DATA 1010
In-class exercises
Samuel S. Watson
10 October 2018

Problem 1

Consider the random variable 1_E which maps each $\omega \in E$ to 1 and each $\omega \in E^c$ to 0. Find the expected value of 1_E.

Solution

We have $\mathbb{E}[1_E] = 1 \cdot \mathbb{P}(E) + 0 \cdot \mathbb{P}(E^c) = \mathbb{P}(E)$.

We can see that random variables *generalize* events: if we identify an event E with its indicator random variable 1_E, then we can perform set operations (for example, intersection of two sets corresponds to multiplication of their indicator random variables), and we can take probabilities of events (by finding the expectations of the corresponding indicator random variables).

This makes sense from the random-experiment point of view: An event is a statement which is either true (1) or false (0) for each outcome. A random variable *also* associates a number to each outcome, but the values are not restricted to $\{0, 1\}$.

Problem 2

The expectation of a random variable need not be finite or even well-defined. Show that the expectation of the random variable which assigns a probability mass of 2^{-n} to the point 2^n (for all $n \geq 1$) is not finite.

Consider a random variable X whose distribution assigns a probability mass of $2^{-|n|-1}$ to each point 2^n for $n \geq 1$ and a probability mass of $2^{-|n|-1}$ to -2^n for each $n \leq -1$. Show that $\mathbb{E}[X]$ is not well-defined. (Note: a sum $\sum_{x \in \mathbb{R}} f(x)$ is not defined if $\sum_{x \in \mathbb{R}\,:\, f(x) > 0} f(x)$ and $\sum_{x \in \mathbb{R}\,:\, f(x) < 0} f(x)$ are equal to ∞ and $-\infty$, respectively.)

Solution

We multiply the probability mass at each point x by the location x and sum to get

$$\sum_{n=1}^{\infty} 2^{-n} 2^n = \sum_{n=1}^{\infty} 1 = \infty.$$

For the second distribution, the positive and negative parts of the are both infinite for the same reason. Therefore, the sum does not make sense and the mean is therefore not well-defined.

Problem 3

Shuffle a standard 52-card deck, and let X be the number of consecutive pairs of cards in the deck which are both red. Find $E[X]$.

Write some code to simulate this experiment and confirm that your answer is correct. Hint: store the deck of undrawn cards as a `Set`, and `pop!` cards from it as you draw. You can draw a random element from a set `S` using `rand(S)`.

Solution

Calculating $\mathbb{E}[X]$ from the distribution of X would be very complicated, because the distribution of X is complicated. Try to calculate, for instance, $\mathbb{P}(X = 10)$.

However, finding the mean is nevertheless manageable: we can write $X = X_2 + \ldots + X_{52}$ where X_j is equal to 1 if the cards in positions $j - 1$ and j are both red and is equal to 0 otherwise. We see that $\mathbb{E}[X_j] = (1/2)(25/51)$, since card j is red with probability $1/2$, and card $j - 1$ is red with conditional probability $25/51$, given that card j is red. So $\mathbb{E}[X] = \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_{52}] = 51(1/2)(25/51) = 25/2$.

To simulate this experiment, we identify the cards using integers 1 to 52. We treat even numbers as red cards and odd numbers as black cards.

```
using Statistics
"""
Put the integers 1 to n in random order
"""
function draw(n)
    deck = Set(1:n)
    drawn_cards = Int64[]
    while !isempty(deck)
        card = rand(deck)
        push!(drawn_cards,card)
        pop!(deck,card)
    end
    drawn_cards
end

"""
Count the number of consecutive pairs of even numbers
in the sequence `cards`
"""
function countpairs(cards)
    pairs = 0
    for i=2:length(cards)
        if iseven(cards[i]) && iseven(cards[i-1])
            pairs += 1
        end
    end
    pairs
end
```

Problem 4

Show that variance satisfies the properties

$$\begin{cases} \mathrm{Var}(aX) = a^2 \,\mathrm{Var}\, X, & \text{for all random variables } X \text{ and real numbers } a \\ \mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y), & \text{if } X \text{ and } Y \text{ are independent random variables} \end{cases}$$

Solution

The first part of the statement follows easily from linearity of expectation

$$\begin{aligned} \mathrm{Var}(aX) &= \mathbb{E}[a^2X^2] - \mathbb{E}[aX]^2 \\ &= a^2\mathbb{E}[X^2] - a^2\mathbb{E}[X]^2 \\ &= a^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2) \\ &= a^2\,\mathrm{Var}(X). \end{aligned}$$

Since $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ by linearity, we have

$$\begin{aligned} \mathrm{Var}(X+Y) &= \mathbb{E}[(X+Y)^2] - (\mathbb{E}[X] + \mathbb{E}[Y])^2 \\ &= \mathbb{E}[X^2 + 2XY + Y^2] - \mathbb{E}[X]^2 - 2\mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[Y]^2. \end{aligned}$$

Rearranging and using linearity of expectation, we get

$$\begin{aligned} \mathrm{Var}(X+Y) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 + \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 + 2(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]) \\ &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]). \end{aligned}$$

The desired result follows because if X and Y are independent, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.