Define a field and a $\sigma-$ field

BILLINGSLEY

Define a probability measure and give some of its properties

BILLINGSLEY

List some properties of probability measures

BILLINGSLEY

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 $\sigma -$ field $\mathcal{F}$ on the set Ω also has:

4. (closure under countable union) if $A_1, \cdots \in \mathcal{F}$, then $\cup_i A_i \in \mathcal{F}$.

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

1. $0 \leq P(A) \leq 1, \forall A \in \mathcal{F}$,
2. $P(\emptyset) = 0$ and $P(\Omega) = 1$,
3. if $A_1, \ldots$ are disjoint and $\cup_i A_i \in \mathcal{F}$, 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 $P(A) \leq P(B)$,
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 $A_1, \cdots \in \mathcal{F}$ and $\cup_i A_i \in \mathcal{F}$, then $P(\cup_i A_i) \leq \sum_i P(A_i)$,
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 $P(A_n) \downarrow P(A)$

Briefly describe the process of showing the Lebesgue measure exists

Billingsley

Define $\limsup_n A_n$ *and* $\liminf_n A_n$ *of sets and give an interpretation and an inequality*

Billingsley

Define measure-theoretic independence (events, collections of events, $\sigma-$ *field)*

Billingsley

Theorem 3.1 (Cartheodory extension theorem): a probability measure on a field can be uniquely extended to the generated $\sigma -$ 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.

$\limsup_n A_n = \cup_n \cap_{k\geq n} A_k$. 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".

$$P(\liminf_n A_n) \leq \liminf_n P(A_n)$$
$$\leq \limsup_n P(A_n) \leq P(\limsup_n A_n)$$

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

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

  for all $2 \leq j \leq n$ and $1 \leq k_1 < \dots < k_n \leq n$.
- A collection of classes $\mathcal{A}_1, \dots, \mathcal{A}_n$ in a $\sigma -$ field $\mathcal{F}$ are independent if for each choice of $A_i \in \mathcal{A}_i$, the collection $\{A_n\}$ is independent.
- Two $\sigma -$ fields $\mathcal{A}$ and $\mathcal{B}$ are independent if for every $A \in \mathcal{A}$ and $B \in \mathcal{B}$, we have $\mu(A \cap B) = \mu(A)\mu(B)$.

State the Borel-Cantelli Lemma 1 and proof

Billingsley

State the Borel-Cantelli Lemma 2 and proof

Billingsley

Define tail $\sigma -$ field and state the Kolmogorov 0-1 law

Billingsley

If $\sum P(A_n) < \infty$, then $P(\limsup_n A_n) = 0$.
Proof: Observe that $\limsup_n A_n \subset \cup_{k\geq m} A_k$ for all m. This implies that

$$P(\limsup_n A_n) \leq P(\cup_{k\geq m} A_k) \leq \sum_{k\geq m} P(A_k).$$

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 $P(\cup_n \cap_{k\geq n} A_k^c) = 0$ 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)

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

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 $(\Omega, \mathcal{F}, P)$, the tail $\sigma-$field associated with the sequence $\{A_n\}$ is

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

The $\limsup_n A_n$ and $\liminf_n A_n$ are events in the tail $\sigma-$field.
The Kolmogorov zero-one law: if $A_1, A_2, \ldots$ are independent, then for each event A in the tail $\sigma-$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 $(\Omega, \mathcal{F})$ 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

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

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

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

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

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

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:

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

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

$$\Pr(|X - m| \geq \alpha) \leq \frac{1}{\alpha}^2 \operatorname{Var}(X)$$

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

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

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

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:

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

Proof: Young's. Here.

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

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

Proof: WLOG $m = 0$. It is enough to show that $\Pr\big(|n^{-1}S_n| \geq \epsilon \text{ i.o}\big) = 0$ for each ϵ.
Let $\mathbb{E}[X_i^2] = \sigma^2$ and $\mathbb{E}[X_i^4] = \xi^4$. By independence, we have

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

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

$$\Pr(|S_n| \geq n\epsilon) \leq Kn^{-2}\epsilon^{-4},$$

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

What is a measurable mapping? What are some properties?

Billingsley

What is the pushforward measure?

Billingsley

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

$$\lim_n \Pr\big(|n^{-1}S_n - m| \geq \epsilon\big) = 0.$$

Proof: By appealing to the strong law, we have

$$\Pr\big(|n^{-1}S_n - m| \geq \epsilon\big) \leq \frac{\mathrm{Var}(S_n)}{n^2\epsilon^2} = \frac{n\mathrm{Var}(X_1)}{n^2\epsilon^2} \to 0.$$

For two measure spaces $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$, a transformation $T : \Omega \to \Omega'$ is measurable $\mathcal{F}/\mathcal{F}'$ iff for all $A \in \mathcal{F}'$, $T^{-1}(A) \in \mathcal{F}$.
If $T^{-1}(A) \in \mathcal{F}$ for each $A \in \mathcal{A}$, where $\mathcal{A}$ generates $\mathcal{F}'$, then T is $\mathcal{F}/\mathcal{F}'$ measurable.
A random vector is measurable iff each component function is measurable.
Continuous functions are measurable. If $f_k : \Omega \to \mathbb{R}$ are measurable $\mathcal{F}$, then $g(f_1(w), \ldots, f_k(w))$ is measurable $\mathcal{F}$ if $g : \mathbb{R}^k \to \mathbb{R}$ is measurable.
Composition of measurable functions is measurable. Sum, sup, lim sup, product are measure-preserving. A limit of measurable functions is measurable if the limit exists everywhere. We can construct a sequence of simple measurable functions that increase to any given measurable function.

Given $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ and a measurable transformation $T : \Omega \to \Omega'$ and a measure μ on $\mathcal{F}$, then $\mu T^{-1}(A') = \mu(T^{-1}(A'))$ is a pushforward measure on $\mathcal{F}'$.

What is the change-of-variables formula?

Billingsley

Say a few things about distribution functions.

Billingsley

How is the integral defined for a non-negative measurable function? Properties?

Billingsley

A measurable function g on Ω' is integrable with respect to the pushforward measure $\mu T^{-1} = T(\mu)$ iff the composition $g \circ T$ is integrable with respect to the measure μ. In that case,

$$\int_{\Omega'} g d(\mu T^{-1}) = \int_{\Omega} g \circ T d\mu$$

A distribution function for a random variable X on $\mathbb{R}$ is $F(x) = \Pr(X \leq x)$. It is non-decreasing, right-continuous (by continuity from above). By continuity from below, $\lim_{y \uparrow x} F(y) = F(x^-) = \Pr(X < x)$.
For every non-decreasing, right-continuous function with $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$, there exists on some probability space a random variable X for F is the distribution function.
If $\lim_n F_n(x) = F(x)$ for all x, then we write $F_n \implies F$ and say that the distributions converge weakly and their corresponding random variables converge weakly.

$\int_\Omega f d\mu = \sup \sum_i (\inf_{w \in A_i} f(w)) \mu(A_i)$ where the sup is taken over all partitions of Ω.
If $f \leq g$ then $\int f \leq \int g$.
If $f_n \uparrow f$ then $\int f_n \uparrow \int f$.
The integral is linear.
If $f = 0$ a.e., then $\int f = 0$. If the measure of the set where f is non-zero is positive, then the integral is positive. If the integral exists, then $f < \infty$ a.e.

Billingsley