ITERATIVE PROBLEM SOLVING

TEAM SCIENTIFIC DRAGONS

Abstract. When we consider a question involving properties of a random variable, it is often more intuitive to consider either the discrete case or continuous case first. For the two examples below, we consider the continuous case first and then try to generalize our result. In the case of the first example, this technique is helpful when generalizing our result. We then show that considering continuous vs discrete at first creates trouble when generalizing the result. It is in these cases that we then are forced to question whether the result does hold in general.

A RETROSPECTIVE DISCUSSION

We consider the following problem:

Problem 1. Show that if $Y$ is a non-negative random variable on a probability space $(\Omega, \mathcal{F}, P)$ such that $E[Y] < \infty$, then there exists a continuous function $f : [0, 1] \rightarrow \mathbb{R}$ such that $f(0) = 0$ and

$$E(|Y| 1_A) \leq f(P(A)) \quad \text{for all } A \in \mathcal{F}$$

We know that our function must somehow rely on the random variable $Y$. Since, for each, $A$, we require that the function evaluated at $P(A)$ be larger than the expected value of $Y$ restricted to that set, our intuition is to look for a function that searches for the greatest expected value on any given set with measure no greater than $P(A)$. With continuity of our resulting function in mind, we come up with the following solution:

Solution. Define, for $\epsilon > 0$ functions

$$h(x) = \sup_{B : P(B) \leq x} E(|Y| 1_B) , \quad f(x) = \frac{1}{\epsilon} \int_x^{x+\epsilon} h(t) dt.$$

The function $h(x)$ is a monotonically increasing function, measurable on $[0, 1]$. Therefore, we know that $f(x)$ is differentiable almost everywhere, hence $f(x)$ is continuous.

By definition of the supremum, we know that

$$E(|Y| 1_A) \leq h(P(A)).$$

Also, for $t \in [x, x + \epsilon]$, $h(x) \leq h(t)$ so that

$$h(x) = \frac{1}{\epsilon} \int_x^{x+\epsilon} h(x) dt \leq \frac{1}{\epsilon} \int_x^{x+\epsilon} h(t) dt = f(x)$$

Hence we have satisfied

$$E(|Y| 1_A) \leq f(P(A)).$$
Moreover, it is clear that \( f(0) = 0 \), therefore \( f \) is a function that satisfies the conditions specified by the problem.

Although, in principle, we have solved the problem, we are left feeling dissatisfied. Our construction of the function \( h(\cdot) \), and subsequently \( f(\cdot) \), certainly gave an upper bound on \( E[Y \mathbb{1}_A] \). However, we never actually located the set on which we achieve such an upper bound; moreover, it is not easy, in general, to find such a set, as the behavior of random variables, in general, is hard to predict.

If we restrict our attention to the case when \( Y \) is a continuous random variable, then we can consider sets of the form

\[
B_t = \{ \omega : |Y(\omega)| \geq t \}.
\]

Note that for \( A \in F \) such that \( P(A) = P(B_t) \), we have

\[
E[Y \mathbb{1}_A] - E[Y \mathbb{1}_{B_t}] = E[Y(\mathbb{1}_A - \mathbb{1}_{B_t})]
= E[Y(\mathbb{1}_{A \setminus B_t} - \mathbb{1}_{B_t \setminus A})]
= E[Y \mathbb{1}_{A \setminus B_t}] - E[Y \mathbb{1}_{B_t \setminus A}]
\leq tP(A \setminus B_t) - tP(B_t \setminus A)
= 0.
\]

Thus we see that

\[
E[Y \mathbb{1}_A] \leq E[Y \mathbb{1}_{B_t}].
\]

Unfortunately, these calculations do not hold in general. In fact, we cold end up finding more than one value of \( t \) for which the above holds. With a little bit of trial and error, we are able to alter the allowable values of \( t \) and arrive at the following:

**Solution.** For each \( t \) define \( B_t = \{ \omega : |Y(\omega)| \geq t \} \). Then we define \( h_1(x) = \sup \{ t : P(B_t) \geq x \} \) and \( h_2(t) = E(|Y| \mathbb{1}_{B_t}) \). Note that if \( Y \) were a continuous random variable, we could ignore taking the supremum as defined in \( h_1(x) \). However, in the case that \( Y \) is discrete, the set \( \{ t : P(B_t) \geq x \} \) may be a set with nonzero measure. Thus the supremum is required.

To guarantee the result function is continuous, for \( \epsilon > 0 \) we let

\[
f(x) = \frac{1}{\epsilon} \int_0^{x+\epsilon} h_2 \circ h_1(t) dt.
\]

We claim that \( f(x) \) satisfies the conditions prescribed by the problem. Clearly, \( f(0) = 0 \). Continuity follows from the measurability of \( h_2 \circ h_1(t) \) and the fact that \( f(x) \) is therefore differentiable almost everywhere.

Moreover, for any \( A \in F \), we let \( t^* = h_1(P(A)) \). Certainly, \( P(B_{t^*}) \geq P(A) \). Therefore, using the monotonicty of \( h_2 \circ h_1 \), we have

\[
E(|Y| \mathbb{1}_A) \leq E(|Y| \mathbb{1}_{B_{t^*}}) = h_2 \circ h_1(P(A)) \leq f(P(A)).
\]

\( \square \)

**A Warning, by example**

First step analysis is not always as fruitful as it was above. Looking at continuous vs. discrete type random variables can be misleading. We present the following to show just that:

**Problem 2.** Suppose the random variables \( X_n \) are non-negative and \( E[X_n] \to \infty \) as \( n \to \infty \). Can you find events \( A_n, n = 1, 2, \ldots \) such that
(1) \( P(A_n) \to 0 \) as \( n \to \infty \), and
(2) \( E[X_n \mathbb{1}_{A_n}] \to \infty \) as \( n \to \infty \).

Like before, we must be careful when considering whether \( X_n \) are continuous or discrete random variables. To begin with, we assume that \( X_n \) are continuous random variables, and as before, we can then try to extend our argument to general random variables. Under such assumption we obtain our first solution.

**Solution.** Since \( EX_n \to +\infty \), we have that \( EX_n > 1 \) when \( n \) is sufficiently large. For simplicity we assume it holds for all \( n \).

For fixed positive integer \( k \), \( P(t < X_n \leq k) \) is a continuous function of \( t \), since we have assumed that \( X_n \) are continuous. Continuity also guarantees that \( P(k < X_n \leq k) = 0 \). Noting that

\[
0 \leq \frac{1}{\sqrt{EX_n}} P(k - 1 < X_n \leq k) \leq P(k - 1 < X_n \leq k)
\]

by the intermediate value theorem, there exists \( a_{n,k} \in [k - 1, k] \) s.t.

\[
P(a_{n,k} < X_n \leq k) = \frac{1}{\sqrt{EX_n}} P(k - 1 < X_n \leq k).
\]

Define sets

\[
A_n := \bigcup_{k=1}^{\infty} \{a_{n,k} < X_n \leq k\}.
\]

Then

\[
P(A_n) \leq \sum_{k=1}^{\infty} P(a_{n,k} < X_n \leq k) = \frac{1}{\sqrt{EX_n}} \sum_{k=1}^{\infty} P(k - 1 < X_n \leq k) \leq \frac{1}{\sqrt{EX_n}} \to 0.
\]

On the other hand,

\[
E(X_n \mathbb{1}_{A_n}) = \sum_{k=1}^{\infty} E(X_n \mathbb{1}_{A_n})
\]

\[
geq \sum_{k=1}^{\infty} a_{n,k} P(a_{n,k} < X_n \leq k) \geq \sum_{k=1}^{\infty} (k - 1) P(a_{n,k} < X_n \leq k) = \frac{1}{\sqrt{EX_n}} \left( \sum_{k=1}^{\infty} k P(k - 1 < X_n \leq k) - \sum_{k=1}^{\infty} P(k - 1 < X_n \leq k) \right) \geq \frac{1}{\sqrt{EX_n}} (EX_n - 1) \to +\infty.
\]

The proof is complete. \( \Box \)
We now try to use the similar argument to extend our result. However, after some thought, we find that the statement is not necessarily true in general in the discrete case. A counterexample shows this.

**Example:**
Construct a probability space as follows. \( F = \{\emptyset, \Omega\}, P(\emptyset) = 0, P(\Omega) = 1. \)
Define \( X_n(\omega) = n, \forall \omega \in F. \)
Then \( EX_n = n \to +\infty. \)
Suppose there exist events \( A_n \) s.t. \( P(A_n) \to 0. \) Since these events are either the empty set of \( \Omega, \) then when \( n \) is sufficiently large we have \( A_n = \emptyset. \)
Therefore when \( n \) is large enough, \( E[X_n \mathbb{1}_{A_n}] = 0. \)

**Conclusion**

What we have seen through these two problems is that in constructing sets, or events, to get desired integral properties, we need to take special care when classifying random variables. Often, it pays off to start with continuous random variables. The change over to any random variable at that point is usually seamless. This was the case when looking at problem 1. However, problem 2 shows that this methodology is not always so fruitful. By counterexample, we were able to see that sometimes, solutions to problems hold only for one part of the dichotomy of continuous and discrete variables. It is thus important to take care when approaching a problem, also making sure to see if ones first idea makes sense for all random variables with prescribed properties, or just some.