STAT 810 Probability Theory I

Chapter 3: Random Variables, Elements, and Measurable Maps

Dr. Dewei Wang
Associate Professor
Department of Statistics
University of South Carolina
deweiwang@stat.sc.edu

Introduction

We will precisely define a random variable. A random variable is a real valued function with domain Ω which has an extra property called **measurability** that allows us to make probability statements about the random variable.

3.1 Inverse Maps

Suppose Ω and Ω' are two sets. Frequently, $\Omega' = \mathbb{R}$. Suppose

$$X : \Omega \mapsto \Omega'$$

meaning X is a function with domain Ω and range Ω'. Then X determines a function

$$X^{-1} : \mathcal{P}(\Omega') \to \mathcal{P}(\Omega)$$

defined by

$$X^{-1}(A') = \{\omega \in \Omega : X(\omega) \in A'\}$$

for $A' \subset \Omega'$.

For example, $\Omega = \{hh, ht, th, tt\}$ collects all possible results of flipping a coin twice, X denotes the number of heads which is a map from Ω to $\Omega' = \{0, 1, 2\}$, where $X(hh) = 2$, $X(ht) = X(th) = 1$, and $X(tt) = 0$.

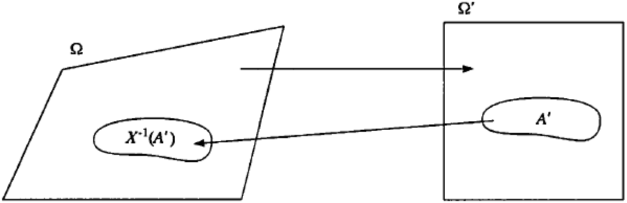

FIGURE 3.1 Inverses

The X^{-1} preserves complementation, union, and intersections as the following properties show. For $A' \subset \Omega'$, $A'_t \subset \Omega'$, and T an arbitrary index set, we have

(i) $X^{-1}(\emptyset) = \emptyset$ and $X^{-1}(\Omega') = \Omega$.

(ii) $X^{-1}(A'^c) = \{X^{-1}(A')\}^c$ or $X^{-1}\{\Omega' \setminus A'\} = \Omega \setminus X^{-1}(A')$.

(iii) $X^{-1}(\cup_{t\in T} A'_t) = \cup_{t\in T} X^{-1}(A'_t)$ and $X^{-1}(\cap_{t\in T} A'_t) = \cap_{t\in T} X^{-1}(A'_t)$.

3.1 Inverse Maps

Notation: If $\mathcal{C}' \in \mathcal{P}(\Omega')$ is a class of subsets of Ω', define

$$X^{-1}(\mathcal{C}') = \{X^{-1}(C') : C' \in \mathcal{C}'\}.$$

Proposition 3.1.1

If $\mathcal{B}'$ is a σ-algebra of subsets of Ω', then $X^{-1}(\mathcal{B}')$ is a σ-algebra of subsets of Ω.

Proposition 3.1.2

If $\mathcal{C}'$ is a class of subsets of Ω' then

$$X^{-1}(\sigma(\mathcal{C}')) = \sigma(X^{-1}(\mathcal{C}'))$$

that is, the inverse image of the σ-algebra generated by $\mathcal{C}'$ in Ω' is the same as the σ-algebra generated in Ω by the inverse image.

Proof of Proposition 3.1.2

From Proposition 3.1.1, $X^{-1}(\sigma(\mathcal{C}'))$ is a σ-algebra, and since $\sigma(\mathcal{C}') \supset \mathcal{C}'$, $X^{-1}(\sigma(\mathcal{C}')) \supset X^{-1}(\mathcal{C}')$. Therefore $X^{-1}(\sigma(\mathcal{C}')) \supset \sigma(X^{-1}(\mathcal{C}'))$. Conversely, define

$$\mathcal{F}' = \{B' \in \mathcal{P}(\Omega') : X^{-1}(B') \in \sigma(X^{-1}(\mathcal{C}'))\} \supset \mathcal{C}'.$$

Then $\mathcal{F}'$ is a σ-algebra since

1. $\Omega' \in \mathcal{F}'$, since $X^{-1}(\Omega') = \Omega \in \sigma(X^{-1}(\mathcal{C}'))$.
2. $A' \in \mathcal{F}'$ implies $A'^c \in \mathcal{F}'$ since $X^{-1}(A'^c) = (X^{-1}(A'))^c$.
3. $B'_n \in \mathcal{F}'$ implies $\cup_n B'_n \in \mathcal{F}'$ since $X^{-1}(\cup_n B'_n) = \cup_n X^{-1}(B'_n) \in \sigma(X^{-1}(\mathcal{C}'))$.

By definition, $X^{-1}(\mathcal{F}') \subset \sigma(X^{-1}(\mathcal{C}'))$ and $\mathcal{C}' \subset \mathcal{F}'$. Because $\mathcal{F}'$ is an algebra, $\sigma(\mathcal{C}') \subset \mathcal{F}'$. Thus $X^{-1}(\sigma(\mathcal{C}')) \subset X^{-1}(\mathcal{F}')$. Done.

3.2 Measurable Maps, Random Elements, Induced Probability Measures

A pair $(\Omega, \mathcal{B})$ consisting of a set and a σ-field of subsets is called a **measurable space**. If $(\Omega, \mathcal{B})$ and $(\Omega', \mathcal{B}')$ are two measurable spaces, then a map $X : \Omega \to \Omega'$ is called measurable if

$$X^{-1}(\mathcal{B}') \subset \mathcal{B}.$$

X is also called a **random element** of Ω'. We will use the notation that

$$X \in \mathcal{B}/\mathcal{B}' \text{ or } X : (\Omega, \mathcal{B}) \mapsto (\Omega', \mathcal{B}').$$

A special case occurs when $(\Omega', \mathcal{B}') = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. In this case, X is called a **random variable**. (e.g., $X = I_A$ is a random variable iff $A \in \mathcal{B}$).

3.2 Measurable Maps, Random Elements, Induced Probability Measures

Let $(\Omega, \mathcal{B}, P)$ be a probability space and suppose

$$X : (\Omega, \mathcal{B}) \mapsto (\Omega', \mathcal{B}')$$

is measurable. Define for $A' \subset \Omega'$,

$$[X \in A'] = X^{-1}(A') = \{\omega : X(\omega) \in A'\}.$$

Define the set function $P \circ X^{-1}$ on $\mathcal{B}'$ by

$$P \circ X^{-1}(A') = P(X^{-1}(A')),$$

$P \circ X^{-1}$ is a probability on $(\Omega', \mathcal{B}')$ called the **induced** probability on the **distribution** of X. Usually, we write

$$P \circ X^{-1}(A) = P[X \in A'].$$

If X is a random variable, then $P \circ X^{-1}$ is the measure induced on $\mathbb{R}$ by the distribution function $P \circ X^{-1}(-\infty, x] = P[X \leq x]$.

Example

Consider the experiment of tossing two die and let

$$\Omega = \{(i, j) : 1 \leq i, j \leq 6\}$$

and Define $X : \Omega \mapsto \{2, 3, \ldots, 12\} = \Omega'$ by $X((i, j)) = i + j$. Then

$$X^{-1}(\{4\}) = [X \in \{4\}] = [X = 4] = \{(1, 3), (3, 1), (2, 2)\} \subset \Omega$$

and

$$X^{-1}(\{2, 3\}) = [X \in \{2, 3\}] = \{(1, 1), (1, 2), (2, 1)\} \subset \Omega.$$

The distribution of X is the probability measure on Ω' specified by

$$P \circ X^{-1}(\{i\}) = P[X = i], \; i \in \Omega'.$$

3.2 Measurable Maps, Random Elements, Induced Probability Measures

We now verify $P \circ X^{-1}$ is a probability measure on $\mathcal{B}'$:

(a) $P \circ X^{-1}(\Omega') = P(\Omega) = 1$

(b) $P \circ X^{-1}(A') = P(X^{-1}(A')) \geq 0$

(c) if $\{A'_n\}$ are disjoint in $\mathcal{B}'$, then $\{X^{-1}(A'_n)\}$ are disjoint in $\mathcal{B}$,

$$\begin{aligned} P \circ X^{-1}(\cup_n A'_n) &= P(X^{-1}(\cup_n A'_n)) = P(\cup_n X^{-1}(A'_n)) \\ &= \sum_n P(X^{-1}(A'_n)) = \sum_n P \circ X^{-1}(A'_n). \end{aligned}$$

Remark: When X is a random element of $\mathcal{B}'$, we can make probability statements about X, since $X^{-1}(B') \in \mathcal{B}$ and the probability measure P knows how to assign probabilities to elements of $\mathcal{B}$. The concept of measurability is logically necessary in order to be able to assign probabilities to sets determined by random elements.

3.2 Measurable Maps, Random Elements, Induced Probability Measures

The definition of measurability makes it seem like we have to check $X^{-1}(A') \in \mathcal{B}$ for every $A' \in \mathcal{B}'$; that is $X^{-1}(\mathcal{B}') \subset \mathcal{B}$. In fact, it usually suffices to check that X^{-1} is well behaved on a smaller class than $\mathcal{B}'$.

Proposition 3.2.1 (Test for measurability)

Suppose $X : \Omega \mapsto \omega'$ where $(\Omega, \mathcal{B})$ and $(\Omega', \mathcal{B}')$ are two measurable spaces. Suppose $\mathcal{C}'$ generates $\mathcal{B}'$; that

$$\mathcal{B}' = \sigma(\mathcal{C}').$$

Then X is measurable iff

$$X^{-1}(\mathcal{C}') \subset \mathcal{B}.$$

Proof of Proposition 3.2.1. if $X^{-1}(\mathcal{C}') \subset \mathcal{B}$, then by minimality $\sigma(X^{-1}(\mathcal{C}')) \subset \mathcal{B}$. However we get

$$X^{-1}(\sigma(\mathcal{C}')) = X^{-1}(\mathcal{B}') = \sigma(X^{-1}(\mathcal{C}')) \subset \mathcal{B},$$

which is the definition of measurability.

Corollary 3.2.1 (Special case of random variables)

The real valued function $X : \Omega \mapsto \mathbb{R}$ is a random variable iff $X^{-1}((-\infty, x]) = [X \leq x] \in \mathcal{B}$, for $x \in \mathbb{R}$.

Proof of Corollary 3.2.1

This follows directly from

$$\sigma((-\infty, x], x \in \mathbb{R}) = \mathcal{B}(\mathbb{R}).$$

Proposition 3.2.2 (Composition) (HW 3-1: prove this proposition)

Let X_1, X_2 be two measurable maps $X_1 : (\Omega_1, \mathcal{B}_1) \mapsto (\Omega_2, \mathcal{B}_2)$ and $X_2 : (\Omega_2, \mathcal{B}_2) \mapsto (\Omega_3, \mathcal{B}_3)$ where $(\Omega_i, \mathcal{B}_i)$, $i = 1, 2, 3$ are measurable spaces. Define

$$X = X_2 \circ X_1 : \Omega_1 \mapsto \Omega_3$$

by $X(\omega) = X_2 \circ X_1(\omega_1) = X_2(X_1(\omega_1)), \quad \omega_1 \in \Omega_1$. Then

$$X = X_2 \circ X_1 \in \mathcal{B}_1/\mathcal{B}_3.$$

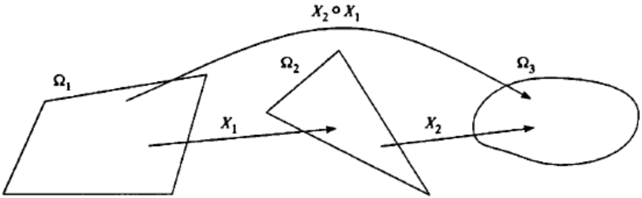

FIGURE 3.2

3.2.2 Random Elements of Metric Spaces

The most common use of the name random elements is when the range is a metric space. Let (S, d) be a metric space with metric d so that $d : S \times S \mapsto \mathbb{R}_+$ satisfies

(i) $d(x, y) \geq 0$.

(ii) $d(x, y) = 0$ iff $x = y$.

(iii) $d(x, y) = d(y, x)$.

(iv) $d(x, z) \leq d(x, y) + d(y, z)$.

Let $\mathcal{O}$ be the class of open subsets of S. Define the Borel σ-algebra $\mathcal{S} = \sigma(\mathcal{O})$. If $X : (\Omega, \mathcal{B}) \mapsto (S, \mathcal{S})$, that is $X \in \mathcal{B}/\mathcal{S}$, then call X a random element of S.

Noteworthy Examples

1. $S = \mathbb{R}$, $d(x, y) = |x - y|$, a random element X of S is called a **random variable**.
2. $S = \mathbb{R}^k$, $d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|_2$ (Euclidean norm), a random element $\boldsymbol{X} = (X_1, \ldots, X_k)$ of S is called a **random vector**.
3. $S = \mathbb{R}^\infty$,

$$d(\boldsymbol{x}, \boldsymbol{y}) = \sum_{k=1}^{\infty} 2^{-k} \left(\frac{\sum_{i=1}^{k} |x_i - y_i|}{1 + \sum_{i=1}^{k} |x_i - y_i|} \right),$$

a random element $\boldsymbol{X} = (X_1, X_2, \ldots)$ of S is called a **random sequence**.

Noteworthy Examples (continued)

4. $S = C[0, \infty)$ be the set of all real valued continuous functions with domain $[0, \infty)$. Define

$$\|x - y\|_m = \sup_{0 \leq t \leq m} |x(t) - y(t)|$$

and

$$d(x, y) = \sum_{m=1}^{\infty} 2^{-m} \left(\frac{\|x - y\|_m}{1 + \|x - y\|_m} \right),$$

a random element $X = X(\cdot)$ of S is called a **random (continuous) function**.

3.2.3 Measurability and Continuity

Proposition 3.2.3

Suppose $(S_i, d_i), i = 1, 2$ are two metric spaces. Let the Borel σ-algebra (generated by open sets) be $\mathcal{S}_i$, $i = 1, 2$. If $X : S_1 \to S_2$ is continuous, then X is measurable: $X \in S_1/S_2$.

Proof

Let $\mathcal{O}_i$ be the class of open subsets of S_i, $i = 1, 2$. If X is continuous, then inverse images of open sets are open, which means that $X^{-1}(\mathcal{O}_2) \subset \mathcal{O}_1 \subset \sigma(\mathcal{O}_1) = \mathcal{S}_1$. So $X \in \mathcal{S}_1/\mathcal{S}_2$ by Proposition 3.2.1.

Corollary 3.2.2

If $\boldsymbol{X} = (X_1, \ldots, X_k)$ is a random vector, and

$$g : \mathbb{R}^k \mapsto \mathbb{R}, \quad g \in \mathcal{B}(\mathbb{R}^k)/\mathcal{B}(\mathbb{R}),$$

then from Proposition 3.2.2, $g(\boldsymbol{X})$ is a random variable. In particular, if g is continuous, then g is measurable and the result holds.

We often consider $g(x_1, \ldots, x_k) = \sum_{i=1}^k x_i$, $k^{-1} \sum_{i=1}^k x_i$, $\sum_{i=1}^k x_i$, $\bigvee_{i=1}^k x_i$, $\prod_{i=1}^k x_i$, $\sum_{i=1}^k x_i^2$, or x_i (projection).

Proposition 3.2.4

$\boldsymbol{X} = (X_1, \ldots, X_k)$ is a random vector, that is a measurable map from $(\Omega, \mathcal{B}) \mapsto (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$, iff X_i is a random variable for each $i = 1, \ldots, k$.

Proposition 3.2.5

$\boldsymbol{X} = (X_1, X_2, \ldots)$ is a random sequence iff X_i is a random variable for each i. Furthermore, iff $(X_1, \ldots, X_k)$ is a random vector for any k.

Proposition 3.2.6

Let $X_1, X_2, \dots$ be random variables defined on $(\Omega, \mathcal{B})$. Then

(i) $\bigvee_n X_n$ and $\bigwedge_n X_n$ are random variables.

(ii) $\liminf_{n\to\infty} X_n$ and $\limsup_{n\to\infty} X_n$ are random variables.

(iii) If $\lim_{n\to\infty} X_n(\omega)$ exists for all ω, then $\lim_{n\to\infty} X_n$ is a random variable.

(iv) The set on which $\{X_n\}$ has a limit is measurable; that is $\{\omega : \lim_{n\to\infty} X_n(\omega) \textit{ exits}\} \in \mathcal{B}$.

(i) $[\bigvee_n X_n \le x] = \cap_n [X_n \le x] \in \mathcal{B}, [\bigwedge_n X_n \le x] = \cup_n [X_n \le x] \in \mathcal{B}$.
(ii) $\liminf_{n\to\infty} X_n = sup_{n\ge 1} \inf_{k\ge n} X_k$, then use (i).
(iii) If $\lim_{n\to\infty} X_n(\omega)$ exists for all ω, then $[\lim_{n\to\infty} X_n \le x] = [\liminf_{n\to\infty} X_n \le x] \in \mathcal{B}$ by (ii).

(iv) Let $\mathbb{Q}$ be the set of all rational real numbers so that $\mathbb{Q}$ is countable. We have

$$\begin{aligned}
&\{\omega : \lim_{n\to\infty} X_n(\omega) \ \textit{exists}\}^c \\
&= \{\omega : \liminf_{n\to\infty} X_n(\omega) < \limsup_{n\to\infty} X_n(\omega)\} \\
&= \ \cup_{r\in\mathbb{Q}} \left[\liminf_{n\to\infty} X_n \le r < \limsup_{n\to\infty} X_n\right] \\
&= \ \cup_{r\in\mathbb{Q}} \left[\liminf_{n\to\infty} X_n \le r\right] \cap \left[\limsup_{n\to\infty} X_n \le r\right]^c \in \mathcal{B}
\end{aligned}$$

since $[\liminf_{n\to\infty} X_n \le r] \in \mathcal{B}$, and $[\limsup_{n\to\infty} X_n \le r] \in \mathcal{B}$.

σ-Algebras Generated by Maps

Let $X : (\Omega, \mathcal{B}) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a random variable. The σ-algebra generated by X, denoted by $\sigma(X)$, is defined as

$$\sigma(X) = X^{-1}(\mathcal{B}(\mathbb{R})) \quad \text{or equivalently} \quad \{[X \in A] : A \in \mathcal{B}(\mathbb{R})\}.$$

This is the σ-algebra generated by information about X, which is away of isolating that information in the probability space that pertains to X.

More generally, suppose $X : (\Omega, \mathcal{B}) \mapsto (\Omega', \mathcal{B}')$. Then we define

$$\sigma(X) = X^{-1}(\mathcal{B}').$$

If $\mathcal{F} \subset \mathcal{B}$ is a sub-σ-algebra of $\mathcal{B}$, we say X is measurable with respect to $\mathcal{F}$, written, $X \in \mathcal{F}$, if $\sigma(X) \subset \mathcal{F}$.

σ-Algebras Generated by Maps

Extreme example: Let $X(\omega) = 810$ for all ω. X can only be 810. $X^{-1}(\{810\}) = \Omega$. Therefore,

$$\sigma(X) = \{\emptyset, \Omega\}.$$

Less extreme example: Let $X(\omega) = I_A$ for some $A \in \mathcal{B}$. X can only take 0 or 1. $X^{-1}(\{0\}) = A^c$ and $X^{-1}(\{1\}) = A$. Thus

$$\sigma(X) = \{\emptyset, \Omega, A, A^c\}.$$

Useful example: Simple function. A random variable is simple if it has a finite range. Suppose the range of X is $\{a_1, \ldots, a_k\}$, where the a's are distinct. Then define

$$A_i = X^{-1}(\{a_i\}) = [X = a_i].$$

Then $\{A_i : i = 1, \ldots, k\}$ partitions Ω, $X = \sum_{i=1}^k a_i I_{A_i}$, and

$$\sigma(X) = \sigma(A_1, \ldots, A_k) = \{\cup_{i \in T} : T \subset \{1, \ldots, k\}\}.$$

σ-Algebras Generated by Maps

In stochastic process theory, we frequently keep track of potential information that can be revealed to us by observing the evolution of a stochastic process by an increasing family of σ-algebras. If $\{X_n : n \geq 1\}$ is a (discrete time) stochastic process, we may define

$$\mathcal{B}_n = \sigma(X_1, \ldots, X_n), \quad n \geq 1.$$

Thus, $\mathcal{B}_n \subset \mathcal{B}_{n+1}$ and we think $\mathcal{B}_n$ as the information potentially available at time n. This is a way of cataloguing what information is contained in the probability model.

σ-Algebras Generated by Maps

Proposition 3.3.1

Suppose X is a random variable and $\mathcal{C}$ is a class of subsets of $\mathbb{R}$, such that $\sigma(\mathcal{C}) = \mathbb{R}$ (e.g., $\mathcal{C} = \{(-\infty, x] : x \in \mathbb{R}\}$), then

$$\sigma(X) = \sigma([X \in \mathcal{B}] : \mathcal{B} \in \mathcal{C}).$$

Proof: We have

$$\begin{aligned}\sigma([X \in B], B \in \mathcal{C}) &= \sigma(X^{-1}(B), B \in \mathcal{C}) \\ &= \sigma(X^{-1}(\mathcal{C})) = X^{-1}(\sigma(\mathcal{C})) \\ &= X^{-1}(\mathcal{B}(\mathbb{R})) = \sigma(X).\end{aligned}$$

Thus $\sigma(X) = (\{[X \leq t] : x \in \mathbb{R}\})$.
(Other HW 3 problems: Section 3.4, Q1-2, Q4-5, Q8, Q11-12, Q14-17, Q19)