

Lecture 2 : Random variables and their distributions

STAT205 *Lecturer: Jim Pitman*

Scribe: *Saurabh Amin* <amins@berkeley.edu>

2.1 Random variables

Let $(\Omega, \mathcal{F})$ and $(S, \mathcal{S})$ be two measurable spaces. A map $X : \Omega \rightarrow S$ is *measurable* or a *random variable* (denoted r.v.) if

$$X^{-1}(A) \equiv \{\omega : X(\omega) \in A\} \in \mathcal{F} \text{ for all } A \in \mathcal{S}$$

One can write $\{X \in A\}$ or $(X \in A)$ as a shorthand for $\{\omega : X(\omega) \in A\} = X^{-1}(A)$. If $(S, \mathcal{S}) = (\mathbb{R}^d, \mathcal{R}^d)$, then X is called a d -dimensional random vector. $\mathcal{R}$ is the Borel σ -field or the σ -field generated by the open subsets of $\mathbb{R}^n$.

An *indicator function* is a classic example of a r.v. where $S = \{0, 1\}$ and $\mathcal{S}$ is the collection of all subsets of S . The indicator function of a set $F \in \mathcal{F}$ is defined as

$$1_F(\omega) = \begin{cases} 1 & \text{if } \omega \in F \\ 0 & \text{if } \omega \notin F \end{cases}$$

If $S = \Omega$, then the identity map on Ω is a r.v. iff $\mathcal{S} \subset \mathcal{F}$.

Fact: The composition of two measurable maps is measurable.

2.2 Generation of σ -field

Let $\mathcal{A}$ be a collection of subsets of Ω . The σ -field generated by $\mathcal{A}$, denoted by $\sigma(\mathcal{A})$, is the smallest σ -field on Ω which contains $\mathcal{A}$.

Let $(X_i, i \in I)$ be a family of mappings of Ω into measurable spaces $(S_i, \mathcal{S}_i)$, $i \in I$. Here, $I \neq \emptyset$ is an arbitrary index set (i.e., possibly uncountable). The σ -field generated by $(X_i, i \in I)$, denoted by $\sigma(X_i, i \in I)$, is the smallest σ -field on Ω with respect to which each X_i is measurable. If we take $\mathcal{A} = \cup_i (X_i^{-1}(\mathcal{S}_i))$, this case reduces

to the previous one. In both the above cases, 'smallest' means the intersection of the collection of σ -fields with the given property.

We now introduce product spaces and product σ -fields. Given $(S, \mathcal{S})$ and index set I , let $\Omega = \prod_i (S_i) = \{(\omega_i, i \in I) : \omega_i \in S_i\}$, where each S_i is a copy of S . We have $\omega = (\omega_i \in I) \in \Omega$ and projection maps $X_i : \Omega \rightarrow S_i$ such that $X_i(\omega) = \omega_i$. The product σ -field $\mathcal{F}$ on Ω is the σ -field generated by the projections, i.e., $\mathcal{F} = \sigma((X_i^{-1}(F_i)) : F_i \in \mathcal{F}_i)$.

2.3 Checking measurability

Theorem 2.1 *Let $(\Omega, \mathcal{F})$ be a measurable space and $X : \Omega \rightarrow S$. If S has the σ -field $\sigma(\mathcal{A})$ for an arbitrary collection of sets $\mathcal{A}$, then X is measurable iff $(X \in \mathcal{A}) \in \mathcal{F}$ for $A \in \mathcal{A}$.*

Proof: We first prove the reverse direction. Since $\{X \in A\} = \{\omega : X(\omega) \in A\} = X^{-1}(A)$, we have

$$\begin{aligned} X^{-1}(A^c) &= (X^{-1}(A))^c \\ X^{-1}\left(\bigcup_i A_i\right) &= \bigcup_i X^{-1}(A_i) \\ X^{-1}\left(\bigcap_i A_i\right) &= \bigcap_i X^{-1}(A_i) \end{aligned}$$

Thus, $X^{-1}(\sigma(\mathcal{A})) = \sigma(X^{-1}(\mathcal{A}))$.

To prove the forward direction, note that the collection $\mathcal{C}$ of subsets of S given by $\mathcal{C} = \{B \subset S : X^{-1}(B) \in \mathcal{F}\}$ is a σ -field which contains $\mathcal{A}$ and hence $\sigma(\mathcal{A})$ which is the σ -field generated by $\mathcal{A}$. ■

Similarly, if S has the σ -field $\sigma(Y_i, i \in I)$, X is measurable iff each $Y_i \circ X$ is measurable.

2.4 Real and extended real random variables

Let S be a metric or topological space. The *Borel σ -field* on S , denoted by $\mathcal{B}(S)$, is the σ -field generated by open subsets of S . If $f : S \rightarrow T$ is a continuous function, then f is measurable from $(S, \mathcal{B}(S))$ to $(T, \mathcal{B}(T))$ by the previous theorem.

If $(S, \mathcal{S}) = (\mathbb{R}, \mathcal{R})$, then some possible choices of $\mathcal{A}$ are $\{(-\infty, x] : x \in \mathbb{R}\}$ or $\{(-\infty, x) : x \in \mathbb{Q}\}$ where $\mathbb{Q}$ = the rationals.

For the real line $\mathbb{R} = (-\infty, \infty)$ and extended real line $\bar{\mathbb{R}} = [-\infty, \infty]$, the Borel σ -fields can be defined as follows.

$$\begin{aligned}\mathcal{B}(\mathbb{R}) &= \sigma\{(-\infty, x], x \in \mathbb{R}\} \\ \mathcal{B}(\bar{\mathbb{R}}) &= \sigma\{[-\infty, x], x \in \bar{\mathbb{R}}\}\end{aligned}$$

Definition 2.2 (Real Random Variable) Let $(\Omega, \mathcal{F})$ be a measurable space. A real random variable (r.r.v.) is a measurable map from Ω to $\mathbb{R}$.

Thus a function X with range $\mathbb{R}$ is a r.v. iff $(X \leq x) \in \mathcal{F}$ for all $x \in \mathbb{R}$ (by theorem 2.1). Similarly, extended real random variables (e.r.r.v.) can be defined on range $\bar{\mathbb{R}}$.

Operations on real numbers are performed pointwise on real-valued functions, e.g.,

$$\begin{aligned}Z = X + Y &\text{ means } Z(\omega) = X(\omega) + Y(\omega) \text{ for all } \omega \in \Omega \\ \text{and } Z = \lim_n Z_n &\text{ means } Z(\omega) = \lim_n Z_n(\omega) \text{ for all } \omega \in \Omega\end{aligned}$$

Notation for real numbers: $x \vee y = \max(x, y)$, $x \wedge y = \min(x, y)$, $x^+ = x \vee 0$, $x^- = -(x \wedge 0)$. Note that $|x| = x^+ + x^-$ and $x = x^+ - x^-$.

Theorem 2.3 If $X_1, X_2, \dots$ are e.r.r.v.'s on $(\Omega, \mathcal{F})$, then they are closed under all limiting operations, i.e.,

$$\inf_n X_n, \sup_n X_n, \liminf_n X_n, \limsup_n X_n$$

are also e.r.r.v.

Proof: Since the infimum of a sequence is $< a$ iff some term is $< a$, we have

$$\left\{ \inf_n X_n < a \right\} = \bigcup_n \{X_n < a\} \in \mathcal{F}$$

The proof for supremum follows similarly.

For limit inferior of X_n , we have

$$\liminf_{n \rightarrow \infty} X_n := \sup_n \left\{ \inf_{m \geq n} X_m \right\} = \uparrow \lim_n \left\{ \inf_{m \geq n} X_m \right\}$$

Now note that $Y_n = \inf_{m \geq n} X_m$ is an e.r.r.v. for each n and so $\sup_n Y_n$ is also an e.r.r.v. The proof for limit superior follows similarly. ■

From the above proof we see that

$$\Omega_0 \equiv \left\{ \omega : \lim_{n \rightarrow \infty} X_n \text{ exists} \right\} = \left\{ \omega : \limsup_{n \rightarrow \infty} X_n - \liminf_{n \rightarrow \infty} X_n = 0 \right\}$$

is a measurable set. If $X_n(\omega)$ converges for all ω , i.e., $\mathbb{P}(\Omega_0) = 1$, we say that X_n converges almost surely to X which is also a e.r.r.v.

Definition 2.4 (Simple Random Variable) X is a simple random variable iff X is a finite linear combination of indicators, i.e., X can be expressed as $X(\omega) = \sum_{i=1}^n c_i 1_{A_i}(\omega)$ where $c_i \in \mathbb{R}$ and $A_i \in \mathcal{F}$. A simple r.v. can only take finitely many values.

Theorem 2.5 Every real r.v. X is a pointwise limit of a sequence of simple r.v.'s, which can be taken to be increasing if $X \geq 0$.

Proof: For $X \geq 0$ let,

$$X_n = \begin{cases} \frac{k-1}{2^n} & \text{on } \left\{ \frac{k-1}{2^n} \leq X < \frac{k}{2^n} \right\}, 0 \leq k \leq n2^n \\ n & \text{on } \{X \geq n\} \end{cases}$$

Then $X_n \uparrow X$. For general X use the decomposition $X = X^+ - X^-$. ■

Corollary 2.6 Let X and Y be real valued r.v.'s. Then so are XY , $X + Y$, $X - Y$, $\min(X, Y)$, $\max(X, Y)$.

Proof: Consider $X_n \uparrow X$ and $Y_n \uparrow Y$. This implies $X_n Y_n \uparrow XY$. Similarly, use the previous theorem to pass from simple case to the more general cases. ■

2.5 Probability distribution on the real line

If X is a real r.v. defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then X induces a probability measure on $\mathbb{R}$ called its *distribution*. Consider $\mu(B) = \mathbb{P}(X \in B) = \mathbb{P}(X^{-1}(B))$ as a function of Borel sets B of $\mathbb{R}$. To show that μ is a probability measure one can observe that for disjoint B_i 's,

$$\mu(\cup_i B_i) = \mathbb{P}(X^{-1}(\cup_i B_i)) = \mathbb{P}(\cup_i X^{-1}(B_i)) = \sum_i \mathbb{P}(X^{-1}(B_i)) = \sum_i \mu(B_i)$$

Commonly, the distribution of a r.v. X on reals is described by its *cumulative distribution function* (cdf), $F(x) = \mathbb{P}(X \leq x)$. In general, for a probability measure induced by r.v. X on $\mathbb{R}$, we can consider a point function $F(x) := F((-\infty, x])$ (by abuse of notation).

Theorem 2.7 *A cdf F of some probability measure on $\mathbb{R}$ has the following properties:*

1. F is an increasing function of x .
2. $\lim_{x \rightarrow \infty} F(x) = 1$ and $\lim_{x \rightarrow -\infty} F(x) = 0$
3. F is right continuous, i.e., $\lim_{y \downarrow x} F(y) = F(x)$

Proof: Refer to Theorem 1.1. in Durrett on page 4. ■

Theorem 2.8 *If F satisfies the properties of Theorem 2.7, then it is the distribution function of some r.v. and there is a unique probability measure on $(\mathbb{R}, \mathcal{R})$ that has $\mu((a, b]) = F(b) - F(a)$ for all a, b .*

Proof: Let $F : \mathbb{R} \rightarrow (0, 1)$ have properties 1, 2, 3 in Theorem 2.5. We construct a r.v. with distribution function F carried by $(\Omega, \mathcal{F}, \mathbb{P}) = ((0, 1), \mathcal{B}(0, 1), Leb)$. Define

$$\begin{aligned} X^+(\omega) &:= \inf\{z : F(z) > \omega\} = \sup\{y : F(y) \leq \omega\} \\ X^-(\omega) &:= \inf\{z : F(z) \geq \omega\} = \sup\{y : F(y) < \omega\} \end{aligned}$$

Figure 2.1 shows cases to consider carefully. We have $(\omega \leq F(c)) \Rightarrow (\omega : X^-(\omega) \leq c)$ by definition. Now, $(z > X^-(\omega)) \Rightarrow (F(z) \geq \omega)$, and so by right continuity of F , $(X^-(\omega) \leq c) \Rightarrow (\omega \leq F(X^-(\omega)) \leq F(c))$. Thus $(\omega \leq F(c)) \Leftrightarrow (X^-(\omega) \leq c)$ so that $\mathbb{P}(X^- \leq c) = F(c)$. The variable X^- therefore has distribution function F , and we call its probability law $\mathcal{L}$. Here, $\mathcal{L}$ is the unique probability measure on $(\mathbb{R}, \mathcal{B})$ such that $\mathcal{L}((-\infty, x]) = F(x), \forall x$.

Now, by definition of X^+ , $(\omega < F(c)) \Rightarrow (X^+(\omega) \leq c)$, So $F(c) \leq \mathbb{P}(X^+ \leq c)$. Since $X^- \leq X^+$, $\{X^- \neq X^+\} = \bigcup_{c \in \mathbb{Q}} \{X^- \leq c < X^+\}$ by the denseness of the rational numbers. However, for every $c \in \mathbb{R}$, $\mathbb{P}(X^- \leq c < X^+) = \mathbb{P}(\{X^- \leq c\} \setminus \{X^+ \leq c\}) \leq F(c) - F(c) = 0$. Since, $\mathbb{Q}$ is countable, the result follows.

Having proved the existence of a r.v. X with distribution function F , the uniqueness can be checked easily by the $\pi - \lambda$ theorem. ■

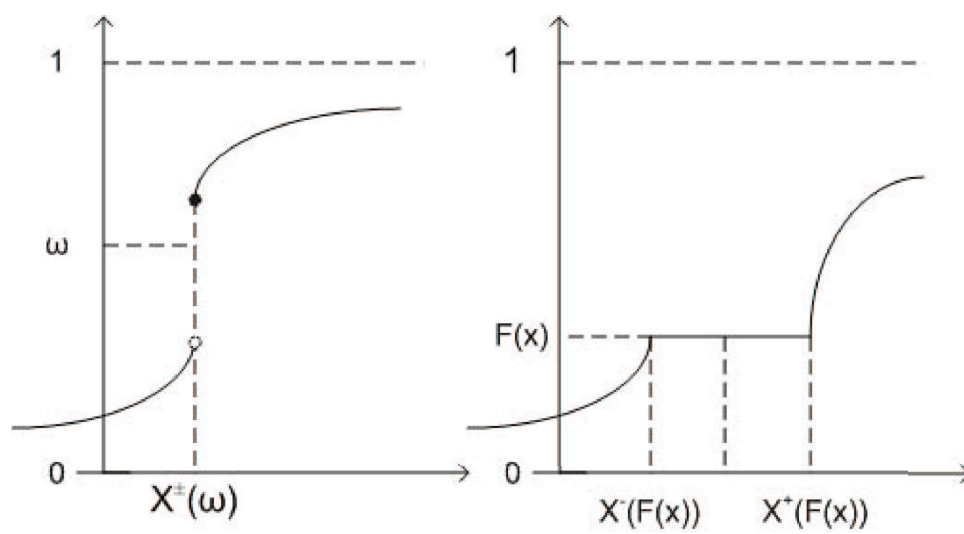


Figure 2.1: An illustration of some of the important cases to consider in the construction of X^- and X^+ .