Probability and Stochastic Processes I - Lecture 5

Michael Evans
University of Toronto
http://www.utstat.utoronto.ca/mikevans/stac62/STAC622023.html

2023

- we start (as always) with a probability model $(\Omega, \mathcal{A}, P)$

- a random variable isn't really "random", it is just a function X defined on Ω and taking values in R^1, namely, $X : \Omega \to R^1$

- think of Ω as a population and $X(\omega)$ is a measurement of some sort taken of ω

- we want to assign probabilities to events like $a \le X(\omega) \le b$, namely, $X(\omega) \in [a, b]$

- but the probabilities are on Ω not R^1 so how do we do this? answer: through *inverse images*

- the inverse image of the set $B \subset R$ under the function $X : \Omega \to R^1$ is given by

$$X^{-1}B = \{\omega \in \Omega : X(\omega) \in B\},$$

and it is the set of ω that get mapped into B by X

Example II.1.1 - suppose $\Omega = \{1, 2, 3, 5, 6\}$ and

$$X(\omega) = \begin{cases} 0.00 & \omega = 1 \\ 0.20 & \omega = 2 \\ 0.30 & \omega = 3 \\ 0.01 & \omega = 4 \\ 0.20 & \omega = 5 \\ 0.20 & \omega = 6 \end{cases}$$

- note that X is not 1-1

if $B = [0, 1]$, then $X^{-1}B = \Omega$
if $B = [0.00, 0.25]$, then $X^{-1}B = \{1, 2, 4, 5, 6\}$,
if $B = \{0\}$, then $X^{-1}B = \{1\}$
if $B = (-\infty, 0)$, then $X^{-1}B = \phi$
■

- X^{-1} has the important property that it preserves Boolean operations

$$\begin{aligned} X^{-1}(B_1 \cup B_2) &= X^{-1}B_1 \cup X^{-1}B_2 \\ X^{-1}(B_1 \cap B_2) &= X^{-1}B_1 \cap X^{-1}B_2 \\ X^{-1}B^c &= (X^{-1}B)^c \end{aligned}$$

Exercise II.1.1 Prove that X^{-1} preserves Boolean operations and if $B_1 \cap B_2 = \phi$, then $X^{-1}B_1$ and $X^{-1}B_2$ are also disjoint.

note - this property holds for any number of Boolean operations, e.g., $X^{-1} \cup_{i=1}^{\infty} B_i = \cup_{i=1}^{\infty} X^{-1} B_i$

- when $X : \Omega \to R^1$ it is natural to assign the probability $P(X^{-1}B)$ to the event $X(\omega) \in B$ since $X(\omega) \in B$ iff $\omega \in X^{-1}B$

- but for this we need to place a restriction on X

Definition II.1 A *random variable (r.v.)* is a function $X : \Omega \to R^1$ with the property that for any $B \in \mathcal{B}^1$ (B is a Borel set in R^1) then $X^{-1}B \in \mathcal{A}$.

- then when X is a r.v. the assignment $P(\text{“}X(\omega) \in B\text{”}) = P(X^{-1}B)$ can be made

Proposition II.1.1 When X is a r.v., then P_X defined on $\mathcal{B}^1$ by $P_X(B) = P(X^{-1}B)$ is a probability measure on $\mathcal{B}^1$ called the *marginal probability measure of* X.

Proof: Clearly $P_X : \mathcal{B}^1 \rightarrow [0, 1]$. Now
(i) $P_X(R^1) = P(X^{-1}R^1) = P(\Omega) = 1$ so P_X is normed and
(ii) if $B_1, B_2, \ldots$ are mutually disjoint elements of $\mathcal{B}^1$, then

$$\begin{aligned} P_X(\cup_{i=1}^{\infty} B_i) &= P(X^{-1} \cup_{i=1}^{\infty} B_i) = P(\cup_{i=1}^{\infty} X^{-1} B_i) \\ &= \sum_{i=1}^{\infty} P(X^{-1}B_i) = \sum_{i=1}^{\infty} P_X(B_i) \end{aligned}$$

and P_X is countably additive. ■

- a r.v. X has an associated probability model $(R^1, \mathcal{B}^1, P_X)$

Example II.1.2 - when $\mathcal{A} = 2^{\Omega}$, then any $X : \Omega \to R^1$ is a r.v. ■

- how do we check whether or not a specific $X : \Omega \to R^1$ is a r.v.?

Proposition II.1.2 If $X^{-1}(a, b] \in \mathcal{A}$ for every $a, b \in R^1$, then X is a r.v.
Proof: Let

$$\mathcal{B}_*^1 = \{B \in \mathcal{B}^1 : X^{-1}B \in \mathcal{A}\}.$$

Since $\phi \in \mathcal{B}^1$ and $X^{-1}\phi = \phi \in \mathcal{A}$ then $\phi \in \mathcal{B}_*^1$. If $B \in \mathcal{B}_*^1$ then $X^{-1}B \in \mathcal{A}$ which implies $(X^{-1}B)^c = X^{-1}B^c \in \mathcal{A}$ and since $B^c \in \mathcal{B}^1$ this implies $B^c \in \mathcal{B}_*^1$. If $B_1, B_2, \ldots \in \mathcal{B}_*^1$ then $X^{-1}B_1, X^{-1}B_2, \ldots \in \mathcal{A}$ which implies that $\cup_{i=1}^{\infty} X^{-1}B_i = X^{-1} \cup_{i=1}^{\infty} B_i \in \mathcal{A}$ and since $\cup_{i=1}^{\infty} B_i \in \mathcal{B}^1$ this implies $\cup_{i=1}^{\infty} B_i \in \mathcal{B}_*^1$. Therefore, $\mathcal{B}_*^1$ is a sub σ-algebra of $\mathcal{B}^1$.

By hypothesis, $(a, b] \in \mathcal{B}_*^1$ for every $a, b \in R^1$ and so $\mathcal{B}^1 \subset \mathcal{B}_*^1$ as $\mathcal{B}^1$ is the smallest containing all the intervals $(a, b]$. Therefore, $\mathcal{B}_*^1 = \mathcal{B}^1$. This implies $X^{-1}B \in \mathcal{A}$ for every $B \in \mathcal{B}^1$ and so X is a r.v. ■

- **note** - $(a, b] = (-\infty, b] \backslash (-\infty, a]$ so "If $X^{-1}(-\infty, b] \in \mathcal{A}$ for every $b \in R^1$, then X is a r.v." is also true.

Example II.1.2

- let $\Omega = R^1, \mathcal{A} = \mathcal{B}^1$

- let $X : \Omega \to R^1$ be given by $X(\omega) = c$ for all ω, so X is constant, is R^1 a r.v.?

- for any $(-\infty, b]$ then

$$X^{-1}(-\infty, b] = \begin{cases} \Omega & \text{if } c \le b \\ \phi & \text{if } c > b \end{cases}$$

so $X^{-1}(-\infty, b] \in \mathcal{A}$ for every b and X is a r.v.

- now consider $X(\omega) = \omega$, for any $(-\infty, b]$ then $X^{-1}(-\infty, b] = (-\infty, b] \in \mathcal{A}$ for every b, so X is a r.v.

- consider $X(\omega) = \omega^2$, for any $(-\infty, b]$ then

$$X^{-1}(-\infty, b] = \begin{cases} [-\sqrt{b}, \sqrt{b}] & \text{if } b \ge 0 \\ \phi & \text{if } b < 0 \end{cases}$$

so $X^{-1}(-\infty, b] \in \mathcal{A}$ for every b and X is a r.v. ■

Exercise II.1.2 For Example II.1.2 show that $X(\omega) = \omega^n$ is a r.v. for any $n \in \mathbb{Z}$.

Example II.1.3 - let $\Omega = \{1, 2, 3, 4\}, \mathcal{A} = \{\phi, \{1, 2\}, \{3, 4\}, \Omega\}$ and $X : \Omega \to R^1$ be given by

$$X(1) = 3, X(2) = 4, X(3) = 5, X(4) = 5$$

- then let $B = \{4\} \in \mathcal{B}^1$ and note that $X^{-1}\{4\} = \{2\}$ and $\{2\} \notin \mathcal{A}$ so X is not a r.v. ■

Example II.1.4 - quite often $\Omega = R^k, \mathcal{A} = \mathcal{B}^k$ and almost any $X : \Omega \to R^1$ that you can think of is a r.v.

- in particular if X is continuous then X is a random variable

- suppose $X(\omega_1, \ldots, \omega_k) = \omega_i =$ *projection on the i-th coordinate*

- then X is continuous and so is a random variable but also

$$\begin{aligned} X^{-1}(-\infty, b] &= \{(\omega_1, \ldots, \omega_k) \in R^k : \omega_i \leq b\} \\ &= R^1 \times \cdots \times (-\infty, b] \times \cdots \times R^1 \in \mathcal{B}^n \end{aligned}$$

and so X is a r.v. ■

Proposition II.1.3 If X, Y are r.v.'s defined on Ω, then (i) $W = X + Y$ is a r.v. and (ii) $W = XY$ is a r.v.

Proof: (i) Let $c_n \in \mathbb{Q}$ be s.t. $c_n \downarrow b$. Suppose $\omega \in W^{-1}(-\infty, b] = \{\omega : X(\omega) + Y(\omega) \leq b\}$. Then $\exists\ q \in \mathbb{Q}$ s.t. $X(\omega) \leq q$, $Y(\omega) \leq c_n - q$ s.t.

$$\omega \in X^{-1}(-\infty, q] \cap Y^{-1}(-\infty, c_n - q] \in \mathcal{A}$$

and putting

$$C_n = \cup_{q \in \mathbb{Q}} \{\omega : X(\omega) \leq q\} \cap \{\omega : Y(\omega) \leq c_n - q\}$$

we have $W^{-1}(-\infty, b] \subset C_n$ for every n and $C_n \in \mathcal{A}$ (since $\mathbb{Q}$ is countable C_n is a countable union of elements of $\mathcal{A}$). Now note that C_n is monotone decreasing so $\lim_{n\to\infty} C_n = \cap_{n=1}^{\infty} C_n = W^{-1}(-\infty, b] \in \mathcal{A}$ and $W = X + Y$ is a r.v. ■

(ii) **Exercise II.1.3.** (Challenge).

Exercise II.1.4. Prove that a polynomial $p(X) = \sum_{i=0}^{n} a_i X^i$ is a r.v. whenever X is a r.v.

Exercise II.1.5. When $a, b, c \in R^1$ and X, Y r.v.'s then prove that $aX + bY + c$ is a r.v.

- later we will need the σ-algebra generated by r.v. X

Proposition II.1.3 When X is a r.v., then

$$\mathcal{A}_X = X^{-1}\mathcal{B}^1 = \{X^{-1}B : B \in \mathcal{B}^1\}$$

is a sub σ-algebra of $\mathcal{A}$ called the *σ-algebra on Ω generated by X*.

Proof: (i) $\phi = X^{-1}\phi \in \mathcal{A}_X$. (ii) If $A_1, A_2, \ldots \in \mathcal{A}_X$, then there exist $B_1, B_2, \ldots \in \mathcal{B}^1$ such that $A_i = X^{-1}B_i$.Then

$$\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty} X^{-1}B_i = X^{-1} \cup_{i=1}^{\infty} B_i \in \mathcal{A}_X$$

since $\cup_{i=1}^{\infty} B_i \in \mathcal{B}^1$. (iii) If $A \in \mathcal{A}_X$, then there exists $B \in \mathcal{B}^1$ such that $A = X^{-1}B$. This implies $A^c = (X^{-1}B)^c = X^{-1}B^c \in \mathcal{A}_X$ since $B^c \in \mathcal{B}^1$. We conclude that $\mathcal{A}_X$ is a sub σ-algebra of $\mathcal{A}$. ■

note also can write $\mathcal{A}_X = \mathcal{A}(\{X^{-1}(a, b] : a, b \in R^1\})$

Definition II.2 A *random vector* is a function $\mathbf{X} : \Omega \to R^k$ with the property that for any $B \in \mathcal{B}^k$ (B is a Borel set in R^k) then $\mathbf{X}^{-1}B = \{\omega : \mathbf{X}(\omega) \in B\} \in \mathcal{A}$. ■

- similar results hold for random vectors as for random variables

- if $\mathbf{X}, \mathbf{Y}$ are random vectors $a, b \in R^1$, then $a\mathbf{X}+b\mathbf{Y}$ is a random vector

- $P_{\mathbf{X}} : \mathcal{B}^k \to [0, 1]$ given by $P_{\mathbf{X}}(B) = P(\mathbf{X}^{-1}B)$ is the marginal probability measure of $\mathbf{X}$

- $\mathcal{A}_{\mathbf{X}} = \mathbf{X}^{-1}\mathcal{B}^k = \{\mathbf{X}^{-1}B : B \in \mathcal{B}^k\} = \mathcal{A}(\{\mathbf{X}^{-1}(\mathbf{a}, \mathbf{b}] : \mathbf{a}, \mathbf{b} \in R^k\})$ is the *σ-algebra on Ω generated by* $\mathbf{X}$

Example II.1.5 - suppose $\Omega = \{1, 2, 3\}, \mathcal{A} = 2^{\Omega}$ and let P be the uniform probability measure

- define X_1, X_2 by

$$\begin{aligned} X_1(1) &= 0, X_1(2) = 0, X_1(3) = 1 \\ X_2(1) &= 1, X_2(2) = 0, X_2(3) = 0 \end{aligned}$$

and let $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} : \Omega \to R^2$ be given by $\mathbf{X}(\omega) = \begin{pmatrix} X_1(\omega) \\ X_2(\omega) \end{pmatrix}$

- then for $B \in \mathcal{B}^2$

$$P_{\mathbf{X}}(B) = \begin{cases} 1 & \text{if } (0,0), (0,1), (1,0) \in B \\ 2/3 & \text{if only two of } (0,0), (0,1), (1,0) \in B \\ 1/3 & \text{if only one of } (0,0), (0,1), (1,0) \in B \\ 0 & \text{if none of } (0,0), (0,1), (1,0) \in B \ \blacksquare \end{cases}$$

Exercise II.1.6. Compute $P_{\mathbf{X}}$ in **Example II.1.5** when $P(\{1\}) = 1/2, P(\{2\}) = 1/3, P(\{3\}) = 1/6$ and

$$\begin{aligned} X_1(1) &= 0, X_1(2) = 0, X_1(3) = 1 \\ X_2(1) &= 1, X_2(2) = 1, X_2(3) = 0. \end{aligned}$$

- we can obtain the Borel sets on R^k from the Borel sets on R^1

Proposition II.1.4 If $B_1, B_2, \ldots, B_k \in \mathcal{B}^1$, then

$$B_1 \times B_2 \times \cdots \times B_k = \{(x_1, \ldots, x_k)' : x_i \in B_i, i = 1, \ldots, k\} \in \mathcal{B}^k.$$

and the smallest σ-algebra on R^k containing all such sets is $\mathcal{B}^k$.

Proof: Consider the sets $R^1 \times \cdots \times B_i \times \cdots \times R^1$ that only restrict the i-th coordinate. Then $\{R^1 \times \cdots \times B_i \times \cdots \times R^1 : B_i \in \mathcal{B}^1\}$ is a sub σ-algebra of $\mathcal{B}^k$ (**Exercise II.1.7**) and so

$$B_1 \times B_2 \times \cdots \times B_k = \cap_{i=1}^k (R^1 \times \cdots \times B_i \times \cdots \times R^1) \in \mathcal{B}^k.$$

Since each k-cell $(\mathbf{a}, \mathbf{b}] = (a_1, b_1] \times \cdots \times (a_k, b_k]$ is of this form there cannot be a smaller σ-algebra on R^k containing all such sets than $\mathcal{B}^k$. ■

Proposition II.1.5 If $X_i : \Omega \to R^1$ is a r.v. for $i = 1, \ldots, k$, then $\mathbf{X} = (X_1, \ldots, X_k)' : \Omega \to R^k$ is a random vector.

Proof: Suppose $B_1, B_2, \ldots, B_k \in \mathcal{B}^1$ so $B_1 \times B_2 \times \cdots \times B_k \in \mathcal{B}^k$ by the previous result. Then

$$\begin{aligned} \mathbf{X}^{-1}(B_1 \times B_2 \times \cdots \times B_k) &= \{\omega : \mathbf{X}(\omega) \in B_1 \times B_2 \times \cdots \times B_k\} \\ &= \{\omega : X_i(\omega) \in B_i \text{ for } i = 1, \ldots, k\} \\ &= \cap_{i=1}^{k} X_i^{-1} B_i \in \mathcal{A}. \end{aligned}$$

Since this implies $\mathbf{X}^{-1}(\mathbf{a}, \mathbf{b}] \in \mathcal{A}$ for every $\mathbf{a}, \mathbf{b} \in R^k$ this implies (as in Prop. II.1.2) that $\mathbf{X}^{-1} B \in \mathcal{A}$ for every $B \in \mathcal{B}^k$ and $\mathbf{X}$ is a random vector. ■