Probability and Stochastic Processes I - Lecture 13

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

2023

- probability model $(\Omega, \mathcal{A}, P)$
- recall the definition of the indicator function for $A \in \mathcal{A}$

$$I_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases} \sim \text{ Bernoulli}(P(A))$$

- some properties of indicator functions

$$\begin{aligned} I_{A^c}(\omega) &= 1 - I_A(\omega), I_{\cap_{i=1}^n A_i} = \prod_{i=1}^n I_{A_i}, \\ I_{\cup_{i=1}^n A_i} &= 1 - \prod_{i=1}^n I_{A_i^c} = 1 - \prod_{i=1}^n (1 - I_{A_i}) \\ &= \sum_{i=1}^n I_{A_i} - \sum_{i<j} I_{A_i} I_{A_j} + \cdots + (-1)^{n+1} \prod_{i=1}^n I_{A_i} \text{ (induction)} \\ &= \sum_{i=1}^n I_{A_i} - \sum_{i<j} I_{A_i \cap A_j} + \cdots + (-1)^{n+1} I_{\cap_{i=1}^n A_i} \end{aligned}$$

Definition III.1.1 If $A_1, \ldots, A_l \in \mathcal{A}$ and $a_1, \ldots, a_l \in R^1$, a function $X : \Omega \to R^1$ given by $X(\omega) = \sum_{i=1}^{l} a_i I_{A_i}(\omega)$ is called a *simple function*. ■

note - a simple function takes only finitely many values and it is a random variable (a sum of r.v.'s is a r.v.) and any r.v. that takes only finitely many values is a simple function (**Exercise III.1.1**)

- let $c_1, \ldots, c_m \in R^1$ be the distinct values taken by simple function X and $C_i = X^{-1}\{c_i\} \in \mathcal{A}$ so $C_i \cap C_j = \phi$ when $i \neq j$, $\cup_{i=1}^{n} C_i = \Omega$ and

$$X(\omega) = \sum_{i=1}^{m} c_i I_{C_i}(\omega)$$

is in canonical form with a discrete distribution

$$p_X(x) = P_X(\{x\}) = P(X^{-1}\{x\}) = \begin{cases} 0 & x \notin \{c_1, \ldots, c_m\} \\ P(C_i) & x = c_i \end{cases}$$

- when $\omega_1, \ldots, \omega_n$ are i.i.d. (independently and identically distributed) P, then

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} X(\omega_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{l} a_j I_{A_j}(\omega_i) = \sum_{j=1}^{l} a_j \left(\frac{1}{n}\sum_{i=1}^{n} I_{A_j}(\omega_i)\right) \to \sum_{j=1}^{l} a_j P(A_j) \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} c_j I_{C_j}(\omega_i) = \sum_{j=1}^{m} c_j \left(\frac{1}{n}\sum_{i=1}^{n} I_{C_j}(\omega_i)\right) \to \sum_{j=1}^{m} c_j P(C_j)
\end{aligned}
$$

as $n \longrightarrow \infty$ so

$$\sum_{j=1}^{l} a_j P(A_j) = \sum_{j=1}^{m} c_j P(C_j)$$

- this leads to the following definition

Definition III.1.2 For a simple function $X = \sum_{i=1}^{l} a_i I_{A_i}$ the *expectation* of X is defined by

$$E(X) = \sum_{i=1}^{l} a_i P(A_i). \blacksquare$$

- if X_1, X_2 are simple functions, then so is $a_0 + a_1 X_1 + a_2 X_2$ for any constants a_0, a_1, a_2 and also $X_1 X_2$ is a simple function

Proposition III.1.1 If X_1, X_2 are simple functions, then
(i) $E(a_0 + a_1 X_1 + a_2 X_2) = a_0 + a_1 E(X_1) + a_2 E(X_2)$,
(ii) if $X_1 \le X_2$, then $E(X_1) \le E(X_2)$,
(iii) if $P(\{\omega : X_1(\omega) \ne X_2(\omega)\}) = 0$, then $E(X_1) = E(X_2)$.

Proof: (i) **Exercise III.1.2**

(ii) Since $X_2 - X_1$ is a nonnegative simple function so distinct values taken are nonnegative which implies, using (i),

$$0 \le E(X_2 - X_1) = E(X_2) - E(X_1).$$

(iii) Suppose $X_1 = \sum_{i=1}^{l} a_i I_{A_i}, X_2 = \sum_{i=1}^{m} b_i I_{B_i}$ are in canonical form. Note that if $P(A_j) = 0$, then

$$E(X_1) = \sum_{i=1}^{l} a_i P(A_i) = \sum_{i \neq j} a_i P(A_i)$$

and similarly for X_2. So assume that $P(A_i) > 0, P(B_j) > 0$ for all i, j. Then for each a_i there exists b_j (and conversely) such that $a_i = b_j$ and A_i and B_j satisfy $P(A_i \cap B_j^c) = P(A_i^c \cap B_j) = 0$ which implies $P(A_i) = P(B_j)$. This gives the result. ■

- now we want to extend the definition of expectation to as many r.v.'s as possible

- suppose X is a nonnegative r.v. and for $i \in \{1, \ldots, n\}, j \in \{1, \ldots, 2^n\}$ let

$$\begin{aligned} A_{i,j,n} &= \{\omega : (i-1) + (j-1)/2^n \leq X(\omega) < (i-1) + j/2^n\} \in \mathcal{A} \\ X_n &= \sum_{i=1}^{n} \sum_{j=1}^{2^n} ((i-1) + (j-1)/2^n) I_{A_{i,j,n}} \end{aligned}$$

and then X_n is a nonnegative simple function satisfying $X_n(\omega) \leq X(\omega)$

- suppose $n \leq n'$,

$$\begin{aligned} &\text{if } X(\omega) \geq n, \text{ then } 0 = X_n(\omega) \leq X_{n'}(\omega), \\ &\text{if } \omega \in A_{i,j,n}, \text{ then } \omega \in A_{i,j',n'} \text{ for some } j' \text{ and } X_n(\omega) \leq X_{n'}(\omega) \end{aligned}$$

- furthermore $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ for all $\omega \in \Omega$

- by Prop. III.1.1(ii) $E(X_n)$ is increasing and so $\lim_{n\to\infty} E(X_n)$ exists (could be ∞) and it makes sense then to define

$$E(X) = \lim_{n\to\infty} E(X_n)$$

provided this limit is the same for any increasing sequence of simple functions X_n satisfying $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ for all $\omega \in \Omega$ and (fact) this is true

- suppose X is a r.v. and define

$$\begin{aligned} X_+(\omega) &= \max\{0, X(\omega)\} \text{ the positive part of } X \\ X_-(\omega) &= \max\{0, -X(\omega)\} \text{ the negative part of } X \end{aligned}$$

so $X = X_+ - X_-$ and for any Borel set $B \subset R^1$

$$X_+^{-1}B = \begin{cases} X^{-1}(B \cap (0,\infty)) & \text{if } 0 \notin B \\ X^{-1}(-\infty, 0] \cup X^{-1}(B \cap (0,\infty)) & \text{if } 0 \in B \end{cases} \in \mathcal{A}$$

so X_+ is a nonnegative r.v. and similarly X_- is a nonnegative r.v.

Definition III.1.3 For a r.v. X define the *expectation* of X by

$$E(X) = E(X_+) - E(X_-)$$

provided at least one of $E(X_+), E(X_-)$ is finite, otherwise $E(X)$ is not defined. ■