1 Introduction

How is this course different from your earlier probability courses? There are some problems that simply can’t be handled with finite-dimensional sample spaces and random variables that are either discrete or have densities.

Example 1 Try to express the strong law of large numbers without using an infinite-dimensional space. Oddly enough, the weak law of large numbers requires only a sequence of finite-dimensional spaces, but the strong law concerns entire infinite sequences.

Example 2 Consider a distribution whose cumulative distribution function (cdf) increases continuously part of the time but has some jumps. Such a distribution is neither discrete nor continuous. How do you define the mean of such a random variable? Is there a way to treat such distributions together with discrete and continuous ones in a unified manner?

General Measures Both of the above examples are accommodated by a generalization of the theories of summation and integration. Indeed, summation becomes a special case of the more general theory of integration. It all begins with a generalization of the concept of “size” of a set.

Example 3 One way to measure the size of a set is to count its elements. All infinite sets would have the same size (unless you distinguish different infinite cardinals).

Example 4 Special subsets of Euclidean spaces can be measured by length, area, volume, etc. But what about sets with lots of holes in them? For example, how large is the set of irrational numbers between 0 and 1?

We will use measures to say how large sets are. First, we have to decide which sets we will measure.
2 σ-fields

Definition 1 (fields and σ-fields) Let Ω be a set. A collection \( F \) of subsets of Ω is called a field if it satisfies

- \( \Omega \in F \),
- for each \( A \in F \), \( A^c \in F \),
- for all \( A_1, A_2 \in F \), \( A_1 \cup A_2 \in F \).

A field \( F \) is a σ-field if, in addition, it satisfies

- for every sequence \( \{A_k\}_{k=1}^\infty \) in \( F \), \( \bigcup_{k=1}^\infty A_k \in F \).

We will define measures on fields and σ-field’s.

Definition 2 (Measurable Space) A set \( \Omega \) together with a σ-field \( F \) is called a measurable space \((\Omega, F)\), and the elements of \( F \) are called measurable sets.

Example 5 (Intervals on \( \mathbb{R}^1 \)) Let \( \Omega = \mathbb{R} \) and define \( U \) to be the collection of all unions of finitely many disjoint intervals of the form \((a, b]\) or \((-\infty, b]\) or \((a, \infty)\) or \((-\infty, \infty)\), together with \( \emptyset \). Then \( U \) is a field.

Example 6 (Power set) Let \( \Omega \) be an arbitrary set. The collection of all subsets of \( \Omega \) is a σ-field. It is denoted \( 2^\Omega \) and is called the power set of \( \Omega \).

Example 7 (Trivial σ-field) Let \( \Omega \) be an arbitrary set. Let \( F = \{\Omega, \emptyset\} \). This is the trivial σ-field.

Exercise 1 Let \( F_1, F_2, \ldots \) be classes of sets in a common space \( \Omega \) such that \( F_n \subset F_{n+1} \) for each \( n \). Show that if each \( F_n \) is a field, then \( \bigcup_{n=1}^\infty F_n \) is also a field.

If each \( F_n \) is a σ-field, then is \( \bigcup_{n=1}^\infty F_n \) also necessarily a σ-field? Think about the following case: \( \Omega \) is the set of nonnegative integers and \( F_n \equiv \sigma(\{\{0\}, \{1\}, \ldots, \{n\}\}) \).
Generated $\sigma$-fields A field is closed under finite set theoretic operations whereas a $\sigma$-field is closed under countable set theoretic operations. In a problem dealing with probabilities, one usually deals with a small class of subsets $\mathcal{A}$, for example the class of subintervals of $(0,1]$. It is possible that if we perform countable operations on such a class $\mathcal{A}$ of sets, we might end up operating on sets outside the class $\mathcal{A}$. Hence, we would like to define a class denoted by $\sigma(\mathcal{A})$ in which we can safely perform countable set-theoretic operations. This class $\sigma(\mathcal{A})$ is called the $\sigma$-field generated by $\mathcal{A}$, and it is defined as the intersection of all the $\sigma$-fields containing $\mathcal{A}$ (exercise: show that this is a $\sigma$-field). $\sigma(\mathcal{A})$ is the smallest $\sigma$-field containing $\mathcal{A}$.

Example 8 Let $\mathcal{C} = \{A\}$ for some nonempty $A$ that is not itself $\Omega$. Then $\sigma(\mathcal{C}) = \{\emptyset, A, A^c, \Omega\}$.

Example 9 Let $\Omega = \mathbb{R}$ and let $\mathcal{C}$ be the collection of all intervals of the form $(a,b]$. Then the field generated by $\mathcal{C}$ is $\mathcal{U}$ from Example 5 while $\sigma(\mathcal{C})$ is larger.

Example 10 (Borel $\sigma$-field) Let $\Omega$ be a topological space and let $\mathcal{C}$ be the collection of open sets. Then $\sigma(\mathcal{C})$ is called the Borel $\sigma$-field. If $\Omega = \mathbb{R}$, the Borel $\sigma$-field is the same as $\sigma(\mathcal{C})$ in Example 9. The Borel $\sigma$-field of subsets of $\mathbb{R}^k$ is denoted $\mathcal{B}^k$.

Exercise 2 Give some examples of classes of sets $\mathcal{C}$ such that $\sigma(\mathcal{C}) = \mathcal{B}^1$.

Exercise 3 Are there subsets of $\mathbb{R}$ which are not in $\mathcal{B}^1$?

3 Measures

Notation 11 (Extended Reals) The extended reals is the set of all real numbers together with $\infty$ and $-\infty$. We shall denote this set $\overline{\mathbb{R}}$. The positive extended reals, denoted $\overline{\mathbb{R}}^+$ is $(0,\infty]$, and the nonnegative extended reals, denoted $\overline{\mathbb{R}}^{+0}$ is $[0,\infty]$.

Definition 3 Let $(\Omega, \mathcal{F})$ be a measurable space. Let $\mu : \mathcal{F} \to \overline{\mathbb{R}}^{+0}$ satisfy

- $\mu(\emptyset) = 0$,
- for every sequence $\{A_k\}_{k=1}^\infty$ of mutually disjoint elements of $\mathcal{F}$, $\mu(\bigcup_{k=1}^\infty A_k) = \sum_{k=1}^\infty \mu(A_k)$.

Then $\mu$ is called a measure on $(\Omega, \mathcal{F})$ and $(\Omega, \mathcal{F}, \mu)$ is a measure space. If $\mathcal{F}$ is merely a field, then a $\mu$ that satisfies the above two conditions whenever $\bigcup_{k=1}^\infty A_k \in \mathcal{F}$ is called a measure on the field $\mathcal{F}$. 
Example 12 Let $\Omega$ be arbitrary with $\mathcal{F}$ the trivial $\sigma$-field. Define $\mu(\emptyset) = 0$ and $\mu(\Omega) = c$ for arbitrary $c > 0$ (with $c = \infty$ possible).

Example 13 (Counting measure) Let $\Omega$ be arbitrary and $\mathcal{F} = 2^\Omega$. For each finite subset $A$ of $\Omega$, define $\mu(A)$ to be the number of elements of $A$. Let $\mu(A) = \infty$ for all infinite subsets. This is called counting measure on $\Omega$.

Definition 4 (Probability measure) Let $(\Omega, \mathcal{F}, P)$ be a measure space. If $P(\Omega) = 1$, then $P$ is called a probability, $(\Omega, \mathcal{F}, P)$ is a probability space, and elements of $\mathcal{F}$ are called events.

Sometimes, if the name of the probability $P$ is understood or is not even mentioned, we will denote $P(E)$ by Pr($E$) for events $E$.

Infinite measures pose a few unique problems. Some infinite measures are just like finite ones.

Definition 5 ($\sigma$-finite measure) Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, and let $\mathcal{C} \subseteq \mathcal{F}$. Suppose that there exists a sequence $\{A_n\}_{n=1}^\infty$ of elements of $\mathcal{C}$ such that $\mu(A_n) < \infty$ for all $n$ and $\Omega = \bigcup_{n=1}^\infty A_n$. Then we say that $\mu$ is $\sigma$-finite on $\mathcal{C}$. If $\mu$ is $\sigma$-finite on $\mathcal{F}$, we merely say that $\mu$ is $\sigma$-finite.

Example 14 Let $\Omega = \mathbb{Z}$ with $\mathcal{F} = 2^\Omega$ and $\mu$ being counting measure. This measure is $\sigma$-finite. Counting measure on an uncountable space is not $\sigma$-finite.

Exercise 4 Prove the claims in Example 14.

3.1 Basic properties of measures

There are several useful properties of measures that are worth knowing.

First, measures are countably subadditive in the sense that

$$\mu \left( \bigcup_{n=1}^\infty A_n \right) \leq \sum_{n=1}^\infty \mu(A_n),$$

for arbitrary sequences $\{A_n\}_{n=1}^\infty$. The proof of this uses a standard trick for dealing with countable sequences of sets. Let $B_1 = A_1$ and let $B_n = A_n \setminus \bigcup_{i=1}^{n-1} B_i$ for $n > 1$. The $B_n$’s are disjoint and have the same finite and countable unions as the $A_n$’s. The proof of Equation 5 relies on the additional fact that $\mu(B_n) \leq \mu(A_n)$ for all $n$.

Next, if $\mu(A_n) = 0$ for all $n$, it follows that $\mu \left( \bigcup_{n=1}^\infty A_n \right) = 0$. This gets used a lot in proofs. Similarly, if $\mu$ is a probability and $\mu(A_n) = 1$ for all $n$, then $\mu \left( \bigcap_{n=1}^\infty A_n \right) = 1.$
Definition 6 (Almost sure/almost everywhere) Suppose that some statement about elements of \( \Omega \) holds for all \( \omega \in A^C \) where \( \mu(A) = 0 \). Then we say that the statement holds almost everywhere, denoted a.e. \([\mu]\). If \( P \) is a probability, then almost everywhere is often replaced by almost surely, denoted a.s. \([P]\).

Example 15 Let \((\Omega, \mathcal{F}, P)\) be a probability space. Let \(\{X_n\}_{n=1}^\infty\) be a sequence of functions from \(\Omega\) to \(\mathbb{R}\). To say that \(X_n\) converges to \(X\) a.s. \([P]\) (denoted \(X_n \overset{a.s.}{\to} X\)) means that there is a set \(A\) with \(P(A) = 0\) and \(\lim_{n \to \infty} X_n(\omega) = X(\omega)\) for all \(\omega \in A^C\).

Proposition 6 (Linearity) If \(\mu_1, \mu_2, \ldots\) are all measures on \((\Omega, \mathcal{F})\) and if \(\{a_n\}_{n=1}^\infty\) is a sequence of positive numbers, then \(\sum_{n=1}^\infty a_n \mu_n\) is a measure on \((\Omega, \mathcal{F})\).

Exercise 7 Prove Proposition 6.

3.2 Monotone sequences of sets and limits of measure

Definition 7 (Monotone sequences of sets) Let \((\Omega, \mathcal{F}, \mu)\) be a measure space. A sequence \(\{A_n\}_{n=1}^\infty\) of elements of \(\mathcal{F}\) is called monotone increasing if \(A_n \subseteq A_{n+1}\) for each \(n\). It is monotone decreasing if \(A_n \supseteq A_{n+1}\) for each \(n\).

Lemma 16 Let \((\Omega, \mathcal{F}, \mu)\) be a measure space. Let \(\{A_n\}_{n=1}^\infty\) be a monotone sequence of elements of \(\mathcal{F}\). Then \(\lim_{n \to \infty} \mu(A_n) = \mu(\lim_{n \to \infty} A_n)\) if either of the following hold:

- the sequence is increasing,
- the sequence is decreasing and \(\mu(A_k) < \infty\) for some \(k\).

Proof: Define \(A_\infty = \lim_{n \to \infty} A_n\). In the first case, write \(B_1 = A_1\) and \(B_n = A_n \setminus A_{n-1}\) for \(n > 1\). Then \(A_n = \bigcup_{k=1}^{n} B_k\) for all \(n\) (including \(n = \infty\)). Then \(\mu(A_n) = \sum_{k=1}^{n} \mu(B_k)\), and

\[
\mu\left(\lim_{n \to \infty} A_n\right) = \mu(A_\infty) = \sum_{k=1}^{\infty} \mu(B_k) = \lim_{n \to \infty} \sum_{k=1}^{n} \mu(B_k) = \lim_{n \to \infty} \mu(A_n).
\]

In the second case, write \(B_n = A_n \setminus A_{n+1}\) for all \(n \geq k\). Then, for all \(n > k\),

\[
A_k \setminus A_n = \bigcup_{i=k}^{n-1} B_i,
\]

\[
A_k \setminus A_\infty = \bigcup_{i=k}^{\infty} B_i.
\]
By the first case,
\[ \lim_{n \to \infty} \mu(A_k \setminus A_n) = \mu \left( \bigcup_{i=k}^{\infty} B_i \right) = \mu(A_k \setminus A_\infty). \]

Because \( A_n \subseteq A_k \) for all \( n > k \) and \( A_\infty \subseteq A_k \), it follows that
\[
\begin{align*}
\mu(A_k \setminus A_n) &= \mu(A_k) - \mu(A_n), \\
\mu(A_k \setminus A_\infty) &= \mu(A_k) - \mu(A_\infty).
\end{align*}
\]

It now follows that \( \lim_{n \to \infty} \mu(A_n) = \mu(A_\infty) \).

Exercise 8 Construct a simple counterexample to show that the condition \( \mu(A_k) < \infty \) is required in the second claim of Lemma 16.

3.3 Uniqueness of Measures

There is a popular method for proving uniqueness theorems about measures. The idea is to define a function \( \mu \) on a convenient class \( C \) of sets and then prove that there can be at most one extension of \( \mu \) to \( \sigma(C) \).

Example 17 Suppose it is given that for any \( a \in \mathbb{R} \),
\[ P((-\infty, a]) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} \exp(-u^2/2) \, du. \]
Does that uniquely define a probability measure on the class of Borel subsets of the line, \( B^1 \)?

Definition 8 (\( \pi \)-system and \( \lambda \)-system) A collection \( A \) of subsets of \( \Omega \) is a \( \pi \)-system if, for all \( A_1, A_2 \in A \), \( A_1 \cap A_2 \in A \). A class \( C \) is a \( \lambda \)-system if
\[
\begin{align*}
&\Omega \in C, \\
&\text{for each } A \in C, \ A^C \in C, \\
&\text{for each sequence } \{A_n\}_{n=1}^{\infty} \text{ of disjoint elements of } C, \bigcup_{n=1}^{\infty} A_n \in C.
\end{align*}
\]

Example 18 The collection of all intervals of the form \((-\infty, a]\) is a \( \pi \)-system of subsets of \( \mathbb{R} \). So too is the collection of all intervals of the form \((a, b]\) (together with \( \emptyset \)). The collection of all sets of the form \( \{(x, y) : x \leq a, y \leq b\} \) is a \( \pi \)-system of subsets of \( \mathbb{R}^2 \). So too is the collection of all rectangles with sides parallel to the coordinate axes.

Some simple results about \( \pi \)-systems and \( \lambda \)-systems are the following.
Proposition 9 If $\Omega$ is a set and $\mathcal{C}$ is both a $\pi$-system and a $\lambda$-system, then $\mathcal{C}$ is a $\sigma$-field.

Proposition 10 Let $\Omega$ be a set and let $\Lambda$ be a $\lambda$-system of subsets. If $A \in \Lambda$ and $A \cap B \in \Lambda$ then $A \cap B^C \in \Lambda$.

Exercise 11 Prove Propositions 9 and 10.

Lemma 19 ($\pi - \lambda$ theorem) Let $\Omega$ be a set and let $\Pi$ be a $\pi$-system and let $\Lambda$ be a $\lambda$-system that contains $\Pi$. Then $\sigma(\Pi) \subseteq \Lambda$.

Proof: Define $\lambda(\Pi)$ to be the smallest $\lambda$-system containing $\Pi$. For each $A \subseteq \Omega$, define $\mathcal{G}_A$ to be the collection of all sets $B \subseteq \Omega$ such that $A \cap B \in \lambda(\Pi)$.

First, we show that $\mathcal{G}_A$ is a $\lambda$-system for each $A \in \lambda(\Pi)$. To see this, note that $A \cap \Omega \in \lambda(\Pi)$, so $\Omega \in \mathcal{G}_A$. If $B \in \mathcal{G}_A$, then $A \cap B \in \lambda(\Pi)$, and Proposition 10 says that $A \cap B^C \in \lambda(\Pi)$, so $B^C \in \mathcal{G}_A$. Finally, $\{B_n\}_{n=1}^\infty \in \mathcal{G}_A$ with the $B_n$ disjoint implies that $A \cap B_n \in \lambda(\Pi)$ with $A \cap B_n$ disjoint, so their union is in $\lambda(\Pi)$. But their union is $A \cap (\bigcup_{n=1}^\infty B_n)$. So $\bigcup_{n=1}^\infty B_n \in \mathcal{G}_A$.

Next, we show that $\lambda(\Pi) \subseteq \mathcal{G}_C$ for every $C \in \lambda(\Pi)$. Let $A, B \in \Pi$, and notice that $A \cap B \in \Pi$, so $B \in \mathcal{G}_A$. Since $\mathcal{G}_A$ is a $\lambda$-system containing $\Pi$, it must contain $\lambda(\Pi)$. It follows that $A \cap C \in \lambda(\Pi)$ for all $C \in \lambda(\Pi)$. If $C \in \lambda(\Pi)$, it then follows that $A \in \mathcal{G}_C$. So, $\Pi \subseteq \mathcal{G}_C$ for all $C \in \lambda(\Pi)$. Since $\mathcal{G}_C$ is a $\lambda$-system containing $\Pi$, it must contain $\lambda(\Pi)$.

Finally, if $A, B \in \lambda(\Pi)$, we just proved that $B \in \mathcal{G}_A$, so $A \cap B \in \lambda(\Pi)$ and hence $\lambda(\Pi)$ is also a $\pi$-system. By Proposition 9, $\lambda(\Pi)$ is a $\sigma$-field containing $\Pi$ and hence must contain $\sigma(\Pi)$. Since $\lambda(\Pi) \subseteq \Lambda$, the proof is complete.

The uniqueness theorem is the following.

Theorem 20 (Uniqueness theorem) Suppose that $\mu_1$ and $\mu_2$ are measures on $(\Omega, \mathcal{F})$ and $\mathcal{F} = \sigma(\Pi)$, for a $\pi$-system $\Pi$. If $\mu_1$ and $\mu_2$ are both $\sigma$-finite on $\Pi$ and they agree on $\Pi$, then they agree on $\mathcal{F}$.

Proof: First, let $C \in \Pi$ be such that $\mu_1(C) = \mu_2(C) < \infty$, and define $\mathcal{G}_C$ to be the collection of all $B \in \mathcal{F}$ such that $\mu_1(B \cap C) = \mu_2(B \cap C)$. It is easy to see that $\mathcal{G}_C$ is a $\lambda$-system that contains $\Pi$, hence it equals $\mathcal{F}$ by Lemma 19. (For example, if $B \in \mathcal{G}_C$, $\mu_1(B^C \cap C) = \mu_1(C) - \mu_1(B \cap C) = \mu_2(C) - \mu_2(B \cap C) = \mu_2(B^C \cap C)$, so $B^C \in \mathcal{G}_C$.)

Since $\mu_1$ and $\mu_2$ are $\sigma$-finite, there exists a sequence $\{C_n\}_{n=1}^\infty \in \Pi$ such that $\mu_1(C_n) = \mu_2(C_n) < \infty$, and $\Omega = \bigcup_{n=1}^\infty C_n$. (Since $\Pi$ is only a $\pi$-system, we cannot assume that the $C_n$ are disjoint.) For each $A \in \mathcal{F}$,

$$\mu_j(A) = \lim_{n \to \infty} \mu_j \left( \bigcup_{i=1}^n [C_i \cap A] \right) \text{ for } j = 1, 2.$$
Since \( \mu_j \left( \bigcup_{i=1}^n [C_i \cap A] \right) \) can be written as a linear combination of values of \( \mu_j \) at sets of the form \( A \cap C_i \), where \( C_i \in \Pi \) is the intersection of finitely many of \( C_1, \ldots, C_n \), it follows from \( A \in \mathcal{G}_C \) that \( \mu_1 \left( \bigcup_{i=1}^n [C_i \cap A] \right) = \mu_2 \left( \bigcup_{i=1}^n [C_i \cap A] \right) \) for all \( n \), hence \( \mu_1(A) = \mu_2(A) \). \( \square \)

Exercise 12 Return to Example 17. You should now be able to answer the question posed there.

Exercise 13 Suppose that \( \Omega = \{a, b, c, d, e\} \) and I tell you the value of \( P(\{a, b\}) \) and \( P(\{b, c\}) \). For which subset of \( \Omega \) do I need to define \( P(\cdot) \) in order to have a unique extension of \( P \) to a \( \sigma \)-field of subsets of \( \Omega \)?

4 Lebesgue Measure and Carathéodory’s Extension Theorem

Let \( F \) be a cdf (nondecreasing, right-continuous, limits equal 0 and 1 at \(-\infty\) and \( \infty \) respectively). Let \( \mathcal{U} \) be the field in Example 5 (unions of finitely many disjoint intervals). Define \( \mu : \mathcal{U} \to [0, 1] \) by \( \mu(A) = \sum_{k=1}^n F(b_k) - F(a_k) \) when \( A = \bigcup_{k=1}^n (a_k, b_k] \) and \( \{(a_k, b_k]\} \) are disjoint. This set-function is well-defined and finitely additive.

Is \( \mu \) countably additive as probabilities are supposed to be? That is, if \( A = \bigcup_{i=1}^\infty A_i \), where the \( A_i \)'s are disjoint, each \( A_i \) is a union of finitely many disjoint intervals, and \( A \) itself is the union of finitely many disjoint intervals \((a_k, b_k]\) for \( k = 1, \ldots, n \), does \( \mu(A) = \sum_{i=1}^\infty \mu(A_i) \)?

First, take the collection of intervals that go into all of the \( A_i \)'s and split them, if necessary, so that each is a subset of at most one of the \((a_k, b_k]\) intervals. Then apply the following result to each \((a_k, b_k]\).

Lemma 21 Let \( (a, b] = \bigcup_{k=1}^\infty (c_k, d_k] \) with the \((c_k, d_k]\)'s disjoint. Then \( F(b) - F(a) = \sum_{k=1}^\infty F(d_k) - F(c_k) \).

Proof: Since \((a, b] \supseteq \bigcup_{k=1}^n (c_k, d_k]\) for all \( n \), it follows that \( F(b) - F(a) \geq \sum_{k=1}^n F(d_k) - F(c_k) \) (because \((c_k, d_k]\)'s are disjoint), hence \( F(b) - F(a) \geq \sum_{k=1}^\infty F(d_k) - F(c_k) \). We need to prove the opposite inequality.

Suppose first that both \( a \) and \( b \) are finite. Let \( \epsilon > 0 \). For each \( k \), there is \( e_k > d_k \) such that

\[
F(d_k) \leq F(e_k) \leq F(d_k) + \frac{\epsilon}{2^k}.
\]

Also, there is \( f > a \) such that \( F(a) \geq F(f) - \epsilon \). Now, the interval \([f, b] \) is compact and \([f, b] \subseteq \bigcup_{k=1}^\infty (c_k, e_k]\). So there are finitely many \((c_k, e_k]\)'s (suppose they are the first \( n \)) such that \([f, b] \subseteq \bigcup_{k=1}^n (c_k, e_k]\). Now,

\[
F(b) - F(a) \leq F(b) - F(f) + \epsilon \leq \epsilon + \sum_{k=1}^n F(e_k) - F(c_k) \leq 2\epsilon + \sum_{k=1}^n F(d_k) - F(c_k).
\]
It follows that $F(b) - F(a) \leq 2\epsilon + \sum_{k=1}^{\infty} F(d_k) - F(c_k)$. Since this is true for all $\epsilon > 0$, it is true for $\epsilon = 0$.

If $-\infty = a < b < \infty$, let $g > -\infty$ be such that $F(g) < \epsilon$. The above argument shows that

$$F(b) - F(g) \leq \sum_{k=1}^{\infty} F(d_k \vee g) - F(c_k \vee g) \leq \sum_{k=1}^{\infty} F(d_k) - F(c_k).$$

Since $\lim_{g \to -\infty} F(g) = 0$, it follows that $F(b) \leq \sum_{k=1}^{\infty} F(d_k) - F(c_k)$. Similar arguments work when $a < b = \infty$ and $-\infty = a < b = \infty$.

In Lemma 21 you can replace $F$ by an arbitrary nondecreasing right-continuous function with only a bit more effort. (See the supplement following at the end of this lecture.)

The function $\mu$ defined in terms of a nondecreasing right-continuous function is a measure on the field $\mathcal{U}$. There is an extension theorem that gives conditions under which a measure on a field can be extended to a measure on the generated $\sigma$-field. Furthermore, the extension is unique.

**Example 22 (Lebesgue measure)** Start with the function $F(x) = x$, form the measure $\mu$ on the field $\mathcal{U}$ and extend it to the Borel $\sigma$-field. The result is called Lebesgue measure, and it extends the concept of “length” from intervals to more general sets.

**Example 23** Every distribution function for a random variable has a corresponding probability measure on the real line.

**Theorem 24 (Caratheodory Extension)** Let $\mu$ be a $\sigma$-finite measure on the field $\mathcal{C}$ of subsets of $\Omega$. Then $\mu$ has a unique extension to a measure on $\sigma(A)$.

**Exercise 14** In this exercise, we prove Theorem 24. Note that the uniqueness of the extension is a direct consequence of Theorem 20. We only need to prove the existence.

First, for each $B \in 2^{\Omega}$, define

$$\mu^*(B) = \inf \sum_{i=1}^{\infty} \mu(A_i),$$

where the inf is taken over all $\{A_i\}_{i=1}^{\infty}$ such that $B \subseteq \bigcup_{i=1}^{\infty} A_i$ and $A_i \in \mathcal{C}$ for all $i$. Since $\mathcal{C}$ is a field, we can assume that the $A_i$’s are mutually disjoint without changing the value of $\mu^*(B)$. Let

$$A = \{B \in 2^{\Omega} : \mu^*(C) = \mu^*(C \cap B) + \mu^*(C \cap B^C), \text{ for all } C \in 2^{\Omega}\}.$$
1. Show that \( \mu^* \) extends \( \mu \), i.e. that \( \mu^*(A) = \mu(A) \) for each \( A \in \mathcal{C} \).
2. Show that \( \mu^* \) is monotone and subadditive.
3. Show that \( \mathcal{C} \subseteq \mathcal{A} \).
4. Show that \( \mathcal{A} \) is a field.
5. Show that \( \mu^* \) is finitely additive on \( \mathcal{A} \).
6. Show that \( \mathcal{A} \) is a \( \sigma \)-field.
7. Show that \( \mu^* \) is countably additive on \( \mathcal{A} \).
Supplement: Measures from Increasing Functions

Lemma 21 deals only with functions $F$ that are cdf’s. Suppose that $F$ is an unbounded nondecreasing function that is continuous from the right. If $-\infty < a < b < \infty$, then the proof of Lemma 21 still applies. Suppose that $(-\infty, b] = \bigcup_{k=1}^{\infty} (c_k, d_k]$ with $b < \infty$ and all $(c_k, d_k]$ disjoint. Suppose that $\lim_{x \to -\infty} F(x) = -\infty$. We want to show that $\sum_{k=1}^{\infty} F(d_k) - F(c_k) = \infty$. If one $c_k = -\infty$, the proof is immediate, so assume that all $c_k > -\infty$. Then there must be a subsequence $\{k_j\}_{j=1}^{\infty}$ such that $\lim_{j \to \infty} c_{k_j} = -\infty$. For each $j$, let $\{(c'_{j,n}, d'_{j,n}]\}_{n=1}^{\infty}$ be the subsequence of intervals that cover $(c_{k_j}, b]$. For each $j$, the proof of Lemma 21 applies to show that

$$F(b) - F(c_{k_j}) = \sum_{n=1}^{\infty} F(d'_{j,n}) - F(c'_{j,n}).$$

(16)

As $j \to \infty$, the left side of Equation 16 goes to $\infty$ while the right side eventually includes every interval in the original collection.

A similar proof works for an interval of the form $(a, \infty)$ when $\lim_{x \to \infty} F(x) = \infty$. A combination of the two works for $(-\infty, \infty)$.