Module 9.4: Indicators

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

Indicator random variables

Let A be an event in a probability space $(\Omega, \Pr)$. The **indicator** random variable of the event A, notation I_A, is defined by

$$I_A(w) = \begin{cases} 1 & \text{if } w \in A \\ 0 & \text{if } w \notin A \end{cases}$$

Note that I_A is a Bernoulli random variable
with success probability $\Pr[I_A = 1] = \Pr[A]$.

As we have shown before, its expectation is $\mathrm{E}[I_A] = \Pr[I_A = 1] = \Pr[A]$.

Number of heads in n coin flips I

Problem. We flip a biased coin n times with heads probability p. Let H be the r.v. that returns the number of heads observed. Compute $\mathrm{E}[H]$.

Answer (first attempt). We will try to use the formula for expectation.

The outcomes are 2^n sequences of length n of $\mathtt{H}$'s and $\mathtt{T}$'s. An outcome with k $\mathtt{H}$'s has probability $p^k q^{n-k}$ where $q = 1 - p$.

There are $\binom{n}{k}$ outcomes with k $\mathtt{H}$'s. Therefore the probability of the event "k heads observed" is $\binom{n}{k} p^k q^{n-k}$.

Using the formula for expectation:

$$\mathrm{E}[H] = \sum_{k=0}^{n} k \cdot \Pr[H = k] = \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k}$$

Now what?

Number of heads in n coin flips II

Answer (second attempt). We are going to use **linearity of expectation**.

In the probability space of the n flips that we just saw, let
H_k be the event "the k'th flip is $\mathtt{H}$" for $k = 1, \ldots, n$ and let
I_k be the **indicator** random variable of the event H_k.

Clearly, $H = I_1 + \cdots + I_n$.

By linearity of expectation $\mathrm{E}[H] = \mathrm{E}[I_1] + \cdots + \mathrm{E}[I_n]$.

We established earlier that $\mathrm{E}[I_k] = \Pr[H_k]$.

Since the flips are independent $\Pr[H_k] = p$.

In conclusion $\mathrm{E}[H] = p + \cdots + p = n \cdot p$.

Interestingly,

$$\sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} = np$$

Balls to bins "on average"

Problem. We throw k balls into n bins. What is the number of balls that end up in Bin 1, "on average"?

Answer. We want $\mathrm{E}[X]$ where X is the random variable that returns the number of balls that end up in Bin 1.

We can express X as $X = I_1 + \cdots + I_k$.

where I_i is the **indicator** r.v. of the event $L_i =$"ball i ends up in Bin 1".

Recall from the discussion that we introduced the model that $\Pr[L_i] = 1/n$.

Therefore $\mathrm{E}[I_i] = \Pr[L_i] = 1/n$.

Using **linearity of expectation** we obtain $\mathrm{E}[X] = (1/n) + \ldots + (1/n) = k/n$.

Activity : Balls in bins and biased coins

The balls into bins problem we just saw can be seen as a particular case of the biased coin one that precedes it.

Indeed for each ball throw consider only two outcomes: (1) falls in Bin 1, and (2) does not fall in Bin 1. Since each of the n bins is equally likely to receive the ball, outcome (1) has probability $1/n$ (while outcome (2) has probability $(n-1)/n$).

Therefore each ball throw is like flipping biased coin with $1/n$ probability of showing heads.

We can then apply the calculation of the expected number of heads in multiple flips of biased coin. Here we have k throws so the expected number of balls in Bin 1 will be $(k)(1/n) = k/n$, the same answer.