Homework 0: Exercises on Classical Probability

Not Due

NOTE: Start by solving the underlined problems. These are recommended for all students. The starred problems are ones that are recommended for students who have already taken graduate probability.

Problem 0.1. Suppose $\Omega$ is a countable space and $p : \Omega \to [0,1]$ is such that $\sum_{\omega \in \Omega} p(\omega) = 1$. For $A \subset \Omega$ let $P(A) = \sum_{\omega \in A} p(\omega)$ with $P(\emptyset) = 0$. Prove that

\[ P(A) \in [0,1], \quad P(\Omega) = 1, \quad \text{and} \quad P(A^c) = 1 - P(A). \]

Furthermore, if $A_n, n \in \mathbb{N}$, are mutually disjoint, i.e. $A_n \cap A_m = \emptyset$ for $m \neq n$, then

\[ P(\bigcup_{n} A_n) = \sum_{n} P(A_n). \]

Problem 0.2. Let $\Omega$ be a countable space. For a given $\omega_0 \in \Omega$, $\delta_{\omega_0}$ is the probability measure that corresponds to setting $p(\omega)$ from the above problem to 1 when $\omega = \omega_0$ and to 0 otherwise. Prove that $\delta_{\omega_0}(A) = \mathbb{1}_A(\omega_0)$. Given a probability measure $P$ that corresponds to an arbitrary $p$, as in the above problem, prove that $P = \sum_{\omega \in \Omega} p(\omega)\delta_{\omega}$. By this we mean that for any $A \subset \Omega$,

\[ P(A) = \sum_{\omega \in A} p(\omega)\delta_{\omega}(A). \]

Problem 0.3. Suppose we are given a function $p : [0,1] \to [0,1]$. Prove that if $\sum_{\omega \in [0,1]} p(\omega) < \infty$, then the set \{ $\omega \in [0,1] : p(\omega) > 0$ \} is at most countable.

Problem 0.4. Prove that the intersection of any family of $\sigma$-algebras is a $\sigma$-algebra. That is, if $\mathcal{A}_i$ are $\sigma$-algebras, for all $i \in I$ ($I$ arbitrary), then $\bigcap_{i \in I} \mathcal{A}_i$ is also a $\sigma$-algebra.

Problem 0.5. Let $\Omega$ be an arbitrary space and let $B_n, n \in \mathbb{N}$, be a partition of it. This means they are mutually disjoint and $\cup_n B_n = \Omega$. Describe the $\sigma$-algebra generated by $\{ B_n : n \in \mathbb{N} \}$. \textbf{Hint:} Start with just two sets $B_1 \cup B_2 = \Omega$, then work out the case of three sets $B_1 \cup B_2 \cup B_3 = \Omega$. 1
Problem 0.6. Prove that the σ-algebra generated by intervals of the form \((a, b], \ a < b\) is the same as the Borel σ-algebra (i.e. the σ-algebra generated by the open sets of \(\mathbb{R}\)).

Problem 0.7. Prove that σ-additivity is equivalent to continuity under increasing sequences (that is, if \(A_n\) is an increasing sequence of sets, then \(P(\bigcup_n A_n) = \lim_{n \to \infty} P(A_n)\)).

Problem 0.7. Prove that for a finite measure, σ-additivity is equivalent to continuity under decreasing sequences (that is, if \(A_n\) is a decreasing sequence of sets, then \(P(\bigcap_n A_n) = \lim_{n \to \infty} P(A_n)\)).

Problem 0.8.* For an interval of the type \((a, b]\) let \(m((a, b]) = b - a\). If \(I\) is a finite union of disjoint such intervals, let \(m(I)\) be the sum of the values of \(m\) on each of the intervals. Prove that this definition is consistent, i.e. if we write \(I\) in two different ways as a disjoint union of intervals, then we get the same value for \(m(I)\). Therefore, we have defined \(m\) on the algebra of finite unions of semi-open intervals. To be able to extend it to the Borel σ-algebra, we need to prove that \(m\) is σ-additive on the algebra. Do that.

Problem 0.9. Identify the interval \([0, 1]\) with the unit circle \(S^1 = \{e^{i\theta} : \theta \in [0, 2\pi)\} \subset \mathbb{C}\). We will construct a set in \(S^1\) that is not Borel-measurable. To this end, for \(z = e^{i\alpha}, w = e^{i\beta} \in S^1\), we say \(z \sim w\) if \(\alpha - \beta \in \mathbb{Q}\). This is clearly an equivalence relation. Using the axiom of choice we can construct a set \(\Lambda\) whose elements are one from each equivalence class. For \(\alpha \in [0, 2\pi) \cap \mathbb{Q}\) let \(\Lambda_\alpha = e^{i\alpha} \Lambda\) be the rotation of \(\Lambda\) by the angle \(\alpha\).

1. Prove that if \(\alpha, \beta \in [0, 2\pi) \cap \mathbb{Q}\) are distinct, then \(\Lambda_\alpha \cap \Lambda_\beta = \emptyset\).

2. Prove that \(S^1\) is the union over \(\alpha \in [0, 2\pi) \cap \mathbb{Q}\) of \(\Lambda_\alpha\).

3. Let \(m\) be the Lebesgue measure on \(S^1\) (equipped with the Borel σ-algebra). Prove that \(m(\Lambda)\) cannot exist. (Hint: use the thing you just proved and the fact that \(m\) is shift-invariant.)

Problem 0.10.* Prove that the Lebesgue measure is shift-invariant.

Problem 0.11. Let \(P\) be a probability measure on \(\mathbb{R}\), equipped with the Borel σ-algebra. Let \(F(x) = P((\infty, x])\). Prove that \(F\) is non-decreasing,
right-continuous, \( F(x) \to 0 \) as \( x \to -\infty \), and \( F(x) \to 1 \) as \( x \to \infty \). Prove that if \( P \) and \( Q \) are two probability measures such that \( P((\neg\infty, x]) = Q((\neg\infty, x]) \) for all \( x \) rational, then \( P = Q \), i.e. \( P(A) = Q(A) \) for any Borel-measurable set \( A \).

**Problem 0.12.** Let \( \Omega_1 = \{1, 2, 3, 4, 5, 6\} \) and \( \Omega_2 = \{H, T\} \). Let \( X : \Omega_1 \to \{0, 1\} \) and \( Y : \Omega_2 \to \{0, 1\} \) be such that \( X(1) = X(3) = X(5) = 0, X(2) = X(4) = X(6) = 1, Y(T) = 0, \) and \( Y(H) = 1 \). Let \( P_1 \) be the probability measure on \( \Omega_1 \) such that \( P_1(\{1\}) = 1/15, P_1(\{2\}) = 2/15, P_1(\{3\}) = P_1(\{5\}) = 2/15, P_1(\{4\}) = P_1(\{6\}) = 4/15 \). Let \( P_2 \) be the probability measure on \( \Omega_2 \) with \( P_2(\{T\}) = 1/3 \) and \( P_2(\{H\}) = 2/3 \). Prove that the distribution of \( X \) under \( P_1 \) is the same as the distribution of \( Y \) under \( \Omega_2 \). This says that if you are to only observe the outcomes of \( X \) and \( Y \), then there is no statistical test that can tell the difference between the two.

**Problem 0.13.** Let \( X \) be a nonnegative bounded random variable. Let \( X_n \) be the sequence of simple random variables that was defined in the lecture, which we used to define \( \int X(\omega)P(d\omega) \). Prove that if \( Y_n \) is a simple random variable and for all \( \omega \), \( Y_n(\omega) \) is non-decreasing in \( n \) and \( Y_n(\omega) \to X(\omega) \), then \( \int Y_n(\omega)P(d\omega) \) has the same limit as \( \int X_n(\omega)P(d\omega) \) had.

**Problem 0.14.** Let \( \Phi : \mathbb{R} \to \mathbb{R} \) be a bounded measurable function and \( X \) be a random variable, then \( \int \Phi(X)dP = \int \Phi(x)P_X(dx) \).

**Problem 0.15.** Let \( Y_n \) be i.i.d. Bernoulli(1/2) random variables. Let \( Y = \sum_{n=1}^{\infty} Y_n/4^n \).

1. Prove that \( P(Y = y) = 0 \) for any \( y \in [0, 1] \). (Therefore, \( Y \) does not have any discrete part.)

2. Prove that there exists a Borel-measurable set \( A \subset [0, 1] \) such that \( P(Y \in A) = 1 \) while \( m(A) = 0 \). (Therefore, \( Y \) does not have a part that is absolutely continuous with respect to Lebesgue measure.)

**Problem 0.16.** Solve the following:

1. A standard normal random variable is a random variable, say \( Z \), with distribution that has the pdf \( \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \). Calculate the pdf of \( Z^2 \).

2. Calculate \( E[Z] \) and \( E[Z^2] \).
3. An exponential random variable with rate $\lambda$ is a random variable with distribution that has the pdf $\lambda e^{-\lambda x}$ when $x > 0$ and 0 otherwise. Calculate $E[X]$ and $E[X^2]$ for such a random variable.

4. A geometric random variable is a random variable with distribution $\sum_{k \in \mathbb{N}} p^k(1-p)\delta_k$, where $p \in (0, 1)$ is some given parameter. Calculate $E[X]$ and $E[X^2]$ for such a random variable.

One can do the same for all sorts of other known distributions, such as a uniform on an interval $[a, b]$, a normal with parameters $\mu$ and $\sigma^2$, a Poisson with parameter $\lambda > 0$, a Bernoulli with parameter $p \in (0, 1)$, a binomial with parameters $n \in \mathbb{N}$ and $p \in (0, 1)$, etc.

Problem 0.17. Prove Jensen’s, Lyapunov’s, Hölder’s, Minkowski’s inequalities.

Problem 0.18. Prove that almost sure convergence implies convergence in probability. Also prove that if $X_n$ converges to $X$ in probability, then there exists a subsequence $n_j$ such that $X_{n_j} \to X$ almost surely.

Problem 0.19.* Read the proofs of the monotone convergence theorem, Fatou’s lemma, and the dominated convergence theorem.

Problem 0.20. Let $f(x, y) = (x^2 - y^2)/(x^2 + y^2)^2$ for $x, y \in (0, 1]$. Prove that

$$
\int_0^1 \left( \int_0^1 f(x, y) \, dx \right) \, dy \neq \int_0^1 \left( \int_0^1 f(x, y) \, dy \right) \, dx.
$$

Problem 0.21. Compute the value of

$$
\int_0^\infty \int_0^\infty e^{-x^2/2-y} \, dy \, dx.
$$

Problem 0.22. Find a setting in which $A, B, C$ are pairwise independent, but not all independent. Find a setting in which $P(A \cap B \cap C) = P(A)P(B)P(C)$ and yet the three events are not independent.

Problem 0.23.* Prove that $X$ and $Y$ are independent if and only if $E[\Phi(X)\Psi(Y)] = E[\Phi(X)]E[\Psi(Y)]$ for all bounded measurable functions $\Phi, \Psi : \mathbb{R} \to \mathbb{R}$. 
Problem 0.24. Suppose you are given a sequence of events \( A_n, n \in \mathbb{N}, \) and \( \sum_{n \in \mathbb{N}} P(A_n) < \infty. \) The event “\( A_n \) happens infinitely often” can be represented as \( \cap_n \cup_{m \geq n} A_m. \) Prove that this event has probability one. **Hint:** Use a union bound to prove that the complementary event has zero probability. This result is called the Borel-Cantelli Lemma.

Problem 0.25. Suppose you are given a sequence of events \( A_n, n \in \mathbb{N}, \) that are independent and such that \( \sum_{n \in \mathbb{N}} P(A_n) = \infty. \) Prove that the event “\( A_n \) happens infinitely often” has probability one. This result is the reverse direction in the Borel-Cantelli Lemma. Note how it needs the events to be independent.

Problem 0.26. Prove the above result under just the assumption that the events are pairwise uncorrelated, i.e., that \( 1_{A_n} \) and \( 1_{A_m} \) are uncorrelated random variables, if \( m \neq n. \) **Hint:** Apply Chebyshev’s inequality to the random variable \( S_n = \sum_{k=1}^{n} 1_{A_k}. \)

Problem 0.27. Suppose \( X_n \) are i.i.d. mean-zero random variables such that \( E[X_1^4] < \infty. \) Let \( S_n = X_1 + \cdots + X_n. \) Compute \( E[S_n^4] \) then apply Chebyshev’s inequality and the Borel-Cantelli lemma to prove that \( S_n/n \to 0 \) almost surely. This is the strong law of large numbers, but under a fourth moment assumption. (With considerably more work, one can remove the fourth moment assumption and prove the same result under just \( E[|X_1|] < \infty. \))

Problem 0.28. Let \( X_1, \ldots, X_n \) be i.i.d. Normal(\( \mu, \sigma^2 \)) random variables. What is the distribution of \( (X_1 + \cdots + X_n - n\mu)/\sqrt{n\sigma^2}? \) How does the central limit theorem work in this case?

Problem 0.29. Prove the central limit theorem for a sequence of i.i.d. Bernoulli(\( p \)) random variables, where \( p \in (0, 1). \) **Hint:** Compute the moment generating function of the object you want the limit of and use Taylor’s expansion to show that it converges to the moment generating function of a standard normal. (In fact, the same proof, but without the computation being so explicit, works for a general distribution, as long as the second moment is finite. And then pushing the proof a bit further leads to the Lindeberg-Feller CLT.)

Problem 0.30. By considering \( L^2(\mathcal{A}) \) as a Hilbert subspace of the Hilbert space \( L^2(\mathcal{F}), \) construct the conditional expectation \( E[X|\mathcal{A}] \) for a
random variable $X \in L^2(\mathcal{F})$ as being the orthogonal projection of $X$ onto $L^2(\mathcal{A})$. Next, use the fact that $L^2(\mathcal{F})$ is dense in $L^1(\mathcal{F})$ and that Jensen’s inequality says that for two random variables $X_1, X_2 \in L^2$,

$$E[|E[X_1|\mathcal{A}] - E[X_2|\mathcal{A}]|] = E[|E[X_1 - X_2|\mathcal{A}]|] \leq E[E[|X_1 - X_2| | \mathcal{A}]] = E[|X_1 - X_2|],$$

to define $E[X|\mathcal{A}]$ for a random variable $X \in L^1(\mathcal{F})$.

**Problem 0.30.** Prove the following two properties:

1. If $X$ is integrable and $\mathcal{A}$-measurable, then $E[X|\mathcal{A}] = X$.

2. If $X, Y$ are integrable and $a, b \in \mathbb{R}$, then $E[aX + bY|\mathcal{A}] = aE[X|\mathcal{A}] + bE[Y|\mathcal{A}]$.

**Problem 0.31.** Prove that if $\sigma(X)$ and $\mathcal{A}$ are independent, then $E[X|\mathcal{A}] = E[X]$.

**Problem 0.32.** Prove that if $(X, Y)$ have pdf $f(x, y)$, then $Y$ has pdf $f_Y(y) = \int f(x, y) \, dx$ and the distribution of $X$, conditional on $Y$, has conditional pdf $f(x|y) = f(x, y)/f_Y(y)$.

**Problem 0.33.** Let $X$ be an exponential random variable with unit rate. Fix two positive numbers $x$ and $y$. Prove that $P(X > x + y|X > x) = P(X > y)$. This shows that conditioning the exponential clock on not having rung by time $x$ and then restarting the count at that point gives statistically the same exponential clock! This is called the memoriless property of the exponential distribution. The same holds for the geometric distribution.