The Balloon Problem

The problem:

We have a dartboard with n balloons. We flip a fair coin. If the coin comes up heads, we randomly throw a dart at the dartboard. The dart pops a balloon if it hits it, unless that balloon is already popped. A dart will always hit a balloon (popped or not popped) when thrown. We repeat this process every time the coin comes up heads. If the coin comes up tails, we stop flipping the coin, we stop throwing darts, and count how many balloons have been popped.

Given n balloons to begin with, what is the probability of popping x ballons.

Solution:

A critical piece of information we will use is the geometric series:

$$\sum_{a=0}^{\infty}(r)^a = \frac{1}{1-r}, \quad \text{Given } |r| < 1$$

What we will be using is equivalent, but will look more like this:

$$\sum_{a=0}^{\infty}\left(\frac{b}{2n}\right)^a = \frac{1}{1-\frac{b}{2n}} = \frac{2n}{2n-b}$$

This is what it will look like for our specific case, so I am noting it here.

Now, consider that each trial has three possible outcomes, which we will list here:

P_1: The coin is heads and we hit a balloon that has not yet been popped.

P_2: The coin is heads and we hit an already popped balloon.

P_3: The coin is tails, and so we cease throwing darts.

Now, let us determine the probability of each of these outcomes. On the k-th trial, the probability of each outcome is:

$P(P_1) = P(H) \cdot \frac{n-p_k}{n}$

$P(P_2) = P(H) \cdot \frac{p_k}{n}$

$P(P_3) = P(T)$

Where $P(H)$ is the probability of the coin landing heads, $P(T)$ is the probability of the coin landing tails, n is the number of balloons that we start with, and p_k is the number of balloons that have been popped at the beginning of trial k. Clearly, $P(H) = P(T) = \frac{1}{2}$.

Suppose we wished to find the probability of popping 1 balloon across all trials. That is, what's the probability of $P(x = 1)$?

Well, we would first have to flip heads, and pop a balloon, (as if we flip tails, we never throw our first dart, and thus, pop 0 balloons).

So, we could have a sequence of outcomes for each trial that is something like:

$P_1P_2P_2P_2P_3$, and have popped exactly one balloon. Note that the number of P_2 outcomes that we achieve does not matter, but we have to start with an outcome of P_1 and end with an outcome of P_3 without ever having a second P_1 outcome. Let's begin listing possible sequences of trials that satisfy $x = 1$.

P_1P_3

$P_1P_2P_3$

$P_1P_2P_2P_3$

$P_1P_2P_2P_2P_3$

$\ldots$

Note that there are infinitely many sequences that result in popping exactly $x = 1$ balloons.

The probability of each sequence of outcomes is the product of the probability of each outcome. So, the probability of P_1P_3 is $P(P_1) \cdot P(P_3) = \left(P(H) \cdot \frac{n-p_1}{n}\right) \cdot \left(P(T)\right)$.

The probability of popping exactly $x = 1$ balloons is equal to the sum of these products. That is,

$$P(x = 1) = P(P_1P_3) + P(P_1P_2P_3) + P(P_1P_2P_2P_3) + P(P_1P_2P_2P_2P_3) + \cdots$$

However, this is the same as

$$P(P_1)P(P_3) + P(P_1)P(P_2)P(P_3) + P(P_1)[P(P_2)]^2P(P_3) + P(P_1)[P(P_2)]^3P(P_3) + \cdots$$

Notably, we can factor out $P(P_1) \cdot P(P_3)$ and get

$$P(P_1)P(P_3) \cdot \sum_{a=0}^{\infty} [P(P_2)]^a$$

Thus, we find that

$$P(x = 1) = P(P_1)P(P_3) \cdot \sum_{a=0}^{\infty} [P(P_2)]^a$$

Now, noting that $P(P_2) = \frac{1}{2} \cdot \frac{p_k}{n}$ will always be less than 1, as $\frac{p_k}{n} \leq 1$, since the number of popped ballons, p_k, on trial k is at most n, we can safely apply the formula for the geometric series. We will actually perform this application when handling the more general case, but this is far enough to give you the idea for now.

For now, let us introduce some notation that will help us moving forward:

Let's describe the sequences that satisfy $x = 1$:

$$P_1[P_2]^{(a)}P_3$$

We will use this notation to describe the sequences that have the property that they begin with P_1, then have some number a occurrences of P_2 and then end with P_3. Note that a has a range from 0 to ∞.

Now, for a more general case, suppose we wish to find $P(x = 3)$. Then we would describe the sequences that satisfy $x = 3$ as

$$P_1[P_2]^{(a)}P_1[P_2]^{(b)}P_1[P_2]^{(c)}P_3$$

Again, where $a, b,$ and c each range from 0 to ∞.

Similarly to before, the probability of this sequence of outcomes occurring is the product of the probabilities of each outcome, so, we would write this as

$$P(x = 3) = P(P_1)\left[\sum_{a=0}^{\infty} P(P_2)\right]P(P_1)\left[\sum_{b=0}^{\infty} P(P_2)\right]P(P_1)\left[\sum_{c=0}^{\infty} P(P_2)\right]P(P_3)$$

There are a few important things to note about this. Firstly, these $P(P_1)'$s and $P(P_2)'$s are not the same. Since the probability of hitting a balloon that hasn't popped yet is different depending on how many balloons have been popped so far, these values will be different. The $P(P_2)$ is the same within each summation notation, but the $P(P_2)$ within the summation notation with an index of a is going to be of a different value than that of the $P(P_2)$ within the summation notation with an index of b. In this case,

$$P(x=3)=\left(\frac{1}{2}\cdot\frac{n-0}{n}\right)\cdot\left(\sum_{a=0}^{\infty}\frac{1}{2}\cdot\frac{1}{n}\right)\cdot\left(\frac{1}{2}\cdot\frac{n-1}{n}\right)\cdot\left(\sum_{b=0}^{\infty}\frac{1}{2}\cdot\frac{2}{n}\right)\cdot\left(\frac{1}{2}\cdot\frac{n-2}{n}\right)$$
$$\cdot\left(\sum_{c=0}^{\infty}\frac{1}{2}\cdot\frac{3}{n}\right)\cdot\left(\frac{1}{2}\right)$$

This is simply by plugging in the probabilities for P_1, P_2, and P_3, noting that the calculation of these probabilities is different depending on where we are in the sequence. That is, the first time we calculate $P(P_1)$ we get $\left(\frac{1}{2}\cdot\frac{n-0}{n}\right)$ as there are 0 popped balloons. The second time we calculate the probability of $P(P_1)$ we find it to be $\left(\frac{1}{2}\cdot\frac{n-1}{n}\right)$, as now there is a popped balloon, namely, the one we popped on our first trial. Similarly, after we've popped the first balloon, $P(P_2)=\left(\frac{1}{2}\cdot\frac{1}{n}\right)$ as we've only popped one balloon so far. After a trials of hitting this same, already popped balloon, we eventually hit a different balloon. Now, with two popped balloons, when we calculate $P(P_2)$ for the b trials where we hit an already popped balloon, we find it to be $P(P_2)=\left(\frac{1}{2}\cdot\frac{2}{n}\right)$.

Finally, we will use the geometric series formula to determine what the probability, exactly, is.

$$P(x=3)=\frac{n}{2n}\cdot\frac{n-1}{2n}\cdot\frac{n-2}{2n}\cdot\frac{1}{2}\cdot\left(\sum_{a=0}^{\infty}\frac{1}{2}\cdot\frac{1}{n}\right)\cdot\left(\sum_{b=0}^{\infty}\frac{1}{2}\cdot\frac{2}{n}\right)\cdot\left(\sum_{c=0}^{\infty}\frac{1}{2}\cdot\frac{3}{n}\right)$$

$$P(x=3)=\frac{1}{2}\cdot\frac{(n)(n-1)(n-2)}{(2n)^3}\cdot\left[\left(\frac{2n}{2n-1}\right)\cdot\left(\frac{2n}{2n-2}\right)\cdot\left(\frac{2n}{2n-3}\right)\right]$$

Noting that we have $(2n)^3$ in the numerator and denominator, and that $n(n-1)(n-2)=\frac{n!}{(n-3)!}$, we find:

$$P(x=3)=\frac{1}{2}\cdot\frac{n!}{(n-3)!\,(2n-1)(2n-2)(2n-3)}$$

It is not difficult to see that, continuing this process with greater values of x will yield similar results. For the most general case, fix $x \in \mathbb{Z}^+$.

So, the sequence which satisfies $P(x)$ can be written:

$$P_1[P_2]^{(a_1)}P_1[P_2]^{(a_2)}P_1[P_2]^{(a_3)} \dots P_1[P_2]^{(a_x)}P_3$$

Note that there will be exactly x occurrences of P_1 in this sequence.

Now, in a similar manner to the previous cases,

$$P(x) = \frac{n}{2n} \cdot \frac{n-1}{2n} \cdot \frac{n-2}{2n} \cdot \dots \cdot \frac{n-(x-1)}{2n}$$
$$\cdot \left[\left(\sum_{a_1=0}^{\infty} \left(\frac{1}{2n}\right)^{a_1}\right) \cdot \left(\sum_{a_2=0}^{\infty} \left(\frac{2}{2n}\right)^{a_2}\right) \cdot \dots \cdot \sum_{a_x=0}^{\infty} \left(\frac{x}{2n}\right)^{a_x}\right] \cdot \left(\frac{1}{2}\right)$$

Now, applying the geometric series formula,

$$P(x) = \left(\frac{1}{2}\right) \cdot \frac{n!}{(n-x)!} \cdot \frac{1}{(2n)^x} \cdot \left[\frac{2n}{2n-1} \cdot \frac{2n}{2n-2} \cdot \frac{2n}{2n-3} \cdot \dots \cdot \frac{2n}{2n-x}\right]$$

Which, again, we have x $2n'$s in the numerator, and x $2n'$s in the denominator, so, canceling those out we end up with:

$$P(x) = \frac{1}{2} \cdot \frac{n!}{(n-x)!} \cdot \frac{1}{(2n-1)(2n-2)(2n-3) \dots (2n-x)}$$

Or,

$$P(x) = \frac{1}{2} \cdot \frac{(n)(n-1)(n-2) \dots \big(n-(x-1)\big)}{(2n-1)(2n-2)(2n-3) \dots (2n-x)}$$

If you prefer.

Another way to view the same problem:

Suppose we have a bag with scrabble pieces in it. The bag starts off with n pieces with an a on them, and n pieces with a z on them. Now, on each trial, we reach into the bag and pull out a piece. Whatever piece we retrieve, we lay it down, adding it to our “word” that we have constructed so far. We also replace that piece that was in the bag with a new piece, one that has a b on it. We repeat this process, forever, until we retrieve a z. When we retrieve a z, we add it to our "word" and cease pulling pieces from the bag. Note that at all times the bag has $2n$ pieces in it, and that n of those pieces are always z's. What is the probability of constructing a "word" with exactly x a's in it.

This problem is equivalent to the balloon problem, but may help view the outcomes in a sequence. Similar to what we found for the balloon problem, we can construct the most general sequence as

$$a[b]^{(k_1)}a[b]^{(k_2)}a[b]^{(k_3)} \dots a[b]^{(k_x)}z$$

Similarly,

$P(a) = P(P_1) = \frac{n-p_k}{2n}$

$P(b) = P(P_2) = \frac{p_k}{2n}$

$P(z) = P(P_3) = \frac{n}{2n} = \frac{1}{2}$

So, similarly,

$$P(x) = \frac{1}{2} \cdot \frac{n!}{(n-x)!} \cdot \frac{1}{(2n-1)(2n-2)(2n-3) \dots (2n-x)}$$