Stat 513

Fall 2025

Problem Set 3

Topic 3: Integration and Expectation

Total: 50 points

Due Sunday, October 5 at 23:59

Problem 1. Cantor Function.

Let $\mathcal{C}$ denote the Cantor set. Recall the definition of the Cantor function:

$$F(x) = \begin{cases} \sum_{i=1}^{\infty} \frac{t_i(x)/2}{2^i} & x \in \mathcal{C} \\ \sup_{y \leq x : y \in \mathcal{C}} c(y) & x \in [0,1] \cap \mathcal{C}^c \end{cases}$$

where t_i is the i-ith ternary digit of x (see PS1). Formally fill in some of the following details from class:

a) Show that the Cantor function is continuous at all $x \in (0,1)$.

b) Show that the Cantor function is not absolutely continuous.

c) Show that $F'(x) = 0$ almost everywhere.

Problem 1 Solution

a) Consider the case where $x \in \mathcal{C}$, and let $\delta > 0$ be arbitrary. Let N be large enough such that $\frac{1}{2^N} < \delta$. We can see that x is contained within the interval

$$[a,b] \equiv \left[\sum_{i=1}^{N} \frac{t_i(x)}{3^i}, \sum_{i=1}^{N} \frac{t_i(x)}{3^i} + \sum_{i=N+1}^{\infty} \frac{2}{3^n} \right]$$

of width $\frac{1}{3^n}$. Since both endpoints are in the Cantor set, we can see that

$$F(b) - F(a) = \sum_{i=N+1}^{\infty} \frac{1}{2^i} = \frac{1}{2^N}.$$

For any $y \in [a,b]$, $|x-y| < \epsilon = \frac{1}{2^N}$ and

$$|F(y) - F(x)| \leq |F(b) - F(a)| < \delta$$

since F is non-decreasing (to see this, note that $x < z$ implies that $\sup_{y \leq x : y \in \mathcal{C}} c(y) \leq \sup_{y \leq z : y \in \mathcal{C}} c(y)$.)

Now consider the case where $x \notin \mathcal{C}$. Let n denote the level at which x is excluded from the points in C_n (equivalently, the first integer such that $t_n(x) = 1$). Then by construction, x is within an interval of width $\frac{1}{3^n}$ that has no points within the Cantor set. As such, $\sup_{y \le x : y \in \mathcal{C}} c(y)$ must be constant on this interval, and as such $|x - y| < \epsilon = \frac{1}{3^n}$ implies that $|F(x) - F(y)| = 0$ which is less than any δ.

b) To show that a function is not absolutely continuous, it suffices to show for some $\delta > 0$, that for every $\epsilon > 0$ there exists a finite collection of intervals such that

$$\sum_{n=1}^{N} |b_i - a_i| < \epsilon \implies \sum_{n=1}^{N} |F(b_i) - F(a_i)| \ge \delta.$$

Let n be large enough such that $\left(\frac{2}{3}\right)^n < \epsilon$. Consider the $N = 2^n$ intervals of width $\frac{1}{3^n}$ within C_n. Then

$$\sum_{n=1}^{N} |b_i - a_i| = \left(\frac{2}{3}\right)^n < \epsilon.$$

However, we can also see that

$$F(b_i) - F(a_i) = \sum_{i=n+1}^{\infty} \frac{1}{2^i} = \frac{1}{2^n}$$

for each i, so

$$\sum_{n=1}^{N} |F(b_i) - F(a_i)| = \left(\frac{2}{2}\right)^n = 1 \equiv \delta$$

c) This follows from the following two facts:

- when $x \notin \mathcal{C}$, then there exists a neighborhood around x such that F is constant on $(x - \epsilon, x + \epsilon)$, and as such $F'(x) = 0$ (see part (a))
- $\mu(\mathcal{C}) = 0$ where μ is the Lebesgue measure, as shown in problem set 1.

Problem 2. Almost everywhere.

Addendum: For the following, you can assume that $\left|\int f d\mu\right| < \infty$ and $\left|\int g d\mu\right| < \infty$

a) Show that if $0 \le f, g$ and $f = g$ almost everywhere (with respect to a measure μ), then

$$\int g d\mu = \int f d\mu.$$

b) Show that if $0 \le f, g$ and $f \le g$ almost everywhere, then

$$\int f d\mu \le \int g d\mu$$

Problem 2 Solution

a) Let's consider first the positive part of the difference, $h^+ = (f - g)^+$. Let A denote the set on which $f = g$, it follows that $h(x) = 0$ when $x \in A$. Consider an arbitrary partition $\{B_i\}_{i=1}^N$. For each term i in the sum

$$\sum_{i=1}^{N} \inf_{x \in B_i} h^+(x)\mu(B_i)$$

if $A \cap B_i \neq \emptyset$, then $\inf_{x \in B_i} h^+(x) = 0$, so the term is zero. Otherwise, $B_i \subset A^c$, so $\mu(B_i) \leq \mu(A^c) = 0$. Since all terms in the sum are zero, it follows that $\int h^+ d\mu = 0$. Using a similar argument for h^-, we can see that $\int (f - g)d\mu = 0$, which implies that $\int f d\mu = \int g d\mu$ as required.

b) Let A denote the set where $f \geq g$. Consider the function $h = f - g$. Since $\int h d\mu = \int h^+ d\mu - \int h^- d\mu$ and we can see that $\int h^- d\mu = 0$ (using the fact that $h^- = 0$ almost everywhere, $\int 0 d\mu = 0$, and the previous part) so it must be that $\int f d\mu \leq \int g d\mu$.

Problem 3. General definition of the integral.

Let $f : \Omega \to \mathbb{R}$ denote a positive function, i.e. $f \geq 0$. Let

$$SF(f) = \left\{ s(x) = \sum_{i=1}^{N} x_i 1\{x \in A_i\} \middle| s \leq f \right\}$$

denote the set of simple functions that are less than f (where $\{A_i\}_{i=1}^N$ denotes a finite partition of Ω). Show that the following two definitions of the integral are equivalent:

a)

$$\int f d\mu = \sup_{s \in SF(f)} \sum_{i=1}^{N} x_i \mu(A_i)$$

b)

$$\int f d\mu = \sup \sum_{i=1}^{N} \left[\inf_{x \in A_i} f(x) \right] \mu(A_i).$$

where the suprenum is over all finite partitions $\{A_i\}_{i=1}^N$ of Ω.

Problem 3 Solution

Because $s(x) = [\inf_{x \in A_i} f(x_i)] 1\{x \in A_i\}$ is a simple function, it follows that (a) $\geq$ (b).

To see that (b) $\geq$ (a), consider a simple function with corresponding partition $\{A_i\}_{i=1}^N$. For any A_i, $s \leq f$ implies that $s \leq \inf_{x \in A_i} f$. As such, for every element in the suprenum of (a), there exists an element in the set of the supremnum of (b) that is greater than or equal to it. As such (b) $\geq$ (a) as required.

Problem 4. Integral of simple functions.

Prove that if f is a (non-negative) simple function, i.e.

$$f(x) = \sum_{i=1}^{N} x_i 1\{x \in A_i\}$$

for a finite partition $\{A_i\}_{i=1}^N$ of Ω, then $\int f d\mu = \sum_{i=1}^N x_i \mu(A_i)$ when using the second definition of Problem 3.

Problem 4 Solution Consider an arbitrary partition $\{B_j\}_{j=1}^M$. We can see that $\{A_i \cap B_j\}_{i,j}$ is also a partition and that

$$\sum_{i=1}^{N} x_i \mu(A_i) = \sum_{i=1}^{N} \sum_{j=1}^{M} \left[\inf_{x \in A_i \cap B_j} f(x) \right] \mu(A_i \cap B_j) \geq \sum_{j=1}^{M} \left[\inf_{x \in B_j} f(x) \right] \mu(B_j)$$

where the first equality is a result of the fact that f is constant over A_i, and the inequality comes from the fact that the infinum over a larger set must be smaller than that over a larger set. As such, the left-hand side term in the above is at least as large as any of the terms in the "sup" of the definition from Problem 3, so they must be equal.