PROBLEMS IN PROBABILITY THEORY

MANJUNATH KRISHNAPUR

Problem 1. Let $\mathcal{F}$ be a σ-algebra of subsets of Ω.

(1) Show that $\mathcal{F}$ is closed under countable intersections ($\bigcap\limits_n A_n$), under set differences ($A \setminus B$), under symmetric differences ($A \Delta B$).

(2) If A_n is a countable sequence of subsets of Ω, the set $\limsup_n A_n$ (respectively $\liminf_n A_n$) is defined as the set of all $\omega \in \Omega$ that belongs to infinitely many (respectively, all but finitely many) of the sets A_n.
    If $A_n \in \mathcal{F}$ for all n, show that $\limsup A_n \in \mathcal{F}$ and $\liminf A_n \in \mathcal{F}$. [**Hint:** First express $\limsup A_n$ and $\liminf A_n$ in terms of A_ns and basic set operations].

(3) If $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$, what are $\limsup A_n$ and $\liminf A_n$?

Problem 2. Let $(\Omega, \mathcal{F})$ be a set with a σ-algebra.

(1) Suppose $\mathbf{P}$ is a probability measure on $\mathcal{F}$. If $A_n \in \mathcal{F}$ and A_n increase to A (respectively, decrease to A), show that $\mathbf{P}(A_n)$ increases to (respectively, decreases to) $\mathbf{P}(A)$.

(2) Suppose $\mathbf{P} : \mathcal{F} \to [0,1]$ is a function such that (a) $\mathbf{P}(\Omega) = 1$, (b) $\mathbf{P}$ is finitely additive, (c) if $A_n, A \in \mathcal{F}$ and A_ns increase to A, then $\mathbf{P}(A_n) \uparrow \mathbf{P}(A)$. Then, show that $\mathbf{P}$ is a probability measure on $\mathcal{F}$.

Problem 3. Suppose S is a π-system and is further closed under complements ($A \in S$ implies $A^c \in S$). Show that S is an algebra.

Problem 4. Let $\mathbf{P}$ be a p.m. on a σ-algebra $\mathcal{F}$ and suppose $S \subseteq \mathcal{F}$ be a π-system. If $A_k \in S$ for $k \leq n$, write $\mathbf{P}(A_1 \cup A_2 \cup \ldots \cup A_n)$ in terms of probabilities of sets in S.

Problem 5. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. Let $\mathcal{G} = \{A \in \mathcal{F} : \mathbf{P}(A) = 0 \text{ or } 1\}$. Show that $\mathcal{G}$ is a σ-algebra.

Problem 6. Suppose $\sigma(S) = \mathcal{F}$ and $\mathbf{P}, \mathbf{Q}$ are two probability measure on $\mathcal{F}$. If $\mathbf{P}(A) = \mathbf{Q}(A)$ for all $A \in S$, is it necessarily true that $\mathbf{P}(A) = \mathbf{Q}(A)$ for all $A \in \mathcal{F}$? If yes, prove it. If not, give a counterexample.

Problem 7. (1) Let $\mathcal{B}$ be the Borel sigma-algebra of $\mathbb{R}$. Show that $\mathcal{B}$ contains all closed sets, all compact sets, all intervals of the form $(a, b]$ and $[a, b)$.

(2) Show that there is a countable family $\mathcal{S}$ of subsets of $\mathbb{R}$ such that $\sigma(\mathcal{S}) = \mathcal{B}_{\mathbb{R}}$.

(3) Let K be the $1/3$-Cantor set. Show that $\mu_*(K) = 0$.

Problem 8. (1) Let X be an arbitrary set. Let S be the collection of all singletons in Ω. Describe $\sigma(S)$.

(2) Let $S = \{(a, b] \cup [-b, -a) : a < b \text{ are real numbers}\}$. Show that $\sigma(S)$ is strictly smaller than the Borel σ-algebra of $\mathbb{R}$.

(3) Suppose S is a collection of subsets of X and a, b are two elements of X such that any set in S either contains a and b both, or contains neither. Let $\mathcal{F} = \sigma(S)$. Show that any set in $\mathcal{F}$ has the same property (either contains both a and b or contains neither).

Problem 9. Let Ω be an infinite set and let $\mathcal{A} = \{A \subseteq \Omega : A \text{ is finite or } A^c \text{ is finite }\}$. Define $\mu : \mathcal{A} \to \mathbb{R}_+$ by $\mu(A) = 0$ if A is finite and $\mu(A) = 1$ if A^c is finite.

(1) Show that $\mathcal{A}$ is an algebra and that μ is finitely additive on $\mathcal{A}$.

(2) Under what conditions does μ extend to a probability measure on $\mathcal{F} = \sigma(\mathcal{A})$?

Problem 10. If $\mathcal{G} \subseteq \mathcal{F}$ are sigma algebras on Ω and $\mathcal{F}$ is countably generated, then is it necessarily true that $\mathcal{G}$ is countably generated? [*Soln:* False. $\mathcal{G} =$Countable-cocountable sigma algebra of $\mathbb{R}$ and $\mathcal{F} = \mathcal{B}_{\mathbb{R}}$.]

Problem 11. Let $A_1, A_2, \ldots$ be a finite or countable partition of a non-empty set Ω (i.e., A_i are pairwise disjoint and their union is Ω). What is the σ-algebra generated by the collection of subsets $\{A_n\}$? What is the algebra generated by the same collection of subsets?

Problem 12. Let $X = [0, 1]^{\mathbb{N}}$ be the countable product of copies of $[0, 1]$. We define two sigma algebras of subsets of X.

(1) Define a metric on X by $d(x, y) = \sup_n |x_n - y_n|$. Let $\mathcal{B}_X$ be the Borel sigma-algebra of (X, d). [**Note:** For those who know topology, it is better to define $\mathcal{B}_X$ as the Borel sigma algebra for the product topology on X. The point is that the metric is flexible. We can take $d(x, y) = \sum_n |x_n - y_n| 2^{-n}$ or many or other things. What matters is only the topology on X.]

(2) Let $\mathcal{C}_X$ be the sigma-algebra generated by the collection of all cylinder sets. Recall that cylinder sets are sets of the form $A = U_1 \times U_2 \times \ldots \times U_n \times \mathbb{R} \times \mathbb{R} \times \ldots$ where U_i are Borel subsets of $[0, 1]$.

Show that $\mathcal{B}_X = \mathcal{C}_X$.

Problem 13. Let μ be the Lebesgue p.m. on the Cartheodary σ-algebra $\bar{\mathcal{B}}$ and let μ_* be the corresponding outer Lebesgue measure defined on all subsets of $[0, 1]$. We say that a subset $N \subseteq [0, 1]$ is a null set if $\mu_*(N) = 0$. Show that

$$\bar{\mathcal{B}} = \{B \cup N : B \in \mathcal{B} \text{ and } N \text{ is null}\}$$

where $\mathcal{B}$ is the Borel σ-algebra of $[0, 1]$.

[**Note:** The point of this exercise is to show how much larger is the Lebesgue σ-algebra than the Borel σ-algebra. The answer is, not much. Up to a null set, every Lebesgue measurable set is a Borel set. However, cardinality-wise, there is a difference. The Lebesgue σ-algebra is in bijection with $2^{\mathbb{R}}$ while the Borel σ-algebra is in bijection with $\mathbb{R}$.]

Problem 14. Suppose $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space. Say that a subset $N \subseteq \Omega$ is $\mathbf{P}$-null if there exists $A \in \mathcal{F}$ with $\mathbf{P}(A) = 0$ and such that $N \subseteq A$. Define $\mathcal{G} = \{A \cup N : A \in \mathcal{F} \text{ and } N \text{ is null}\}$.

(1) Show that $\mathcal{G}$ is a σ-algebra.

(2) For $A \in \mathcal{G}$, write $A = B \cup N$ with $b \in \mathcal{F}$ and a null set N, and define $\mathbf{Q}(A) = \mathbf{P}(B)$. Show that $\mathbf{Q}$ is well-defined, that $\mathbf{Q}$ is a probability measure on $\mathcal{G}$ and $\mathbf{Q}\Big|_{\mathcal{F}} = \mathbf{P}$.

[**Note:** $\mathcal{G}$ is called the $\mathbf{P}$-completion of $\mathcal{F}$. It is a cheap way to enlarge the σ-algebra and extend the measure to the larger σ-algebra. Another description of the extended σ-algebra is $\mathcal{G} = \{A \subseteq \Omega : \exists B, C \in \mathcal{F} \text{ such that } B \subseteq A \subseteq C \text{ and } \mathbf{P}(B) = \mathbf{P}(C)\}$. Combined with the previous problem, we see that the Lebesgue σ-algebra is just the completion of the Borel σ-algebra under the Lebesgue measure. However, note that completion depends on the probability measure (for a discrete probability measure on $\mathbb{R}$, the completion will be the power set σ-algebra!). For this reason, we prefer to stick to the Borel σ-algebra and not bother to extend it.]

Problem 15. Follow these steps to obtain Sierpinski's construction of a non-measurable set. Here μ_* is the outer Lebesgue measure on $\mathbb{R}$.

(1) Regard $\mathbb{R}$ as a vector space over $\mathbb{Q}$ and choose a basis H (why is it possible?).

(2) Let $A_0 = H \cup (-H) = \{x : x \in H \text{ or } -x \in H\}$. For $n \geq 1$, define $A_n := A_{n-1} - A_{n-1}$ (may also write $A_n = A_{n-1} + A_{n-1}$ since A_0 is symmetric about 0). Show that $\bigcup\limits_{n \geq 0} \bigcup\limits_{q \geq 1} \frac{1}{q} A_n = \mathbb{R}$ where $\frac{1}{q} A_n$ is the set $\{\frac{x}{q} : x \in A_n\}$.

(3) Let $N := \min\{n \geq 0 : \mu_*(A_n) > 0\}$ (you must show that N is finite!). If A_N is measurable, show that $\cup_{n \geq N+1} A_n = \mathbb{R}$.

(4) Get a contradiction to the fact that H is a basis and conclude that A_N cannot be measurable.

[**Remark:** If you start with H which has zero Lebesgue measure, then $N \geq 1$ and $A := E_{N-1}$ is a Lebesgue measurable set such that $A + A$ is not Lebesgue measurable! That was the motivation for Sierpinski. To find such a basis H, show that the Cantor set spans $\mathbb{R}$ and then choose a basis H contained inside the Cantor set.]

Problem 16. We saw that for a Borel probability measure μ on $\mathbb{R}$, the pushforward of Lebesgue measure on $[0,1]$ under the map $F_\mu^{-1} : [0,1] \to \mathbb{R}$ (as defined in lectures) is precisely μ. This is also a practical tool in simulating random variables. We assume that a random number generator gives us uniform random numbers from $[0,1]$. Apply the above idea to simulate random numbers from the following distributions (in matlab/mathematica or a program of your choice) a large number of times and compare the histogram to the actual density/mass function.

(1) Uniform distribution on $[a,b]$, (2) Exponential(λ) distribution, (3) Cauchy distribution, (4) Poisson(λ) distribution. What about the normal distribution?

Problem 17. Let $\Omega = X = \mathbb{R}$ and let $T : \Omega \to X$ be defined by $T(x) = x$. We give a pair of σ-algebras, $\mathcal{F}$ on Ω and $\mathcal{G}$ on X by taking $\mathcal{F}$ and $\mathcal{G}$ to be one of $2^{\mathbb{R}}$ or $\mathcal{B}_{\mathbb{R}}$ or $\{\emptyset, \mathbb{R}\}$. Decide for each of the nine pairs, whether T is measurable or not.

Problem 18. (1) Define $T : \Omega \to \mathbb{R}^n$ by $T(\omega) = (\mathbf{1}_{A_1}(\omega), \ldots, \mathbf{1}_{A_n}(\omega))$ where $A_1, \ldots, A_n$ are given subsets of Ω. What is the smallest σ-algebra on Ω for which T becomes a random variable?

(2) Suppose $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space and assume that $A_k \in \mathcal{F}$. Describe the push-forward measure $\mathbf{P} \circ T^{-1}$ on $\mathbb{R}^n$.

Problem 19. For $k \geq 0$, define the functions $r_k : [0,1) \to \mathbb{R}$ by writing $[0,1) = \bigsqcup_{0 \leq j < 2^k} I_j^{(k)}$ where $I_j^{(k)}$ is the dyadic interval $[j2^{-k}, (j+1)2^{-k})$ and setting

$$r_k(x) = \begin{cases} -1 & \text{if } x \in I_j^{(k)} \text{ for odd } j, \\ +1 & \text{if } x \in I_j^{(k)} \text{ for even } j. \end{cases}$$

Fix $n \geq 1$ and define $T_n : [0,1) \to \{-1,1\}^n$ by $T_n(x) = (r_0(x), \ldots, r_{n-1}(x))$. Find the push-forward of the Lebesgue measure on $[0,1)$ under T_n

Problem 20. Let $\mathcal{G}$ be the countable-cocountable sigma algebra on $\mathbb{R}$. Define the probability measure μ on $\mathcal{G}$ by $\mu(A) = 0$ if A is countable and $\mu(A) = 1$ if A^c is countable. Show that μ is *not* the push-forward of Lebesgue measure on $[0,1]$, i.e., there does not exist a measurable function $T : [0,1] \mapsto \Omega$ (w.r.t. the σ-algebras $\mathcal{B}$ and $\mathcal{G}$) such that $\mu = \lambda \circ T^{-1}$.

Problem 21. (1) If $T : \mathbb{R}^n \to \mathbb{R}^m$, show that T is Borel measurable if it is (a) continuous or (b) right continuous or (c) lower semicontinuous or (d) non-decreasing (take $m = n = 1$ for the last one).

(2) If $\mathbb{R}^n$ and $\mathbb{R}^m$ are endowed with the Lebesgue sigma-algebra, show that even if T is continuous, it need not be measurable! Just do this for $n = m = 1$.

Problem 22. Show that composition of random variables is a random variable. Show that real-valued random variables on a given $(\Omega, \mathcal{F})$ are closed under linear combinations, under multiplication, under countable suprema (or infima) and under limsup (or liminf) of countable sequences.

Problem 23. Let $\mu_n = \frac{1}{n}\sum\limits_{k=1}^{n} \delta_{k/n}$ and let μ be the uniform p.m. on $[0,1]$. Show directly by definition that $d(\mu_n, \mu) \to 0$ as $n \to \infty$.

Problem 24 (Change of variable for densities). (1) Let μ be a p.m. on $\mathbb{R}$ with density f by which we mean that its CDF $F_\mu(x) = \int_{-\infty}^{x} f(t)dt$ (you may assume that f is continuous, non-negative and the Riemann integral $\int_{\mathbb{R}} f = 1$). Then, find the (density of the) push forward measure of μ under (a) $T(x) = x + a$ (b) $T(x) = bx$ (c) T is any increasing and differentiable function.

(2) If X has $N(\mu, \sigma^2)$ distribution, find the distribution of $(X - \mu)/\sigma$.

Problem 25. (1) Let $X = (X_1, \ldots, X_n)$. Show that X is an $\mathbb{R}^d$-valued r.v. if and only if $X_1, \ldots, X_n$ are (real-valued) random variables. How does $\sigma(X)$ relate to $\sigma(X_1), \ldots, \sigma(X_n)$?

(2) Let $X : \Omega_1 \to \Omega_2$ be a random variable. If $X(\omega) = X(\omega')$ for some $\omega, \omega' \in \Omega_1$, show that there is no set $A \in \sigma(X)$ such that $\omega \in A$ and $\omega' \notin A$ or vice versa. [**Extra!** If $Y : \Omega_1 \to \Omega_2$ is another r.v. which is measurable w.r.t. $\sigma(X)$ on Ω_1, then show that Y is a function of X].

Problem 26 (Lévy metric). (1) Show that the Lévy metric on $\mathcal{P}(\mathbb{R}^d)$ defined in class is actually a metric.

(2) Show that under the Lévy metric, $\mathcal{P}(\mathbb{R}^d)$ is a complete and seperable metric space.

Problem 27 (Lévy-Prohorov metric). If (X, d) is a metric space, let $\mathcal{P}(X)$ denote the space of Borel probability measures on X. For $\mu, \nu \in \mathcal{P}(X)$, define

$$D(\mu, \nu) = \inf\{r \geq 0 : \mu(A_r) + r \geq \nu(A) \text{ and } \nu(A_r) + r \geq \mu(A) \text{ for all closed sets } A\}.$$

Here $A_r = \{y \in X : d(x, y) \leq r \text{ for some } x \in A\}$ is the closed r-neighbourhood of A.

(1) Show that D is a metric on $\mathcal{P}(X)$.

(2) When X is $\mathbb{R}^d$, show that this agrees with the definition of Lévy metric given in class (i.e., for any μ_n, μ, we have that $\mu_n \to \mu$ in both metrics or neither).

Problem 28 (Lévy metric). Let $\mathcal{P}([-1,1]) \subseteq \mathcal{P}(\mathbb{R})$ be the set of all Borel probability measures μ such that $\mu([-1,1]) = 1$. For $\epsilon > 0$, find a finite ϵ-net for $\mathcal{P}([-1,1])$. [*Note:* Recall that an ϵ-net means a subset such that every element of $\mathcal{P}([-1,1])$ is within ϵ distance of some element of the subset. Since $\mathcal{P}([-1,1])$ is compact, we know that a finite ϵ-net exists for all $\epsilon > 0$.]

Problem 29. On the probabiity space $([0,1], \mathcal{B}, \mu)$, for $k \geq 1$, define the functions

$$X_k(t) := \begin{cases} 0 & \text{if } t \in \bigcup\limits_{j=0}^{2^{k-1}-1} [\frac{2j}{2^k}, \frac{2j+1}{2^k}). \\ 1 & \text{if } t \in \bigcup\limits_{j=0}^{2^{k-1}-1} [\frac{2j+1}{2^k}, \frac{2j+2}{2^k}) \text{ or } t = 1. \end{cases}$$

(1) For any $n \geq 1$, what is the distribution of X_n?

(2) For any fixed $n \geq 1$, find the joint distribution of $(X_1, \ldots, X_n)$.

[**Note:** $X_k(t)$ is just the k^{th} digit in the binary expansion of t. Dyadic rationals have two binary expansions, and we have chosen the finite expansion (except at $t = 1$)].

Problem 30 (Coin tossing space). Continuing with the previous example, consider the mapping $X : [0,1] \to \{0,1\}^{\mathbb{N}}$ defined by $X(t) = (X_1(t), X_2(t), \ldots)$. With the Borel σ-algebra on $[0,1]$ and the σ-algebra generated by cylinder sets on $\{0,1\}^{\mathbb{N}}$, show that X is a random variable and find the push-foward of the Lebesue measure under X.

Problem 31 (Equivalent conditions for weak convergence). Show that the following statements are equivalent to $\mu_n \xrightarrow{d} \mu$ (you may work in $\mathcal{P}(\mathbb{R})$).

(1) $\limsup_{n\to\infty} \mu_n(F) \leq \mu(F)$ if F is closed.

(2) $\liminf_{n\to\infty} \mu_n(G) \geq \mu(G)$ if G is open.

(3) $\limsup_{n\to\infty} \mu_n(A) = \mu(A)$ if $A \in \mathcal{F}$ and $\mu(\partial A) = 0$.

Problem 32. Fix $\mu \in \mathcal{P}(\mathbb{R})$. For $s \in \mathbb{R}$ and $r > 0$, let $\mu_{r,s} \in \mathcal{P}(\mathbb{R})$ be defined as $\mu_{r,s}(A) = \mu(rA + s)$ where $rA + s = \{rx + s : x \in A\}$. For which $R \subseteq (0, \infty)$ and $S \subseteq \mathbb{R}$ is it true that $\{\mu_{r,s} : r \in R,\ s \in S\}$ a tight family? [*Remark:* If not clear, just take μ to be the Lebesgue measure on $[0, 1]$.]

Problem 33. (1) Show that the family of Normal distributions $\{N(\mu, \sigma^2) : \mu \in \mathbb{R}$ and $\sigma^2 > 0\}$ is not tight.

(2) For what $A \subseteq \mathbb{R}$ and $B \subseteq (0, \infty)$ is the restricted family $\{N(\mu, \sigma^2) : \mu \in A$ and $\sigma^2 \in B\}$ tight?

Problem 34. (1) Show that the family of exponential distributions $\{\text{Exp}(\lambda) : \lambda > 0\}$ is not tight.

(2) For what $A \subseteq \mathbb{R}$ is the restricted family $\{\text{Exp}(\lambda) : \lambda > 0\}$ tight?

Problem 35. Suppose $\mu_n, \mu \in \mathcal{P}(\mathbb{R})$ and that the distribution function of μ is continuous. If $\mu_n \xrightarrow{d} \mu$, show that $F_{\mu_n}(t) - F_\mu(t) \to 0$ uniformly over $t \in \mathbb{R}$. [*Restatement:* When F_μ is continuous, convergence to μ in Lévy-Prohorov metric also implies convergence in Kolmogorov-Smirnov metric.]

Problem 36. Show that the statement in the previous problem cannot be quantified. That is,

Given any $\epsilon_n \downarrow 0$ (however fast) and $\delta_n \downarrow 0$ (however slow), show that there is some μ_n, μ with F_μ continuous, such that $d_{LP}(\mu_n, \mu) \le \epsilon_n$ and $d_{KS}(\mu_n, \mu) \ge \delta_n$.

Problem 37. Consider the family of Normal distributions, $\{N(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0\}$. Show that the map $(\mu, \sigma^2) \to N(\mu, \sigma^2)$ from $\mathbb{R} \times \mathbb{R}_+$ to $\mathcal{P}(\mathbb{R})$ is continuous. (Complicated way of saying that if $(\mu_n, \sigma_n^2) \to (\mu, \sigma^2)$, then $N(\mu_n, \sigma_n^2) \xrightarrow{d} N(\mu, \sigma^2)$).

Do the same for other natural families if distributions, (1) Exp(λ), (2) Uniform$[a, b]$, (3) Bin(n, p) (fix n and show continuity in p), (4) Pois(λ).

Problem 38. Suppose μ_n, μ are discrete probability measures supported on $\mathbb{Z}$ having probability mass functions $(p_n(k))_{k\in\mathbb{Z}}$ and $(p(k))_{k\in\mathbb{Z}}$. Show that $\mu_n \xrightarrow{d} \mu$ if and only if $p_n(k) \to p(k)$ for each $k \in \mathbb{Z}$.

Problem 39. Given a Borel p.m. μ on $\mathbb{R}$, show that it can be written as a convex combination $\alpha\mu_1 + (1-\alpha)\mu_2$ with $\alpha \in [0,1]$, where μ_1 is a purely atomic Borel p.m and μ_2 is a Borel p.m with no atoms.

Problem 40. Let F be a CDF on $\mathbb{R}$.

(1) Show that F can have at most countably many discontinuity points. [**Hint:** How many jumps of size more than $1/10$ can it have?]

(2) Give example of a CDF that has a dense set of discontinuity points.

Problem 41. Let X be a random variable with distribution μ and X_n are random variables defined as follows. If μ_n is the distribution of X_n, in each case, show that $\mu_n \stackrel{d}{\to} \mu$ as $n \to \infty$.

(1) (Truncation). $X_n = (X \wedge n) \vee (-n)$.

(2) (Discretization). $X_n = \frac{1}{n}\lfloor nX \rfloor$.

Problem 42. Consider the space $X = [0,1]^{\mathbb{N}} := \{\mathbf{x} = (x(1), x(2), \dots) : 0 \le x(i) \le 1 \text{ for each } i \in \mathbb{N}\}$. Define the metric $d(\mathbf{x}, \mathbf{y}) = \sup_i \frac{|x(i)-y(i)|}{i}$.

(1) Show that $\mathbf{x}_n \to \mathbf{x}$ in (X, d) if and only if $x_n(i) \to x(i)$ for each i, as $n \to \infty$.

[**Note:** What matters is this pointwise convergence criterion, not the specific metric. The resulting topology is called *product topology*. The same convergence would hold if we had defined the metric as $d(\mathbf{x}, \mathbf{y}) = \sum_i 2^{-i}|x(i) - y(i)|$ or $d(\mathbf{x}, \mathbf{y}) = \sum_i i^{-2}|x(i) - y(i)|$ etc., But not the metric $\sup_i |x(i) - y(i)|$ as convergence in this metric is equivalent to uniform convergence over all $i \in \mathbb{N}$].

(2) Show that X is compact.

[**Note:** What is this problem doing here? The purpose is to reiterate a key technique we used in the proof of Helly's selection principle!]

Problem 43. Recall the Cantor set $C = \bigcap_n K_n$ where $K_0 = [0,1]$, $K_1 = [0,1/3] \cup [2/3,1]$, etc. In general, $K_n = \bigcup_{1 \le j \le 2^n} [a_{n,j}, b_{n,j}]$ where $b_{n,j} - a_{n,j} = 3^{-n}$ for each j.

(1) Let μ_n be the uniform probability measure on K_n. Describe its CDF F_n.

(2) Show that F_n converges uniformly to a CDF F.

(3) Let μ be the probability measure with CDF equal to F. Show that $\mu(C) = 1$.

Problem 44. Let $\mu \in \mathcal{P}(\mathbb{R})$.

(1) For any $n \geq 1$, define a new probability measure by $\mu_n(A) = \mu(n.A)$ where $n.A = \{nx : x \in A\}$. Does μ_n converge as $n \to \infty$?

(2) Let μ_n be defined by its CDF

$$F_n(t) = \begin{cases} 0 & \text{if } t < -n, \\ F(t) \text{ if } -n \leq t < n, \\ 1 & \text{if } t \geq n. \end{cases}$$

Does μ_n converge as $n \to \infty$?

(3) In each of the cases, describe μ_n in terms of random variables. That is, if X has distribution μ, describe a transformation $T_n(X)$ that has the distribution μ_n.

Problem 45 (Bernoulli convolutions). For any $\lambda > 1$, define $X_\lambda : [0,1] \to \mathbb{R}$ by $X(\omega) = \sum_{k=1}^{\infty} \lambda^{-k} X_k(\omega)$. Check that X_λ is measurable, and define $\mu_\lambda = \mu X_\lambda^{-1}$. Show that for any $\lambda > 2$, show that μ_λ is singular w.r.t. Lebesgue measure.

Problem 46. For $p = 1, 2, \infty$, check that $\|X - Y\|_p$ is a metric on the space $L^p := \{[X] : \|X\|_p < \infty\}$ (here $[X]$ denotes the equivalence class of X under the above equivalence relation).

Problem 47. (1) Give an example of a sequence of r.v.s X_n such that $\liminf \mathbf{E}[X_n] < \mathbf{E}[\liminf X_n]$.

(2) Give an example of a sequence of r.v.s X_n such that $X_n \overset{a.s.}{\to} X$, $\mathbf{E}[X_n] = 1$, but $\mathbf{E}[X] = 0$.

Problem 48 (Alternate construction of Cantor measure). Let $K_1 = [0, 1/3] \cup [2/3, 1]$, $K_2 = [0, 1/9] \cup [2/9, 3/9] \cup [6/9, 7/9] \cup [8/9, 1]$, etc., be the decreasing sequence of compact sets whose intersection is K. Observe that K_n is a union of 2^n intervals each of length 3^{-n}. Let μ_n be the p.m. which is the "renormalized Lebesgue measure" on K_n. That is, $\mu_n(A) := 3^n 2^{-n} \mu(A \cap K_n)$ for $A \in \mathcal{B}_\mathbb{R}$. Then each μ_n is a Borel p.m. Show that $\mu_n \overset{d}{\to} \mu$, the Cantor measure (which was defined differently in class).

Problem 49 (A quantitative characterization of absolute continuity). Suppose $\mu \ll \nu$. Then, show that given any $\epsilon > 0$, there exists $\delta > 0$ such that $\nu(A) < \delta$ implies $\mu(A) < \epsilon$. (The converse statement is obvious but worth noticing). [**Hint**: Argue by contradiction].

Problem 50. Suppose $f : [a, b] \to \mathbb{R}$ is a Borel measurable function. Then, show that $g(x) := \int_0^x f(u)du$ is a continuous function on $[0, 1]$. [**Note:** It is in fact true that g is differentiable at almost every x and that $g' = f$ $a.s.$, but that is a more sophisticated fact, called *Lebesgue's differentiation theorem*. In this course, we only need Lebesgue integration, not differentiation. The latter may be covered in your measure theory class].

Problem 51. (**Differentiating under the integral**). Let $f : [a, b] \times \mathbb{R} \to \mathbb{R}$, satisfy the following assumptions.

(1) $x \to f(x, \theta)$ is Borel measurable for each θ.

(2) $\theta \to f(x, \theta)$ is continuously differentiable for each x.

(3) $f(x, \theta)$ and $\frac{\partial f}{\partial \theta}(x, \theta)$ are uniformly bounded functions of (x, θ).

Then, justify the following "differentiation under integral sign" (including the fact that the integrals here make sense).

$$\frac{d}{d\theta} \int_a^b f(x, \theta)dx = \int_a^b \frac{\partial f}{\partial \theta}(x, \theta)\ dx$$

[**Hint:** Remember that derivative is the limit of difference quotients, $h'(t) = \lim_{\epsilon \to 0} \frac{h(t+\epsilon)-h(t)}{\epsilon}$.

Problem 52. (1) Let $X \geq 0$ be a r.v on $(\Omega, \mathcal{F}, \mathbf{P})$ with $0 < \mathbf{E}[X] < \infty$. Then, define $\mathbf{Q}(A) = \mathbf{E}[X \mathbf{1}_A]/\mathbf{E}[X]$ for any $A \in \mathcal{F}$. Show that $\mathbf{Q}$ is a probability measure on $\mathcal{F}$. Further, show that for any bounded random variable Y, we have $\mathbf{E}_{\mathbf{Q}}[Y] = \frac{\mathbf{E}[YX]}{\mathbf{E}[X]}$.

(2) If μ and ν are Borel probability measures on the line with continuous densities f and g (respectively) w.r.t. Lebesgue measure. Under what conditions can you assert that μ has a density w.r.t ν? In that case, what is that density?

Problem 53. For $p = 1, 2, \infty$, check that $\|X - Y\|_p$ is a metric on the space $L^p := \{[X] : \|X\|_p < \infty\}$ (here $[X]$ denotes the equivalence class of X under the equivalence relation $X \sim Y$ if $\mathbf{P}(X = Y) = 1$).

Problem 54. Let X be a non-negative random variable.

(1) Show that $\mathbf{E}[X] = \int_0^\infty \mathbf{P}\{X > t\}dt$ (in particular, if X is a non-negative integer valued, then $\mathbf{E}[X] = \sum_{n=1}^\infty \mathbf{P}(X \geq n)$).

(2) Show that $\mathbf{E}[X^p] = \int_0^\infty pt^{p-1}\mathbf{P}\{X \geq t\}dt$ for any $p > 0$.

Problem 55. Let X be a non-negative random variable. If $\mathbf{E}[X]$ is finite, show that $\sum_{n=1}^\infty \mathbf{P}\{X \geq an\}$ is finite for any $a > 0$. Conversely, if $\sum_{n=1}^\infty \mathbf{P}\{X \geq an\}$ for some $a > 0$, show that $\mathbf{E}[X]$ is finite.

Problem 56. Show that the values $\mathbf{E}[f \circ X]$ as f varies over the class of all smooth (infinitely differentiable), compactly supported functions determine the distribution of X.

Problem 57. (i) Express the mean and variance of of $aX + b$ in terms of the same quantities for X (a, b are constants).

(ii) Show that $\text{Var}(X) = \mathbf{E}[X^2] - \mathbf{E}[X]^2$.

Problem 58. Compute mean, variance and moments (as many as possible!) of the Normal(0,1), exponential(1), Beta(p,q) distributions.

Problem 59. (1) Suppose $X_n \geq 0$ and $X_n \to X$ a.s. If $\mathbf{E}[X_n] \to \mathbf{E}[X]$, show that $\mathbf{E}[|X_n - X|] \to 0$.

(2) If $\mathbf{E}[|X|] < \infty$, then $\mathbf{E}[|X|\mathbf{1}_{|X|>A}] \to 0$ as $A \to \infty$.

Problem 60. (1) Suppose (X, Y) has a continuous density $f(x, y)$. Find the density of X/Y. Apply to the case when (X, Y) has the *standard bivariate normal distribution* with density $f(x, y) = (2\pi)^{-1} \exp\{-\frac{x^2+y^2}{2}\}$.

(2) Find the distribution of $X + Y$ if (X, Y) has the standard bivariate normal distribution.

(3) Let $U = \min\{X, Y\}$ and $V = \max\{X, Y\}$. Find the density of (U, V).

Problem 61. Let $\mu_n, \mu \in \mathcal{P}(\mathbb{R}^n)$. Show that $\mu_n \stackrel{d}{\to} \mu$ if and only if $\int f d\mu_n \to \int f d\mu$ for every $f \in C_b(\mathbb{R})$. What if we only assume $\int f d\mu_n \to \int f d\mu$ for all $f \in C_c(\mathbb{R}^n)$ - can we conclude that $\mu_n \stackrel{d}{\to} \mu$?

Problem 62. Let $\mu_n, \mu \in \mathcal{P}(\mathbb{R}^n)$ having densities f_n, f with respect to Lebesgue measure. If $f_n \to f$ a.e. (w.r.t. Lebesgue measure), show that $\mu_n \stackrel{d}{\to} \mu$.

Problem 63 (Moment matrices). Let $\mu \in \mathcal{P}(\mathbb{R})$ and let $\alpha_k = \int x^k d\mu(x)$ (assume that all moments exist). Then, for any $n \geq 1$, show that the matrix $(\alpha_{i+j})_{0 \leq i,j \leq n}$ is non-negative definite. [**Suggestion:** First solve $n = 1$].

Problem 64. Let X be a non-negative random variable with all moments (i.e., $\mathbf{E}[X^p] < \infty$ for all $p < \infty$). Show that $\log \mathbf{E}[X^p]$ is a convex function of p.

Problem 65. (1) Let $\mu_n, \mu \in \mathcal{P}(\mathbb{R}^d)$. Assume that μ_n has density f_n and μ has density f w.r.t Lebesgue measure on $\mathbb{R}^n$. If $f_n(t) \to f(t)$ for all t, then show that $\mu_n \xrightarrow{d} \mu$.

(2) Show that $N(\mu_n, \sigma_n^2) \xrightarrow{d} N(\mu, \sigma)$ if and only if $\mu_n \to \mu$ and $\sigma_n^2 \to \sigma^2$.

Problem 66. (1) Let $X \sim \Gamma(\alpha, 1)$ and $Y \sim \Gamma(\alpha', 1)$ be independent random variables on a common probability space. Find the distribution of $\frac{X}{X+Y}$.

(2) If U, V are independent and have uniform([0,1]) distribution, find the distribution of $U + V$.

Problem 67. Let $\Omega = \{1, 2, \ldots, n\}$. For a probability measure $\mathbf{P}$ on Ω, we define it "entropy" $H(\mathbf{P}) := -\sum_{k=1}^{n} p_k \log p_k$ where $p_k = \mathbf{P}\{k\}$ and it is understood that $x \log x = 0$ if $x = 0$. Show that among all probability measures on Ω, the uniform probability measure (the one with $p_k = \frac{1}{n}$ for each k) is the unique maximizer of entropy.

Problem 68. (1) If $\mu_n \ll \nu$ for each n and $\mu_n \xrightarrow{d} \mu$, then is it necessarily true that $\mu \ll \nu$? If $\mu_n \perp \nu$ for each n and $\mu_n \xrightarrow{d} \mu$, then is it necessarily true that $\mu \perp \nu$? In either case, justify or give a counterexample.

(2) Suppose X, Y are independent (real-valued) random variables with distribution μ and ν respectively. If μ and ν are absolutely continuous w.r.t Lebesgue measure, show that the distribution of $X + Y$ is also absolutely continuous w.r.t Lebesgue measure.

Problem 69. Suppose $\{\mu_\alpha : \alpha \in I\}$ and $\{\nu_\beta : \alpha \in J\}$ are two families of Borel probability measures on $\mathbb{R}$. If both these families are tight, show that the family $\{\mu_\alpha \otimes \nu_\beta : \alpha \in I,\ \beta \in J\}$ is also tight.

Problem 70. Let X be a non-negative random variable. If $\mathbf{E}[X] \leq 1$, then show that $\mathbf{E}[X^{-1}] \geq 1$.

Problem 71. On the probabiity space $([0,1], \mathcal{B}, \mu)$, for $k \geq 1$, define the functions

$$X_k(t) := \begin{cases} 0 & \text{if } t \in \bigcup\limits_{j=0}^{2^{k-1}-1} [\frac{2j}{2^k}, \frac{2j+1}{2^k}). \\ 1 & \text{if } t \in \bigcup\limits_{j=0}^{2^{k-1}-1} [\frac{2j+1}{2^k}, \frac{2j+2}{2^k}) \text{ or } t = 1. \end{cases}$$

(1) For any $n \geq 1$, what is the distribution of X_n?

(2) For any fixed $n \geq 1$, find the joint distribution of $(X_1, \ldots, X_n)$.

[**Note:** $X_k(t)$ is just the k^{th} digit in the binary expansion of t. Dyadic rationals have two binary expansions, and we have chosen the finite expansion (except at $t = 1$)].

Problem 72. If $A \in \mathcal{B}(\mathbb{R}^2)$ has positive Lebesgue measure, show that for some $x \in \mathbb{R}$ the set $A_x := \{y \in \mathbb{R} : (x, y) \in A\}$ has positive Lebesgue measure in $\mathbb{R}$.

Problem 73 (A quantitative characterization of absolute continuity). Suppose $\mu \ll \nu$. Then, show that given any $\epsilon > 0$, there exists $\delta > 0$ such that $\nu(A) < \delta$ implies $\mu(A) < \epsilon$. (The converse statement is obvious but worth noticing). [**Hint**: Argue by contradiction].

Problem 74. Let $Z_1, \ldots, Z_n$ be i.i.d $N(0, 1)$ and write $\mathbf{Z}$ for the vector with components $Z_1, \ldots, Z_n$. Let A be an $m \times n$ matrix and let μ be a vector in $\mathbb{R}^m$. Then the m-dimensional random vector $\mathbf{X} = \mu + A\mathbf{Z}$ is said to have distribution $N_m(\mu, \Sigma)$ where $\Sigma = AA^t$ ('Normal distribution with mean vector μ and covariance matrix Σ').

(1) If $m \leq n$ and A has rank m, show that $\mathbf{X}$ has density $(2\pi)^{-\frac{m}{2}} \exp\{-\frac{1}{2}\mathbf{x}^t A^{-1}\mathbf{x}\}$ w.r.t Lebesgue measure on $\mathbb{R}^m$. In particular, note that the distribution depends only on μ and AA^t. (**Note:** If $m > n$ or if rank$(A) < m$, then satisfy yourself that $\mathbf{X}$ has no density w.r.t Lebesgue measure on $\mathbb{R}^m$ - you do not need to submit this).

(2) Check that $\mathbf{E}[X_i] = \mu_i$ and $\text{Cov}(X_i, X_j) = \Sigma_{i,j}$.

(3) What is the distribution of (i) $(X_1, \ldots, X_k)$, for $k \leq n$? (ii) $B\mathbf{X}$, where B is a $p \times m$ matrix? (iii) $X_1 + \ldots + X_m$?

Problem 75. (1) If X, Y are independent random variables, show that $\text{Cov}(X, Y) = 0$.

(2) Give a counterexample to the converse by giving an infinite sequence of random variables $X_1, X_2, \ldots$ such that $\text{Cov}(X_i, X_j) = 0$ for any $i \neq j$ but such that X_i are not independent.

(3) Suppose $(X_1, \ldots, X_m)$ has (joint) normal distribution (see the first question). If $\text{Cov}(X_i, X_j) = 0$ for all $i \leq k$ and for all $j \geq k + 1$, then show that $(X_1, \ldots, X_k)$ is independent of $(X_{k+1}, \ldots, X_m)$.

Problem 76. (1) Suppose $2 \leq k < n$. Give an example of random variables $X_1, \ldots, X_n$ such that any subset of k of these random variables are independent but no subset of $k + 1$ of them is independent.

(2) Suppose $(X_1, \ldots, X_n)$ has a multivariate Normal distribution. Show that if X_i are pairwise independent, then they are independent.

Problem 77. Show that it is not possible to define uncountably many independent Ber(1/2) random variables on the probability space $([0,1],\mathcal{B},\lambda)$.

Problem 78. Let X_i, $i\geq 1$ be random variables on a common probability space. Let $f:\mathbb{R}^{\mathbb{N}}\to\mathbb{R}$ be a measurable function (with product sigma algebra on $\mathbb{R}^{\mathbb{N}}$ and Borel sigma algebra on $\mathbb{R}$) and let $Y=f(X_1,X_2,\ldots)$. Show that the distribution of Y depends only on the joint distribution of $(X_1,X_2,\ldots)$ and not on the original probability space. [**Hint:** We used this to say that if X_i are independent Bernoulli random variables, then $\sum_{i\geq 1}X_i2^{-i}$ has uniform distribution on $[0,1]$, irrespective of the underlying probability space.]

Problem 79. Suppose $(X_1,\ldots,X_n)$ has density f (w.r.t Lebesgue measure on $\mathbb{R}^n$).

(1) If $f(x_1,\ldots,x_n)$ can be written as $\prod_{k=1}^n g_k(x_k)$ for some one-variable functions g_k, $k\leq n$. Then show that $X_1,\ldots,X_n$ are independent. (Don't assume that g_k is a density!)

(2) If $X_1,\ldots,X_n$ are independent, then $f(x_1,\ldots,x_n)$ can be written as $\prod_{k=1}^n g_k(x_k)$ for some one-variable densities $g_1,\ldots,g_n$.

Problem 80. Among all $n!$ permutations of $[n]$, pick one at random with uniform probability. Show that the probability that this random permutation has no fixed points is at most $\frac{1}{2}$ for any n.

Problem 81. Suppose each of $r=\lambda n$ balls are put into n boxes at random (more than one ball can go into a box). If N_n denotes the number of empty boxes, show that for any $\delta>0$, as $n\to\infty$,

$$\mathbf{P}\left(\left|\frac{N_n}{n}-e^{-\lambda}\right|>\delta\right)\to 0$$

Problem 82. Let X_n be i.i.d random variables such that $\mathbf{E}[|X_1|]<\infty$. Define the random power series $f(z)=\sum_{k=0}^{\infty}X_nz^n$. Show that almost surely, the radius of convergence of f is equal to 1. [**Note:** Recall from Analysis class that the radius of convergence of a power series $\sum c_nz^n$ is given by $(\limsup|c_n|^{\frac{1}{n}})^{-1}$].

Problem 83. (1) Let X be a real values random variable with finite variance. Show that $f(a):=\mathbf{E}[(X-a)^2]$ is minimized at $a=\mathbf{E}[X]$.

(2) What is the quantity that minimizes $g(a)=\mathbf{E}[|X-a|]$? [**Hint:** First consider X that takes finitely many values with equal probability each].

Problem 84 (Existence of Markov chains). Let S be a countable set (with the power set sigma algebra). Two ingredients are given: *A transition matrix*, that is, a function $p : S \times S \to [0,1]$ be a function such that $p(x,\cdot)$ is a probability mass function on S for each $x \in S$. (1) *An initial distribution*, that is a probability mass function μ_0 on S.

For $n \geq 0$ define the probability measure ν_n on S^{n+1} (with the product sigma algebra) by

$$\nu_n(A_0 \times A_1 \times \ldots \times A_n) = \sum_{(x_0,\ldots,x_n)\in A_0\times\ldots\times A_n} \mu_0(x_0) \prod_{j=0}^{n-1} p(x_j, x_{j+1}).$$

Show that ν_n form a consistent family of probability distributions and conclude that a Markov chain with initial distribution μ_0 and transition matrix p exists.

Problem 85. Show that it is not possible to define uncountably many independent $\text{Ber}(1/2)$ random variables on the probability space $([0,1], \mathcal{B}, \lambda)$.

Problem 86. Let $(\Omega_i, \mathcal{F}_i, \mathbf{P}_i)$, $i \in I$, be probability spaces and let $\Omega = \times_i \Omega_i$ with $\mathcal{F} = \otimes_i \mathcal{F}_i$ and $\mathbf{P} = \otimes_i \mathbf{P}_i$. If $A \in \mathcal{F}$, show that for any $\epsilon > 0$, there is a cylinder set B such that $\mathbf{P}(A \Delta B) < \epsilon$.

Problem 87. Let ξ, ξ_n be i.i.d. random variables with $\mathbf{E}[\log_+ \xi] < \infty$ and $\mathbf{P}(\xi = 0) < 1$.

(1) Show that $\limsup_{n\to\infty} |\xi_n|^{\frac{1}{n}} = 1$ $a.s.$

(2) Let c_n be (non-random) complex numbers. Show that the radius of convergence of the random power series $\sum_{n=0}^{\infty} c_n \xi_n z^n$ is almost surely equal to the radius of convergence of the non-random power series $\sum_{n=0}^{\infty} c_n z^n$.

Problem 88. (Ergodicity of product measure). This problem guides you to a proof of a different zero-one law.

(1) Consider the product measure space $(\mathbb{R}^{\mathbb{Z}}, \mathcal{B}(\mathbb{R}^{\mathbb{Z}}), \otimes_{\mathbb{Z}} \mu)$ where $\mu \in \mathcal{P}(\mathbb{R})$. Define $\tau : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}}$ by $(\tau\omega)_n = \omega_{n+1}$. Let $\mathcal{I} = \{A \in \mathcal{B}(\mathbb{R}^{\mathbb{Z}}) : \tau(A) = A\}$. Then, show that $\mathcal{I}$ is a sigma-algebra (called the invariant sigma algebra) and that every event in $\mathcal{I}$ has probability equal to 0 or 1.

(2) Let $X_n, n \geq 1$ be i.i.d. random variables on a common probability space. Suppose $f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ is a measurable function such that $f(x_1, x_2, \ldots) = f(x_2, x_3, \ldots)$ for any $(x_1, x_2, \ldots) \in \mathbb{R}^{\mathbb{N}}$. Then deduce from the first part that the random variable $f(X_1, X_2, \ldots)$ is a constant, $a.s.$

[**Hint:** Approximate A by cylinder sets. Use translation by τ^m to show that $\mathbf{P}(A) = \mathbf{P}(A)^2$.]

Problem 89. Consider the invariant sigma algebra and the tail sigma algebra. Show that neither is contained in the other.

Problem 90. (Chung-Erdös inequality).

(1) Let A_i be events in a probability space. Show that
$$\mathbf{P}\left\{\bigcup_{k=1}^{n} A_k\right\} \geq \frac{(\sum_{k=1}^{n}\mathbf{P}(A_k))^2}{\sum_{k,\ell=1}^{n}\mathbf{P}(A_k\cap A_\ell)}$$

(2) Place r_m balls in m bins at random and count the number of empty bins Z_m. Fix $\delta > 0$. If $r_m > (1+\delta)m\log m$, show that $\mathbf{P}(Z_m > 0) \to 0$ while if $r_m < (1-\delta)m\log m$, show that $\mathbf{P}(Z_m > 0) \to 1$.

Problem 91. Give example of an infinite sequence of pairwise independent random variables for which Kolmogorov's zero-one law fails.

Problem 92. Let X_i, $i \in I$ be random variables on a probability space. Suppose that for some $p > 0$ and $M < \infty$ we have $\mathbf{E}[|X_i|^p] \leq M$ for all $i \in I$. Show that the family $\{X_i : i \in I\}$ is tight (by which we mean that $\{\mu_{X_i} : i \in I\}$ is tight, where μ_{X_i} is the distribution of X_i).

Problem 93. Let X_i be i.i.d. random variables with zero mean and finite variance. Let $S_n = X_1 + \ldots + X_n$. Show that the collection $\{\frac{1}{\sqrt{n}}S_n : n \geq 1\}$ is tight. [**Note:** Tightness is essential for convergence in distribution. In the case at hand, convergence in distribution to $N(0,1)$ is what is called central limit theorem. We shall see it later.]

Problem 94. Among all $n!$ permutations of $[n]$, pick one at random with uniform probability. Show that the probability that this random permutation has no fixed points is at most $\frac{1}{2}$ for any n.

Problem 95. Suppose each of $r = \lambda n$ balls are put into n boxes at random (more than one ball can go into a box). If N_n denotes the number of empty boxes, show that for any $\delta > 0$, as $n \to \infty$,
$$\mathbf{P}\left(\left|\frac{N_n}{n} - e^{-\lambda}\right| > \delta\right) \to 0$$

Problem 96. Let X_n be i.i.d random variables such that $\mathbf{E}[|X_1|] < \infty$. Define the random power series $f(z) = \sum_{k=0}^{\infty} X_n z^n$. Show that almost surely, the radius of convergence of f is equal to 1. [**Note:** Recall from Analysis class that the radius of convergence of a power series $\sum c_n z^n$ is given by $(\limsup |c_n|^{\frac{1}{n}})^{-1}$].

Problem 97. (1) Let X be a real values random variable with finite variance. Show that $f(a) := \mathbf{E}[(X-a)^2]$ is minimized at $a = \mathbf{E}[X]$.

(2) What is the quantity that minimizes $g(a) = \mathbf{E}[|X-a|]$? [**Hint:** First consider X that takes finitely many values with equal probability each].

Problem 98. Let X_i be i.i.d. Cauchy random variables with density $\frac{1}{\pi(1+t^2)}$. Show that $\frac{1}{n}S_n$fils the weak law of large numbers by completing the following steps.

(1) Show that $t\mathbf{P}\{|X_1| > t\} \to c$ for some constant c.

(2) Show that if $\delta > 0$ is small enough, then $\mathbf{P}\{|\frac{1}{n-1}S_{n-1}| \geq \delta\} + \mathbf{P}\{|\frac{1}{n-1}S_{n-1}| \geq \delta\}$ does not go to 0 as $n \to \infty$ [*Hint:* Consider the possibility that $|X_n| > 2\delta n$].

(3) Conclude that $\frac{1}{n}S_n \stackrel{P}{\to} 0$. [*Extra:* With a little more effort, you can try showing that there does not exist deterministic numbers a_n such that $\frac{1}{n}S_n - a_n \stackrel{P}{\to} 0$].

Problem 99. Let X_n, X be random variables on a common probability space.

(1) If $X_n \stackrel{P}{\to} X$, show that some subsequence $X_{n_k} \stackrel{a.s.}{\to} X$.

(2) If every subsequence of X_n has a further subsequence that converges almost surely to X, show that $X_n \stackrel{P}{\to} X$.

Problem 100. For $\mathbb{R}^d$-valued random vectors X_n, X, the notions of convergence almost surely, in probability and in distribution are well-defined. If $X_n = (X_{n,1}, \ldots, X_{n,d})$ and $X = (X_1, \ldots, X_d)$, which of the following is true? Justify or give counterexamples.

(1) $X_n \stackrel{a.s.}{\to} X$ if and only if $X_{n,k} \stackrel{a.s.}{\to} X_k$ for $1 \leq k \leq d$.

(2) $X_n \stackrel{P}{\to} X$ if and only if $X_{n,k} \stackrel{P}{\to} X_k$ for $1 \leq k \leq d$.

(3) $X_n \stackrel{d}{\to} X$ if and only if $X_{n,k} \stackrel{d}{\to} X_k$ for $1 \leq k \leq d$.

Problem 101. Let X_n, Y_n, X, Y be random variables on a common probability space.

(1) If $X_n \stackrel{P}{\to} X$ and $Y_n \stackrel{P}{\to} Y$ (all r.v.s on the same probability space), show that $aX_n + bY_n \stackrel{P}{\to} aX + bY$ and $X_nY_n \stackrel{P}{\to} XY$. [**Hint:** You could try showing more generally that $f(X_n, Y_n) \to f(X, Y)$ for any continuous $f : \mathbb{R}^2 \to \mathbb{R}$.]

(2) If $X_n \stackrel{P}{\to} X$ and $Y_n \stackrel{d}{\to} Y$ (all on the same probability space), then show that $X_nY_n \stackrel{d}{\to} XY$.

Problem 102. Let X_n, Y_n, X, Y be random variables on a common probability space.

(1) Suppose that X_n is independent of Y_n for each n (no assumptions about independence across n). If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$, then $(X_n, Y_n) \xrightarrow{d} (U, V)$ where $U \stackrel{d}{=} X$, $V \stackrel{d}{=} Y$ and U, V are independent. Further, $aX_n + bY_n \xrightarrow{d} aU + bV$.

(2) Give counterexample to show that the previous statement is false if the assumption of independence of X_n and Y_n is dropped.

Problem 103. For $\mathbb{R}^d$-valued random vectors X_n, X, we say that $X_n \xrightarrow{P} X$ if $\mathbf{P}(\|X_n - X\| > \delta) \to 0$ for any $\delta > 0$ (here you may take $\|\cdot\|$ to denote the usual norm, but any norm on $\mathbb{R}^d$ gives the same definition).

(1) If $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{P} Y$, show that $(X_n, Y_n) \xrightarrow{P} (X, Y)$.

(2) If $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{P} Y$, show that $X_n + Y_n \xrightarrow{P} X + Y$ and $\langle X_n, Y_n \rangle \xrightarrow{P} XY$. **[Hint:** Show more generally that $f(X_n, Y_n) \xrightarrow{P} f(X, Y)$ for any continuous function f by using the previous problem for random vectors].

Problem 104. (1) If X_n, Y_n are independent random variables on the same probability space and $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$, then $(X_n, Y_n) \xrightarrow{d} (U, V)$ where $U \stackrel{d}{=} X$, $V \stackrel{d}{=} Y$ and U, V are independent.

(2) If $X_n \xrightarrow{d} X$ and $Y_n - X_n \xrightarrow{P} 0$, then show that $Y_n \xrightarrow{d} X$.

Problem 105. Show that the sequence $\{X_n\}$ is tight if and only if $c_n X_n \xrightarrow{P} 0$ whenever $c_n \to 0$.

Problem 106. Suppose X_n are i.i.d with $\mathbf{E}[|X_1|^4] < \infty$. Show that there is some constant C (depending on the distribution of X_1) such that $\mathbf{P}\left(|n^{-1}S_n - \mathbf{E}[X_1]| > \delta\right) \le Cn^{-2}$. (What is your guess if we assume $\mathbf{E}[|X_1|^6] < \infty$? You don't need to show this in the homework).

Problem 107. (1) (**Skorokhod's representation theorem**) If $X_n \xrightarrow{d} X$, then show that there is a probability space with random variables Y_n, Y such that $Y_n \stackrel{d}{=} X_n$ and $Y \stackrel{d}{=} X$ and $Y_n \xrightarrow{a.s.} Y$. **[Hint:** Try to construct Y_n, Y on the canonical probability space $([0,1], \mathcal{B}, \mu)$]

(2) If $X_n \xrightarrow{d} X$, and $f : \mathbb{R} \to \mathbb{R}$ is continuous, show that $f(X_n) \xrightarrow{d} f(X)$. **[Hint:** Use the first part]

Problem 108. Suppose X_i are i.i.d with the Cauchy distribution (density $\pi^{-1}(1+x^2)^{-1}$ on $\mathbb{R}$). Note that X_1 is not integrable. Then, show that $\frac{S_n}{n}$ does not converge in probability to any constant. [**Hint:** Try to find the probability $\mathbf{P}(X_1 > t)$, and then use it].

Problem 109. Let $\{X_i\}_{i\in I}$ be a family of r.v on $(\Omega, \mathcal{F}, \mathbf{P})$.

(1) If $\{X_i\}_{i\in I}$ is uniformly integrable, then show that $\sup_i \mathbf{E}|X_i| < \infty$. Give a counterexample to the converse statement.

(2) Suppose $h : \mathbb{R}_+ \to \mathbb{R}_+$ is a non-decreasing function that goes to infinity and $\sup_i \mathbf{E}[|X_i| h(|X_i|)] < \infty$. Show that $\{X_i\}_{i\in I}$ is uniformly integrable. In particular, if $\sup_i \mathbf{E}[|X_i|^p] < \infty$ for some $p > 1$, then $\{X_i\}$ is uniformly integrable.

Problem 110. Let X_n be i.i.d with $\mathbf{P}(X_1 = +1) = \mathbf{P}(X_1 = -1) = \frac{1}{2}$. Show that for any $\gamma > \frac{1}{2}$,

$$\frac{S_n}{n^\gamma} \overset{a.s.}{\to} 0.$$

[**Remark:** Try to imitate the proof of SLLN under fourth moment assumption. If you write the proof correctly, it should go for any random variable which has moments of all orders. You do not need to show this for the homework].

Problem 111. Let X_n be independent real-valued random variables.

(1) Show by example that the event $\{\sum X_n$ converges to a number in [1,3]$\}$ can have probability strictly between 0 and 1.

(2) Show that the event $\{\sum X_n$ converges to a finite number$\}$ has probability zero or one.

Problem 112. Let X_n be i.i.d exponential(1) random variables.

(1) If b_n is a sequence of numbers that converge to 0, show that $\limsup b_n X_n$ is a constant (a.s.). Find a sequence b_n so that $\limsup\ b_n X_n = 1$ a.s.

(2) Let M_n be the maximum of $X_1, \ldots, X_n$. If $a_n \to \infty$, show that $\limsup \frac{M_n}{a_n}$ is a constant (a.s.). Find a_n so that $\limsup \frac{M_n}{a_n} = 1$ (a.s.).

[**Remark:** Can you do the same if X_n are i.i.d N(0,1)? Need not show this for the homework, but note that the main ingredient is to find a simple expression for $\mathbf{P}(X_1 > t)$ asymptotically as $t \to \infty$].

Problem 113. Let X_n be i.i.d real valued random variables with common distribution μ. For each n, define the random probabilty measure μ_n as $\mu_n := \frac{1}{n}\sum_{k=1}^{n}\delta_{X_k}$. Let F_n be the CDF of μ_n. Show that

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \overset{a.s.}{\to} 0 \quad a.s.$$

Problem 114. Let X_n be independent and $\mathbf{P}(X_n = n^a) = \frac{1}{2} = \mathbf{P}(X_n = -n^a)$ where $a > 0$ is fixed. For what values of a does the series $\sum X_n$ converge a.s.? For which values of a does the series converge absolutely, a.s.?

Problem 115. (**Random series**) Let X_n be i.i.d $N(0,1)$ for $n \geq 1$.

(1) Show that the random series $\sum X_n \frac{\sin(n\pi t)}{n}$ converges a.s., for any $t \in \mathbb{R}$.

(2) Show that the random series $\sum X_n \frac{t^n}{\sqrt{n!}}$ converges for all $t \in \mathbb{R}$, a.s.

[**Note:** The location of the phrase "a.s" is all important here. Let A_t and B_t denote the event that the series converges for the fixed t in the first or second parts of the question, respectively. Then, the first part is asking you to show that $\mathbf{P}(A_t) = 1$ for each $t \in \mathbb{R}$, while the second part is asking you to show that $\mathbf{P}(\cap_{t\in\mathbb{R}} B_t) = 1$. It is also true (and very important!) that $\mathbf{P}(\cap_{t\in\mathbb{R}} A_t) = 1$ but showing that is not easy.]

Problem 116. Suppose X_n are i.i.d random variables with finite mean. Which of the following assumptions guarantee that $\sum X_n$ converges a.s.?

(1) (i) $\mathbf{E}[X_n] = 0$ for all n and (ii) $\sum \mathbf{E}[X_n^2 \wedge 1] < \infty$.

(2) (i) $\mathbf{E}[X_n] = 0$ for all n and (ii) $\sum \mathbf{E}[X_n^2 \wedge |X_n|] < \infty$.

Problem 117. (**Large deviation for Bernoullis**). Let X_n be i.i.d Ber(1/2). Fix $p > \frac{1}{2}$.

(1) Show that $\mathbf{P}(S_n > np) \leq e^{-np\lambda}\left(\frac{e^{\lambda}+1}{2}\right)^n$ for any $\lambda > 0$.

(2) Optimize over λ to get $\mathbf{P}(S_n > np) \leq e^{-nI(p)}$ where $I(p) = -p\log p - (1-p)\log(1-p)$. (Observe that this is the *entropy* of the Ber(p) measure introduced in the first class test).

(3) Recall that $S_n \sim$ Binom$(n, 1/2)$, to write $\mathbf{P}(S_n = \lceil np \rceil)$ and use Stirling's approximation to show that

$$\mathbf{P}(S_n \geq np) \geq \frac{1}{\sqrt{2\pi np(1-p)}} e^{-nI(p)}.$$

(4) Deduce that $\mathbf{P}(S_n \geq np) \approx e^{-nI(p)}$ for $p > \frac{1}{2}$ and $\mathbf{P}(S_n < np) \approx e^{-nI(p)}$ for $p < \frac{1}{2}$ where the notation $a_n \approx b_n$ means $\frac{\log a_n}{\log b_n} \to 1$ as $n \to \infty$ (i.e., asymptotic equality on the logarithmic scale).

Problem 118. Carry out the same program for i.i.d exponential(1) random variables and deduce that $\mathbf{P}(S_n > nt) \approx e^{-nI(t)}$ for $t > 1$ and $\mathbf{P}(S_n < nt) \approx e^{-nI(t)}$ for $t < 1$ where $I(t) := t - 1 - \log t$.

Problem 119. Let $Y_1, \ldots, Y_n$ be independent random variables. A random variable τ taking values in $\{1, 2, \ldots, n\}$ is called a *stopping time* if the event $\{\tau \le k\} \in \sigma(Y_1, \ldots, Y_k)$ for all k (equivalently $\{\tau = k\} \in \sigma(Y_1, \ldots, Y_k)$ for all k).

(1) Which of the following are stopping times? $\tau_1 := \min\{k \le n : S_k \in A\}$ (for some fixed $A \subseteq \mathbb{R}$). $\tau_2 := \max\{k \le n : S_k \in A\}$. $\tau_3 := \min\{k \le n : S_k = \max\limits_{j \le n} S_j\}$. In the first two cases set $\tau = n$ if the desired event does not occur.

(2) Assuming each X_k has zero mean, show that $\mathbf{E}[S_\tau] = 0$ for any stopping time τ. Assuming that each X_k has zero mean and finite variance, show that $\mathbf{E}[S_1^2] \le \mathbf{E}[S_\tau^2] \le \mathbf{E}[S_n^2]$ for any stopping time τ.

(3) Give examples of random τ that are not stopping times and for which the results in the second part of the question fail.

Problem 120. Let X_k be independent random variables with zero mean and unit variance. Assume that $\mathbf{E}[|X_k|^{2+\delta}] \le M$ for some $\delta < 0$ and $M < \infty$. Show that S_n is asymptotically normal.

Problem 121. Let X_k be i.i.d. random variables with zero mean and unit variance. Let $0 < a_1 < a_2 < \ldots$ be given numbers. Find sufficient conditions on $(a_i)_i$ such that S_n is asymptotically normal.

Problem 122. Fix $\alpha > 0$.

(1) If X, Y are i.i.d. random variables such that $\frac{X+Y}{2^{\frac{1}{\alpha}}} \stackrel{d}{=} X$, then show that X must have characteristic function $\varphi_X(\lambda) = e^{-c|\lambda|^\alpha}$ for some constant c.

(2) Show that for $\alpha = 2$ we get $N(0, \sigma^2)$ and for $\alpha = 1$ we get symmetric Cauchy.

[*Note:* Only for $0 < \alpha \le 2$ is $e^{-c|\lambda|^\alpha}$ a characteristic function. Hence a distribution with the desired property exists only for this range of α].

Problem 123. Let X_k be independent Ber(p_k) random variables. If Var(S_n) stays bounded, show that S_n cannot be asymptotically normal.

Problem 124 (*Weak law using characteristic functions*). Let X_k be i.i.d. random variables having characteristic function φ.

(1) If $\varphi'(0) = i\mu$, show that the characteristic function of S_n/n converges to the characteristic function of δ_μ. Conclude that weak law holds for S_n/n.

(2) If $\frac{1}{n}S_n \xrightarrow{P} \mu$ for some μ, then show that φ is differentiable at 0 and $\varphi'(0) = i\mu$.

Problem 125. Find the characteristic functions of the distributions with the given densities. (1) $e^{-|x|}$ for $x \in \mathbb{R}$, (2) $\frac{1}{2}\left(1 - \frac{|x|}{2}\right)_+$.

Problem 126 (*Multidimensional central limit theorem*). Let X_n be i.i.d. $\mathbb{R}^d$-valued random vectors with zero mean and covariance matrix Σ. Let $S_n = X_1 + \ldots + X_n$. Show that $\frac{1}{\sqrt{n}}S_n \xrightarrow{d} N_d(0, \Sigma)$ using the replacement principle. Assume (for convenience) that third moments are finite (i.e., $\mathbf{E}[\|X_1\|^3] < \infty$).

Problem 127. [3 marks each] For each of the following statements, state whether they are true or false, and justify or give counterexample accordingly.

(1) If μ, ν are Borel probability measures on $\mathbb{R}$ and $\mu \ll \nu$, then either $\nu \perp \mu$ or $\nu \ll \mu$.

(2) If $\sum\limits_n X_n$ converges a.s. and $\mathbf{P}(Y_n = X_n) = 1 - \frac{1}{n^2}$. Then $\sum\limits_n Y_n$ converges a.s.

(3) If $\{X_n\}$ is an L^2 bounded sequence of random variables, and $\mathbf{E}[X_n] = 1$ for all n, then X_n cannot converge to zero in probability.

(4) If $X_n \xrightarrow{d} X$, then $X_n^2 \xrightarrow{d} X^2$.

(5) Suppose X_n are independent with $\mathbf{E}[X_n] = 0$ and $\sum \mathrm{Var}(X_n) = \infty$. Then, almost surely $\sum X_n$ does not converge.

(6) Suppose X_n, Y_n are random variables such that $|X_n| \leq |Y_n|$ for all n. If $\sum Y_n$ converges almost surely, then $\sum X_n$ converges almost surely.

Problem 128. [2 marks+4 marks + 4 marks] Let X, Y be random variables on a common probability space. Assume that both X and Y have finite variance.

(1) Show that $\mathbf{E}[(X-a)^2]$ is minimized uniquely at $a = \mathbf{E}[X]$.

(2) Find values of a, b that minimize $f(a,b) = \mathbf{E}\left[(Y - a - bX)^2\right]$. Are they unique?

(3) Suppose $\mathbf{P}(X = k) = \frac{1}{10}$ for $k = 1, 2 \ldots, 10$. At what value(s) of a is $\mathbf{E}[|X - a|]$ minimized? Is the minimizer unique?

Problem 129. [10 marks] Let $G_1, G_2, \ldots$ be i.i.d Geometric(p) random variables (this means $\mathbf{P}(G_1 = k) = p(1-p)^{k-1}$ for $k \geq 1$). Let $X_1, X_2, \ldots$ be i.i.d random variables with $\mathbf{E}[|X_1|] < \infty$. Define $N_k := G_1 + G_2 + \ldots + G_k$. Show that as $k \to \infty$,

$$\frac{X_1 + X_2 + \ldots + X_{N_k}}{k} \xrightarrow{P} \frac{1}{p}\mathbf{E}[X_1]$$

Problem 130. [5 marks+5 marks] Let U_k, V_k be i.i.d Uniform([0,1]) random variable.

(1) Show that $\sum\limits_k U_k^{\frac{1}{k}} - V_k^{\frac{1}{k}}$ converges a.s.

(2) Let $S_n = U_1 + U_2^2 + \ldots + U_n^n$. Show that S_n satisfies a CLT. In other words, find a_n, b_n such that $\frac{S_n - a_n}{b_n} \xrightarrow{d} N(0,1)$.

Problem 131. [5 marks+5 marks] Let $\mathbf{Z}^{(n)} = (Z_1^{(n)}, \ldots, Z_n^{(n)})$ be a point sampled uniformly from the sphere S^{n-1} (this means that $\mathbf{P}(\mathbf{Z}^{(n)} \in A) = \text{area}(A)/\text{area}(S^{n-1})$ for any Borel set $A \subseteq S^{n-1}$).

(1) Find the density of $Z_1^{(n)}$.

(2) Using (1) or otherwise, show that $\sqrt{n}Z_1^{(n)} \xrightarrow{d} N(0,1)$ as $n \to \infty$.

[**Hint:** One way to generate $\mathbf{Z}^{(n)}$ is to sample $X_1, \ldots, X_n$ i.i.d N(0,1) and to set $\mathbf{Z}^{(n)} = \frac{1}{\|X\|}(X_1, \ldots, X_n)$ where $\|X\| = \sqrt{X_1^2 + X_2^2 + \ldots + X_n^2}$. You may assume this fact without having to justify it].

Problem 132. [5 marks+5 marks]

(1) Let μ be a probability measure on $\mathbb{R}$ with characteristic function $\hat{\mu}(t)$. Then, show that for any $t_1, t_2, \ldots, t_n \in \mathbb{R}$, the $n \times n$ matrix A with entries $a_{i,j} = \hat{\mu}(t_i - t_j)$ is non-negative definite.

(2) Suppose $|\hat{\mu}(t_0)| = 1$ for some $t_0 \neq 0$. Then, μ is supported on a lattice, that is, $\mu(a\mathbb{Z} + b) = 1$ for some $a, b \in \mathbb{R}$. [**Hint:** Use part (1) with $n = 2$ and appropriate t_1, t_2].

Problem 133. [10 marks] Let $X_1, X_2, \ldots$ be i.i.d Bernoulli$\left(\frac{1}{2}\right)$ random variables. For each $n \geq 1$, define L_n to be the longest run of ones in $(X_1, \ldots, X_n)$, that is,

$$L_n := \max\{k : \exists j \leq n - k \text{ such that } X_{j+1} = X_{j+2} = \ldots = X_{j+k} = 1\}.$$

Prove that $\frac{L_n}{\log n} \xrightarrow{P} \frac{1}{\log 2}$.

Problem 1. Show that it is not possible to define uncountably many independent, non-constant random variables on $([0,1], \mathcal{B}, \lambda)$.

Solution. Let X_i, $i \in I$, be independent, non-constant random variables. Let $\mathcal{F}_I = \sigma\{X_i : i \in I\}$. Let $H = L^2([0,1], \mathcal{B}, \lambda)$ and $H_I = L^2([0,1], \mathcal{F}_I, \lambda)$. Then, H_I is a closed subspace of H and since H is separable, so is H_I.

Let $\{Z_1, Z_2, \ldots\}$ be a countable dense subset in H_I. Then, each Z_n is $\mathcal{F}_I$ measurable and hence $Z_n \in \mathcal{F}_{I_n}$ for some countable set $I_n \subseteq I$. Therefore, $Z_n \in H_{I_n}$. If $J = \cup_n I_n$, then it follows that $H_I = H_J$.

Now if $i \in I \setminus J$, then $X_i \in H_J$, X_i is independent of all random variables in H_J and X_i is not constant. ■