1 Course Info

2.1 Probability spaces (cont)

2.2 Distributions

2.3 Random variables
Course syllabus is available from my homepage: https://faculty.math.illinois.edu/~rsong/561s22/561s222.html


You do need a copy of this book. Most of the homework assignment will be from this book.

Office Hours: MWF: noon-12:50 pm in 227 CAB. I will also be on Zoom during this time.
Course syllabus is available from my homepage: https://faculty.math.illinois.edu/~rsong/561s22/561s222.html


You do need a copy of this book. Most of the homework assignment will be from this book.

Office Hours: MWF: noon-12:50 pm in 227 CAB. I will also be on Zoom during this time.
Course syllabus is available from my homepage: https://faculty.math.illinois.edu/~rsong/561s22/561s222.html


You do need a copy of this book. Most of the homework assignment will be from this book.

Office Hours: MWF: noon-12:50 pm in 227 CAB. I will also be on Zoom during this time.
Course syllabus is available from my homepage: https://faculty.math.illinois.edu/~rsong/561s22/561s222.html


You do need a copy of this book. Most of the homework assignment will be from this book.

Office Hours: MWF: noon-12:50 pm in 227 CAB. I will also be on Zoom during this time.
Outline

1. Course Info
2. 1.1 Probability spaces (cont)
3. 1.2 Distributions
4. 1.3 Random variables
Measures on $\mathbb{R}$

Measures on $(\mathbb{R}, \mathcal{R})$ are defined via Stieltjes functions, i.e., real-valued functions with the following properties:

(i) $F$ is non-decreasing;
(ii) $F$ is right-continuous.

Theorem 1.1.4

Associated with each Stieltjes function $F$, there is a unique measure $\mu$ on $(\mathbb{R}, \mathcal{R})$ with

$$\mu((a, b]) = F(b) - F(a), \quad \text{for all } a \leq b.$$

When $F(x) = x$ for all $x \in \mathbb{R}$, the resulting measure is called the Lebesgue measure on $\mathbb{R}$. 
Measures on $\mathbb{R}$

Measures on $(\mathbb{R}, \mathcal{R})$ are defined via Stieltjes functions, i.e., real-valued functions with the following properties:

(i) $F$ is non-decreasing;
(ii) $F$ is right-continuous.

Theorem 1.1.4

Associated with each Stieltjes function $F$, there is a unique measure $\mu$ on $(\mathbb{R}, \mathcal{R})$ with

$$\mu((a, b]) = F(b) - F(a), \quad \text{for all } a \leq b.$$
Measures on $\mathbb{R}$

Measures on $(\mathbb{R}, \mathcal{R})$ are defined via Stieltjes functions, i.e., real-valued functions with the following properties:

(i) $F$ is non-decreasing;
(ii) $F$ is right-continuous.

Theorem 1.1.4

Associated with each Stieltjes function $F$, there is a unique measure $\mu$ on $(\mathbb{R}, \mathcal{R})$ with

$$
\mu((a, b]) = F(b) - F(a), \quad \text{for all } a \leq b.
$$

When $F(x) = x$ for all $x \in \mathbb{R}$, the resulting measure is called the Lebesgue measure on $\mathbb{R}$. 
Last time, we talked about the following theorem:

**Theorem 1.1.9**

Let $S$ be a semi-algebra and let $\mu$ be defined on $S$ with $\mu(\emptyset) = 0$. Suppose (i) if $S \in S$ is a finite disjoint union of sets $A_i \in S$, then $\mu(S) = \sum_i \mu(A_i)$; (ii) If $S_i, S \in S$ with $S = \bigcup_{i=1}^{\infty} S_i$, then $\mu(S) \leq \sum_{i=1}^{\infty} \mu(S_i)$. Then $\mu$ has a unique extension $\overline{\mu}$ which is a measure on $\overline{S}$. If $\overline{\mu}$ is $\sigma$-finite, then there is a unique extension $\nu$ that is a measure on $\sigma(S)$.

We will use Theorem 1.1.9 to give a proof of Theorem 1.1.4. Before doing this, we first present a lemma.
Last time, we talked about the following theorem:

**Theorem 1.1.9**

Let $S$ be a semi-algebra and let $\mu$ be defined on $S$ with $\mu(\emptyset) = 0$. Suppose (i) if $S \in S$ is a finite disjoint union of sets $A_i \in S$, then $\mu(S) = \sum_i \mu(A_i)$; (ii) If $S_i, S \in S$ with $S = \bigcup_{i=1}^{\infty} S_i$, then $\mu(S) \leq \sum_{i=1}^{\infty} \mu(S_i)$. Then $\mu$ has a unique extension $\overline{\mu}$ which is a measure on $\overline{S}$. If $\overline{\mu}$ is $\sigma$-finite, then there is a unique extension $\nu$ that is a measure on $\sigma(S)$.

We will use Theorem 1.1.9 to give a proof of Theorem 1.1.4. Before doing this, we first present a lemma.
Last time, we talked about the following theorem:

**Theorem 1.1.9**

Let $S$ be a semi-algebra and let $\mu$ be defined on $S$ with $\mu(\emptyset) = 0$. Suppose (i) if $S \in S$ is a finite disjoint union of sets $A_i \in S$, then $\mu(S) = \sum_i \mu(A_i)$; (ii) If $S_i, S \in S$ with $S = +\bigcup_{i=1}^{\infty} S_i$, then $\mu(S) \leq \sum_{i=1}^{\infty} \mu(S_i)$. Then $\mu$ has a unique extension $\overline{\mu}$ which is a measure on $\overline{S}$. If $\overline{\mu}$ is $\sigma$-finite, then there is a unique extension $\nu$ that is a measure on $\sigma(S)$.

We will use Theorem 1.1.9 to give a proof of Theorem 1.1.4. Before doing this, we first present a lemma.
Lemma 1.1.10

Suppose that condition (i) in Theorem 1.1.9 holds, that is, if $S \in \mathcal{S}$ is a finite disjoint union of sets $A_i \in \mathcal{S}$, then $\mu(S) = \sum_i \mu(A_i)$.

(a) If $A, B_i \in \overline{\mathcal{S}}$ with $A = \bigcup_{i=1}^n B_i$, then $\bar{\mu}(A) = \sum_{i=1}^n \bar{\mu}(B_i)$.

(b) If $A, B_i \in \overline{\mathcal{S}}$ with $A \subset \bigcup_{i=1}^n B_i$, then $\bar{\mu}(A) \leq \sum_{i=1}^n \bar{\mu}(B_i)$.

Proof

(a) If $B_i = +_j S_{ij}$, $i = 1, \ldots, n$, $S_{ij} \in \mathcal{S}$, then $A = +_j S_{ij}$. Thus by definition

$$\bar{\mu}(A) = \sum_{i,j} \mu(S_{ij}) = \sum_i \bar{\mu}(B_i).$$
Lemma 1.1.10

Suppose that condition (i) in Theorem 1.1.9 holds, that is, if \( S \in \mathcal{S} \) is a finite disjoint union of sets \( A_i \in \mathcal{S} \), then \( \mu(S) = \sum_i \mu(A_i) \).

(a) If \( A, B_i \in \mathcal{S} \) with \( A = \bigcup_{i=1}^n B_i \), then \( \mu(A) = \sum_{i=1}^n \mu(B_i) \).

(b) If \( A, B_i \in \mathcal{S} \) with \( A \subset \bigcup_{i=1}^n B_i \), then \( \mu(A) \leq \sum_{i=1}^n \mu(B_i) \).

Proof

(a) If \( B_i = \bigcup_{j=1}^n S_{ij} \), \( i = 1, \ldots, n, S_{ij} \in \mathcal{S} \), then \( A = \bigcup_{i,j} S_{ij} \). Thus by definition

\[
\mu(A) = \sum_{i,j} \mu(S_{ij}) = \sum_i \mu(B_i).
\]
Proof (cont)

(b) We start with the case \( n = 1 \), \( B_1 = B \), \( B = A + (B \cap A^c) \) and \( B \cap A^c \in \mathcal{S} \). In this case,

\[
\bar{\mu}(A) \leq \bar{\mu}(A) + \bar{\mu}(B \cap A^c) = \bar{\mu}(B).
\]

Now suppose \( n > 1 \). Let \( F_k = B_1^c \cap \cdots \cap B_{k-1}^c \cap B_k \), \( k = 1, \ldots, n \). Then

\[
\bigcup_i B_i = F_1 + \cdots + F_n
\]

\[
A = A \cap (\bigcup_i B_i) = (A \cap F_1) + \cdots + (A \cap F_n).
\]

Using (a), and (b) with \( n = 1 \), we get

\[
\bar{\mu}(A) = \sum_{i=1}^{n} \bar{\mu}(A \cap F_k) \leq \sum_{i=1}^{n} \bar{\mu}(F_k) = \sum_{i} \bar{\mu}(B_i).
\]
Proof of Theorem 1.1.4

Let $S$ be the semi-algebra

$$\{(a, b] : -\infty \leq a \leq b \leq \infty\}.$$

We define $\mu$ on $S$ by

$$\mu((a, b]) = F(b) - F(a), \quad -\infty \leq a \leq b \leq \infty.$$ 

If $(a, b] = \bigcup_{i=1}^{n} (a_i, b_i]$, then after relabeling we must have $a_1 = a, b_n = b$ and $a_i = b_{i-1}$ for $i = 2, \ldots n$. So Condition (i) in Theorem 1.1.9 holds. To check (ii), suppose first that $-\infty < a < b < \infty$ and $(a, b] \subset \bigcup_i (a_i, b_i]$ where (without loss of generality) $-\infty < i < b_i < \infty$ for all $i$. For any $\epsilon > 0$, pick $\delta > 0$ so that $F(a + \delta) < F(a) + \epsilon$ and pick $\eta_i > 0$ so that

$$F(b_i + \eta_i) < F(b_i) + 2^{-i}\epsilon.$$ 

The open intervals $\{(a_i, b_i + \eta_i]\}$ cover $[a + \delta, b]$, so there is a finite cover $\{(\alpha_j, \beta_j), 1 \leq j \leq J\}$. Since $(a + \delta, b] \subset \bigcup_{j=1}^{J} (\alpha_j, \beta_j]$,
Proof of Theorem 1.1.4 (cont)

so Lemma 1.1.10.(b) implies

\[ F(b) - F(a + \delta) \leq \sum_{j=1}^{J} (F(\beta_j) - F(\alpha_j)) \leq \sum_{i=1}^{\infty} (F(b_i + \eta_i) - F(a_i)). \]

By the choice of \(\delta\) and \(\eta_i\), we have

\[ F(b) - F(a) \leq 2\epsilon + \sum_{i=1}^{\infty} (F(b_i) - F(a_i)). \]

Since \(\epsilon\) is arbitrary, we have proved the result in the case \(-\infty < a < b < \infty\). To remove this restriction, note that if \((a, b] \subset \bigcup_i (a_i, b_i]\) and \((A, B] \subset (a, b]\) with \(-\infty < A < B < \infty\), then

\[ F(B) - F(A) \leq \sum_{i=1}^{\infty} (F(b_i) - F(a_i)). \]

The above holds for any finite \((A, B] \subset (a, b]\), the conclusion follows.
Measure on $\mathbb{R}^d$

Suppose $F : \mathbb{R}^d \mapsto \mathbb{R}$ is such that

(i) $x \leq y$ (means: $x_i \leq y_i$ for all $i = 1, \ldots, d$) $\Rightarrow F(x) \leq F(y)$;
(ii) $F$ is right continuous, i.e., $\lim_{y \downarrow x} F(y) = F(x)$ for all $x \in \mathbb{R}^d$;
(iii) For any

$$A = (a_1, b_1] \times \cdots \times (a_d, b_d], \quad -\infty < a_i < b_i < \infty,$$

$\Delta_A F \geq 0$, where

$$\Delta_A F = \sum_{v \in V} \text{sgn}(v) F(v),$$

$V = \{a_1, b_1\} \times \cdots \times \{a_d, b_d\}$ and $\text{sgn}(v) = (-1)^\# \text{ of } a' \text{'s in } v$.

Theorem 1.1.11

Suppose $F : \mathbb{R}^d \mapsto \mathbb{R}$ satisfies (i)–(iii) above. Then there is a unique measure $\mu$ on $(\mathbb{R}^d, \mathcal{R}^d)$ so that $\mu(A) = \Delta_A F$ for all finite rectangles.
Measure on $\mathbb{R}^d$

Suppose $F : \mathbb{R}^d \mapsto \mathbb{R}$ is such that
(i) $x \leq y$ (means: $x_i \leq y_i$ for all $i = 1, \ldots, d$) $\Rightarrow F(x) \leq F(y)$;
(ii) $F$ is right continuous, i.e., $\lim_{y \downarrow x} F(y) = F(x)$ for all $x \in \mathbb{R}^d$;
(iii) For any

$$A = (a_1, b_1] \times \cdots \times (a_d, b_d], \quad -\infty < a_i < b_i < \infty,$$

$\Delta_A F \geq 0$, where

$$\Delta_A F = \sum_{v \in V} \text{sgn}(v) F(v),$$

$V = \{a_1, b_1\} \times \cdots \times \{a_d, b_d\}$ and $\text{sgn}(v) = (-1)^\#$ of $a$'s in $v$.

Theorem 1.1.11

Suppose $F : \mathbb{R}^d \mapsto \mathbb{R}$ satisfies (i)–(iii) above. Then there is a unique
measure $\mu$ on $(\mathbb{R}^d, \mathcal{R}^d)$ so that $\mu(A) = \Delta_A F$ for all finite rectangles.
If in the theorem above, $F$ is a $[0, 1]$-valued function with $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$, then the resulting measure $\mu$ is a probability measure.

**Example**

Suppose $F(x) = \prod_{i=1}^{d} F_i(x_i)$, where each $F_i$ is a Stieltjes function on $\mathbb{R}$. Then

$$\Delta_A F = \prod_{i=1}^{d} (F_i(b_i) - F_i(a_i)).$$

If $F_i(x_i) = x_i$ for all $i = 1, \ldots, d$ and all $x_i \in \mathbb{R}$, the resulting measure is called the Lebesgue measure on $\mathbb{R}^d$. 
If in the theorem above, $F$ is a $[0, 1]$-valued function with $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$, then the resulting measure $\mu$ is a probability measure.

**Example**

Suppose $F(x) = \prod_{i=1}^{d} F_i(x_i)$, where each $F_i$ is a Stieltjes function on $\mathbb{R}$. Then

$$\Delta_{\mathcal{A}} F = \prod_{i=1}^{d} (F_i(b_i) - F_i(a_i)).$$

If $F_i(x_i) = x_i$ for all $i = 1, \ldots, d$ and all $x_i \in \mathbb{R}$, the resulting measure is called the Lebesgue measure on $\mathbb{R}^d$. 
If in the theorem above, $F$ is a $[0, 1]$-valued function with $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to +\infty} F(x) = 1$, then the resulting measure $\mu$ is a probability measure.

Example

Suppose $F(x) = \prod_{i=1}^{d} F_i(x_i)$, where each $F_i$ is a Stieltjes function on $\mathbb{R}$. Then

$$\Delta_A F = \prod_{i=1}^{d} (F_i(b_i) - F_i(a_i)).$$

If $F_i(x_i) = x_i$ for all $i = 1, \ldots, d$ and all $x_i \in \mathbb{R}$, the resulting measure is called the Lebesgue measure on $\mathbb{R}^d$. 
Outline

1. Course Info
2. 1.1 Probability spaces (cont)
3. 1.2 Distributions
4. 1.3 Random variables
Suppose \((\Omega, \mathcal{F}, P)\) is a probability space. A function \(X : \Omega \mapsto \mathbb{R}\) is called a random variable on \((\Omega, \mathcal{F}, P)\) if, for any \(B \in \mathcal{B}\), \(X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}\).

When we want to emphasize the \(\sigma\)-field, we will say that \(X\) is \(\mathcal{F}\)-measurable or \(X \in \mathcal{F}\).

If \((\Omega, \mathcal{F}, P)\) is a discrete probability space, then any function \(X : \Omega \mapsto \mathbb{R}\) is a random variable.
Suppose \((\Omega, \mathcal{F}, P)\) is a probability space. A function \(X : \Omega \mapsto \mathbb{R}\) is called a random variable on \((\Omega, \mathcal{F}, P)\) if, for any \(B \in \mathcal{R}\),

\[X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}.\]

When we want to emphasize the \(\sigma\)-field, we will say that \(X\) is \(\mathcal{F}\)-measurable or \(X \in \mathcal{F}\).

If \((\Omega, \mathcal{F}, P)\) is a discrete probability space, then any function \(X : \Omega \mapsto \mathbb{R}\) is a random variable.
Suppose \((\Omega, \mathcal{F}, P)\) is a probability space. A function \(X : \Omega \mapsto \mathbb{R}\) is called a random variable on \((\Omega, \mathcal{F}, P)\) if, for any \(B \in \mathcal{R}\),
\[X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}.
\]

When we want to emphasize the \(\sigma\)-field, we will say that \(X\) is \(\mathcal{F}\)-measurable or \(X \in \mathcal{F}\).

If \((\Omega, \mathcal{F}, P)\) is a discrete probability space, then any function \(X : \Omega \mapsto \mathbb{R}\) is a random variable.
Here is a cheap, but useful, way to get random variables. For any $A \in \mathcal{F}$,

$$X(\omega) = 1_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \text{otherwise} \end{cases}$$

is a random variable. The indicator of $A$.

If $X$ is a random variable, then

$$\mu(A) = P(X \in A), \quad A \in \mathcal{R}$$

defines a probability measure on $(\mathbb{R}, \mathcal{R})$. It is called the distribution of $X$.

The distribution of a random variable is usually described by giving its distribution function: $F(x) = P(X \leq x), \ x \in \mathbb{R}$. 
Here is a cheap, but useful, way to get random variables. For any $A \in \mathcal{F}$,
\[
    X(\omega) = 1_A(\omega) = \begin{cases} 
        1, & \omega \in A, \\
        0, & \text{otherwise}
    \end{cases}
\]
is a random variable. The indicator of $A$.

If $X$ is a random variable, then
\[
    \mu(A) = P(X \in A), \quad A \in \mathcal{R}
\]
defines a probability measure on $(\mathbb{R}, \mathcal{R})$. It is called the distribution of $X$.

The distribution of a random variable is usually described by giving its distribution function: $F(x) = P(X \leq x)$, $x \in \mathbb{R}$. 
Here is a cheap, but useful, way to get random variables. For any $A \in \mathcal{F}$,

$$X(\omega) = 1_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \text{otherwise} \end{cases}$$

is a random variable. The indicator of $A$.

If $X$ is a random variable, then

$$\mu(A) = P(X \in A), \quad A \in \mathcal{R}$$

defines a probability measure on $(\mathbb{R}, \mathcal{R})$. It is called the distribution of $X$.

The distribution of a random variable is usually described by giving its distribution function: $F(x) = P(X \leq x), \ x \in \mathbb{R}$. 
Theorem 1.2.1

The distribution function $F$ of any random variable $X$ satisfies the following properties:

(i) $F$ is non-decreasing;

(ii) $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$;

(iii) $F$ is right-continuous;

(iv) for any $x \in \mathbb{R}$, $P(X < x) = F(x -)$;

(v) for any $x \in \mathbb{R}$, $P(X = x) = F(x) - F(x -)$.

Theorem 1.2.2

If $F$ is a function on $\mathbb{R}$ satisfying (i)–(iii) of the theorem above, then it is the distribution function of some random variable.
Theorem 1.2.1

The distribution function $F$ of any random variable $X$ satisfies the following properties:

(i) $F$ is non-decreasing;
(ii) $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$;
(iii) $F$ is right-continuous;
(iv) for any $x \in \mathbb{R}$, $P(X < x) = F(x -)$;
(v) for any $x \in \mathbb{R}$, $P(X = x) = F(x) - F(x -)$.

Theorem 1.2.2

If $F$ is a function on $\mathbb{R}$ satisfying (i)–(iii) of the theorem above, then it is the distribution function of some random variable.
Proof of Theorem 1.2.2

Let $\Omega = (0, 1)$, $\mathcal{F}$ the Borel subsets of $(0, 1)$ and $P$ the Lebesgue measure. For $\omega \in (0, 1)$, define

$$X(\omega) = \sup\{y : F(y) < \omega\}.$$ 

To prove the theorem, it suffices to show

$$\{\omega : X(\omega) \leq x\} = \{\omega : \omega \leq F(x)\}.$$ 

If $\omega \leq F(x)$, then $X(\omega) \leq x$ since $x \notin \{y : F(y) < \omega\}$.

If $\omega > F(x)$, then, since $F$ is right-continuous, $\exists \epsilon > 0$ such that $F(x + \epsilon) < \omega$, which implies $X(\omega) \geq x + \epsilon > x$. 

If two random variables $X$ and $Y$ induce the same probability measure on $(\mathbb{R}, \mathcal{R})$, then we say that $X$ and $Y$ are equal in distribution (or identically distributed) and write $X \overset{d}{=} Y$.

Two random variables $X$ and $Y$ have the same distribution if and only if they have the same distribution function.

Review concepts from undergraduate probability: continuous random variables; absolutely continuous random variables, density function; binomial random variables; Poisson random variables; geometric random variables; negative binomial random variables; uniform random variables; normal random variables; exponential random variables; Gamma random variables.
If two random variables $X$ and $Y$ induce the same probability measure on $(\mathbb{R}, \mathcal{R})$, then we say that $X$ and $Y$ are equal in distribution (or identically distributed) and write $X \overset{d}{=} Y$.

Two random variables $X$ and $Y$ have the same distribution if and only if they have the same distribution function.

Review concepts from undergraduate probability: continuous random variables; absolutely continuous random variables, density function; binomial random variables; Poisson random variables; geometric random variables; negative binomial random variables; uniform random variables; normal random variables; exponential random variables; Gamma random variables.
If two random variables $X$ and $Y$ induce the same probability measure on $(\mathbb{R}, \mathcal{R})$, then we say that $X$ and $Y$ are equal in distribution (or identically distributed) and write $X \overset{d}{=} Y$.

Two random variables $X$ and $Y$ have the same distribution if and only if they have the same distribution function.

Review concepts from undergraduate probability: continuous random variables; absolutely continuous random variables, density function; binomial random variables; Poisson random variables; geometric random variables; negative binomial random variables; uniform random variables; normal random variables; exponential random variables; Gamma random variables.
Outline

1. Course Info
2. 1.1 Probability spaces (cont)
3. 1.2 Distributions
4. 1.3 Random variables
In this section, we review some properties of random variables. Since most of the results are true for measurable maps from \((\Omega, \mathcal{F})\) to an arbitrary measurable space \((S, \mathcal{S})\), we will develop our results in this generality.

A function \(X : \Omega \mapsto S\) is said to be a *measurable map* from \((\Omega, \mathcal{F})\) to \((S, \mathcal{S})\) if

\[
X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}, \quad \forall B \in \mathcal{S}.
\]

If \((S, \mathcal{S}) = (\mathbb{R}, \mathcal{R})\), then \(X\) reduces to a random variable.
If \((S, \mathcal{S}) = (\mathbb{R}^d, \mathcal{R}^d)\), then \(X\) reduces to a random vector.
In this section, we review some properties of random variables. Since most of the results are true for measurable maps from $(\Omega, \mathcal{F})$ to an arbitrary measurable space $(S, \mathcal{S})$, we will develop our results in this generality.

A function $X : \Omega \mapsto S$ is said to be a measurable map from $(\Omega, \mathcal{F})$ to $(S, \mathcal{S})$ if

$$X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}, \quad \forall B \in \mathcal{S}.\$$

If $(S, \mathcal{S}) = (\mathbb{R}, \mathcal{B})$, then $X$ reduces to a random variable.

If $(S, \mathcal{S}) = (\mathbb{R}^d, \mathcal{B}^d)$, then $X$ reduces to a random vector.
In this section, we review some properties of random variables. Since most of the results are true for measurable maps from \((\Omega, \mathcal{F})\) to an arbitrary measurable space \((S, \mathcal{S})\), we will develop our results in this generality.

A function \(X : \Omega \mapsto S\) is said to be a measurable map from \((\Omega, \mathcal{F})\) to \((S, \mathcal{S})\) if
\[
X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}, \quad \forall B \in \mathcal{S}.
\]

If \((S, \mathcal{S}) = (\mathbb{R}, \mathcal{R})\), then \(X\) reduces to a random variable.
If \((S, \mathcal{S}) = (\mathbb{R}^d, \mathcal{R}^d)\), then \(X\) reduces to a random vector.
**Theorem 1.3.1**

If \( \{ \omega : X(\omega) \in A \} \in \mathcal{F} \) for all \( A \in \mathcal{A} \) and \( \mathcal{A} \) generates \( S \), then \( X \) is measurable.

**Proof**

It is easy to check that

\[
\{ X \in \bigcup_i B_i \} = \bigcup_i \{ X \in B_i \} \quad \{ X \in B^c \} = \{ X \in B \}^c.
\]

Thus \( \mathcal{B} = \{ B : \{ X \in B \} \in \mathcal{F} \} \) is a \( \sigma \)-field. Since \( \mathcal{B} \supset \mathcal{A} \) and \( \mathcal{A} \) generates \( S \), \( \mathcal{B} \supset S \).

\[
\{ \{ X \in B \} : X \in S \}
\]

is a \( \sigma \)-field. It is the smallest \( \sigma \)-field on \( \Omega \) that makes \( X \) a measurable map. It is called the \( \sigma \)-field generated by \( X \) and denoted by \( \sigma(X) \).
Theorem 1.3.1

If \( \{ \omega : X(\omega) \in A \} \in \mathcal{F} \) for all \( A \in \mathcal{A} \) and \( \mathcal{A} \) generates \( S \), then \( X \) is measurable.

Proof

It is easy to check that

\[
\{ X \in \bigcup_i B_i \} = \bigcup_i \{ X \in B_i \} \quad \{ X \in B^c \} = \{ X \in B \}^c.
\]

Thus \( \mathcal{B} = \{ B : \{ X \in B \} \in \mathcal{F} \} \) is a \( \sigma \)-field. Since \( \mathcal{B} \supset \mathcal{A} \) and \( \mathcal{A} \) generates \( S \), \( \mathcal{B} \supset S \).

\[
\{ \{ X \in B \} : X \in S \}
\]

is a \( \sigma \)-field. It is the smallest \( \sigma \)-field on \( \Omega \) that makes \( X \) a measurable map. It is called the \( \sigma \)-field generated by \( X \) and denoted by \( \sigma(X) \).
Theorem 1.3.1

If \( \{ \omega : X(\omega) \in A \} \in \mathcal{F} \) for all \( A \in \mathcal{A} \) and \( \mathcal{A} \) generates \( S \), then \( X \) is measurable.

Proof

It is easy to check that
\[
\{ X \in \bigcup_i B_i \} = \bigcup_i \{ X \in B_i \} \quad \{ X \in B^c \} = \{ X \in B \}^c.
\]
Thus \( \mathcal{B} = \{ B : \{ X \in B \} \in \mathcal{F} \} \) is a \( \sigma \)-field. Since \( \mathcal{B} \supset \mathcal{A} \) and \( \mathcal{A} \) generates \( S \), \( \mathcal{B} \supset S \).

\[
\{ \{ X \in B \} : X \in S \}
\]
is a \( \sigma \)-field. It is the smallest \( \sigma \)-field on \( \Omega \) that makes \( X \) a measurable map. It is called the \( \sigma \)-field generated by \( X \) and denoted by \( \sigma(X) \).