Lecture #21: Tail Bounds
What about Countably Infinite Sample Spaces?

Introduction

Once again, suppose that Ω is a sample space, $\mathsf{P} : \Omega \to \mathbb{R}$ is a probability distribution for Ω, and $X : \Omega \to \mathbb{R}$ is a random variable.

The definition of the ***expected value*** of the random variable X, with respect to P, included an additional condition, namely, that the sum

$$\sum_{\sigma \in \Omega} \mathsf{P}(\sigma) \times |X(\sigma)|$$

is finite — that is, "less than $+\infty$". While the condition is not stated over, and over again, it is "implicitly" required every time the expected value of a random variable is introduced — because the "expected value" of the random variable is only defined when the condition is satisfied.

- For example, Claim #2 (Linearity of Random Variables) concerns random variables

  $$X_1, X_2, \dots, X_n : \Omega \to \mathbb{R},$$

  as well as random variables

  $$Y = X_1 + X_2 + \cdots + X_n$$

  and

  $$Z = aX_1 + b$$

  for real numbers a and b. It is "implicitly" required that

  $$\sum_{\sigma \in \Omega} P(\sigma) \times |X_i(\sigma)|$$

is finite, for every integer i such that $1 \leq i \leq n$ — because the expected values $\mathsf{E}[X_i]$, for $1 \leq i \leq n$, are not guaranteed to be defined, otherwise. If this condition is satisfied then it turns out that both

$$\sum_{\sigma \in \Omega} \mathsf{P}(\sigma) \times |Y(\sigma)| \quad \text{and} \qquad \sum_{\sigma \in \Omega} \mathsf{P}(\sigma) \times |Z(\sigma)|$$

are finite as well — and the relationships in the claim can be established, even when Ω is countably infinite.

This condition is trivially satisfied if the sample space, Ω is finite — then (since X and P are required to be well-defined total functions from Ω to $\mathbb{R}$) the sum is a sum of finitely many terms, and its value must certainly be fixed and finite, as well.

Lecture #21 introduced several results that could be used to establish “tail bounds”, that is, bounds on the probabilities that the values of random variables exceeded given thresholds. Most of these results were only stated for the case that the sample space, Ω, is finite.

We we are often dealing with experiments with finite sample spaces, there also interesting experiments — including one in which we repeatedly toss a fair coin until we see “Heads” — where the sample space is countably infinite, instead. It would be helpful to have results that can be applied in these situations too.

The goal of this document is to explain why additional technical conditions must be checked when we are working with countably infinite sample spaces, and to identify situations where these conditions are trivially satisfied (so that we can apply results in much the same was as we can for finite sample spaces).

Why Infinite Series Complicate Things

Consider the *alternating harmonic series*

$$\sum_{i=1}^{+\infty} \frac{(-1)^{i+1}}{i} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + - \dots$$

One way to approach this is to define (for $\in \mathbb{N}$)

$$S_n = \sum_{i=1}^{n} \frac{(-1)^{i+1}}{i}$$

and to think of the above series as

$$\lim_{n \to +\infty} S_n.$$

This limit exists and is equal to $\ln 2$.

We could also *reorder the terms in the series*. Suppose that, every time we included the next *positive term*, we included the *next two negative terms* after it:

$$1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + - - ...$$

This series includes *exactly the same* set of terms as the previous one did. It simply lists them in different order. If you set T_n to be the sum of the first n terms in the above series, then you can check that

$$\lim_{n\to+\infty} T_n$$

also exists — but this limit is $\frac{1}{2}\ln 2$ instead of $\ln 2$.

Indeed, it can be proved that there is a way to reorder the terms in this infinite series, so that you are still including the same set of terms, and the limit of the series is ***whatever real number you might want it to be***. Proofs that are straightforward and correct, for finite sample spaces and finite seres, can *fall apart* for infinite sample spaces, and infinite series, because you are changing the order of the terms in your series as you go from one line of your argument to the next.

Now, if we take the sum of the *absolute values* of the terms in our original series, then the series that we get is the *harmonic series*

$$\sum_{i=1}^{+\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \dots$$

— which is ***divergent***: That is, if we set

$$H_n = \sum_{i=1}^{n} \tfrac{1}{n}$$

then it can be shown that $\lim\limits_{n\to+\infty} H_n = +\infty$ — this series does not converge to any real number, at all.

An infinite series

$$\sum_{\mu\in\Omega} \alpha_\mu$$

is said to be ***absolutely convergent*** if the series

$$\sum_{\mu\in\Omega} |\alpha_\mu|$$

(including the *absolute values* of the terms in the original series) converges to a real number (instead of to $+\infty$).

So, Things are Not So Bad

It turns out that if a series is absolutely convergent then the original series also converges to a real number and, furthermore, this value is well-defined — you cannot change this value by changing the order in which the terms in the series are included. The proofs that were correct for finite series are — generally — correct for absolutely convergent series too.

So, that pesky technical condition in the definition of "expected value" is just saying that we are only going to define and use these values when it makes (mathematical) sense to do so — that is, when the values we want to work with really ***are*** well-defined, and well-behaved. Various ***results*** that were stated in the preparatory reading for Lecture #20 still hold for countably infinite sample spaces when expected values of random variables are defined.

- In particular, if Ω is a countably infinite sample space with probability distribution $\mathsf{P} : \Omega \to \mathbb{R}$, and $X_1, X_2, \dots, X_n : \Omega \to \mathbb{R}$ are random variables such that each of the series

  $$\sum_{\mu \in \Omega} X_i(\mu) \times \mathsf{P}(\mu)$$

  is absolutely convergent — so that the expected value of X_i is defined — for $1 \leq i \leq n$, then the series

  $$\sum_{\mu \in \Omega} (X_1(\mu) + X_2(\mu) + \cdots + X_n(\mu)) \times \mathsf{P}(\mu)$$

  is also absolutely convergent — so that the expected value of the random variable $X_1 + X_2 + \cdots + X_n$ is also defined — and

  $$\mathsf{E}[X_1 + X_2 + \cdots + X_n] = \mathsf{E}[X_1] + \mathsf{E}[X_2] + \cdots + \mathsf{E}[X_n]$$

  — as claimed for the case that Ω is finite, in part (a) of Claim #2 from Lecture #20.

- If Ω is a countably infinite sample space with probability distribution $\mathsf{P} : \Omega \to \mathbb{R}$ $X : \Omega \to \mathbb{R}$ is a random variable such that the series

  $$\sum_{\mu \in \Omega} X(\mu) \times \mathsf{P}(\mu)$$

  is absolutely convergent — so that the expected value of the random variable X is defined — and $a, b \in \mathbb{R}$, then the series

  $$\sum_{\mu \in \Omega} (a \cdot X(\mu) + b) \cdot \mathsf{P}(\mu)$$

  is also absolutely convergent — so that the expected value of the random variable $a \cdot X + b$ is defined — and

  $$\mathsf{E}[a \cdot X + b] = a \cdot \mathsf{E}[X] + b$$

  — as is claimed for the case that Ω is finite, in part (b) of Claim #2 from Lecture #20.

Claims and their proofs, in Lecture #21, can be extended as follows.

- If Ω is a countably infinite sample space with a probability distribution $\mathsf{P} : \Omega \to \mathbb{R}$, $X : \Omega \to V$ is a random variable (for some set $V \subseteq \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$ is a total function and the series
$$\sum_{\mu \in \Omega} g(X(\mu)) \times \mathsf{P}(\mu)$$
is absolutely convergent — so that
$$\sum_{\mu \in \Omega} |g(X(\mu) \times \mathsf{P}(\mu)|$$
is finite — then
$$\mathsf{E}[g(X)] = \sum_{w \in g(V)} w \times \mathsf{P}(g(X) = w)$$
— as is claimed in Lemma #1 from Lecture #21 for the case that Ω is finite — and (since it only involves reordering terms in the series) the same proof can be used without change.

- Suppose that Ω is a countably infinite sample space with probability distribution $\mathsf{P} : \Omega \to \mathbb{R}$. Let $X : \Omega \to V$ be a random variable (so that $V \subseteq \mathbb{R}$) and let $a, b \in V$.

  (i) If $\mathsf{P}(X = b) = 1$ then
  $$\sum_{\mu \in \Omega} |X(\mu)| = |b|,$$
  so that the series
  $$\sum_{\mu \in \Omega} X(\mu)$$
  is absolutely convergent, and the expected value of X is defined. Furthermore,
  $$\mathsf{E}[X] = b$$
  as claimed in Lemma #2(i), for the case that the sample space Ω is finite.

  (ii) If $\mathsf{P}(a < X \leq b) = 1$ then
  $$\sum_{\mu \in \Omega} |X(\mu)| \leq \max(|a|, |b|),$$
  so that the series
  $$\sum_{\mu \in \Omega} X(\mu)$$
  is absolutely convergent, and the expected value of X is defined. Furthermore,
  $$a \leq \mathsf{E}[X] \leq b,$$
  which is *almost* what is claimed in Lemma #2(ii), for the case that Ω is finite.

(iii) Suppose that $g, h : \mathbb{R} \to \mathbb{R}$ are total functions and that the series

$$\sum_{\mu \in \Omega} |g(X(\mu))| \qquad \text{and} \qquad \sum_{\mu \in \Omega} |h(X(\mu))|$$

are both finite — so that the series

$$\sum_{\mu \in \Omega} g(X(\mu)) \qquad \text{and} \qquad \sum_{\mu \in \Omega} h(X(\mu))$$

are both absolutely convergent, and the expected values of $g(X)$ and $h(X)$ are defined. Then the series

$$\sum_{\mu \in \Omega} (g+h)(X(\mu))$$

is also absolutely convergent, so that the expected value of $(g+h)(X)$ is defined, and

$$\mathsf{E}[(g+h)(X)] = \mathsf{E}[g(X) + h(X)] = \mathsf{E}[g(X)] + \mathsf{E}[h(X)],$$

as is claimed in Lemma #2(iii), for the case that Ω is finite.

- The ***Basic Inequality*** — given as Theorem #3 in Lecture #20 — also holds when Ω is a countably infinite sample space, instead of finite, the other conditions given in the result all hold, and the series

$$\sum_{\mu \in \Omega} h(X(\mu)) \times \mathsf{P}(\mu)$$

  is absolutely convergent — so that the expected value of the random variable $h(X)$ is defined.

- ***Markov's Inequality*** also holds for a random variable X, over a countably infinite sample space Ω with probability distribution $\mathsf{P} : \Omega \to \mathbb{R}$, provided that the series

$$\sum_{\mu \in \Omega} |X(\mu)| \times \mathsf{P}(\mu)$$

  is finite — in which case this series is also absolutely convergent, and the expected values of the random variables X and $|X|$ are both defined.

- Let Ω be a countably infinite sample space with probability distribution $\mathsf{P} : \Omega \to \mathbb{R}$, let $X : \Omega \to \mathbb{R}$ be a random variable such that the series

$$\sum_{\mu \in \Omega} |X(\mu) \times \mathsf{P}(\mu)|$$

  and

$$\sum_{\mu \in \Omega} X^2(\mu) \times \mathsf{P}(\mu)$$

are both finite — so that the expected values of the random variables X and X^2 are both defined. Then

$$\mathsf{var}(X) = \sum_{\mu \in \Omega} (X(\mu) - \mathsf{E}[X])^2 \times \mathsf{P}(\mu)$$

is also finite, and

$$\mathsf{var}(X) = \mathsf{E}[X^2] - \mathsf{E}[X]^2$$

— as claimed, for the case that Ω is finite, in Theorem #5 in Lecture #21.

- Chebyshev's Inequality includes the expected value of a *non-negative* random variable, namely, the square X^2 — and the inequality will hold (when the sample space is countably infinite) as long as $\mathsf{E}[X^2]$ is finite,

- Cantelli's Inequality involves the ***variance*** of a random variable X — and this will also be an infinite series if the sample space, Ω, is countably infinite. Once again, though, this is an infinite series whose terms are all greater than or equal to zero: The series will be "well behaved", with a well defined value, as long as the series has a finite value. If it does the Cantelli's Inequality can be applied just as it can when the sample space is finite.

- The conditions that must be satisfied, in order for the Chernoff bound to be applicable at all, guarantee that that any infinite series that must be worked with are absolutely convergent. Thus no additional conditions must be checked to state (and use) a version of this bound for countably infinite sample spaces.

With all that noted — you should be careful when working with values that are defined as summations, when the sample space is countably infinite: You might actually be working with an infinite series and, as noted above, its "value" will generally not be well-defined, at all, unless the series is absolutely convergent.