

Expected Number of Trials until Success

Consider the following famous puzzle.

In a country, all families want a boy. They keep having babies till a boy is born. What is the expected ratio of boys and girls in the country?

This puzzle can be easily solved if we know following interesting result in probability and expectation.

If probability of success is p in every trial, then expected number of trials until success is $1/p$

Proof: Let R be a random variable that indicates number of trials until success.

The expected value of R is sum of following infinite series

$$E[R] = 1*p + 2*(1-p)*p + 3*(1-p)^2*p + 4*(1-p)^3*p + \dots$$

Taking ' p ' out

$$E[R] = p[1 + 2*(1-p) + 3*(1-p)^2 + 4*(1-p)^3 + \dots] \text{ ---->(1)}$$

Multiplying both sides with ' $(1-p)$ ' and subtracting

$$(1-p)*E[R] = p[1*(1-p) + 2*(1-p)^2 + 3*(1-p)^3 + \dots] \text{ --->(2)}$$

Subtracting (2) from (1), we get

$$p*E[R] = p[1 + (1-p) + (1-p)^2 + (1-p)^3 + \dots]$$

Canceling p from both sides

$$E[R] = [1 + (1-p) + (1-p)^2 + (1-p)^3 + \dots]$$

Above is an infinite geometric progression with ratio $(1-p)$. Since $(1-p)$ is less than 1, we can apply sum formula.

$$\begin{aligned} E[R] &= 1/[1 - (1-p)] \\ &= 1/p \end{aligned}$$

Solution of Boys/Girls ratio puzzle:

Let us use the above result to solve the puzzle. In the given puzzle, probability of success in every trial is $1/2$ (assuming that girls and boys are equally likely).

Let p be probability of having a baby boy.

$$\begin{aligned} \text{Number of kids until a baby boy is born} &= 1/p \\ &= 1/(1/2) \\ &= 2 \end{aligned}$$

Since expected number of kids in a family is 2, ratio of boys and girls is 50:50.

Let us discuss another problem that uses above result.

Coupon Collector Problem:

Suppose there are n types of coupons in a lottery and each lot contains one coupon (with probability $1/n$ each). How many lots have to be bought (in expectation) until we have at least one coupon of each type.

The solution of this problem is also based on above result.

Let X_i be the number of lots bought before i 'th new coupon is collected.

Note that X_1 is 1 as the first coupon is always a new coupon (not collected before).

Let 'p' be probability that 2nd coupon is collected in next buy. The value of p is $(n-1)/n$. So the number of trials needed before 2nd new coupon is picked is $1/p$ which means $n/(n-1)$. [This is where we use above result]

Similarly, the number of trials needed before 3rd new coupon is collected is $n/(n-2)$

Using **Linearity of expectation**,

we can say that the total number of expected trials =

$$\begin{aligned} & 1 + n/(n-1) + n/(n-2) + n/(n-3) + \dots + n/2 + n/1 \\ &= n[1/n + 1/(n-1) + 1/(n-2) + 1/(n-3) + \dots + 1/2 + 1/1] \\ &= n * H_n \end{aligned}$$

Here H_n is n -th **Harmonic number**

Since $\text{Log}n \leq H_n \leq \text{Log}n + 1$, we need to buy around $n \text{Log}n$ lots to collect all n coupons.

Exercise:

- 1) A 6 faced fair dice is thrown until a '5' is seen as result of dice throw. What is the expected number of throws?
- 2) What is the ratio of boys and girls in above puzzle if probability of a baby boy is $1/3$?

Reference:

http://www.cse.iitd.ac.in/~mohanty/col106/Resources/linearity_expectation.pdf

<http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-042j-mathematics-for-computer-science-fall-2010/video-lectures/lecture-22-expectation-i/>