5.4 Indicator random variables and the fundamental bridge

Unit 5: Averages

Adapted from Blitzstein-Hwang Chapters 4, 5, and 10.

This section is devoted to *indicator random variables*, which we have encountered previously but will treat in much greater detail here. In particular, we will show that indicator r.v.s are an extremely useful tool for calculating expected values. Recall from the previous chapter that the indicator r.v. I_A (or $I(A)$) for an event A is defined to be 1 if A occurs and 0 otherwise. So I_A is a Bernoulli random variable, where success is defined as "A occurs" and failure is defined as "A does not occur". Some useful properties of indicator r.v.s are summarized below.

> THEOREM 5.4.1 (INDICATOR R.V. PROPERTIES).
>
> Let A and B be events. Then the following properties hold.
>
> 1. $(I_A)^k = I_A$ for any positive integer k.
> 2. $I_{A^c} = 1 - I_A$.
> 3. $I_{A\cap B} = I_A I_B$.
> 4. $I_{A\cup B} = I_A + I_B - I_A I_B$.

Proof

Property 1 holds since $0^k = 0$ and $1^k = 1$ for any positive integer k. Property 2 holds since $1 - I_A$ is 1 if A does not occur and 0 if A occurs. Property 3 holds since $I_A I_B$ is 1 if both I_A and I_B are 1, and 0 otherwise. Property 4 holds since

$$I_{A\cup B} = 1 - I_{A^c\cap B^c} = 1 - I_{A^c} I_{B^c} = 1 - (1 - I_A)(1 - I_B) = I_A + I_B - I_A I_B.$$

Indicator r.v.s provide a link between probability and expectation; we call this fact the *fundamental bridge*.

> THEOREM 5.4.2 (FUNDAMENTAL BRIDGE BETWEEN PROBABILITY AND EXPECTATION).

There is a one-to-one correspondence between events and indicator r.v.s, and the probability of an event A is the expected value of its indicator r.v. I_A:

$$P(A) = E(I_A).$$

Proof

For any event A, we have an indicator r.v. I_A. This is a one-to-one correspondence since A uniquely determines I_A and vice versa (to get from I_A back to A, we can use the fact that $A = \{s \in S : I_A(s) = 1\}$). Since $I_A \sim \mathrm{Bern}(p)$ with $p = P(A)$, we have $E(I_A) = P(A)$.

The fundamental bridge connects events to their indicator r.v.s, and allows us to express *any* probability as an expectation.

Conversely, the fundamental bridge is also extremely useful in many expected value problems. We can often express a complicated discrete r.v. whose distribution we don't know as a sum of indicator r.v.s, which are extremely simple. The fundamental bridge lets us find the expectation of the indicators; then, using linearity, we obtain the expectation of our original r.v.

Recognizing problems that are amenable to this strategy and then defining the indicator r.v.s takes practice, so it is important to study a lot of examples and solve a lot of problems. In applying the strategy to a random variable that counts the number of [noun]s, we should have an indicator for each potential [noun]. This [noun] could be a person, place, or thing; we will see examples of all three types.

Example 5.4.3 (Putnam problem).

A permutation $a_1, a_2, \ldots, a_n$ of $1, 2, \ldots, n$ has a *local maximum* at j if $a_j > a_{j-1}$ and $a_j > a_{j+1}$ (for $2 \leq j \leq n-1$; for $j = 1$, a local maximum at j means $a_1 > a_2$ while for $j = n$, it means $a_n > a_{n-1}$). For example, $4, 2, 5, 3, 6, 1$ has 3 local maxima, at positions 1, 3, and 5. The Putnam exam (a famous, hard math competition, on which the median score is often a 0) from 2006 posed the following question: for $n \geq 2$, what is the average number of local maxima of a random permutation of $1, 2, \ldots, n$, with all $n!$ permutations equally likely?

Solution

This problem can be solved quickly using indicator r.v.s, symmetry, and the fundamental bridge. Let $I_1, \ldots, I_n$ be indicator r.v.s, where I_j is 1 if there is a local maximum at position j, and 0 otherwise. We are interested in the expected value of $\sum_{j=1}^{n} I_j$. For $1 < j < n$, $EI_j = 1/3$ since having a local maximum at j is equivalent to a_j being the largest of a_{j-1}, a_j, a_{j+1}, which has probability $1/3$ since all orders are equally likely. For $j = 1$ or $j = n$, we have $EI_j = 1/2$ since then there is only one neighbor. Thus, by linearity,

$$E\left(\sum_{j=1}^{n} I_j\right) = 2 \cdot \frac{1}{2} + (n-2) \cdot \frac{1}{3} = \frac{n+1}{3}.$$

Learn About Verified Certificates