5. Indicator variables

Problem 5. Indicator variables

5/6 points (graded)

Consider a sequence of $n+1$ independent tosses of a biased coin, at times $k=0,1,2,\ldots,n$. On each toss, the probability of Heads is p, and the probability of Tails is $1-p$.

A reward of one unit is given at time k, for $k\in\{1,2,\ldots,n\}$, if the toss at time k resulted in Tails and the toss at time $k-1$ resulted in Heads. Otherwise, no reward is given at time k.

Let R be the sum of the rewards collected at times $1,2,\ldots,n$.

We will find $\mathbf{E}[R]$ and $\mathsf{Var}(R)$ by carrying out a sequence of steps. Express your answers below in terms of p and/or n using standard notation (available through the "STANDARD NOTATION" button below.) Remember to write "*" for all multiplications and to include parentheses where necessary.

We first work towards finding $\mathbf{E}[R]$.

1. Let I_k denote the reward (possibly 0) given at time k, for $k\in\{1,2,\ldots,n\}$. Find $\mathbf{E}[I_k]$.

$\mathbf{E}[I_k]=$ p*(1-p) ✔ **Answer:** p*(1-p)

$p\cdot(1-p)$

2. Using the answer to part 1, find $\mathbf{E}[R]$.

$\mathbf{E}[R]=$ n*p*(1-p) ✔ **Answer:** n*p*(1-p)

$n\cdot p\cdot(1-p)$

The variance calculation is more involved because the random variables $I_1,I_2,\ldots,I_n$ are not independent. We begin by computing the following values.

3. If $k\in\{1,2,\ldots,n\}$, then

$\mathbf{E}[I_k^2]=$ p*(1-p) ✔ **Answer:** p*(1-p)

$p\cdot(1-p)$

4. If $k \in \{1, 2, \ldots, n-1\}$, then

$\mathbf{E}[I_k I_{k+1}] =$ 0 ✔ **Answer:** 0

0

5. If $k \geq 1$, $\ell \geq 2$, and $k + \ell \leq n$, then

$\mathbf{E}[I_k I_{k+\ell}] =$ p^2*(1-p)^2 ✔ **Answer:** p^2*(1-p)^2

$p^2 \cdot (1-p)^2$

6. Using the results above, calculate the numerical value of $\mathsf{Var}(R)$, assuming that $p = 3/4$, $n = 10$.

$\mathsf{Var}(R) =$ 4506/256 ✖ **Answer:** 0.890625

STANDARD NOTATION

Solution:

1. Since I_k is a Bernoulli indicator variable and the tosses are independent, we have

$$\mathbf{E}[I_k] = \mathbf{P}(I_k = 1) = \mathbf{P}(\text{Tails at time } k \text{ and Heads at time } k-1) = p(1-p).$$

2. The total reward over all the tosses, R, is the sum of all the I_k's, for $k = 1, 2, \ldots, n$. By linearity of expectations, we have

$$\mathbf{E}[R] = \mathbf{E}\left[\sum_{k=1}^{n} I_k\right] = \sum_{k=1}^{n} \mathbf{E}[I_k] = np(1-p).$$

3. Since I_k can be only 0 or 1, $\mathbf{E}[I_k^2] = \mathbf{E}[I_k] = p(1-p)$.

4. $I_k I_{k+1}$ equals 1 if $I_k = 1$ and $I_{k+1} = 1$, i.e., if a reward was given at time k **and** at time $k+1$. Otherwise, $I_k I_{k+1}$ equals 0. But I_k and I_{k+1} cannot both equal 1: $I_k = 1$ implies that the toss at time k resulted in Tails, while $I_{k+1} = 1$ implies that the toss at time k resulted in Heads. Hence, it is not possible to obtain a reward at consecutive times k and $k+1$. Therefore, $\mathbf{E}[I_k I_{k+1}] = 0$.

5. Part 4 above considered the rewards at two consecutive times. We now consider the rewards at two times that are at least 2 periods apart. Since the reward at time k depends only on the tosses at times k and $k-1$, the rewards at times that are at least 2 periods apart depend on

different, non-overlapping pairs of coin tosses, and hence I_k and $I_{k+\ell}$ are independent for $\ell \geq 2$. Therefore, $\mathbf{E}[I_k I_{k+\ell}] = \mathbf{E}[I_k]\mathbf{E}[I_{k+\ell}] = p^2(1-p)^2$ for the values of k and ℓ specified in the problem statement for this part.

6. From Part 2, we have already calculated $\mathbf{E}[R]$. We now find $\mathbf{E}[R^2]$ and use the identity $\mathsf{Var}(R) = \mathbf{E}[R^2] - (\mathbf{E}[R])^2$.

$$\mathbf{E}[R^2] = \mathbf{E}\left[\left(\sum_{k=1}^{n} I_k\right)\left(\sum_{m=1}^{n} I_m\right)\right] = \mathbf{E}\left[\sum_{k=1}^{n}\sum_{m=1}^{n} I_k I_m\right] = \sum_{k=1}^{n}\sum_{m=1}^{n} \mathbf{E}[I_k I_m]$$

There are n^2 terms in this double summation. We can divide them into three groups:

1. There are n terms where $k = m$. From Part 3, we know that $\mathbf{E}[I_k I_m] = p(1-p)$ for this case.

2. There are $n-1$ terms where $k = m+1$ and another $n-1$ terms where $m = k+1$. From Part 4, we know that $\mathbf{E}[I_k I_m] = 0$ for these cases.

3. The remaining $n^2 - n - 2(n-1) = n^2 - 3n + 2$ terms are those where k and m differ by at least 2. From Part 5, we know that $\mathbf{E}[I_k I_m] = p^2(1-p)^2$ for these cases.

Putting these cases together, we have

$$\mathbf{E}[R^2] = n \cdot p(1-p) + 2(n-1) \cdot 0 + (n^2 - 3n + 2) \cdot p^2(1-p)^2.$$

Therefore,

$$\begin{aligned} \mathsf{Var}(R) &= \mathbf{E}[R^2] - (\mathbf{E}[R])^2 \\ &= np(1-p) + (n^2 - 3n + 2)p^2(1-p)^2 - n^2p^2(1-p)^2 \\ &= np(1-p) - (3n-2)p^2(1-p)^2. \end{aligned}$$

When $p = 3/4$ and $n = 10$, we obtain $\mathsf{Var}(R) = 57/64 = 0.890625$.

提交 You have used 5 of 5 attempts

Answers are displayed within the problem

讨论

Topic: Unit 4 / Problem Set / 5. Indicator variables

显示讨论

Learn About Verified Certificates