A categorical view of conditional expectation

Prakash Panangaden

School of Computer Science
McGill University

Applied Category Theory - 8th April 2020
1 Introduction
Outline

1 Introduction

2 Some functional analysis
Outline

1. Introduction
2. Some functional analysis
3. Cones
4. Cones of measures and functions
Outline

1. Introduction
2. Some functional analysis
3. Cones
4. Cones of measures and functions
5. Conditional expectation
Outline

1. Introduction
2. Some functional analysis
3. Cones
4. Cones of measures and functions
5. Conditional expectation
6. The Arena: Two Categories
Joint work with

Chaput, Danos and Plotkin


The idea of functorializing conditional expectation is due to Vincent.
Conditional probability

- Probability triple: $(\Omega, \Sigma, p)$. Random process produces outcome distributed according to $p$. 
Conditional probability

- Probability triple: \((\Omega, \Sigma, p)\). Random process produces outcome distributed according to \(p\).
- If \(\omega \in \Omega\) is the outcome and \(A \in \Sigma\), then the probability that \(\omega \in A\) is \(p(A)\).
Conditional probability

- Probability triple: \((\Omega, \Sigma, p)\). Random process produces outcome distributed according to \(p\).
- If \(\omega \in \Omega\) is the outcome and \(A \in \Sigma\), then the probability that \(\omega \in A\) is \(p(A)\).
- If we *know* that \(\omega \in B\) then we reassess the probability of \(\omega \in A\): \(P(A|B) = p(A \cap B)/p(B)\).
Conditional probability

- Probability triple: $(\Omega, \Sigma, p)$. Random process produces outcome distributed according to $p$.
- If $\omega \in \Omega$ is the outcome and $A \in \Sigma$, then the probability that $\omega \in A$ is $p(A)$.
- If we know that $\omega \in B$ then we reassess the probability of $\omega \in A$: $P(A|B) = p(A \cap B)/p(B)$.
- If we know whether $\omega$ is in $B$ or $B^c$ we can compute $P(A|B)$ and $P(A|B^c)$. Define $f(\omega) = P(A|B)$ if $\omega \in B$ and $f(\omega) = P(A|B^c)$ if $\omega \in B^c$. 
Conditional probability

- Probability triple: $(\Omega, \Sigma, p)$. Random process produces outcome distributed according to $p$.
- If $\omega \in \Omega$ is the outcome and $A \in \Sigma$, then the probability that $\omega \in A$ is $p(A)$.
- If we know that $\omega \in B$ then we reassess the probability of $\omega \in A$: $P(A|B) = p(A \cap B)/p(B)$.
- If we know whether $\omega$ is in $B$ or $B^c$ we can compute $P(A|B)$ and $P(A|B^c)$. Define $f(\omega) = P(A|B)$ if $\omega \in B$ and $f(\omega) = P(A|B^c)$ if $\omega \in B^c$.
- If we are given a countable partition $\{B_i|B_i \in \Sigma\}$ of $\Omega$ we can define a function $f : \Omega \to [0, 1]$ such that $f(\omega) = P(A|B_i)$ if $\omega \in B_i$. 
Conditional probability

- Probability triple: \((\Omega, \Sigma, p)\). Random process produces outcome distributed according to \(p\).
- If \(\omega \in \Omega\) is the outcome and \(A \in \Sigma\), then the probability that \(\omega \in A\) is \(p(A)\).
- If we know that \(\omega \in B\) then we reassess the probability of \(\omega \in A\): \(P(A|B) = \frac{p(A \cap B)}{p(B)}\).
- If we know whether \(\omega\) is in \(B\) or \(B^c\) we can compute \(P(A|B)\) and \(P(A|B^c)\). Define \(f(\omega) = P(A|B)\) if \(\omega \in B\) and \(f(\omega) = P(A|B^c)\) if \(\omega \in B^c\).
- If we are given a countable partition \(\{B_i|B_i \in \Sigma\}\) of \(\Omega\) we can define a function \(f: \Omega \rightarrow [0, 1]\) such that \(f(\omega) = P(A|B_i)\) if \(\omega \in B_i\).
- If we are given \(\Lambda \subset \Sigma\) and for every \(B \in \Lambda\) we know whether \(\omega \in B\) we can define the random variable \(P[A||\Lambda]\) which is \(\Lambda\)-measurable and

\[
\forall B \in \Lambda \int_B P[A||\Lambda] \, dp = P(A \cap B).
\]
Conditional expectation

- Sample space: $(\Omega, \Sigma, p)$, random variable $f : \Omega \rightarrow \mathbb{R}$. 

Why isn't $E[f|\Lambda]$ just $f$?

Because it is only $\Lambda$-measurable; so much “smoother.”
Conditional expectation

- Sample space: \((\Omega, \Sigma, p)\), random variable \(f : \Omega \rightarrow \mathbb{R}\).
- Expectation value of \(f\): \(\int f \, dp\).

We want to revise our expectation based on new information. Given a sub-\(\sigma\)-algebra \(\Lambda \subset \Sigma\) we define \(E[f|\Lambda]\) as a \(\Lambda\)-measurable function such that \(\forall B \in \Lambda, \int_B E[f|\Lambda] \, dp = \int_B f \, dp\).

Why isn't \(E[f|\Lambda]\) just \(f\)?

Because it is only \(\Lambda\)-measurable; so much "smoother."
Conditional expectation

- Sample space: \((\Omega, \Sigma, p)\), random variable \(f : \Omega \rightarrow \mathbb{R}\).
- Expectation value of \(f\): \(\int f \, dp\).
- We want to revise our expectation based on new information.
Conditional expectation

- Sample space: \((\Omega, \Sigma, p)\), random variable \(f : \Omega \to \mathbb{R}\).
- Expectation value of \(f\): \(\int f\,dp\).
- We want to revise our expectation based on new information.
- Given a sub-\(\sigma\)-algebra \(\Lambda \subset \Sigma\) we define \(E[f||\Lambda]\) as a \(\Lambda\)-measurable function such that

\[
\forall B \in \Lambda, \int_B E[f||\Lambda]\,dp = \int_B f\,dp.
\]
Conditional expectation

- Sample space: \((\Omega, \Sigma, p)\), random variable \(f : \Omega \rightarrow \mathbb{R}\).
- Expectation value of \(f\): \(\int f \, dp\).
- We want to revised our expectation based on new information.
- Given a sub-\(\sigma\)-algebra \(\Lambda \subset \Sigma\) we define \(E[f \mid \Lambda]\) as a \(\Lambda\)-measurable function such that

\[
\forall B \in \Lambda, \int_B E[f \mid \Lambda] \, dp = \int_B f \, dp.
\]

- Why isn’t \(E[f \mid \Lambda]\) just \(f\)?

Because it is only \(\Lambda\)-measurable; so much "smoother."
Conditional expectation

- Sample space: \((\Omega, \Sigma, p)\), random variable \(f : \Omega \to \mathbb{R}\).
- Expectation value of \(f\): \(\int f \, dp\).
- We want to revise our expectation based on new information.
- Given a sub-\(\sigma\)-algebra \(\Lambda \subset \Sigma\) we define \(E[f|\Lambda]\) as a \(\Lambda\)-measurable function such that

\[
\forall B \in \Lambda, \int_B E[f|\Lambda] \, dp = \int_B f \, dp.
\]

- Why isn't \(E[f|\Lambda]\) just \(f\)?
- Because it is only \(\Lambda\) measurable; so much “smoother.”
A norm on a vector space $V$ is a function $\| \cdot \| : V \rightarrow \mathbb{R}^{\geq 0}$ such that:

1. $\| v \| = 0$ iff $v = 0$
2. $\| r \cdot v \| = |r| \| v \|$
3. $\| x + y \| \leq \| x \| + \| y \|$

The norm induces a metric: $d(u, v) = \| u - v \|$ and, hence, a topology. This topology is called the norm topology.

If $V$ is complete in this metric it is called a Banach space.
A norm on a vector space $V$ is a function $\| \cdot \| : V \rightarrow \mathbb{R}^{\geq 0}$ such that:

1. $\|v\| = 0$ iff $v = 0$
Banach spaces

A norm on a vector space $V$ is a function $||\cdot|| : V \rightarrow \mathbb{R}^{\geq 0}$ such that:

1. $||v|| = 0$ iff $v = 0$
2. $||r \cdot v|| = |r| ||v||$ and
A norm on a vector space $V$ is a function $\| \cdot \| : V \rightarrow \mathbb{R}^{\geq 0}$ such that:

1. $\|v\| = 0$ iff $v = 0$
2. $\|r \cdot v\| = |r| \|v\|$ and
3. $\|x + y\| \leq \|x\| + \|y\|$. 

The norm induces a metric:

$$d(u, v) = \|u - v\|$$

and, hence, a topology. This topology is called the norm topology. If $V$ is complete in this metric it is called a Banach space.
A *norm* on a vector space $V$ is a function $\| \cdot \| : V \rightarrow \mathbb{R}^{\geq 0}$ such that:

1. $\|v\| = 0$ iff $v = 0$
2. $\|r \cdot v\| = |r| \|v\|$ and
3. $\|x + y\| \leq \|x\| + \|y\|$.

The norm induces a metric: $d(u, v) = \|u - v\|$ and, hence, a topology. This topology is called the *norm topology*. 
A norm on a vector space $V$ is a function $\| \cdot \| : V \to \mathbb{R}^{\geq 0}$ such that:

1. $\|v\| = 0$ iff $v = 0$
2. $\|r \cdot v\| = |r| \|v\|$ and
3. $\|x + y\| \leq \|x\| + \|y\|$.

The norm induces a metric: $d(u, v) = \|u - v\|$ and, hence, a topology. This topology is called the norm topology.

If $V$ is complete in this metric it is called a Banach space.
A linear map $T : U \rightarrow V$ is \textit{bounded} if there exists a positive real number $\alpha$ such that $\forall u \in U, \|Tu\| \leq \alpha \|u\|$.
A linear map $T : U \to V$ is **bounded** if there exists a positive real number $\alpha$ such that $\forall u \in U, \|Tu\| \leq \alpha \|u\|$. 

A linear map between normed spaces is **continuous** iff it is **bounded**.
A linear map $T: U \to V$ is **bounded** if there exists a positive real number $\alpha$ such that $\forall u \in U, \|Tu\| \leq \alpha \|u\|$. 

A linear map between normed spaces is **continuous** iff it is **bounded**.

Given a bounded linear map between normed spaces $T: U \to V$ we define $\|T\| = \sup \{\|Tu\| \mid u \in U, \|u\| \leq 1\}$. 

This is a norm on the space of bounded linear maps and is called the **operator norm**.
A linear map \( T : U \rightarrow V \) is \textit{bounded} if there exists a positive real number \( \alpha \) such that \( \forall u \in U, \| Tu \| \leq \alpha \| u \| \).

A linear map between normed spaces is \textit{continuous} iff it is \textit{bounded}.

Given a bounded linear map between normed spaces \( T : U \rightarrow V \) we define \( \| T \| = \sup \{ \| Tu \| \mid u \in U, \| u \| \leq 1 \} \).

This is a norm on the space of bounded linear maps and is called the \textit{operator norm}.
A linear map $T : U \rightarrow V$ is bounded if there exists a positive real number $\alpha$ such that $\forall u \in U, \|Tu\| \leq \alpha \|u\|$. 

A linear map between normed spaces is continuous iff it is bounded. 

Given a bounded linear map between normed spaces $T : U \rightarrow V$ we define $\|T\| = \sup \{\|Tu\| \mid u \in U, \|u\| \leq 1\}$. 

This is a norm on the space of bounded linear maps and is called the operator norm. 

With this norm the space of bounded linear maps between Banach spaces forms a Banach space.
Duality for Banach spaces

The space of bounded (= continuous) linear maps from $V$, a Banach space, to $\mathbb{R}$ is itself a Banach space, called the *dual space*, $V^*$. 
The space of bounded (= continuous) linear maps from $V$, a Banach space, to $\mathbb{R}$ is itself a Banach space, called the \textit{dual space}, $V^*$. 

For any two vector spaces $U, V$ we say that they are in \textit{algebraic duality} if there is a bilinear form $\langle \cdot, \cdot \rangle : U \times V \to \mathbb{R}$ such that spaces of functionals $\langle \cdot, V \rangle$ and $\langle U, \cdot \rangle$ separates points of $U$ and $V$. 

Panangaden (McGill University) 
Overview and Cones 
8 April 2020 8 / 45
The space of bounded (= continuous) linear maps from $V$, a Banach space, to $\mathbb{R}$ is itself a Banach space, called the dual space, $V^*$. 

For any two vector spaces $U, V$ we say that they are in algebraic duality if there is a bilinear form $\langle \cdot, \cdot \rangle : U \times V \to \mathbb{R}$ such that spaces of functionals $\langle \cdot, V \rangle$ and $\langle U, \cdot \rangle$ separates points of $U$ and $V$.

We say two Banach spaces are in duality if $\langle \cdot, V \rangle \subseteq U^*$ and $\langle U, \cdot \rangle \subseteq V^*$. 
The space of bounded (= continuous) linear maps from $V$, a Banach space, to $\mathbb{R}$ is itself a Banach space, called the dual space, $V^*$. For any two vector spaces $U, V$ we say that they are in algebraic duality if there is a bilinear form $\langle \cdot, \cdot \rangle : U \times V \to \mathbb{R}$ such that spaces of functionals $\langle \cdot, V \rangle$ and $\langle U, \cdot \rangle$ separates points of $U$ and $V$. We say two Banach spaces are in duality if $\langle \cdot, V \rangle \subseteq U^*$ and $\langle U, \cdot \rangle \subseteq V^*$. For $V$ a Banach space, the spaces $V$ and $V^*$ are in duality.
The space of bounded (= continuous) linear maps from $V$, a Banach space, to $\mathbb{R}$ is itself a Banach space, called the dual space, $V^*$.

For any two vector spaces $U, V$ we say that they are in algebraic duality if there is a bilinear form $\langle \cdot, \cdot \rangle : U \times V \rightarrow \mathbb{R}$ such that spaces of functionals $\langle \cdot, V \rangle$ and $\langle U, \cdot \rangle$ separates points of $U$ and $V$.

We say two Banach spaces are in duality if $\langle \cdot, V \rangle \subseteq U^*$ and $\langle U, \cdot \rangle \subseteq V^*$.

For $V$ a Banach space, the spaces $V$ and $V^*$ are in duality.

The bilinear form is $\langle v, \phi \rangle = \phi(v)$. 

The space of bounded (= continuous) linear maps from $V$, a Banach space, to $\mathbb{R}$ is itself a Banach space, called the dual space, $V^*$.

For any two vector spaces $U, V$ we say that they are in algebraic duality if there is a bilinear form $\langle \cdot, \cdot \rangle : U \times V \to \mathbb{R}$ such that spaces of functionals $\langle \cdot, V \rangle$ and $\langle U, \cdot \rangle$ separates points of $U$ and $V$.

We say two Banach spaces are in duality if $\langle \cdot, V \rangle \subseteq U^*$ and $\langle U, \cdot \rangle \subseteq V^*$.

For $V$ a Banach space, the spaces $V$ and $V^*$ are in duality.

The bilinear form is $\langle v, \phi \rangle = \phi(v)$.

There is a canonical injection $\iota : V \to V^{**}$; if this is an isometry we say that the Banach space $V$ is reflexive.
Duality for Banach spaces

- The space of bounded (= continuous) linear maps from \( V \), a Banach space, to \( \mathbb{R} \) is itself a Banach space, called the dual space, \( V^* \).

- For any two vector spaces \( U, V \) we say that they are in algebraic duality if there is a bilinear form \( \langle \cdot, \cdot \rangle : U \times V \to \mathbb{R} \) such that spaces of functionals \( \langle \cdot, V \rangle \) and \( \langle U, \cdot \rangle \) separates points of \( U \) and \( V \).

- We say two Banach spaces are in duality if \( \langle \cdot, V \rangle \subseteq U^* \) and \( \langle U, \cdot \rangle \subseteq V^* \).

- For \( V \) a Banach space, the spaces \( V \) and \( V^* \) are in duality.

- The bilinear form is \( \langle v, \phi \rangle = \phi(v) \).

- There is a canonical injection \( \iota : V \to V^{**} \); if this is an isometry we say that the Banach space \( V \) is reflexive.

- Infinite dimensional Banach spaces are not necessarily reflexive.
Duality for Banach spaces

- The space of bounded (= continuous) linear maps from $V$, a Banach space, to $\mathbb{R}$ is itself a Banach space, called the *dual space*, $V^*$.

- For any two vector spaces $U, V$ we say that they are in *algebraic duality* if there is a bilinear form $\langle \cdot, \cdot \rangle : U \times V \rightarrow \mathbb{R}$ such that spaces of functionals $\langle \cdot, V \rangle$ and $\langle U, \cdot \rangle$ separates points of $U$ and $V$.

- We say two Banach spaces are in *duality* if $\langle \cdot, V \rangle \subseteq U^*$ and $\langle U, \cdot \rangle \subseteq V^*$.

- For $V$ a Banach space, the spaces $V$ and $V^*$ are in duality.

- The bilinear form is $\langle v, \phi \rangle = \phi(v)$.

- There is a canonical injection $\iota : V \rightarrow V^{**}$; if this is an isometry we say that the Banach space $V$ is reflexive.

- Infinite dimensional Banach spaces are not necessarily reflexive.

- Finite dimensional Banach spaces are always reflexive.
If \((X, \Sigma, \mu)\) is a measure space we can define integration on \(X\): we write \(\int_X f \, dx\). We say that \(f\) is *integrable* if this is finite.
If \((X, \Sigma, \mu)\) is a measure space we can define integration on \(X\): we write \(\int_X f \, dx\). We say that \(f\) is \textit{integrable} if this is finite.

If two functions agree everywhere except on a set of \(\mu\)-measure 0 their integrals will be equal.
If $(X, \Sigma, \mu)$ is a measure space we can define integration on $X$: we write $\int_X f \, dx$. We say that $f$ is *integrable* if this is finite.

If two functions agree everywhere except on a set of $\mu$-measure 0 their integrals will be equal.

We define two functions to be equivalent if they are $\mu$-almost everywhere the same and we actually work with equivalence classes.

The integral defines a norm on these equivalence classes and gives the Banach space $L_1(X, \Sigma, \mu)$ or just $L_1(\mu)$.

The space $L_p(\mu)$ is the space obtained by defining the norm $\|f\|_p = \left(\int |f|^p \, d\mu\right)^{1/p}$, where $0 < p < \infty$.

The infinity norm of a measurable function $f$ is $\|f\|_\infty = \inf\{M > 0 | |f(x)| \leq M \text{ for } \mu\text{-almost all } x\}$. The collection of all equivalence classes of measurable functions $f$ with $\|f\|_\infty < \infty$ with the norm just defined is the space $L_\infty(\mu)$.

These are all Banach spaces.
\(L_p\) spaces

- If \((X, \Sigma, \mu)\) is a measure space we can define integration on \(X\): we write \(\int_X f \, dx\). We say that \(f\) is integrable if this is finite.
- If two functions agree everywhere except on a set of \(\mu\)-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are \(\mu\)-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space \(L_1(X, \Sigma, \mu)\) or just \(L_1(\mu)\).
If \((X, \Sigma, \mu)\) is a measure space we can define integration on \(X\): we write \(\int_X f \, dx\). We say that \(f\) is integrable if this is finite.

If two functions agree everywhere except on a set of \(\mu\)-measure 0 their integrals will be equal.

We define two functions to be equivalent if they are \(\mu\)-almost everywhere the same and we actually work with equivalence classes.

The integral defines a norm on these equivalence classes and gives the Banach space \(L_1(X, \Sigma, \mu)\) or just \(L_1(\mu)\).

The space \(L_p(\mu)\) is the space obtained by defining the norm

\[
\|f\|_p = \left( \int |f|^p \, d\mu \right)^{\frac{1}{p}}, \text{ where } 0 < p < \infty.
\]
If \((X, \Sigma, \mu)\) is a measure space we can define integration on \(X\): we write \(\int_X f \, dx\). We say that \(f\) is integrable if this is finite.

If two functions agree everywhere except on a set of \(\mu\)-measure 0 their integrals will be equal.

We define two functions to be equivalent if they are \(\mu\)-almost everywhere the same and we actually work with equivalence classes.

The integral defines a norm on these equivalence classes and gives the Banach space \(L_1(X, \Sigma, \mu)\) or just \(L_1(\mu)\).

The space \(L_p(\mu)\) is the space obtained by defining the norm

\[
\|f\|_p = \left( \int_X |f|^p \, d\mu \right)^{\frac{1}{p}}, \text{ where } 0 < p < \infty.
\]

The \textit{infinity norm} of a measurable function \(f\) is

\[
\|f\|_\infty = \inf \{ M > 0 \ | \ |f(x)| \leq M \text{ for } \mu - \text{almost all } x \}.
\]
\(L_p\) spaces

- If \((X, \Sigma, \mu)\) is a measure space we can define integration on \(X\): we write \(\int_X f \, dx\). We say that \(f\) is integrable if this is finite.
- If two functions agree everywhere except on a set of \(\mu\)-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are \(\mu\)-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space \(L_1(X, \Sigma, \mu)\) or just \(L_1(\mu)\).
- The space \(L_p(\mu)\) is the space obtained by defining the norm
  \[\|f\|_p = \left( \int |f|^p \, d\mu \right)^{1/p}, \text{ where } 0 < p < \infty.\]
- The infinity norm of a measurable function \(f\) is
  \[\|f\|_\infty = \inf \{M > 0 \mid |f(x)| \leq M \text{ for } \mu - \text{almost all } x\}.
- The collection of all equivalence classes of measurable functions \(f\) with \(\|f\|_\infty < \infty\) with the norm just defined is the space \(L_\infty(\mu)\).
If \((X, \Sigma, \mu)\) is a measure space we can define integration on \(X\): we write \(\int_X f \, dx\). We say that \(f\) is integrable if this is finite.

If two functions agree everywhere except on a set of \(\mu\)-measure 0 their integrals will be equal.

We define two functions to be equivalent if they are \(\mu\)-almost everywhere the same and we actually work with equivalence classes.

The integral defines a norm on these equivalence classes and gives the Banach space \(L_1(X, \Sigma, \mu)\) or just \(L_1(\mu)\).

The space \(L_p(\mu)\) is the space obtained by defining the norm

\[
\|f\|_p = \left( \int |f|^p \, d\mu \right)^{\frac{1}{p}}, \text{ where } 0 < p < \infty.
\]

The infinity norm of a measurable function \(f\) is

\[
\|f\|_\infty = \inf \{ M > 0 \mid |f(x)| \leq M \text{ for } \mu - \text{almost all } x \}.
\]

The collection of all equivalence classes of measurable functions \(f\) with \(\|f\|_\infty < \infty\) with the norm just defined is the space \(L_\infty(\mu)\).

These are all Banach spaces.
If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then $L_p$ and $L_q$ are duals of each other!
Duality for $L_p$ spaces

- If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then $L_p$ and $L_q$ are duals of each other!
- However, $L_1$ and $L_\infty$ are not duals.
Duality for \( L_p \) spaces

- If \( 1 < p, q < \infty \) and \( \frac{1}{p} + \frac{1}{q} = 1 \) then \( L_p \) and \( L_q \) are duals of each other!
- However, \( L_1 \) and \( L_\infty \) are not duals.
- The dual of \( L_1 \) is \( L_\infty \) but not the other way around!
Duality for $L_p$ spaces

- If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then $L_p$ and $L_q$ are duals of each other!
- However, $L_1$ and $L_\infty$ are not duals.
- The dual of $L_1$ is $L_\infty$ but not the other way around!
- We will switch to a cone view and the situation will be much improved.
What are cones?

- Want to combine linear structure with order structure.
What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order \( \leq \) we have a natural notion of \textit{positive} and \textit{negative} vectors: \( x \geq 0 \) is positive.
What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include $0$. 

We define a cone $C$ in a vector space $V$ to be a set with exactly these conditions. 

Unfortunately for us, many of the structures that we want to look at are cones but are not part of any obvious vector space: e.g. the measures on a space. We could artificially embed them in a vector space, for example, by introducing signed measures.
What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of positive and negative vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.
- Then for any positive $v \in P$ and positive real $r$, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$. 
What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order \( \leq \) we have a natural notion of positive and negative vectors: \( x \geq 0 \) is positive.
- What properties do the positive vectors have? Say \( P \subset V \) are the positive vectors, we include 0.
- Then for any positive \( v \in P \) and positive real \( r, rv \in P \). For \( u, v \in P \) we have \( u + v \in P \) and if \( v \in P \) and \( -v \in P \) then \( v = 0 \).
- We define a cone \( C \) in a vector space \( V \) to be a set with exactly these conditions.
What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order \( \leq \) we have a natural notion of positive and negative vectors: \( x \geq 0 \) is positive.
- What properties do the positive vectors have? Say \( P \subset V \) are the positive vectors, we include 0.
- Then for any positive \( v \in P \) and positive real \( r \), \( rv \in P \). For \( u, v \in P \) we have \( u + v \in P \) and if \( v \in P \) and \( -v \in P \) then \( v = 0 \).
- We define a cone \( C \) in a vector space \( V \) to be a set with exactly these conditions.
- Any cone defines a order by \( u \leq v \) if \( v - u \in C \).
What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of positive and negative vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include $0$.
- Then for any positive $v \in P$ and positive real $r$, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$.
- We define a cone $C$ in a vector space $V$ to be a set with exactly these conditions.
- Any cone defines a order by $u \leq v$ if $v - u \in C$.
- Unfortunately for us, many of the structures that we want to look at are cones but are not part of any obvious vector space: e.g. the measures on a space.
What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.
- Then for any positive $v \in P$ and positive real $r$, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$.
- We define a cone $C$ in a vector space $V$ to be a set with exactly these conditions.
- Any cone defines a order by $u \leq v$ if $v - u \in C$.
- Unfortunately for us, many of the structures that we want to look at are cones but are not part of any obvious vector space: e.g. the measures on a space.
- We could artificially embed them in a vector space, for example, by introducing signed measures.
Definition of Cones

A **cone** is a commutative monoid \((V, +, 0)\) with an action of \(\mathbb{R}^{\geq 0}\). Multiplication by reals distributes over addition and the following cancellation law holds:

\[
\forall u, v, w \in V, v + u = w + u \Rightarrow v = w.
\]

The following strictness property also holds:

\[
v + w = 0 \Rightarrow v = w = 0.
\]

Note that every cone comes with a natural order.

An order on a cone

If \(u, v \in V\), a cone, one says \(u \leq v\) if and only if there is an element \(w \in V\) such that \(u + w = v\).
Normed cones

Definition of a normed cone

A **normed cone** $C$ is a cone with a function $\| \cdot \| : C \rightarrow \mathbb{R}^{\geq 0}$ satisfying the usual conditions:

- $\| v \| = 0$ if and only if $v = 0$
- $\forall r \in \mathbb{R}^{\geq 0}, v \in C$, $\| r \cdot v \| = r \| v \|$
- $\| u + v \| \leq \| u \| + \| v \|$
- $u \leq v \Rightarrow \| u \| \leq \| v \|.$

Normally one uses norms to talk about convergence of Cauchy sequences. But without negation how can we talk about Cauchy sequences?
Normed cones

Definition of a normed cone

A **normed cone** $C$ is a cone with a function $\| \cdot \| : C \rightarrow \mathbb{R}^{\geq 0}$ satisfying the usual conditions:

- $\|v\| = 0$ if and only if $v = 0$
- $\forall r \in \mathbb{R}^{\geq 0}, v \in C, \|r \cdot v\| = r \|v\|$
- $\|u + v\| \leq \|u\| + \|v\|$
- $u \leq v \Rightarrow \|u\| \leq \|v\|$. 

Normally one uses norms to talk about convergence of Cauchy sequences. But without negation how can we talk about Cauchy sequences?

We can write $u_i - u_j$ when we really mean the (unique) $w$ such that $u_j + w = u_i$; needs $u_j \leq u_i$. So, in the case that we have an increasing sequence we can define Cauchy sequence in, more or less, the usual way.
Completeness

However, order-theoretic concepts can be used instead.

Complete normed cones

An \textbf{ω-complete normed cone} is a normed cone such that if \( \{a_i \mid i \in I\} \) is an increasing sequence with \( \{||a_i||\} \) bounded then the lub \( \bigvee_{i \in I} a_i \) exists and \( \bigvee_{i \in I} ||a_i|| = || \bigvee_{i \in I} a_i || \).

Convergence in the sense of norm and in the order theory sense match.
However, order-theoretic concepts can be used instead.

**Complete normed cones**

An *ω-complete normed cone* is a normed cone such that if \( \{ a_i \mid i \in I \} \) is an increasing sequence with \( \{ \| a_i \| \} \) bounded then the lub \( \bigvee_{i \in I} a_i \) exists and \( \bigvee_{i \in I} \| a_i \| = \| \bigvee_{i \in I} a_i \| \).

Convergence in the sense of norm and in the order theory sense match.

**Selinger’s lemma**

Suppose that \( u_i \) is an *ω*-chain with a l.u.b. in an *ω*-complete normed cone and \( u \) is an upper bound of the \( u_i \). Suppose furthermore that \( \lim_{i \to \infty} \| u - u_i \| = 0 \). Then \( u = \bigvee_i u_i \).
Completeness

However, order-theoretic concepts can be used instead.

**Complete normed cones**

An \( \omega \)-**complete normed cone** is a normed cone such that if \( \{a_i \mid i \in I\} \) is an increasing sequence with \( \{\|a_i\|\} \) bounded then the lub \( \bigvee_{i \in I} a_i \) exists and \( \bigvee_{i \in I} \|a_i\| = \| \bigvee_{i \in I} a_i \| \).

Convergence in the sense of norm and in the order theory sense match.

**Selinger’s lemma**

Suppose that \( u_i \) is an \( \omega \)-chain with a l.u.b. in an \( \omega \)-complete normed cone and \( u \) is an upper bound of the \( u_i \). Suppose furthermore that \( \lim_{i \to \infty} \|u - u_i\| = 0 \). Then \( u = \bigvee_i u_i \).

Here we are writing \( u - u_i \) informally. We really mean \( w_i \) where \( u_i + w_i = u \).
Continuous maps

An $\omega$-continuous linear map between two cones is one that preserves least upper bounds of countable chains.
Maps between cones

Continuous maps

An \textbf{\textit{\(\omega\)-continuous}} linear map between two cones is one that preserves least upper bounds of countable chains.

Bounded maps

A \textit{bounded} linear map of normed cones \(f : C \rightarrow D\) is one such that for all \(u\) in \(C\), \(||f(u)|| \leq K||u||\) for some real number \(K\). Any linear continuous map of complete normed cones is bounded.
Maps between cones

Continuous maps

An \( \omega \text{-continuous} \) linear map between two cones is one that preserves least upper bounds of countable chains.

Bounded maps

A \textit{bounded} linear map of normed cones \( f : C \to D \) is one such that for all \( u \) in \( C \), \( \|f(u)\| \leq K\|u\| \) for some real number \( K \). Any linear continuous map of complete normed cones is bounded.

Norm of a bounded map

The norm of a bounded linear map \( f : C \to D \) is defined as
\[
\|f\| = \sup\{\|f(u)\| : u \in C, \|u\| \leq 1\}.
\]
A category of normed cones

The ambient category

The $\omega$-complete normed cones, along with $\omega$-continuous linear maps, form a category which we shall denote $\omega\text{CC}$. Isomorphisms in this category are always isometries.
A category of normed cones

The ambient category

The $\omega$-complete normed cones, along with $\omega$-continuous linear maps, form a category which we shall denote $\omega\text{CC}$.

The subcategory of interest

we define the subcategory $\omega\text{CC}_1$: the norms of the maps are all bounded by 1. Isomorphisms in this category are always isometries.
Dual cone

Given an $\omega$-complete normed cone $C$, its dual $C^*$ is the set of all $\omega$-continuous linear maps from $C$ to $\mathbb{R}_+$. We define the norm on $C^*$ to be the operator norm.
Dual cones

Given an $\omega$-complete normed cone $C$, its dual $C^*$ is the set of all $\omega$-continuous linear maps from $C$ to $\mathbb{R}_+$. We define the norm on $C^*$ to be the operator norm.

Basic facts

$C^*$ is an $\omega$-complete normed cone as well, and the cone order corresponds to the point wise order.
The duality functor

In $\omega\mathbf{CC}$, the dual operation becomes a contravariant functor. If $f : C \to D$ is a map of cones, we define $f^* : D^* \to C^*$ as follows: given a map $L$ in $D^*$, we define a map $f^*L$ in $C^*$ as $f^*L(u) = L(f(u))$. 
How does this compare with Banach spaces?

This dual is stronger than the dual in usual Banach spaces, where we only require the maps to be bounded. For instance, it turns out that the dual to $L^+_{\infty}(X)$ (to be defined later) is isomorphic to $L^+_1(X)$, which is not the case with the Banach space $L_{\infty}(X)$. 
If $\mu$ is a measure on $X$, then one has the well-known Banach spaces $L_1$ and $L_\infty$. 
If $\mu$ is a measure on $X$, then one has the well-known Banach spaces $L_1$ and $L_\infty$.

These can be restricted to cones by considering the $\mu$-almost everywhere positive functions.
If $\mu$ is a measure on $X$, then one has the well-known Banach spaces $L_1$ and $L_\infty$.

These can be restricted to cones by considering the $\mu$-almost everywhere positive functions.

We will denote these cones by $L_1^+(X, \Sigma, \mu)$ and $L_\infty^+(X, \Sigma)$. 
If \( \mu \) is a measure on \( X \), then one has the well-known Banach spaces \( L_1 \) and \( L_\infty \).

These can be restricted to cones by considering the \( \mu \)-almost everywhere positive functions.

We will denote these cones by \( L_1^+ (X, \Sigma, \mu) \) and \( L_\infty^+ (X, \Sigma) \).

These are complete normed cones.
Let $(X, \Sigma, p)$ be a measure space with finite measure $p$. We denote by $\mathcal{M} \ll p (X)$, the cone of all measures on $(X, \Sigma, p)$ that are absolutely continuous with respect to $p$. If $q$ is such a measure, we define its norm to be $q(X)$. $\mathcal{M} \ll p (X)$ is also an $\omega$-complete normed cone. The cones $\mathcal{M} \ll p (X)$ and $L^+_{\infty}(X, \Sigma, p)$ are isometrically isomorphic. We write $\mathcal{M}_p \text{UB}(X)$ for the cone of all measures on $(X, \Sigma)$ that are uniformly less than a multiple of the measure $p$: $q \in \mathcal{M}_p \text{UB}(X)$ means that for some real constant $K > 0$ we have $q \leq Kp$. The cones $\mathcal{M}_p \text{UB}(X)$ and $L^+_{\infty}(X, \Sigma, p)$ are isomorphic.
Cones that we use II

- Let $(X, \Sigma, p)$ be a measure space with finite measure $p$. We denote by $M \ll p(X)$, the cone of all measures on $(X, \Sigma, p)$ that are absolutely continuous with respect to $p$.
- If $q$ is such a measure, we define its norm to be $q(X)$. 
Let $(X, \Sigma, p)$ be a measure space with finite measure $p$. We denote by $\mathcal{M} \ll p(X)$, the cone of all measures on $(X, \Sigma, p)$ that are absolutely continuous with respect to $p$

If $q$ is such a measure, we define its norm to be $q(X)$.

$\mathcal{M} \ll p(X)$ is also an $\omega$-complete normed cone.
Cones that we use II

Let \((X, \Sigma, p)\) be a measure space with finite measure \(p\). We denote by \(\mathcal{M} \ll p(X)\), the cone of all measures on \((X, \Sigma, p)\) that are absolutely continuous with respect to \(p\).

If \(q\) is such a measure, we define its norm to be \(q(X)\).

\(\mathcal{M} \ll p(X)\) is also an \(\omega\)-complete normed cone.

The cones \(\mathcal{M} \ll p(X)\) and \(L^+_1(X, \Sigma, p)\) are isometrically isomorphic in \(\omega\text{CC}\).
Let \((X, \Sigma, p)\) be a measure space with finite measure \(p\). We denote by \(\mathcal{M} \ll p(X)\), the cone of all measures on \((X, \Sigma, p)\) that are absolutely continuous with respect to \(p\).

If \(q\) is such a measure, we define its norm to be \(q(X)\).

\(\mathcal{M} \ll p(X)\) is also an \(\omega\)-complete normed cone.

The cones \(\mathcal{M} \ll p(X)\) and \(L_1^+(X, \Sigma, p)\) are isometrically isomorphic in \(\omega\text{CC}\).

We write \(\mathcal{M}_\text{UB}^p(X)\) for the cone of all measures on \((X, \Sigma)\) that are uniformly less than a multiple of the measure \(p\): \(q \in \mathcal{M}_\text{UB}^p\) means that for some real constant \(K > 0\) we have \(q \leq Kp\).
Let \((X, \Sigma, p)\) be a measure space with finite measure \(p\). We denote by \(\mathcal{M} \ll p(X)\), the cone of all measures on \((X, \Sigma, p)\) that are absolutely continuous with respect to \(p\).

If \(q\) is such a measure, we define its norm to be \(q(X)\).

\(\mathcal{M} \ll p(X)\) is also an \(\omega\)-complete normed cone.

The cones \(\mathcal{M} \ll p(X)\) and \(L_1^+(X, \Sigma, p)\) are isometrically isomorphic in \(\omega\text{CC}\).

We write \(\mathcal{M}_\text{UB}^p(X)\) for the cone of all measures on \((X, \Sigma)\) that are uniformly less than a multiple of the measure \(p\): \(q \in \mathcal{M}_\text{UB}^p\) means that for some real constant \(K > 0\) we have \(q \leq Kp\).

The cones \(\mathcal{M}_\text{UB}^p(X)\) and \(L_\infty^+(X, \Sigma, p)\) are isomorphic.
A Riesz-like theorem

The dual of the cone $L_{\infty}^+(X, \Sigma, p)$ is isometrically isomorphic to $\mathcal{M}^{\ll p}(X)$. 

Corollary

Since $\mathcal{M}^{\ll p}(X)$ is isometrically isomorphic to $L^1_\infty(X)$, an immediate corollary is that $L^+_{\infty}(X, \Sigma, p)$ is isometrically isomorphic to $L^1_{\infty}(X)$, which is of course false in general in the context of Banach spaces.
A Riesz-like theorem

The dual of the cone $L^+_\infty(X, \Sigma, p)$ is isometrically isomorphic to $\mathcal{M}^{\ll p}(X)$.

Corollary

Since $\mathcal{M}^{\ll p}(X)$ is isometrically isomorphic to $L^+_1(X)$, an immediate corollary is that $L^+_\infty(X)^*$ is isometrically isomorphic to $L^+_1(X)$, which is of course false in general in the context of Banach spaces.
Another Riesz-like theorem

The dual of the cone \( L_1^+(X, \Sigma, p) \) is isometrically isomorphic to \( \mathcal{M}_UB^p(X) \).
Another Riesz-like theorem

The dual of the cone $L_1^+(X, \Sigma, p)$ is isometrically isomorphic to $\mathcal{M}^p_{UB}(X)$.

Corollary

$\mathcal{M}^p_{UB}(X)$ is isometrically isomorphic to $L_1^+(X)$, hence immediate corollary is that $L_1^+,\ast(X)$ is isometrically isomorphic to $L_1^+(X)$. 
The pairing

**Pairing function**

There is a map from the product of the cones $L_{\infty}^+(X,p)$ and $L_1^+(X,p)$ to $\mathbb{R}^+$ defined as follows:

$$\forall f \in L_{\infty}^+(X,p), g \in L_1^+(X,p) \quad \langle f, g \rangle = \int fg dp.$$
The pairing

Pairing function

There is a map from the product of the cones $L^+_\infty(X, p)$ and $L^+_1(X, p)$ to $\mathbb{R}^+$ defined as follows:

$$\forall f \in L^+_\infty(X, p), g \in L^+_1(X, p) \quad \langle f, g \rangle = \int fg dp.$$  

This map is bilinear and is continuous and $\omega$-continuous in both arguments; we refer to it as the pairing.
This pairing allows one to express the dualities in a very convenient way. For example, the isomorphism between $L^+_\infty(X, p)$ and $(L^+_1(X, p))^*$ sends $f \in L^+_\infty(X, p)$ to $\lambda g. \langle f, g \rangle = \lambda g. \int fg dp$. 
We fix a probability triple \((X, \Sigma, p)\) and focus on six spaces of cones that are based on them. They break into two natural groups of three isomorphic spaces. The first three spaces are:

**A1** \(\mathcal{M}_{\ll p}(X)\) - the cone of all measures on \((X, \Sigma, p)\) that are absolutely continuous with respect to \(p\),
Summary of cones

We fix a probability triple $(X, \Sigma, p)$ and focus on six spaces of cones that are based on them. They break into two natural groups of three isomorphic spaces. The first three spaces are:

A1 $\mathcal{M}^{<p}(X)$ - the cone of all measures on $(X, \Sigma, p)$ that are absolutely continuous with respect to $p$,

A2 $L^+_1(X, p)$ - the cone of integrable almost-everywhere positive functions,
We fix a probability triple $(X, \Sigma, p)$ and focus on six spaces of cones that are based on them. They break into two natural groups of three isomorphic spaces. The first three spaces are:

**A1** $\mathcal{M}^{\ll p}(X)$ - the cone of all measures on $(X, \Sigma, p)$ that are absolutely continuous with respect to $p$,

**A2** $L^+_1(X, p)$ - the cone of integrable almost-everywhere positive functions,

**A3** $L^{+,*}_\infty(X, p)$ - the dual cone of the the cone of almost-everywhere positive bounded measurable functions.
The next group of three isomorphic spaces are:

\begin{itemize}
  \item \textbf{B1} \quad \mathcal{M}^p_{UB}(X) - the cone of all measures that are uniformly less than a multiple of the measure $p$,
\end{itemize}
The next group of three isomorphic spaces are:

B1 $\mathcal{M}_\text{UB}^p(X)$ - the cone of all measures that are uniformly less than a multiple of the measure $p$,

B2 $L_\infty^+(X, p)$ - the cone of almost-everywhere positive functions in the normed vector space $L_\infty(X, p)$,
The next group of three isomorphic spaces are:

B1  $\mathcal{M}^p_{UB}(X)$ - the cone of all measures that are uniformly less than a multiple of the measure $p$,

B2  $L^+_\infty(X, p)$ - the cone of almost-everywhere positive functions in the normed vector space $L_\infty(X, p)$,

B3  $L^+_1(X, p)$ - the dual of the cone of almost-everywhere positive functions in the normed vector space $L_1(X, p)$. 
The spaces defined in A1, A2 and A3 are dual to the spaces defined in B1, B2 and B3 respectively. The situation may be depicted in the diagram

\[
\begin{align*}
\mathcal{M}^p (X) & \xleftarrow{\sim} L_1^+ (X, p) & \xrightarrow{\sim} L_\infty^+ (X, p) \\
\mathcal{M}^p_{UB} & \xleftarrow{\sim} L_\infty^+ (X, p) & \xrightarrow{\sim} L_1^+ (X, p)
\end{align*}
\]

where the vertical arrows represent dualities and the horizontal arrows represent isomorphisms.
Given \((X, \Sigma, p)\) and \((Y, \Lambda)\) and a measurable function \(f : X \rightarrow Y\) we obtain a measure \(q\) on \(Y\) by \(q(B) = p(f^{-1}(B))\). This is written \(M_f(p)\) and is called the image measure of \(p\) under \(f\).
Given \((X, \Sigma, p)\) and \((Y, \Lambda)\) and a measurable function \(f : X \rightarrow Y\) we obtain a measure \(q\) on \(Y\) by \(q(B) = p(f^{-1}(B))\). This is written \(M_f(p)\) and is called the *image measure* of \(p\) under \(f\).

We say that a measure \(\nu\) is **absolutely continuous** with respect to another measure \(\mu\) if for any measurable set \(A\), \(\mu(A) = 0\) implies that \(\nu(A) = 0\). We write \(\nu \ll \mu\).
Given \((X, \Sigma, p)\) and \((Y, \Lambda)\) and a measurable function \(f : X \rightarrow Y\) we obtain a measure \(q\) on \(Y\) by \(q(B) = p(f^{-1}(B))\). This is written \(M_f(p)\) and is called the \textit{image measure} of \(p\) under \(f\).

2. We say that a measure \(\nu\) is \textbf{absolutely continuous} with respect to another measure \(\mu\) if for any measurable set \(A\), \(\mu(A) = 0\) implies that \(\nu(A) = 0\). We write \(\nu \ll \mu\).

3. For \textit{finite} measures \(\nu\), \(\nu \ll \mu\) is equivalent to:

\[
\forall \varepsilon > 0, \exists \delta > 0, \text{ s.t. } \forall A \text{ with } \mu(A) \leq \delta, \nu(A) \leq \varepsilon.
\]
The Radon-Nikodym Theorem

The Radon-Nikodym theorem is a central result in measure theory allowing one to define a “derivative” of a measure with respect to another measure.

**Radon-Nikodym**

If \( \nu \ll \mu \), where \( \nu, \mu \) are finite measures on a measurable space \((X, \Sigma)\) there is a positive measurable function \( h \) on \( X \) such that for every measurable set \( B \)

\[
\nu(B) = \int_B h \, d\mu.
\]

The function \( h \) is defined uniquely up to a set of \( \mu \)-measure 0. The function \( h \) is called the Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \); we denote it by \( \frac{d\nu}{d\mu} \). Since \( \nu \) is finite, \( \frac{d\nu}{d\mu} \in L_1^+(X, \mu) \).
Given an (almost-everywhere) positive function $f \in L_1(X, p)$, we let $f \cdot p$ be the measure which has density $f$ with respect to $p$. Two identities that we get from the Radon-Nikodym theorem are:

1. Given $q \ll p$, we have $dq \cdot dp = q$.
2. Given $f \in L_1^+(X, p)$, $df \cdot dp = f$.

These two identities just say that the operations $(\cdot) \cdot p$ and $d(\cdot) \cdot dp$ are inverses of each other as maps between $L_1^+(X, p)$ and $M \ll p(X)$, the space of finite measures on $X$ that are absolutely continuous with respect to $p$. 
Given an (almost-everywhere) positive function \( f \in L_1(X, p) \), we let \( f \cdot p \) be the measure which has density \( f \) with respect to \( p \).

Two identities that we get from the Radon-Nikodym theorem are:
Given an (almost-everywhere) positive function $f \in L_1(X, p)$, we let $f \cdot p$ be the measure which has density $f$ with respect to $p$.

Two identities that we get from the Radon-Nikodym theorem are:

- given $q \ll p$, we have $\frac{dq}{dp} \cdot p = q$. 

These two identities just say that the operations $(\cdot)p$ and $d(\cdot)dp$ are inverses of each other as maps between $L_1^+(X, p)$ and $\mathcal{M} \ll p(X)$, the space of finite measures on $X$ that are absolutely continuous with respect to $p$. 

---

Panangaden (McGill University)
Notation for Radon-Nikodym

1. Given an (almost-everywhere) positive function \( f \in L_1(X, p) \), we let \( f \cdot p \) be the measure which has density \( f \) with respect to \( p \).

2. Two identities that we get from the Radon-Nikodym theorem are:
   - given \( q \ll p \), we have \( \frac{dq}{dp} \cdot p = q \).
   - given \( f \in L_1^+(X, p) \), \( \frac{df \cdot p}{dp} = f \).
Given an (almost-everywhere) positive function \( f \in L_1(X, p) \), we let 
\( f \cdot p \) be the measure which has density \( f \) with respect to \( p \).

Two identities that we get from the Radon-Nikodym theorem are:

- given \( q \ll p \), we have \( \frac{dq}{dp} \cdot p = q \).
- given \( f \in L_1^+(X, p) \), \( \frac{df}{dp} \cdot p = f \).

These two identities just say that the operations \( (\cdot) \cdot p \) and \( \frac{d(\cdot)}{dp} \) are inverses of each other as maps between \( L_1^+(X, p) \) and \( \mathcal{M} \ll p(X) \) the space of finite measures on \( X \) that are absolutely continuous with respect to \( p \).
The expectation $\mathbb{E}_p(f)$ of a measurable function $f$ is the average computed by $\int f \, dp$ and therefore it is just a number.
1. The expectation $\mathbb{E}_p(f)$ of a measurable function $f$ is the average computed by $\int f \, dp$ and therefore it is just a number.

2. The *conditional* expectation is not a mere number but a random variable.
The expectation $\mathbb{E}_p(f)$ of a measurable function $f$ is the average computed by $\int f \, dp$ and therefore it is just a number.

The *conditional* expectation is not a mere number but a random variable.

It is meant to measure the expected value in the presence of additional information.
The expectation $\mathbb{E}_p(f)$ of a measurable function $f$ is the average computed by $\int f \, dp$ and therefore it is just a number.

The *conditional* expectation is not a mere number but a random variable.

It is meant to measure the expected value in the presence of additional information.

The additional information takes the form of a sub-$\sigma$ algebra, say $\Lambda$, of $\Sigma$. The experimenter knows, for every $B \in \Lambda$, whether the outcome is in $B$ or not.
The expectation $\mathbb{E}_p(f)$ of a measurable function $f$ is the average computed by $\int f \, dp$ and therefore it is just a number.

The *conditional* expectation is not a mere number but a random variable.

It is meant to measure the expected value in the presence of additional information.

The additional information takes the form of a sub-$\sigma$ algebra, say $\Lambda$, of $\Sigma$. The experimenter knows, for every $B \in \Lambda$, whether the outcome is in $B$ or not.

Now she can recompute the expectation values given this information.
Formalizing conditional expectation

- It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.
It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.

**Kolmogorov**

Let $(X, \Sigma, p)$ be a measure space with $p$ a finite measure, $f$ be in $L_1(X, \Sigma, p)$ and $\Lambda$ be a sub-$\sigma$-algebra of $\Sigma$, then there exists a $g \in L_1(X, \Lambda, p)$ such that for all $B \in \Lambda$

$$\int_B f \, dp = \int_B g \, dp.$$
Formalizing conditional expectation

- It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.

Kolmogorov

Let $(X, \Sigma, p)$ be a measure space with $p$ a finite measure, $f$ be in $L_1(X, \Sigma, p)$ and $\Lambda$ be a sub-$\sigma$-algebra of $\Sigma$, then there exists a $g \in L_1(X, \Lambda, p)$ such that for all $B \in \Lambda$

$$\int_B f \, dp = \int_B g \, dp.$$ 

- This function $g$ is usually denoted by $\mathbb{E}(f|\Lambda)$. 
Formalizing conditional expectation

- It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.

**Kolmogorov**

Let \((X, \Sigma, p)\) be a measure space with \(p\) a finite measure, \(f\) be in \(L_1(X, \Sigma, p)\) and \(\Lambda\) be a sub-\(\sigma\)-algebra of \(\Sigma\), then there exists a \(g \in L_1(X, \Lambda, p)\) such that for all \(B \in \Lambda\)

\[
\int_B f\,dp = \int_B g\,dp.
\]

- This function \(g\) is usually denoted by \(\mathbb{E}(f|\Lambda)\).
- We clearly have \(f \cdot p \ll p\) so the required \(g\) is simply \(\frac{df \cdot p}{dp|_{\Lambda}}\), where \(p|_{\Lambda}\) is the restriction of \(p\) to the sub-\(\sigma\)-algebra \(\Lambda\).
Properties of conditional expectation

1. The point of requiring $\Lambda$-measurability is that it “smooths out” variations that are too rapid to show up in $\Lambda$. 
Properties of conditional expectation

1. The point of requiring $\Lambda$-measurability is that it “smooths out” variations that are too rapid to show up in $\Lambda$.

2. The conditional expectation is linear, increasing with respect to the pointwise order.
Properties of conditional expectation

1. The point of requiring $\Lambda$-measurability is that it “smooths out” variations that are too rapid to show up in $\Lambda$.

2. The conditional expectation is *linear, increasing* with respect to the pointwise order.

3. It is defined uniquely $p$-almost everywhere.
We define two categories $\text{Rad}_\infty$ and $\text{Rad}_1$ that will be needed for the functorial definition of conditional expectation.
Where the action happens

- We define two categories $\text{Rad}_\infty$ and $\text{Rad}_1$ that will be needed for the functorial definition of conditional expectation.
- This will allow for $L_\infty$ and $L_1$ versions of the theory.
Where the action happens

- We define two categories $\text{Rad}_\infty$ and $\text{Rad}_1$ that will be needed for the functorial definition of conditional expectation.
- This will allow for $L_\infty$ and $L_1$ versions of the theory.
- Going between these versions by duality will be very useful.
The “infinity” category

The category $\mathbf{Rad}_\infty$ has as objects probability spaces, and as arrows $\alpha : (X, p) \to (Y, q)$, measurable maps such that $M_\alpha(p) \leq Kq$ for some real number $K$.

The reason for choosing the name $\mathbf{Rad}_\infty$ is that $\alpha \in \mathbf{Rad}_\infty$ maps to $d/dqM_\alpha(p) \in L_\infty^+(Y, q)$. 
The category $\text{Rad}_1$ has as objects probability spaces and as arrows $\alpha : (X, p) \rightarrow (Y, q)$, measurable maps such that $M_\alpha(p) \ll q$. The reason for choosing the name $\text{Rad}_1$ is that $\alpha \in \text{Rad}_1$ maps to $dq/M_\alpha(p) \in L^+_{\infty}(Y, q)$. The fact that the category $\text{Rad}_\infty$ embeds in $\text{Rad}_1$ reflects the fact that $L^+_{\infty}$ embeds in $L^+_1$. 
The “one” category

Rad$_1$

The category Rad$_1$ has as objects probability spaces and as arrows $\alpha : (X, p) \rightarrow (Y, q)$, measurable maps such that $M_\alpha(p) \ll q$.

The reason for choosing the name Rad$_1$ is that $\alpha \in$ Rad$_1$ maps to $d/dqM_\alpha(p) \in L^+_1(Y, q)$.
Rad$_1$

The category **Rad$_1$** has as objects probability spaces and as arrows $\alpha : (X, p) \rightarrow (Y, q)$, measurable maps such that $M_{\alpha}(p) \ll q$.

1. The reason for choosing the name **Rad$_1$** is that $\alpha \in \text{Rad}_1$ maps to $d/dqM_{\alpha}(p) \in L_1^+(Y, q)$.

2. The fact that the category **Rad$_\infty$** embeds in **Rad$_1$** reflects the fact that $L_\infty^+$ embeds in $L_1^+$. 
Recall the isomorphism between $L^+_{\infty}(X,p)$ and $L^+_1(X,p)$ mediated by the pairing function:

$$f \in L^+_{\infty}(X,p) \mapsto \lambda g : L^+_1(X,p). \langle f, g \rangle = \int f g dp.$$
Now, precomposition with $\alpha$ in $\text{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y, q)$ to $L_1^+(X, p)$. 
Now, precomposition with $\alpha$ in $\text{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y, q)$ to $L_1^+(X, p)$.

Dually, given $\alpha \in \text{Rad}_1 : (X, p) \to (Y, q)$ and $g \in L^+_\infty(Y, q)$ we have that $P_\infty(\alpha)(g) \in L^+_\infty(X, p)$. 

Thus the subscripts on the two precomposition functors describe the target categories.
Now, precomposition with $\alpha$ in $\text{Rad}_{\infty}$ gives a map $P_1(\alpha)$ from $L_1^+(Y, q)$ to $L_1^+(X, p)$.

Dually, given $\alpha \in \text{Rad}_1 : (X, p) \rightarrow (Y, q)$ and $g \in L_\infty^+(Y, q)$ we have that $P_\infty(\alpha)(g) \in L_\infty^+(X, p)$.

Thus the subscripts on the two precomposition functors describe the target categories.
1. Now, precomposition with $\alpha$ in $\text{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y, q)$ to $L_1^+(X, p)$.

2. Dually, given $\alpha \in \text{Rad}_1 : (X, p) \to (Y, q)$ and $g \in L_\infty^+(Y, q)$ we have that $P_\infty(\alpha)(g) \in L_\infty^+(X, p)$.

3. Thus the subscripts on the two precomposition functors describe the *target* categories.

4. Using the $*$-functor we get a map $(P_1(\alpha))^*$ from $L_1^+,^*(X, p)$ to $L_1^+,^*(Y, q)$ in the first case and
Now, precomposition with $\alpha$ in $\text{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y, q)$ to $L_1^+(X, p)$.

Dually, given $\alpha \in \text{Rad}_1 : (X, p) \rightarrow (Y, q)$ and $g \in L_\infty^+(Y, q)$ we have that $P_\infty(\alpha)(g) \in L_\infty^+(X, p)$.

Thus the subscripts on the two precomposition functors describe the target categories.

Using the $*$-functor we get a map $(P_1(\alpha))^*$ from $L_1^+,*(X, p)$ to $L_1^+,*(Y, q)$ in the first case and
dually we get $(P_\infty(\alpha))^*$ from $L_\infty^+,*(X, p)$ to $L_\infty^+,*(Y, q)$. 
The **functor** $\mathbb{E}_\infty(\cdot)$ is a functor from $\text{Rad}_\infty$ to $\omega\text{CC}$ which, on objects, maps $(X, p)$ to $L^+_\infty(X, p)$ and on maps is given as follows:
The functor $E_\infty(\cdot)$ is a functor from $\text{Rad}_\infty$ to $\omega\text{CC}$ which, on objects, maps $(X, p)$ to $L_\infty^+(X, p)$ and on maps is given as follows:

Given $\alpha : (X, p) \to (Y, q)$ in $\text{Rad}_\infty$ the action of the functor is to produce the map $E_\infty(\alpha) : L_\infty^+(X, p) \to L_\infty^+(Y, q)$ obtained by composing $(P_1(\alpha))^*$ with the isomorphisms between $L_1^+,*$ and $L_\infty^+$.
The **functor** $E_\infty(\cdot)$ is a functor from $\text{Rad}_\infty$ to $\omega\text{CC}$ which, on objects, maps $(X, p)$ to $L_\infty^+(X, p)$ and on maps is given as follows:

Given $\alpha : (X, p) \rightarrow (Y, q)$ in $\text{Rad}_\infty$ the action of the functor is to produce the map $E_\infty(\alpha) : L_\infty^+(X, p) \rightarrow L_\infty^+(Y, q)$ obtained by composing $(P_1(\alpha))^*$ with the isomorphisms between $L_1^+,^* \rightarrow L_\infty^+,^*$.
Consequences

1. It is an immediate consequence of the definitions that for any $f \in L_\infty^+(X, p)$ and $g \in L_1(Y, q)$

$$\langle \mathbb{E}_\infty(\alpha)(f), g \rangle_Y = \langle f, P_1(\alpha)(g) \rangle_X.$$
It is an immediate consequence of the definitions that for any $f \in L_1^+(X, p)$ and $g \in L_1(Y, q)$

$$
\langle \mathbb{E}_\infty(\alpha)(f), g \rangle_Y = \langle f, P_1(\alpha)(g) \rangle_X.
$$
It is an immediate consequence of the definitions that for any $f \in L_\infty^+(X, p)$ and $g \in L_1(Y, q)$

$$\langle \mathbb{E}_\infty(\alpha)(f), g \rangle_Y = \langle f, P_1(\alpha)(g) \rangle_X.$$ 

One can informally view this functor as a “left adjoint” in view of this proposition.
It is an immediate consequence of the definitions that for any \( f \in L^+_\infty(X, p) \) and \( g \in L_1(Y, q) \)

\[
\langle \mathbb{E}_\infty(\alpha)(f), g \rangle_Y = \langle f, P_1(\alpha)(g) \rangle_X.
\]

One can informally view this functor as a “left adjoint” in view of this proposition.

Note that since we started with \( \alpha \) in \( \text{Rad}_\infty \) we get the expectation value as a map between the \( L^+_{\infty} \) cones.
The functor $\mathbb{E}_1(\cdot)$ is a functor from $\text{Rad}_1$ to $\omega\text{CC}$ which maps the object $(X, p)$ to $L^+_1(X, p)$ and on maps is given as follows:

Given $\alpha : (X, p) \rightarrow (Y, q)$ in $\text{Rad}_1$ the action of the functor is to produce the map $\mathbb{E}_1(\alpha) : L^+_1(X, p) \rightarrow L^+_1(Y, q)$ obtained by composing $(P_\infty(\alpha))^*$ with the isomorphisms between $L^+_\infty$ and $L^+_1$ as shown in the diagram below:

\[
\begin{array}{ccc}
L^+_\infty & \xleftarrow{\text{(P}_\infty(\alpha))^*} & L^+_1 \\
\downarrow & & \downarrow \mathbb{E}_1(\alpha) \\
L^+_\infty & \xrightarrow{\text{E}_1(\alpha)} & L^+_1
\end{array}
\]
Another “adjoint”

Once again we have an “adjointness” statement; this time it is a right adjoint.

Right adjoint

Given $f \in L_\infty^+(Y, q)$ and $g \in L_1^+(X, p)$ we have

$$\langle f, \mathbb{E}_1(\alpha)(g) \rangle_Y = \langle P_\infty(\alpha)(f), g \rangle_X.$$
Given $\alpha \in \text{Rad}_\infty[(X, p), (Y, q)]$ we have

(a) $\mathbb{E}_1(\alpha)(f \circ \alpha) = \mathbb{E}_\infty(\alpha)(1_X)f$, for $f \in L_1^+(Y, q)$ and

(b) $\mathbb{E}_\infty(\alpha)(f \circ \alpha) = \mathbb{E}_1(\alpha)(1_X)f$, for $f \in L_\infty^+(Y, q)$. 
This is a piece pulled out of a larger work on approximating Markov processes.
This is a piece pulled out of a larger work on approximating Markov processes.

Instead of compressing the state space we compressed the $\sigma$-algebra and used the conditional expectation to define approximate transition kernels.
Why?

- This is a piece pulled out of a larger work on approximating Markov processes.
- Instead of compressing the state space we compressed the $\sigma$-algebra and used the conditional expectation to define approximate transition kernels.
- But that is the subject of a different talk.