EXPECTATION = INTEGRATION (=PROBABILITY)

Simple functions

$$X : (\Omega, \mathcal{B}, P) \xrightarrow{meas} (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

is called a *simple function* if

$$X(\omega) = \sum_{i=1}^{k} a_i 1_{A_i}(\omega)$$

where

$$A_i \in \mathcal{B},\ A_i \cap A_j = \emptyset,\ \sum_{i=1}^{k} A_i = \Omega$$

ie. $\{A_i : i \in \{1, \ldots, k\}\}$ is a *partition* of Ω.

σ-algebra generated by a simple function

is composed of all finite unions of A_i

$$\sigma(X) = \left\{ \sum_{i \in I} A_i : I \subset \{1, \ldots, k\} \right\}$$

$$\mathcal{E} = \{\text{all simple functions on } (\Omega, \mathcal{B})\}$$

$\mathcal{E}$ is a *vector (linear) space*

$$X \in \mathcal{E} \implies aX \in \mathcal{E}$$

$$X, Y \in \mathcal{E} \implies X + Y \in \mathcal{E}$$

Proof, whiteboard

Other properties

$$X, Y \in \mathcal{E} \implies XY \in \mathcal{E}$$

$$X, Y \in \mathcal{E} \implies \max(X, Y) \in \mathcal{E}$$

$$X, Y \in \mathcal{E} \implies \min(X, Y) \in \mathcal{E}$$

Proof, whiteboard

Approximation by a monotone sequence of simple functions

(*Theorem 5.1.1*)

$$X : (\Omega, \mathcal{B}, P) \stackrel{meas}{\longrightarrow} (\mathbb{R}, \mathcal{B}(\mathbb{R})) \iff \exists\{X_n \in \mathcal{E}\} : X_n(\omega) \uparrow X(\omega),\ \omega \in \Omega$$

Proof

"$\Longleftarrow$"

$X_n(\omega)$ *measurable and* $\lim X_n = \limsup X_n$

$$\Longrightarrow \forall B \in \mathcal{B}(\mathbb{R}) :\ X^{-1}(B) = \limsup X_n^{-1}(B) = \bigcap_n \bigcup_{k \geq n} X_n^{-1}(B) \in \mathcal{B}$$

$\square$

“$\Longrightarrow$”

$X(\omega)$ measurable $\Longrightarrow$

$$X_n := \sum_{k=1}^{n2^n} \left(\frac{k-1}{2^n}\right) 1_{\left[\frac{k-1}{2^n} \leq X < \frac{k}{2^n}\right]} + n 1_{[X \geq n]}$$

are measurable since $[\omega : \frac{k-1}{2^n} \leq X(\omega) < \frac{k}{2^n}]$ and $[X(\omega) \geq n]$ are all $\mathcal{B}$-measurable and so

$$X_n \in \mathcal{E}$$

$$X_{n+1} \geq X_n$$

draw a picture!

$$X(\omega) < \infty \implies \forall \omega,\ \exists n : |X(\omega) - X_n(\omega)| \ \leq \ 2^{-n} \longrightarrow 0$$

$$X(\omega) = \infty \implies X_n(\omega) = n \longrightarrow \infty$$

□

Expectations (Lebesgue-Stieltjes integrals) of simple functions

$$X(\omega) = \sum_{i=1}^{k} a_i 1_{A_i}(\omega)$$

Define

$$\forall X \in \mathcal{E} : E(X) \equiv \int X(\omega) dP(\omega) := \sum_{i=1}^{k} a_i P(A_i)$$

Properties (proofs using whiteboard)

- $E(1) = 1$
- $E(1_A) = P(A)$
- $\mathcal{E} \ni X \geq 0 \implies E(X) \geq 0$
- $X, Y \in \mathcal{E} \implies E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$
- $X, Y \in \mathcal{E},\ X \geq Y \implies E(X) \geq E(Y)$
- $X, X_n \in \mathcal{E},\ X_n \uparrow\downarrow X \implies E(X_n) \uparrow\downarrow E(X)$

This latter is true if

$$X_n \downarrow 0 \implies E(X_n) \downarrow 0$$

But

$$\begin{aligned} 0 &\leq X_n(\omega) \leq \sup_{\omega\in\Omega} X_1(\omega) := K \\ \implies & \forall\varepsilon > 0 : 0 \leq E(X_n) = E(X_n 1_{[X_n>\varepsilon]} + X_n 1_{[X_n\leq\varepsilon]}) \leq KP(X_n > \varepsilon) + \varepsilon \end{aligned}$$

and

$$\begin{aligned} [X_n &> \varepsilon] \downarrow \emptyset \implies P(X_n > \varepsilon) \downarrow 0 \\ &\implies \limsup_n E(X_n) \leq \varepsilon \end{aligned}$$

since ε is arbitrary, $E(X_n) \downarrow 0$

□

Extension of the definition

Step 1, extension to arbitrary non-negative rv's (measurable functions)

$$\mathcal{E}_+ = \{X \geq 0,\ X \text{ simple}\}$$

$$\bar{\mathcal{E}}_+ = \{X \geq 0 :\ X : (\Omega, \mathcal{B}) \xrightarrow{meas} (\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))\}$$

For $X \in \bar{\mathcal{E}}_+$:

$P[X = \infty] > 0 \implies E(X) = \infty$

$P[X = \infty] > 0 \implies$ we can find

$$\{X_n \in \mathcal{E}_+\} : X_n(\omega) \uparrow X(\omega),\ \omega \in \Omega$$

and since sequence $E(X_n)$ is nondecreasing, it has a limit and we can define

$$E(X) = \lim_n E(X_n).$$

Proposition 5.2.1.

If

$$\{X_n \in \mathcal{E}_+\} : X_n(\omega) \uparrow X(\omega),\ \omega \in \Omega$$

$$\{Y_n \in \mathcal{E}_+\} : Y_n(\omega) \uparrow X(\omega),\ \omega \in \Omega$$

then

$$\lim_n E(X_n) = \lim_n E(Y_n)$$

so that $E(X) = \lim_n E(X_n)$ is well defined.

Structure of the proof (whiteboard)

Proof of the implication

$$\lim_n \uparrow X_n \leq \lim_n \uparrow Y_n \implies \lim_n \uparrow E(X_n) \leq \lim_n \uparrow E(Y_n)$$

Define simple functions

$$\min(X_n, Y_m) \in \mathcal{E}_+$$

notice that

$$\lim_m \uparrow \min(X_n, Y_m) = X_n \in \mathcal{E}_+$$

since

$$\lim_m Y_m \geq \lim_m X_m \geq X_n$$

which is also a simple function.

For simple $\mathcal{E}_+$ functions we know that

$$\lim_m \uparrow E[\min(X_n, Y_m)] = E[\lim_m \uparrow \min(X_n, Y_m)] = E(X_n)$$

so by monotonicity

$$E(X_n) = \lim_m \uparrow E[\min(X_n, Y_m)] \leq \lim_m E(Y_m)$$

and so also

$$\lim_n \uparrow E(X_n) \leq \lim_m \uparrow E(Y_m)$$

$\square$

Properties of Expectations of $X \in \bar{\mathcal{E}}_+$

1.

$$E(X) \in [0, \infty]$$

$$X, Y \in \bar{\mathcal{E}}_+,\ X \leq Y \implies E(X) \leq E(Y)$$

2.

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y), \ \alpha, \beta \geq 0$$

Indeed, select $X_n \uparrow X$, $Y_n \uparrow Y$ and $X_n, Y_n \in \mathcal{E}_+$.

$$\begin{aligned} E(cX) &= \lim_n E(cX_n) \\ &= \lim_n cE(X_n) \text{ linearity on } \mathcal{E}_+ \\ &= c \lim_n E(X_n) \\ &= cE(X) \end{aligned}$$

$$\begin{aligned} E(X+Y) &= \lim_n E(X_n + Y_n) \\ &= \lim_n [E(X_n) + E(Y_n)] \text{ linearity on } \mathcal{E}_+ \\ &= \lim_n E(X_n) + \lim_n E(Y_n) \\ &= E(X) + E(Y) \end{aligned}$$

3. *Monotone Convergence Theorem (MCT)*:

$$0 \leq X_n \uparrow X,\ X_n, X \in \bar{\mathcal{E}}_+ \implies E(X_n) \uparrow E(X)$$

or equivalently

$$E\left(\lim_n \uparrow X_n\right) = \lim_n \uparrow E(X_n)$$

Proof

X_n approximated by a nondecreasing sequence of simple functions

$$Y_m^{(n)} \uparrow X_n,\ m \longrightarrow \infty$$

Find a sequence of simple functions approximating X

$$Z_m \uparrow X,\ m \longrightarrow \infty$$

Let us try

$$Z_m(\omega) = \max_{n=1,\ldots,m} [Y_m^{(n)}(\omega)]$$

$\{Z_m\}$ is non-decreasing, since

$$\begin{aligned} Z_m(\omega) &\leq \max_{n=1,\ldots,m} [Y_{m+1}^{(n)}(\omega)] \\ &\leq \max_{n=1,\ldots,m+1} [Y_{m+1}^{(n)}(\omega)] = Z_{m+1}(\omega) \end{aligned}$$

(A)

$$Y_m^{(n)} \leq \max_{j=1,\ldots,m} [Y_m^{(j)}] = Z_m$$

(B)

$$Z_m \leq \max_{j=1,\ldots,m} X_j = X_m$$

(C)

$$\Longrightarrow Y_m^{(n)} \leq Z_m \leq X_m$$

By taking limits first wrt m and then wrt n, we obtain

(D)

$$X = \lim_n X_n = \lim_m Z_m$$

ie. $\{Z_m\}$ is the desired simple-function approximation of X.

(E) By monotonicity of expectation on $\mathcal{E}$ and by (C)

$$\begin{aligned} E(X_n) &= \lim_m \uparrow E(Y_m^{(n)}) \\ &\leq \lim_m \uparrow E(Z_m) \\ &\leq \lim_m \uparrow E(X_m) \end{aligned}$$

However $\{Z_m\}$ is the simple-function approximation of X, so by definition of expectation on $\bar{\mathcal{E}}_+$

$$E(X) = E(\lim_m \uparrow Z_m) = \lim_m \uparrow E(Z_m)$$

Hence from (E)

$$E(X_n) \leq E(X) = \lim_m \uparrow E(Z_m) \leq \lim_m \uparrow E(X_m)$$

and

$$\lim_n \uparrow E(X_n) \leq E(X) \leq \lim_m \uparrow E(X_m)$$

□

Extension to random variables of arbitrary sign

Define

$$X^+(\omega) = \max(X(\omega), 0), \; X^-(\omega) = \max(-X(\omega), 0)$$

Properties

$$X^+(\omega), X^-(\omega) \geq 0$$

$$X(\omega) = X^+(\omega) - X^-(\omega)$$

$$|X(\omega)| = X^+(\omega) + X^-(\omega)$$

$$X \in \mathcal{B}/\mathcal{B}(\mathbb{R}) \iff X^+, X^-, |X| \in \mathcal{B}/\mathcal{B}(\mathbb{R})$$

X is *quasi-integrable* if $E(X^+) < \infty$ or $E(X^-) < \infty$ (at least one),

$$E(X) := E(X^+) - E(X^-)$$

In this case, $E(X)$ may be equal to $-\infty$ or ∞.

X is *integrable* if $E(X^+) < \infty$ and $E(X^-) < \infty$ (both),

$$\begin{aligned} E(X) &: = E(X^+) - E(X^-) \\ &\Longrightarrow \quad E(|X|) = E(X^+) + E(X^-) < \infty \end{aligned}$$

In this case, $E(X) < \infty$.

X is *not integrable* if $E(X^+) = \infty$ and $E(X^-) = \infty$ (both),

In this case, $E(X)$ *does not exist*

$L_1(P) = \{\text{rv } X : E(|X|) < \infty\}$

space of (absolutely) integrable functions

Example (heavy tails) (whiteboard)

Properties of expectations

1. If X is integrable, then $P[|X| = \infty] = 0$ (explain)

2. If $E(X)$ exists, then

$$E(cX) = cE(X)$$

If either

$$E(X^+) < \infty,\ E(Y^+) < \infty$$

or

$$E(X^-) < \infty,\ E(Y^-) < \infty$$

then

$$E(X + Y) = E(X) + E(Y)$$

Proof (whiteboard)

3.

$$X \leq Y \implies E(X) \leq E(Y)$$

4. *Generalization of MCT*

$$X_n \in L_1 \implies [\ X_n \uparrow\downarrow X \implies E(X_n) \uparrow\downarrow E(X)]$$

Proof.

Case $X_n \downarrow X$ (the other case in Resnick).

Observe

$$X_n^+ \downarrow X^+, \text{ so that } E(X^+) < \infty$$

Take

$$X_n^- = -X_n + X_n^+$$

$$0 \leq X_n^- = -X_n + X_n^+ \leq -X_n + X_1^+ \uparrow -X + X_1^+$$

therefore by the MCT

$$0 \leq E(-X_n + X_1^+) \uparrow E(-X + X_1^+)$$

By integrability of X_n and X_1^+ we have

$$E(-X_n + X_1^+) = E(-X_n) + E(X_1^+)$$

Also, $E(X^+) < \infty$ and $X_1 \in L_1$, so

$$E(-X + X_1^+) = E(-X) + E(X_1^+)$$

and by comparison of right-hand sides and taking limits

$$\lim_n E(X_n) = E(X)$$

□

Variance and covariance (self-study)

Modulus inequality

$$|E(X)| = \left|E(X^+) - E(X^-)\right| \leq E(X^+)+E(X^-) = E(X^+ + X^-) = E(|X|)$$

Markov inequality

$$P[|X| \geq \lambda] \leq \frac{1}{\lambda} E(|X|)$$

Proof

$$1_{[|X| \geq \lambda]}(\omega) \leq \frac{|X(\omega)|}{\lambda} 1_{[|X| \geq \lambda]}(\omega) \leq \frac{|X(\omega)|}{\lambda}$$

□

Chebyshov inequality

$$P[|Y - E(Y)| \geq \alpha] \leq \frac{1}{\alpha^2} V(Y)$$

is obtained from Markov inequality by taking

$$\begin{aligned} X &= [Y - E(Y)]^2 \\ \lambda &= \alpha^2 \end{aligned}$$

□

Weak law of large numbers (WLLN)

$$\begin{aligned} \{X_n, n &\geq& 1\} \text{ iid} : E(X_1) = \mu, V(X_1) < \infty \\ &\Longrightarrow& \lim_n P\left[\left|\frac{1}{n}\left(\sum_{i=1}^n X_i\right) - \mu\right| \geq \epsilon\right] = 0 \end{aligned}$$

Proof

$$\begin{aligned} E\left[\frac{1}{n}\left(\sum_{i=1}^n X_i\right)\right] &=& \frac{1}{n}\sum_{i=1}^n E\left(X_i\right) = \mu \\ &\Longrightarrow& P\left[\left|\frac{1}{n}\left(\sum_{i=1}^n X_i\right) - \mu\right| \geq \epsilon\right] \leq \frac{1}{\epsilon^2} V\left[\frac{1}{n}\left(\sum_{i=1}^n X_i\right)\right] \\ &=& \frac{1}{(n\epsilon)^2}\sum_{i=1}^n V\left(X_i\right) = \frac{1}{n\epsilon^2} V\left(X_1\right) \overset{n\longrightarrow\infty}{\longrightarrow} 0 \end{aligned}$$

□ WLLN and frequency-based definition of probability (explain)

Limits and Integrals

Thm 5.3.1. MCT, demonstrated before

$$0 \leq X_n \uparrow X,\ X_n, X \in \bar{\mathcal{E}}_+ \implies E(X_n) \uparrow E(X)$$

Special case, MCT for series, take

$$X_n = \sum_{i=1}^{n} \xi_i,\ X = \sum_{i=1}^{\infty} \xi_i,\ \xi_i \geq 0$$

which implies

$$\begin{aligned} E(X) &= E\left(\sum_{i=1}^{\infty} \xi_i\right) = E\left(\lim_n \uparrow \sum_{i=1}^{n} \xi_i\right) \stackrel{MCT}{=} \lim_n \uparrow E\left(\sum_{i=1}^{n} \xi_i\right) \\ &= \lim_n \uparrow \sum_{i=1}^{n} E\left(\xi_i\right) = \sum_{i=1}^{\infty} E\left(\xi_i\right) \end{aligned}$$

Fatou Lemmas and Bounded Convergence Theorem

"Old" Fatou Lemmas for probabilities

$$\begin{aligned} P(\liminf_{n\to\infty} A_n) &\le \liminf_{n\to\infty} P(A_n) \\ &\le \limsup_{n\to\infty} P(A_n) \le P(\limsup_{n\to\infty} A_n) \end{aligned}$$

"New" Fatous Lemmas are more general$\{X_n, n \ge 1\}$

$$\begin{aligned} \{X_n, n &\ge 1\},\ X_n \ge Z,\ Z \in L_1(P) \\ &\Longrightarrow E(\liminf_{n\to\infty} X_n) \le \liminf_{n\to\infty} E(X_n) \end{aligned}$$

If we take

$$X_n(\omega) = 1_{A_n}(\omega)$$

then

$$\begin{aligned} P(\lim_{n\to\infty}\inf A_n) &= E(1_{\lim_{n\to\infty}\inf A_n}) = E(\lim_{n\to\infty}\inf 1_{A_n}) \\ &= E(\lim_{n\to\infty}\inf X_n) \leq \lim_{n\to\infty}\inf E(X_n) \\ &= \lim_{n\to\infty}\inf E(1_{A_n}) = \lim_{n\to\infty}\inf P(A_n) \end{aligned}$$

so the “new” has the “old” as a special case!

Proof (of the "new"): Enough to prove it for $Z = 0$ (otherwise, take $X_n - Z$ instead of X_n)

$$\begin{aligned} E(\liminf_{n\to\infty} X_n) &= E[\lim_{n\to\infty} \uparrow (\inf_{k\geq n} X_k)] \\ \stackrel{MCT}{=} \lim_{n\to\infty} &\uparrow E[(\inf_{k\geq n} X_k)] \leq \lim_{n\to\infty} \uparrow [\inf_{k\geq n} E(X_k)] \text{ (explain)} \\ &= \liminf_{n\to\infty} E(X_n) \end{aligned}$$

$\square$

By taking negatives, we prove a dual version

$$\begin{aligned} \{X_n, n &\geq 1\},\ X_n \leq Z,\ Z \in L_1(P) \\ &\Longrightarrow E(\lim_{n\to\infty} \sup X_n) \geq \lim_{n\to\infty} \sup E(X_n) \end{aligned}$$

Finally, as a consequence, if

$$\begin{aligned}\{X_n, n &\geq 1\},\ Y \leq X_n \leq Z,\ Z \in L_1(P) \\ &\Longrightarrow E(\lim_{n\to\infty}\inf X_n) \leq \lim_{n\to\infty}\inf E(X_n) \leq \lim_{n\to\infty}\sup E(X_n) \leq E(\lim_{n\to\infty}\sup X_n)\end{aligned}$$

which yields the

Dominated Convergence Theorem (DCT)

$$\begin{aligned}X_n &\longrightarrow X,\ \exists Z \in L_1(P) : |X_n| \leq Z \\ &\Longrightarrow E(X_n) \longrightarrow E(X) \text{ and } E(|X_n - X|) \longrightarrow 0\end{aligned}$$

Counterexample, take $X_n(\omega) = 2^{2n} 1_{(0,2^{-n})}(\omega)$ (explain)

Proof of the DCT

$$|X_n| \leq Z \iff -Z \leq X_n \leq Z$$

Also, since $X_n \longrightarrow X$

$$\lim_{n\to\infty} \inf X_n = \lim_{n\to\infty} \sup X_n = \lim_{n\to\infty} X_n = X$$

So

$$\begin{aligned} & E(X) \\ = \; & E(\lim_{n\to\infty} \inf X_n) \leq \lim_{n\to\infty} \inf E(X_n) \leq \lim_{n\to\infty} \sup E(X_n) \leq E(\lim_{n\to\infty} \sup X_n) \\ = \; & E(X) \end{aligned}$$

and by sandwich argument

$$\lim_{n\to\infty} \inf E(X_n) = \lim_{n\to\infty} \sup E(X_n) = \lim_{n\to\infty} E(X_n) = E(X)$$

The final part is a bit trickier. But we have

$$|X_n| \leq Z \text{ and } |X| \leq Z \implies |X_n - X| \leq 2Z$$

so that the first part implies

$$\lim_n E(|X_n - X|) = E(\lim_n |X_n - X|) = E(0) = 0$$

□

Indefinite Integrals

$$X : (\Omega, \mathcal{B}, P) \xrightarrow{meas} (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$
$$X \in L_1(P)$$

$$\int_A X(\omega)dP(\omega) := E(X1_A)$$

Properties, assuming $X \geq 0$ (and $E(X) < \infty$)

(1)

$$0 \leq \int_A X(\omega)dP(\omega) \leq E(X)$$

to see this, notice

$$0 \leq X(\omega)1_A(\omega) \leq X(\omega)$$

and use monotonicity of expectations (integrals)

(2)

$$\int_A X(\omega)dP(\omega) = 0 \iff P(A \cap [X > 0]) = 0$$

First of all

$$\int_{[X>n]} X(\omega)dP(\omega) \downarrow 0, \text{ as } n \longrightarrow \infty$$

Indeed, define

$$X_n = X1_{[X>n]}$$

It holds

$$[X > n] \downarrow \emptyset \iff X_n(\omega) = X1_{[X>n]}(\omega) \downarrow 0 \text{ pointwise}$$

But also

$$|X_n| \leq X \in L_1(P)$$

and the assertion follows by DCT.

Now, suppose $P(A_n) \longrightarrow 0$. We have

$$\begin{aligned}
\int_{A_n} X(\omega)dP(\omega) &= \int_{[X\leq M]\cap A_n} X(\omega)dP(\omega) + \int_{[X>M]\cap A_n} X(\omega)dP(\omega) \\
&\leq M \int 1_{A_n}(\omega)dP(\omega) + \int_{[X>M]} X(\omega)dP(\omega) \\
&= MP(A_n) + \int_{[X>M]} X(\omega)dP(\omega) \\
&\stackrel{n\longrightarrow\infty}{\longrightarrow} 0 + \int_{[X>M]} X(\omega)dP(\omega)
\end{aligned}$$

But, for any $\varepsilon > 0$, there is M large enough such that

$$\int_{[X>M]} X(\omega)dP(\omega) < \varepsilon$$

so that $\int_{A_n} X(\omega)dP(\omega) \stackrel{n\longrightarrow\infty}{\longrightarrow} 0$.

Proof of property (2) $P(A \cap [X > 0]) = 0$ implies

$$\int_A X dP = \int_{A\cap[X>0]} X dP + \int_{A\cap[X=0]} X dP = 0 + \int_{A\cap[X=0]} 0 dP = 0$$

Suppose $\int_A X dP = 0$ and $P(A \cap [X > 0]) > 0$. If

$$\exists k : P(A_k) = P(A \cap [X > 1/k]) > 0$$

then

$$\int_A X dP \geq \frac{1}{k} P(A \cap [X > 1/k]) > 0$$

a contradiction.

Therefore,

$$\forall k : P(A_k) = 0$$

But, since $A_n \uparrow A$

$$\begin{aligned} \int_A X dP &= E(X1_A) \stackrel{DCT}{=} \lim_k E(X1_{A_k}) \\ &= \lim_k \int_{A_k} X dP = 0 \end{aligned}$$

□

Properties (3)-(5) are easier.

Transformation Theorem and Densitites

Setup

$$\begin{aligned} T &: (\Omega, \mathcal{B}, P) \xrightarrow{meas} (\Omega', \mathcal{B}') \\ P' &= P \circ T^{-1} \text{ is a probability measure on } \mathcal{B}' \\ P'(A') &= P[T^{-1}(A')],\ A' \in \mathcal{B}' \end{aligned}$$

We can define two related random variables

$$\begin{aligned} X' &: (\Omega', \mathcal{B}', P') \xrightarrow{meas} (\mathbb{R}, \mathcal{B}(\mathbb{R})) \\ X' \circ T &: (\Omega, \mathcal{B}, P) \xrightarrow{meas} (\mathbb{R}, \mathcal{B}(\mathbb{R})) \end{aligned}$$

Theorem 5.5.1

(i) If $X' \geq 0$

$$\int_{\Omega} X'[T(\omega)]P(d\omega) = \int_{\Omega'} X'(\omega')P'(d\omega')$$

or

$$E(X' \circ T) = E'(X')$$

(ii)

$$X' \in L_1(P') \iff X' \circ T \in L_1(P)$$

then

$$\int_{T^{-1}(A')} X'[T(\omega)]P(d\omega) = \int_{A'} X'(\omega')P'(d\omega')$$

Proof by simple function approximation ... (try to verify for X' being a simple function ?)

A reminder

$$X : (\Omega, \mathcal{B}, P) \stackrel{meas}{\longrightarrow} (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

Distribution *measure* of X

$$\begin{aligned} F(A) &=: P[X^{-1}(A)] = P[\omega : X(\omega) \in A] \\ &= P'[X \in A],\ A \in \mathcal{B}(\mathbb{R}) \end{aligned}$$

$$F = P \circ X^{-1}$$

Distribution *function* of X

$$F(x) =: F((-\infty, x]) = P'[X \leq x]$$

Special case of transformation

$$\begin{aligned}(\Omega', \mathcal{B}') &\longmapsto (\mathbb{R}, \mathcal{B}(\mathbb{R})),\ \omega' \longmapsto x \\ X'(\omega') &\longmapsto X'(x) =: x \\ T(\omega) &\longmapsto X(\omega) \\ P'(A') &\longmapsto P \circ X^{-1}(A') =: F(A')\end{aligned}$$

drawing (...)

We have $E(X' \circ T) = E'(X')$ based on Thm 5.5.1, which can be written as

$$\int_\Omega \underbrace{X'[T(\omega)]}_{T(\omega) \longmapsto X(\omega)} P(d\omega) = \int_{\underbrace{\Omega'}_{\mathbb{R}}} \underbrace{X'(\omega')}_{\omega' \longmapsto x} \underbrace{P'(d\omega')}_{F(dx)}$$

$$E(X) = \int_\Omega X(\omega) P(d\omega) = \int_{\mathbb{R}} x F(dx)$$

Also

$$E[g(x)] = \int_{\mathbb{R}} g(x) F(dx)$$

Densities

Define a random vector X

$$X : (\Omega, \mathcal{B}, P) \stackrel{meas}{\longrightarrow} (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$$

with distribution measure

$$F(A) = P \circ X^{-1}(A),\ A \in \mathcal{B}(\mathbb{R}^k)$$

and distribution function

$$F(x_1, x_2, \ldots, x_k) =: F((-\infty, x_1] \times (-\infty, x_2] \times \cdots \times (-\infty, x_k])$$

Generally

$$E[g(x)] = \int_{\mathbb{R}^k} g(x) F(dx), \quad g : (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k)) \stackrel{meas}{\longrightarrow} (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$$

Suppose F is AC, ie.

$$F(A) = \int_A \underbrace{f(x)}_{\text{distribution density}} dx$$

Proposition 5.5.2

If $F(A)$ is AC, then

$$E[g(x)] = \int_{\mathbb{R}^k} g(x) f(x) dx$$

To prove, start from indicators, proceed to simple functions and then use MCT.

Section 5.6. - Self-study

Riemann vs. Lebesgue Theorem 5.6.1 (w/o proof)

Integral comparison Lemma 5.6.1 (w/o proof)

Product Spaces

$$\Omega_1 \times \Omega_2 =: \{(\omega_1, \omega_2) : \omega_i \in \Omega_i, i = 1, 2\}$$

Examples

$$\begin{aligned} \mathbb{R}^2 &= \mathbb{R} \times \mathbb{R} \\ & (-\infty, 0] \times (0, \infty) \end{aligned}$$

Coordinate (projection) maps

$$\pi_i(\omega_1, \omega_2) =: \omega_i$$

Section of A at ω_1

$$A \in \Omega_1 \times \Omega_2$$

$$A_{\omega_1} = \{\omega_2 : (\omega_1, \omega_2) \in A\} \subset \Omega_2$$

Properties of sections

(i)

$$(A^c)_{\omega_i} = (A_{\omega_i})^c$$

(ii)

$$\begin{aligned} (\bigcup_\alpha A_\alpha)_{\omega_i} &= \bigcup_\alpha (A_\alpha)_{\omega_i} \\ (\bigcap_\alpha A_\alpha)_{\omega_i} &= \bigcap_\alpha (A_\alpha)_{\omega_i} \end{aligned}$$

Sections of functions

$$X : \Omega_1 \times \Omega_2 \longrightarrow S$$

$$X_{\omega_1} : \Omega_2 \longrightarrow S;\ X_{\omega_1}(\omega_2) =: X(\omega_1, \omega_2)$$

Properties (i)

$$(1_A)_{\omega_i} = (1)_{A_{\omega_i}}$$

(ii)

$$S = \mathbb{R}^k \Longrightarrow (X_1 + X_2)_{\omega_i} = (X_1)_{\omega_i} + (X_2)_{\omega_i}$$

(iii)

$$S \text{ metric space } \Longrightarrow \lim_n (X_n)_{\omega_i} = (\lim_n X_n)_{\omega_i}$$

Rectangles

$$\underbrace{A_1 \times A_2}_{\text{sides}} \subset \Omega_1 \times \Omega_2$$

If $(\Omega_i, \mathcal{B}_i)$ measurable and $A_i \in \mathcal{B}_i$, then $A_1 \times A_2$ called a *measurable rectangle*

Class $RECT$ is a semialgebra since

(i) $\emptyset, \Omega \in RECT$

(ii) $RECT$ is a π-system (prove)

(iii) If $A \in RECT$, then there exist (prove) finite n and disjoint $C_1, \ldots C_n \in RECT$ such that

$$A^c = \sum_{i=1}^{n} C_i$$

Product σ-algebra

$$\mathcal{B}_1 \times \mathcal{B}_2 =: \sigma(RECT)$$

Lemma 5.7.1. Sections of measurable sets are measurable (why ???)

$$A \in \mathcal{B}_1 \times \mathcal{B}_2 \Longrightarrow A_{\omega_1} \in \mathcal{B}_2$$

Corollary. Sections of measurable functions are measurable, since

$$X_{\omega_1}^{-1}(A') = [X^{-1}(A')]_{\omega_1}$$

Probabilities on product spaces

Transition functions

$$K(\omega_1, A_2) : \Omega_1 \times \mathcal{B}_2 \longrightarrow [0, 1]$$

If *(i)* $\forall \omega_1,\ K(\omega_1, \cdot)$ is a probability measure on $\mathcal{B}_2$

(ii) $\forall A_2,\ K(\cdot, A_2)$ is measurable $(\Omega_1, \mathcal{B}_1) \longrightarrow ([0,1], \mathcal{B}[0,1])$, ie. it is a rv "wrt ω_1"

And what is this?

$$E[1_{A_1}(\omega_1) K(\omega_1, A_2)] = \int_{A_1} K(\omega_1, A_2) \qquad \underbrace{P_1(d\omega_1)}_{\text{Prob measure on } (\Omega_1, \mathcal{B}_1)}$$

Theorem 3.8.1.

$$\begin{aligned} P(A_1 \times A_2) &= \int_{A_1} K(\omega_1, A_2) P_1(d\omega_1) \\ A_1 \times A_2 &\in RECT \end{aligned}$$

uniquely defines a probability measure on

$$(\Omega_1 \times \Omega_2, \sigma(RECT))$$

Follows instantly from the Extension Theorem, since $RECT$ is a semi-algebra.

Product measures

Suppose that

$$\forall \omega_1,\ K(\omega_1, A_2) = P_2(A_2)$$

Then

$$\begin{aligned} P(A_1 \times A_2) &= \int_{A_1} K(\omega_1, A_2) P_1(d\omega_1) \\ &= \int_{A_1} P_2(A_2) P_1(d\omega_1) \\ &= \int_{A_1} P_1(d\omega_1) \cdot P_2(A_2) = P_1(A_1) P_2(A_2) \\ &=: (P_1 \times P_2)(A_1 \times A_2) \end{aligned}$$

Product measures yield independent σ-fields!

Define σ-fields

$$\begin{aligned} \mathcal{B}_1^\# &= \{A_1 \times \Omega_2 : A_1 \in \mathcal{B}_1\} \\ \mathcal{B}_2^\# &= \{\Omega_1 \times A_2 : A_2 \in \mathcal{B}_2\} \end{aligned}$$

$\mathcal{B}_1^\# \underline{\|} \mathcal{B}_2^\#$ wrt $P_1 \times P_2$, since

$$(P_1 \times P_2)[\underbrace{(A_1 \times \Omega_2) \cap (\Omega_1 \times A_2)}_{=A_1 \times A_2}] = (P_1 \times P_2)(A_1 \times A_2)$$

$$= P_1(A_1)P_2(A_2) = [P_1(A_1)P_2(\Omega_2)][P_1(\Omega_1)P_2(A_2)]$$

$$= (P_1 \times P_2)(A_1 \times \Omega_2) \cdot (P_1 \times P_2)(\Omega_1 \times A_2)$$

Construction of rv's independent wrt product measures

Framework

$$(\Omega_1 \times \Omega_2, \mathcal{B}_1 \times \mathcal{B}_2 \equiv \sigma(RECT), P_1 \times P_2)$$

Define "univariate" rv's

$$X_i : (\Omega_i, \mathcal{B}_i) \xrightarrow{meas} (\mathbb{R}, \mathcal{B}(\mathbb{R})),\ i = 1, 2$$

and their "trivial bivariate extensions"

$$\begin{aligned} X_1^{\#}(\omega_1, \omega_2) &= X_1(\omega_1),\ \text{"independent" of } \omega_2 \\ X_2^{\#}(\omega_1, \omega_2) &= X_2(\omega_2),\ \text{"independent" of } \omega_1 \end{aligned}$$

$X_1^\# \perp\!\!\!\perp X_2^\#$ wrt $P = P_1 \times P_2$, since

$$
\begin{aligned}
P(X_1^\# &\le x,\ X_2^\# \le y) = (P_1 \times P_2)[(\omega_1,\omega_2) : X_1(\omega_1) \le x,\ X_2(\omega_2) \le y] \\
&= (P_1 \times P_2)\{[\omega_1 : X_1(\omega_1) \le x] \times [\omega_2 : X_2(\omega_2) \le y]\} \\
&= P_1[\omega_1 : X_1(\omega_1) \le x] \cdot P_2[\omega_2 : X_2(\omega_2) \le y] \\
&= P_1[\omega_1 : X_1(\omega_1) \le x] P_2[\Omega_2] \cdot P_1[\Omega_1] P_2[\omega_2 : X_2(\omega_2) \le y] \\
&= P[(\omega_1,\omega_2) : X_1^\#(\omega_1,\omega_2) \le x] \cdot P[(\omega_1,\omega_2) : X_2^\#(\omega_1,\omega_2) \le y] \\
&= P[X_1^\# \le x] \cdot P[X_2^\# \le y]
\end{aligned}
$$

Fubini Theorem 5.9.2

Assume

$$
\begin{aligned}
P &= P_1 \times P_2 \\
X &: (\Omega \equiv \Omega_1 \times \Omega_2, \mathcal{B}_1 \times \mathcal{B}_2 \equiv \sigma(RECT), P_1 \times P_2) \xrightarrow{meas} (\mathbb{R}, \mathcal{B}(\mathbb{R})) \\
X &\geq 0 \text{ or } X \in L_1(P)
\end{aligned}
$$

Then

$$
\begin{aligned}
\int_{\Omega_1 \times \Omega_2} X(\omega) P(d\omega) &= \int_{\Omega_1} [\int_{\Omega_2} X_{\omega_1}(\omega_2) P_2(d\omega_2)] P_1(d\omega_1) \\
&= \int_{\Omega_2} [\int_{\Omega_1} X_{\omega_2}(\omega_1) P_1(d\omega_1)] P_2(d\omega_2)
\end{aligned}
$$

and, for the case of $X \geq 0$, all three integrals are finite/infinite at the same time.

Example 5.9.2

$X_1, X_2 \in L^2(P)$ and $X_1 \perp\!\!\!\perp X_2 \Longrightarrow E(X_1 X_2) = E(X_1)E(X_2)$

Proof. Denote $X = (X_1, X_2)$ and $g(x_1, x_2) = x_1 x_2$

First show that the "joint" distribution of $X = (X_1, X_2)$ is a product measure. Indeed define

$$F = P \circ X^{-1} \;\; (= \text{“old” } P', \text{ remember? })$$

We have

$$\begin{aligned} F(A_1 \times A_2) &= P[X^{-1}(A_1 \times A_2)] = P[\omega : (X_1(\omega),\ X_2(\omega)) \in A_1 \times A_2] \\ &= P[\omega : X_1(\omega) \in A_1,\ X_2(\omega) \in A_2] \\ &= P[\omega : X_1(\omega) \in A_1] P[\omega : X_2(\omega) \in A_2] = F_1(A_1)F(A_2) \\ &= (F_1 \times F_2)(A_1 \times A_2) \end{aligned}$$

where step 3 follows from independence of the two rv's.

Now

$$\begin{aligned}
E(X_1X_2) &= E(g(X)) = \int_{\mathbb{R}^2} g(x)F(dx) \\
&= \int_{\mathbb{R}} [\int_{\mathbb{R}} x_1x_2F_1(dx_1)]F_2(dx_2) \\
&= \int_{\mathbb{R}} x_2[\int_{\mathbb{R}} x_1F_1(dx_1)]F_2(dx_2) \\
&= [\int_{\mathbb{R}} x_1F_1(dx_1)] \int_{\mathbb{R}} x_2F_2(dx_2) \\
&= E(X_1)E(X_2)
\end{aligned}$$

Section 5.9 Fubini

Theorem 5.9.2 (w/o proof)

Examples 5.9.1-3