

A field on a set Ω is a collection of subsets \mathcal{F} such that:

- 1. (at least contains two sets) \emptyset , $\Omega \in \mathcal{F}$,
- 2. (closer under complement) if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$,
- 3. (closure under finite union) if $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$
- A σ field \mathcal{F} on the set Ω also has:
 - 4. (closure under countable union) if $A_1, \dots \in \mathcal{F}$, then $\bigcup_i A_i \in \mathcal{F}$.

A probability measure on a set Ω with field \mathcal{F} is a function $P: \mathcal{F} \to [0, \infty)$ with:

- 1. $0 \le P(A) \le 1, \forall A \in \mathcal{F},$
- 2. $P(\emptyset) = 0$ and $P(\Omega) = 1$,
- 3. if A_1, \ldots are disjoint and $\bigcup_i A_i \in \mathcal{F}$, then

$$P(\cup_i A_i) = \sum_i P(A_i)$$

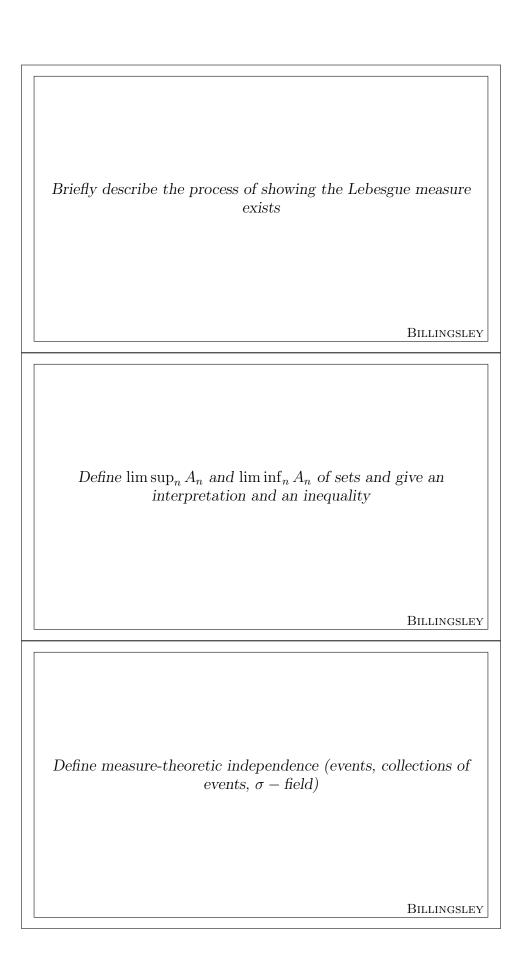
A probability measure P on Ω with field \mathcal{F} has

- 1. (monotonicity), if $A \subset B$, then P(A) < P(B),
- 2. (inclusion-exclusion) $P(A \cup B) = P(A) + P(B) P(A \cap B)$ and more generally

$$P(\cup A_n) = \sum_i P(A_i) - \sum_{i < j} P(A_i \cap A_j) + \sum_{i < j < k} P(A_i \cap A_j \cap A_k)$$

$$+\cdots+(-1)^{n+1}P(A_1\cap\ldots A_n),$$

- 3. (countably subadditive) if $A_1, \dots \in \mathcal{F}$ and $\bigcup_i A_i \in \mathcal{F}$, then $P(\bigcup_i A_i) \leq \sum_i P(A_i)$,
- 4. (continuous from below) if $A_1 \subset A_2 \cdots \subset A$, then $P(A_n) \uparrow P(A)$
- 5. (continuous from above) if $A_1 \supset A_2 \cdots \supset A$, then $P(A_n) \downarrow P(A)$



Theorem 3.1 (Cartheodory extension theorem): a probability measure on a field can be uniquely extended to the generated σ – field if the measure is σ -finite.

Hence, to construct the Lebesgue measure, first we define the Lebesgue measure that assigns to half-open intervals the interval length, second we verify that this is a well-defined measure on the Borel field, and then we apply theorem 3.1.

Proving theorem 3.1 is involved. Also, studying $\sigma(\mathcal{B}_0)$ is necessary, where \mathcal{B}_0 is the field of finite unions and intersections of intervals.

 $\limsup_n A_n = \bigcup_n \cap_{k \geq n} A_k$. If w is in LHS, then for every n, there exists some $k \geq n$ so that $w \in A_k$, hence w is in infinitely many of the A_n . "Infinitely often".

 $\liminf_n A_n = \cap_n \cup_{k \geq n} A_k$. If w is in LHS, then there exists n such that for all $k \geq n$, $w \in A_k$ for all k. Hence, w is in all but finitely many A_n . "Eventually".

$$P(\liminf_{n} A_n) \le \liminf_{n} P(A_n)$$

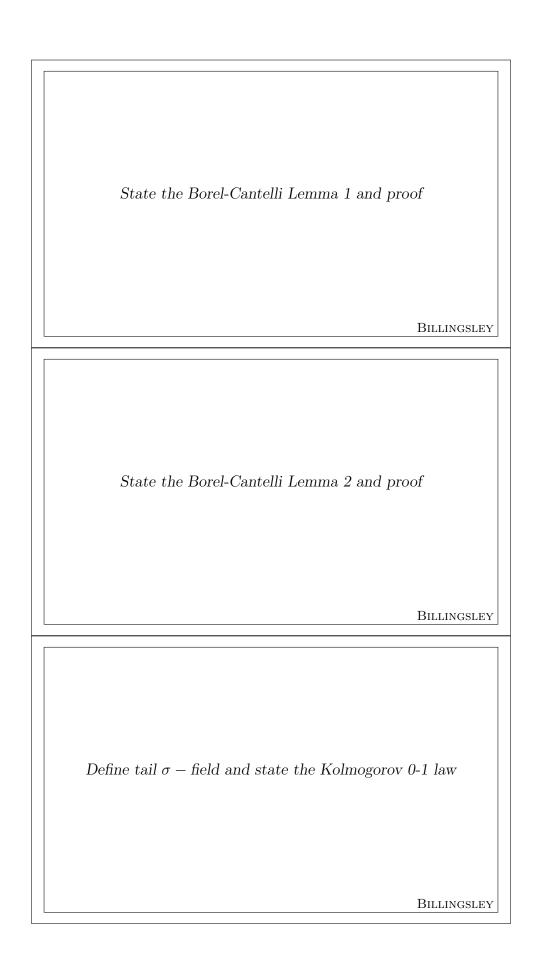
$$\le \lim \sup_{n} P(A_n) \le P(\limsup_{n} A_n)$$

- Two events A and B are independent if $P(A \cap B) = P(A)P(B)$.
- A collection of events $\{A_1, \ldots, A_n\}$ are independent if

$$P(A_{k_1} \cap \dots A_{k_i}) = P(A_{k_1}) \dots P(A_{k_i})$$

for all $2 \le j \le n$ and $1 \le k_1 < \dots < k_n \le n$.

- A collection of classes A_1, \ldots, A_n in a σ field \mathcal{F} are independent if for each choice of $A_i \in \mathcal{A}_i$, the collection $\{A_n\}$ is independent.
- Two σ fields \mathcal{A} and \mathcal{B} are independent if for every $A \in \mathcal{A}$ and $B \in \mathcal{B}$, we have $\mu(A \cap B) = \mu(A)\mu(B)$.



If $\sum P(A_n) < \infty$, then $P(\limsup_n A_n) = 0$.

Proof: Observe that $\limsup_n A_n \subset \bigcup_{k \geq m} A_k$ for all m. This implies that

$$P(\limsup_{n} A_n) \le P(\bigcup_{k \ge m} A_k) \le \sum_{k \ge m} P(A_k).$$

Since this holds for arbitrary m and the right hand side sum goes to 0 if the infinite sum converges, the lemma follows.

If $\{A_n\}$ are independent and $\sum P(A_n) = \infty$ then $P(\limsup_n A_n) = 1$. Proof: It is enough to show that $P(\bigcup_n \cap_{k \geq n} A_k^c) = 0$ for which it is enough to show that $P(\cap_{k \geq n} A_k^c) = 0$ for all k. Note that $1 - x \leq e^{-x}$, then (by independence)

$$P(\bigcap_{k=n}^{n+j} A_k^c) = \prod_{k=n}^{n+j} 1 - P(A_k) \le \exp\{-\sum_{k=n}^{n+j} P(A_k)\}.$$

But since the sum diverges, as $j \to \infty$, the RHS goes to 0, hence

$$P(\cap_{k=n}^{\infty}A_k^c)=\lim_j P(\cap_{k=n}^{n+j}A_k^c)=0$$

Given a sequence of events A_1, A_2, \ldots in a probability space (Ω, \mathcal{F}, P) , the tail σ – field associated with the sequence $\{A_n\}$ is

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(A_n, A_{n+1}, \dots).$$

The $\limsup_n A_n$ and $\liminf_n A_n$ are events in the tail σ – field.

The Kolmogorov zero-one law: if A_1, A_2, \ldots are independent, then for each event A in the tail σ – field, P(A) is either 0 or 1.

Simple random variables	
	Billingsley
State and prove the Markov's inequality	
	Billingsley
State and prove Chebyshev's inequality	
	Billingsley

A random variables X on (Ω, \mathcal{F}) is simple iff it can be written as

$$X(w) = \sum_{i} x_i I_{A_i}$$

for some finite set of x_i and $A_i \in \mathcal{F}$.

Simple random variables X_n converge to X with probability 1 ($\lim_n X_n = X$) iff $\forall \epsilon > 0$,

$$P(|X_n - X| > \epsilon \text{ i.o.}) = 0$$

which, if the above holds, implies that

$$\lim_{n} P(|X_n - X| > \epsilon) = 0.$$

Note that

$$\{\lim_n X_n = X\}^c = \cup_{\epsilon} \{|X_n - X| \ge \epsilon \text{ i.o.}\} = \cup_{\epsilon} \cup_n \cap_{k \ge n} \{|X_n - X| \ge \epsilon\}.$$

Markov's inequality: For a random variable X, nonnegative, then for positive α , we have

$$P(X \ge \alpha) \le \frac{1}{\alpha} \mathbb{E}[X].$$

Proof: Note that for any convex f and any set A, we have that

$$\min_{x \in A} f(x) \mathbf{1}_A \le E[X \mathbf{1}_A] \le E[X]$$

Hence, with f(x) = x and $A = [\alpha, \infty)$, the result follows. If we use $f(x) = |x|^k$, then we have for positive α :

$$\Pr(|X| \ge \alpha) \le \frac{1}{\alpha}^k \mathbb{E}[|X|^k]$$

Chebyshev's inequality: for a random variable X, we have

$$\Pr(|X - m| \ge \alpha) \le \frac{1}{\alpha} \operatorname{Var}(X)$$

Proof: Applying Markov's inequality with k=2 and subtracting $m=\mathbb{E}[X]$, we obtain the desired result.

State and prove Jensen's inequality (finite case)
BILLINGSLEY
State and prove Holder's inequality Billingsley
DILLINGSLEY
State and prove the strong law of large numbers Billingsley

Jensen's inequality says that for a convex function $\phi(x)$ and a random variable X, we have

$$\mathbb{E}[\phi(X)] \ge \phi(\mathbb{E}[X])$$

Proof: the proof follows by induction (base case follows form convexity; induction step follows from grouping n of the sum terms together). More details here.

Suppose that $\frac{1}{p} + \frac{1}{q} = 1$ for p, q > 1. Then:

$$\mathbb{E}[|XY|] \le \mathbb{E}[|X|^p]^{\frac{1}{p}} \mathbb{E}[|Y|^q]^{\frac{1}{q}}$$

Proof: Young's. Here.

If X_n are iid and $\mathbb{E}[X_n] = m$, then

$$\Pr\Bigl(\lim_n n^{-1} S_n = m\Bigr) = 1$$

Proof: WLOG m=0. It is enough to show that $\Pr(|n^{-1}S_n| \ge \epsilon \text{ i.o}) = 0$ for each ϵ .

Let $\mathbb{E}[X_i^2] = \sigma^2$ and $\mathbb{E}[X_i^4] = \xi^4$. By independence, we have

$$\mathbb{E}[S_n^4] = n\xi^4 + 3n(n-1)\sigma^4 \le Kn^2$$

where K does not depend on n. By Markov's inequality for k = 4,

$$\Pr(|S_n| \ge n\epsilon) \le K n^{-2} \epsilon^{-4},$$

so the result follows by the first Borel-Cantelli lemma (the event probs are summable, hence the lim sup is 0).

State and prove the weak law of large numbers
BILLINGSLEY

If X_n are iid and $\mathbb{E}[X_n] = m$, then for all ϵ

$$\lim_{n} \Pr(|n^{-1}S_n - m| \ge \epsilon) = 0.$$

Proof: By appealing to the strong law, we have

$$\Pr(|n^{-1}S_n - m| \ge \epsilon) \le \frac{\operatorname{Var}(S_n)}{n^2 \epsilon^2} = \frac{n\operatorname{Var}(X_1)}{n^2 \epsilon^2} \to 0.$$