MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.436J/15.085J | Fall 2008
Lecture 15 | 10/31/2007

Recitation 8

1 Some common mistakes on the exam

- While its true that

$$\{\sup_n X_n \leq c\} = \cap_{n\geq 1} \{X_n \leq c\},$$

  it is **not** true that

$$\{\sup_n X_n < c\} = \cap_{n\geq 1} \{X_n < c\}.$$

  Indeed, what if $X_n(\omega) = c - 1/n$?

- In what sense is a random variable defined by its cdf? If a random variables X, Y have the same cdf $F(x)$, then

$$P(X \in (a,b]) = F(b) - F(a) = P(Y \in (a,b]),$$

  and since the Borel σ-field is generated by intervals $(\cdot,\cdot]$, it follows that

$$P(X \in B) = P(Y \in B),$$

  for any Borel set B. So the cdf $F(x)$ uniquely determines the probabilities $P(X \in B)$.

  On the other hand, if X, Y have the same cdf, this says nothing about X, Y as mappings from their sample spaces to R - they might not even be defined on the same space! Indeed, consider two examples.

  Let $\Omega_1 = [0,1], \mathcal{F}_1 = B([0,1])$, and let P_1 be the Lebesgue measure μ. Consider the mapping $X(\omega) = \omega$. If $x < 0$, then $P_1(X \leq x) = 0$; if $x > 1$, then $P_1(X \leq x) = 1$; and if $x \in [0,1]$, then $P_1(X \leq x) = P_1([0,x]) = x$. In other words, the cdf of X is the cdf of the uniform distribution on $[0,1]$.

Now let $\Omega_2 = [0,2], \mathcal{F}_2 = \mathcal{B}([0,2])$, and let $P_2 = (1/2)\mu$ (we must have $P_2(\Omega_2) = 1$, which is why we included the normalizing factor of $1/2$). Now define $X(\omega) = \omega/2$. Then, if $x < 0$, then $P_2(X \leq x) = 0$; if $x > 1$, then $P_2(X \leq x) = 1$; and if $x \in [0,1]$, then $P(X \leq x) = P_2([0,2x]) = (1/2)2x = x$. In other words, the cdf of X is again the cdf of the uniform distribution on $[0,1]$.

2 On the connection between uniform and coin toss models

Let's continue our discussion from the previous recitation on the connection between the uniform random variable an an infinite sequences of coin tosses. We had previously argued that if we generate a uniform random number on $[0,1]$, write it out in binary, and let X_i be its i'th digit, that $P(X_i = 0) = P(X_i = 1) = 1/2$, and further that X_i are independent. In other words, the sequence $X_1, X_2, \ldots$ corresponds to an infinite sequence of coin tosses.

Let's now argue the other way, that if we view an infinite sequence of coin tosses by interpreting the i'th toss as the i'th binary digit of a number, the result will be a uniform random number in $[0,1]$.

- Let $X_1, X_2, \ldots$ be an infinite iid sequence of binary random variables with $P(X_i = 0) = P(X_i = 1) = 1/2$. Define

$$Z = \sum_{i=1}^{\infty} \frac{X_i}{2^i}.$$

- We argue that U is a uniform random variable in $[0,1]$. Indeed,

$$P(Z \in [\frac{k}{2^i}, \frac{k+1}{2^i}]) = \frac{1}{2^i},$$

  since this event corresponds to the first i tosses coming up a certain way; here $k/2^i, (k+1)/2^i$ being numbers in $[0,1]$; Hence we have

$$P(Z \in [\frac{l}{2^i}, \frac{u}{2^i}] = \frac{l-u}{2^i},$$

  by additivity [1]. So, whenever I is a finite union of intervals whose end-

[1] Don't be bothered by the fact that we haven't taken care to make sure our intervals

$$[\frac{k}{2^i}, \frac{k+1}{2^i}],$$

don't overlap. Since $P(Z \in a) = 0$ for any a (since $Z = a$ means one of at most two sequences have occurred, and each sequence occurs with probability 0), worrying that the intervals overlap at the endpoints is a non-issue.

points are rational numbers with denominators which are powers of 2,

$$P(Z \in I) = l(I),$$

where l stands for the length (i.e. lebesgue measure). Since the set of such I is a field, Caratheodory's theorem implies that Z is the uniform random variable.

3 Singular random variables

- Not all random variables are continuous or discrete. Indeed, consider the following random variable. Toss a coin. If it lands on heads, $X = 0$. If it lands on tails, let X be a uniform random variable on $[0, 1]$. The resulting random variable X is neither continuous nor discrete.

- The above random variable, however, is a mixture of continuous and discrete random variables, in the sense there is a countable collection of points with positive mass, and once you remove them the random variable has a density. However, some random variables are stranger than that, being neither continuous, nor discrete, nor a mixture of them. We describe one such random variable next.

- Similar to the previous section, define

$$Z' = \sum_{i=1}^{\infty} \frac{2X_i}{3^i}.$$

  In words, we generate an infinite sequence of 0s and 2s, and view them as the digits of the ternary expansion of a number. Intuitively, we "uniformly" generate a random number with no 1s in its ternary expansion.

- Let C be the support of this random variable, that is, the set of numbers in $[0, 1]$ with no 1s in their ternary expansion. Observe that if C_i is the set of numbers with i'th ternary digit not equal to 1 then

$$C = \cap_{n \geq 1} C_n.$$

- Observe that

$$l(\cap_{i=1}^{n} C_i) = \frac{2}{3} l(\cap_{i=1}^{n-1} C_i),$$

  since intersection with C_n amounts to removing the middle third out of every interval in $\cap_{i=1}^{n-1} C_i$. A consequence of this is that

$$l(C) = 0.$$

- So, Z' is concentrated on a set of measure 0. This implies it can't be continuous. Indeed, if it were, then it would have some density $f_{Z'}$ and then
$$P(Z' \in C) = \int_C f_{Z'}(x)dx = 0,$$
since C has measure 0, which can't be.

- On the other hand,
$$P(Z' = a) = 0,$$
for all $a \in [0, 1]$. This is because every number has at most two ternary expansions, each of which has probability 0 of occurring. So Z' can't be discrete.

- Moreover, Z' is not a mixture of continuous and discrete random variables; there is no sequence of positive masses we can remove so that the result would have a density.

Finally, an aside. The set C is typically used as an example of an uncountable set having measure 0. We've already argued above that C has length 0. Its not hard to see that C is uncountable, since we can take the ternary expansion of every number in C, replace every 2 with a 1, and view the result as a binary expansion. This is a one-to-one map[2]from C to $[0, 1]$, and we know the latter is uncountable.

[2]There is the issue of some numbers having two ternary expansions; if we adopt the convention of never using the expansion which ends in an infinite sequence of 0s, the map will be one-to-one.

6.436J / 15.085J Fundamentals of Probability
Fall 2008