. That is,

$$P\{t_0X_i + X_j = \alpha\} = 1,$$

or equivalently

$$P\{X_j = \alpha - t_0X_i\} = 1.$$

□

Proposition 0.41. *For real constants* α_i *and* β_i, $i = 1, 2, \ldots, n$,

$$\mathbb{C}\left\{\sum_{i=1}^{n} \alpha_i X_i, \sum_{j=1}^{n} \beta_j X_j\right\} = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \beta_j \sigma_{ij}.$$

Proof. Applying linearity of expectation, we have

$$\begin{aligned}
\mathbb{C}\left\{\sum_{i=1}^{n} \alpha_i X_i, \sum_{j=1}^{n} \beta_j X_j\right\} &= \mathbb{E}\left\{\left(\sum_{i=1}^{n} \alpha_i X_i\right)\left(\sum_{j=1}^{n} \beta_j X_j\right)\right\} - \mathbb{E}\left\{\sum_{i=1}^{n} \alpha_i X_i\right\}\mathbb{E}\left\{\sum_{j=1}^{n} \beta_j X_j\right\} \\
&= \mathbb{E}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \beta_j X_i X_j\right\} - \mathbb{E}\left\{\sum_{i=1}^{n} \alpha_i X_i\right\}\mathbb{E}\left\{\sum_{j=1}^{n} \beta_j X_j\right\} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \beta_j \mathbb{E}(X_i X_j) - \left\{\sum_{i=1}^{n} \alpha_i \mathbb{E}(X_i)\right\}\left\{\sum_{j=1}^{n} \beta_j \mathbb{E}(X_j)\right\} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \beta_j \mathbb{E}(X_i X_j) - \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \beta_j \mathbb{E}(X_i)\mathbb{E}(X_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \beta_j \mathbb{C}(X_i, X_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \beta_j \sigma_{ij}.
\end{aligned}$$

□

Proposition 0.42. *For real constants* α_i, $i = 1, 2, \ldots, n$,

$$\mathbb{V}\left\{\sum_{i=1}^{n} \alpha_i X_i\right\} = \sum_{i=1}^{n} \alpha_i^2 \sigma_i^2 + 2 \sum_{1 \le i < j \le n} \alpha_i \alpha_j \sigma_{ij}.$$

Proof. By definition, $\mathbb{V}(X) = \mathbb{C}(X, X)$ for any random variable X. Thus, by the previous proposition,

$$\begin{aligned}\mathbb{V}\left\{\sum_{i=1}^{n}\alpha_i X_i\right\} &= \mathbb{C}\left\{\sum_{i=1}^{n}\alpha_i X_i, \sum_{i=1}^{n}\alpha_i X_i\right\} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j \mathbb{C}(X_i, X_j) \\ &= \sum_{i=1}^{n}\alpha_i^2 \mathbb{C}(X_i, X_i) + 2\sum_{1\le i<j\le n}\alpha_i\alpha_j\mathbb{C}(X_i, X_j) \\ &= \sum_{i=1}^{n}\alpha_i^2 \mathbb{V}(X_i) + 2\sum_{1\le i<j\le n}\alpha_i\alpha_j\mathbb{C}(X_i, X_j) \\ &= \sum_{i=1}^{n}\alpha_i^2\sigma_i^2 + 2\sum_{1\le i<j\le n}\alpha_i\alpha_j\sigma_{ij}.\end{aligned}$$

□

Proposition 0.43. *Let $X_1, \ldots, X_n$ be independent random variables. Furthermore, suppose that the constants α_i are restricted to belong to $[0, 1]$ and must satisfy $\sum_{i=1}^{n}\alpha_i = 1$. Finally, letting $s = \sum_{i=1}^{n}\sigma_i^{-1}$,*

$$\mathbb{V}\left\{\sum_{i=1}^{n}\alpha_i X_i\right\} \ge ns^{-2}$$

with equality if and only if $\alpha_i = (\sigma_i s)^{-1}$ for all i.

Proof. Since the X_i are independent, $\sigma_{ij} = 0$ for all $i \neq j$ (Resnick, 5.9.2). Thus, the expression for the variance reduces to

$$\mathbb{V}\left\{\sum_{i=1}^{n}\alpha_i X_i\right\} = \sum_{i=1}^{n}(\alpha_i\sigma_i)^2.$$

Now, since we require $\sum_{i=1}^{n}\alpha_i = 1$, the variance will be minimized precisely when all the terms in its summation are equal. To that end, tentatively set $\alpha_i = \sigma_i^{-1}$ for all i. Thus, $\alpha_i\sigma_i = 1$ for all i, and so all the terms are equal, as desired. Now,

$$\begin{aligned}\sum_{i=1}^{n}\alpha_i &= \sum_{i=1}^{n}\sigma_i^{-1} \\ &= s.\end{aligned}$$

To ensure that the α_i indeed sum to 1, we scale them all by a factor of s^{-1}.

Thus, we take instead $\alpha_i = (\sigma_i s)^{-1}$ for all i. Using these values, we have

$$\begin{aligned}\mathbb{V}\left\{\sum_{i=1}^{n} \alpha_i X_i\right\} &= \sum_{i=1}^{n} (\alpha_i \sigma_i)^2 \\ &= \sum_{i=1}^{n} ((\sigma_i s)^{-1} \sigma_i)^2 \\ &= \sum_{i=1}^{n} s^{-2} \\ &= n s^{-2}.\end{aligned}$$

□

Problem 8

For $i = 1, 2$, let $(\Omega_i, \mathfrak{B}_i, P_i)$ be probability spaces. Define $\Omega = \Omega_1 \times \Omega_2$ and $\mathfrak{B} = \mathfrak{B}_1 \otimes \mathfrak{B}_2 = \sigma(RECTS)$, where

$$RECTS = \{B_1 \times B_2 \mid B_1 \in \mathfrak{B}_1, B_2 \in \mathfrak{B}_2\}.$$

Let $P = P_1 \times P_2$ be the product probability measure so that, for $B_1 \times B_2 \in RECTS$, we have $P(B_1 \times B_2) = P_1(B_1)P_2(B_2)$. Define the class of subsets

$$\mathfrak{C} = \left\{B \subset \Omega \mid \int_{\Omega} 1_B(\omega_1, \omega_2) dP(\omega_1, \omega_2) = \int_{\Omega_1} Y(\omega_1) dP_1(\omega_1)\right\},$$

where $Y(\omega_1) = \int_{\Omega_2} 1_B(\omega_1, \omega_2) dP_2(\omega_2)$.

Proposition 0.44. *The class $RECTS$ is a subset of the class $\mathfrak{C}$.*

Proof. Let $B = B_1 \times B_2$ belong to $RECTS$. We have

$$\begin{aligned}\int_{\Omega} 1_B(\omega_1, \omega_2) dP(\omega_1, \omega_2) &= \int_B dP(\omega_1, \omega_2) \\ &= P(B).\end{aligned}$$

At the same time, we have

$$\begin{aligned}\int_{\Omega_1} \int_{\Omega_2} 1_B(\omega_1, \omega_2) dP_2(\omega_2) dP_1(\omega_1) &= \int_{\Omega_1} \int_{\Omega_2} 1_{B_1}(\omega_1) 1_{B_2}(\omega_2) dP_2(\omega_2) dP_1(\omega_1) \\ &= \int_{\Omega_1} \int_{B_2} 1_{B_1}(\omega_1) dP_2(\omega_2) dP_1(\omega_1) \\ &= \int_{\Omega_1} 1_{B_1}(\omega_1) P_2(B_2) dP_1(\omega_1) \\ &= \int_{B_1} P_2(B_2) dP_1(\omega_1) \\ &= P_1(B_1) P_2(B_2) \\ &= P(B).\end{aligned}$$

Hence, for all $B \in RECTS$,

$$\int_{\Omega} 1_B(\omega_1, \omega_2) dP(\omega_1, \omega_2) = \int_{\Omega_1} Y(\omega_1) dP_1(\omega_1),$$

and so $RECTS \subseteq \mathfrak{C}$. □

Proposition 0.45. *The class $\mathfrak{C}$ is a λ-system.*

Proof. We have immediately that $\Omega = \Omega_1 \times \Omega_2 \in RECTS \subseteq \mathfrak{C}$, so $\Omega \in \mathfrak{C}$.

Next, let $B \in \mathfrak{C}$. We have

$$\begin{aligned}
\int_{\Omega} 1_{B^c}(\omega_1, \omega_2) dP(\omega_1, \omega_2) &= \int_{B^c} dP(\omega_1, \omega_2) \\
&= P(B^c) \\
&= 1 - P(B) \\
&= 1 - \int_{\Omega_1} \int_{\Omega_2} 1_{B_{\omega_1}}(\omega_2) dP_1(\omega_1) && \text{(since } B \in \mathfrak{C}) \\
&= \int_{\Omega_1} \int_{\Omega_2} 1 - 1_{B_{\omega_1}}(\omega_2) dP_1(\omega_1) \\
&= \int_{\Omega_1} \int_{\Omega_2} 1_{(B_{\omega_1})^c}(\omega_2) dP_1(\omega_1) \\
&= \int_{\Omega_1} \int_{\Omega_2} 1_{(B^c)_{\omega_1}}(\omega_2) dP_1(\omega_1) \\
&= \int_{\Omega_1} \int_{\Omega_2} 1_{B^c}(\omega_1, \omega_2) dP_2(\omega_2) dP_1(\omega_1).
\end{aligned}$$

Finally, let $\{B_n \mid n = 1, 2, \dots\}$ be a collection of disjoint elements of $\mathfrak{C}$. We have

$$\begin{aligned}
\int_{\Omega} 1_{\sum_{n=1}^{\infty} A_n} dP(\omega_1, \omega_2) &= \int_{\sum_{n=1}^{\infty} A_n} dP(\omega_1, \omega_2) \\
&= P\left(\sum_{n=1}^{\infty} A_n\right) \\
&= \sum_{n=1}^{\infty} P(A_n) \\
&= \sum_{n=1}^{\infty} \int_{\Omega_1} \int_{\Omega_2} 1_{(A_n)_{\omega_1}}(\omega_2) dP_1(\omega_1) && \text{(since } A_n \in \mathfrak{C} \text{ for all } n) \\
&= \int_{\Omega_1} \int_{\Omega_2} \sum_{n=1}^{\infty} 1_{(A_n)_{\omega_1}}(\omega_2) dP_1(\omega_1) && \text{(by MCT)} \\
&= \int_{\Omega_1} \int_{\Omega_2} 1_{(\sum_{n=1}^{\infty} A_n)_{\omega_1}}(\omega_2) dP_1(\omega_1) \\
&= \int_{\Omega_1} \int_{\Omega_2} 1_{\sum_{n=1}^{\infty} A_n}(\omega_1, \omega_2) dP_2(\omega_2) dP_1(\omega_1).
\end{aligned}$$

Therefore, $\mathfrak{C}$ is a λ-system. □

Proposition 0.46. *For every $B \in \mathfrak{B}$,*

$$\int_{\Omega} 1_B(\omega_1, \omega_2) dP(\omega_1, \omega_2) = \int_{\Omega_1} \left\{ \int_{\Omega_2} 1_B(\omega_1, \omega_2) dP_2(\omega_2) \right\} dP_1(\omega_1).$$

Proof. We have shown $RECTS \subseteq \mathfrak{C}$ and that $\mathfrak{C}$ is a λ-system. If we can show also that $RECTS$ is a π-system, then Dynkin's Theorem gives $\mathfrak{B} = \sigma(RECTS) \subset \mathfrak{C}$, from which the conclusion follows.

To finish the proof, let $B_1 \times B_2$ and $B_1' \times B_2'$ belong to $RECTS$. It follows immediately that

$$(B_1 \times B_2) \cap (B_1' \times B_2') = (B_1 \cap B_1') \times (B_2 \cap B_2').$$

Since $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are closed under intersections, $B_1 \cap B_1' \in \mathfrak{B}_1$ and $B_2 \cap B_2' \in \mathfrak{B}_2$, and so $(B_1 \cap B_1') \times (B_2 \cap B_2') \in RECTS$. □

To establish the more general result where 1_B in part c is replaced with any $\mathfrak{B}$-measurable positive random variable X, we first establish the result for simple functions of the form $X_n = \sum_{i=1}^{n} a_i 1_{B_i}$, where $B_i \in \mathfrak{B}$ for all i. The result for simple functions follows readily from the linearity of the integral. Since each X_n is positive, we can take a sequence $X_n \uparrow X$. By hypothesis,

$$\int_{\Omega} X_n(\omega_1, \omega_2) dP(\omega_1, \omega_2) = \int_{\Omega_1} \left\{ \int_{\Omega_2} X_n(\omega_1, \omega_2) dP_2(\omega_2) \right\} dP_1(\omega_1)$$

for all n. Applying the Monotone Convergence Theorem, we can conclude that

$$\int_{\Omega} X_n(\omega_1, \omega_2) dP(\omega_1, \omega_2) \uparrow \int_{\Omega} X(\omega_1, \omega_2) dP(\omega_1, \omega_2)$$

and

$$\int_{\Omega_1} \left\{ \int_{\Omega_2} X_n(\omega_1, \omega_2) dP_2(\omega_2) \right\} dP_1(\omega_1) \uparrow \int_{\Omega_1} \left\{ \int_{\Omega_2} X(\omega_1, \omega_2) dP_2(\omega_2) \right\} dP_1(\omega_1),$$

from which it follows that

$$\int_{\Omega} X(\omega_1, \omega_2) dP(\omega_1, \omega_2) = \int_{\Omega_1} \left\{ \int_{\Omega_2} X(\omega_1, \omega_2) dP_2(\omega_2) \right\} dP_1(\omega_1).$$

Problem 10

Suppose that X and Y are independent random variables and let $h : \mathfrak{R}^2 \to [0, \infty)$ be a measurable function such that $\mathbb{E}\{h^2(X, Y)\} < \infty$. Define

$$g(x) = \mathbb{E}\{h(x, Y)\} \text{ and } k(x) = \mathbb{V}\{h(x, Y)\}.$$

Proposition 0.47. *The functions g and k are both measureable on $\mathfrak{R} \to \mathfrak{R}$.*

Proof. Define $\hat{h}(x,\omega) = h(x, Y(\omega))$. Since h and Y are measurable, $\hat{h}$ is measurable, as it is defined by the composition of two measurable functions. Hence, we can take a collection $\{\hat{h}_n\}$ of simple functions with $\hat{h}_n \uparrow \hat{h}$. Define now $g_n(x) = \int_\Omega \hat{h}_n(x,\omega)dP(\omega)$ for $n = 1, 2, \ldots$. By the Monotone Convergence Theorem, $g_n \uparrow g$. To conclue that g is measurable, it remains to show that each g_n is simple. To that end, observe that

$$\begin{aligned}
g_n(x) &= \int_\Omega \hat{h}_n(x,\omega)dP(\omega) \\
&= \int_\Omega \sum_{j=1}^{k} a_j 1_{A_j}(x,\omega)dP(\omega) \qquad \text{(constants } a_j \text{ and } \{A_j\} \text{ a partition of } \mathfrak{R}) \\
&= \sum_{j=1}^{k} \int_\Omega a_j 1_{A_j}(x,\omega)dP(\omega) \qquad \text{(by MCT)} \\
&= \sum_{j=1}^{k} \int_\Omega a_j 1_{A_j}(x) 1_{A_j}(\omega)dP(\omega) \\
&= \sum_{j=1}^{k} a_j P(A_j) 1_{A_j}(x),
\end{aligned}$$

and so g_n is simple.

For $k(x)$, we have

$$\begin{aligned}
k(x) &= \mathbb{V}(\hat{h}) \\
&= \mathbb{E}(\hat{h}^2) - \mathbb{E}(\hat{h})^2 \\
&= \mathbb{E}(\hat{h}^2) - g^2.
\end{aligned}$$

Now, since $\hat{h}$ is measureable, $\hat{h}^2$ is measurable. Following the same argument as above, we find that $\mathbb{E}(\hat{h}^2)$ is measurable, and so k is measurable. □

Proposition 0.48. *For g and h as defined above,*

$$\mathbb{E}\{g(X)\} = \mathbb{E}\{h(X,Y)\}.$$

Proof. Suppose X is a random variable on Ω_1 with probability measure P_1 and Y is a random variable on Ω_2 with probability measure P_2. Finally, let P be the probability measure on $\Omega = \Omega_1 \times \Omega_2$ induced by P_1 and P_2. In order to make use of Fubini's Theorem later in the proof, we must establish first that $P = P_1 \times P_2$. To that end, observe that for any measurable sets $A \subset \Omega_1$ and

$B \subset \Omega_2$,

$$
\begin{aligned}
P(A \times B) &= \int_\Omega 1_{A\times B} dP \\
&= \int_\Omega 1_A \cdot 1_B dP \\
&= \int_\Omega 1_A dP \cdot \int_\Omega 1_B dP && \text{(since } X \text{ and } Y \text{ are independent)} \\
&= \int_{\Omega_1} 1_A(\omega_1) dP_1(\omega_1) \cdot \int_{\Omega_2} 1_B(\omega_2) dP_2(\omega_2) \\
&= P_1(A) \cdot P_2(B).
\end{aligned}
$$

Now,

$$
\begin{aligned}
\mathbb{E}(g(X)) &= \int_{\Omega_1} g(X(\omega_1)) dP_1(\omega_1) \\
&= \int_{\Omega_1} \int_{\Omega_2} h(X(\omega_1), Y(\omega_2)) dP_2(\omega_2) dP_1(\omega_1) \\
&= \int_\Omega h(X(\omega_1), Y(\omega_2)) dP(\omega_1, \omega_2) && \text{(by Fubini's Theorem)} \\
&= \mathbb{E}(h(X,Y)).
\end{aligned}
$$

□

Proposition 0.49. *For g, h, and k as defined above,*

$$\mathbb{V}\{g(X)\} + \mathbb{E}\{k(X)\} = \mathbb{V}\{h(X,Y)\}.$$

Proof. We have

$$
\begin{aligned}
&\mathbb{V}(g(X)) + \mathbb{E}(k(X)) \\
&= \int_{\Omega_1} g(X(\omega_1))^2 dP_1(\omega_1) - \left[\int_{\Omega_1} g(X(\omega_1)) dP_1(\omega_1)\right]^2 + \int_{\Omega_1} k(X(\omega_1)) dP_1(\omega_1) \\
&= \int_{\Omega_1} g(X(\omega_1))^2 + k(X(\omega_1)) dP_1(\omega_1) - \left[\int_{\Omega_1} g(X(\omega_1)) dP_1(\omega_1)\right]^2 \\
&= \int_{\Omega_1} \left(\int_{\Omega_2} h(X(\omega_1), Y(\omega_2)) dP_2(\omega_2)\right)^2 + \int_{\Omega_2} h(X(\omega_1), Y(\omega_2))^2 dP_2(\omega_2) \\
&\qquad - \left(\int_{\Omega_2} h(X(\omega_1), Y(\omega_2)) dP_2(\omega_2)\right)^2 dP_1(\omega_1) - \left[\int_{\Omega_1}\int_{\Omega_2} h(X(\omega_1), Y(\omega_2)) dP_2(\omega_2) dP_1(\omega_1)\right]^2 \\
&= \int_{\Omega_1}\int_{\Omega_2} h(X(\omega_1), Y(\omega_2))^2 dP_2(\omega_2) dP_1(\omega_1) - \left[\int_{\Omega_1}\int_{\Omega_2} h(X(\omega_1), Y(\omega_2)) dP_2(\omega_2) dP_1(\omega_1)\right]^2 \\
&= \mathbb{E}(h(X,Y)^2) - \mathbb{E}(h(X,Y))^2 \\
&= \mathbb{V}(h(X,Y)).
\end{aligned}
$$

□