Probability

The theory of probability is a system for making more educated guesses or hypotheses.

Take for instance the following

$$P(A) = \frac{n_A}{n} \quad (0)$$

or

$$P_A = \frac{n_A}{n}$$

where n_A = our estimate of the expected number of observations that will yield event A.

n = total number of observations performed for the experiment

This formula will read given n observations we expect n_A number of observations to yield the result A.

For example given a fair coin (same chance of getting head or tail)

For 30 observations → $n = 30$.

We would rationally expect 15 heads & 15 tails.

or written in statistics

$$P(H) = \frac{15 \text{ Heads}}{30 \text{ observations}} = 0.5 \qquad P(T) = \frac{15 \text{ tails}}{30 \text{ observations}} = 0.5$$

This can also be read as: there is a 0.5 probability or 50% chance that if I toss a coin I am going to get a head to show up.

So, say we do this experiment in real life, it should work right?

Experiment 1:

5 trials of 30 tossing observations each. ($N=30$)

what do we get?

Trial 1	Trial 2	Trial 3	Trial 4	Trial 5.
12 Heads (18 Tails)	14 Heads (16 Tails)	15 Heads (15 Tails)	17 Heads (13 Tails)	16 Heads (14 Tails)

But yet it doesn't why?

because probability is <u>not</u> telling you that you would get the expected value everytime, it is telling you that the <u>most likely</u> value you will get will be the expected value (The value with the greatest probability)

now what about if we average the values over all the trials

$$\frac{12+14+15+17+16}{30+30+30+30+30} = \frac{74}{150} = 0.493 \quad \text{so } P'(H) = 0.493$$

then we see our probability looks closer to what we would expect or $P(H) = 0.5$. thus it is only as we let N or the number of observations approach infinity ($N \to \infty$) that we will get a perfect probability match. However for large values of N the probability will be close enough to our expected value & will be sufficient for practical purposes.

now we have another question, can we know the probability of say k events (e.g. 2 heads) occurring in say a total of N. tosses.

yes we can! to understand the process for doing this let's do a quick experiment

Experiment 2.

3 tosses, what combinations are possible?

Experiment 2: (results)

Combo#	toss 1	toss 2.	toss 3.
1	h	h	h
2	h	h	t
3	h	t	h
4	h	t	t
5	t	h	h
6	t	h	t
7	t	t	h
8	t	t	t

So as can be seen out of the total number of combinations (8) there are only 3 combinations (#2, #3, #5) that will give us perfectly 2 heads.

Thus the probability of getting 2 heads in 3 observations (3 "coin" tosses) would be

$$P_h(k,n) = P_h(2,3) = \frac{3}{8} = 0.375 \text{ (or } 37.5\%)$$

This is read as the probability of getting exactly 2 heads in 3 coin tosses is 0.375.

This tells us that

$$P_a(k,n) = \frac{\text{number of possible combinations with outcome "a" repeated "k" times}}{\text{total number of possible combinations}}$$

also as can be seen we could have also come up with the total number of combinations by: $2 \times 2 \times 2 = 2^3 = \text{states}^{\text{tosses}} = s^n$ (tosses)

faces (2)

This is also written as:

for example 2.

$$P_a(k,n) = \frac{\binom{n}{k}}{s^t} = \frac{\frac{n!}{k!(n-k)!}}{s^t} = \frac{\frac{3!}{2!(3-2)!}}{2^3} = \frac{\frac{6}{2}}{8} = \frac{3}{8} \checkmark \quad (1)$$

$\binom{n}{k}$ are also sometimes called the binomial coefficients & can be computed also from the binomial expansions or the pascal's triangle

e.g. $(a+b)^n = (a+b)^3 = \underset{(k=0)}{a^3} + \underset{(k=1)}{3a^2b} + \underset{(k=2)}{3ab^2} + \underset{(k=3)}{b^3}$ $\therefore \binom{n}{k} = 3 \checkmark$

binomial expansion

note that using this logic we can solve for the probability of any phenomenon that supports 2 states (with equal probability i.e. $P_t = P_h$) as the outcome (e.g. coin tossing, win/loss, binary combinations ... etc)

for example. →

for example someone "clever" might ask
what is the probability of getting one head if we throw two coins 2 times?
but after thinking about it for a bit you realize, this is the same think we have been doing.
To make it clear, you reword the problem:
This is the same as asking:
what is the probability of getting exactly one head if you throw a single coin 4 times.
& this just

$$P_a(k,n) = \frac{\text{number of possible combinations with outcome "h" only once}}{\text{total number of possible combinations}}$$

$$P_h(1,4) = \frac{\frac{n!}{k!(n-k)!}}{f^t} = \frac{\frac{4!}{1!(3)!}}{2^4} = \frac{\frac{24}{6}}{16} = \frac{\text{4 possible combinations}}{\text{16 total combinations}}$$

or also just the probability of getting a head in the first toss of 2 coins × the probability of getting a head in the second toss of 2 coins or:

$$P_h(1,4) = P_{h_1} \times P_{h_2} = \frac{\overset{(ht,th)}{\text{2 possible combinations}}}{\text{4 total cominations}} \times \frac{\overset{(ht,th)}{\text{2 possible combinations}}}{\text{4 total cominations}}$$

$$= \frac{\text{4 possible combinations}}{\text{16 total combinations.}}$$

then they might say what about the probability of getting at least one head in 3 single coin tosses
but you breaking up the problem into little pieces realize this is straightforward

Prob of at least 1 head = Probability of getting exactly one head
+ Probability of getting exactly 2 heads
+ Probability of getting exactly 3 heads

$$= P_h(1,3) + P_h(2,3) + P_h(3,3)$$

and this is easy, we have done similar problems already !!

there is also another way to arrive to formula (1) intuitively that leads us to our next subject.

to see this let's try to arrive with the results of "experiment 2" in another way.

we know we want to find out:

the probability of getting exactly 2 heads to show up. over 3 tosses

we also now that there are three possible combinations (#2, #3, #5) that could get us our desired outcome

However, the thing we don't know is the probability of getting one of these combinations to show up.

But this is easy to calculate! i.e.

Probability of getting combination #2

$P_{\#2} = P_h \times P_h \times P_t = 0.5 \times 0.5 \times 0.5 = 0.125$

Probability of getting combination #3

$P_{\#3} = P_h \times P_t \times P_h = 0.5 \times 0.5 \times 0.5 = 0.125$

Probability of getting combination #5

$P_{\#5} = P_t \times P_h \times P_h = 0.5 \times 0.5 \times 0.5 = 0.125$

So then we can express the probability of getting exactly 2 heads as.

$P_h(2,3) =$ 3 possible combinations $\times$ 0.125 probability of occurrence per combination.

$= 3 \times 0.125 = 0.375$

which is exactly what we got through the other technique!

So we can re-write formula 1 as:

$$P_a(k,n) = \binom{n}{k}(\text{probability of event a})^k \times (\text{probability of event b})^{n-k}$$

$$= \binom{n}{k} P_a^k \times P_b^{n-k}$$ (this is called the bernoulli equation or binomial probability

now the situation gets a bit more interesting is a coin is loaded or if a game is partially fixed

That is, if we let the probability of winning a game be different than the probability of lossing a game.
(or the probability of obtaining a head be different than the probability of obtaining a tail, same thing)

what would happen then? Let's do a quick experiment

Experiment 3.

let the prob of winning a game in the next try be 0.6
& the probability of lossing the game be 0.4
& let's play 3 games. again:

we want to know the probability that I will win exactly 2 games in the 3 tries

i.e. $n = 3$ (3 games)
$K = 2$ (2 games won)

now, clearly the combinations displayed in the previous page would remain unchanged, what would change is that some of this combinations would now occurr more often than others.

Experiment 2: (results)

Combo#	toss 1	toss 2.	toss 3.
1	W	W	W
2	W	W	L
3	W	L	W
4	W	L	L
5	L	W	W
6	L	W	L
7	L	L	W
8	L	L	L

i.e. Intuitively we know that combination number 1 would occur very often & combination 8 would occur very little, same follows for say combination 3 (more often) & combination 4 (less often)

but we already have the tools to solve this.

we have 3 possible combinations that could give us 2 heads, so we can calculate their probabilities just like before.

#2 $P_{h_1} = P_W \times P_W \times P_L = 0.6 \times 0.6 \times 0.4 = 0.144$

#3 $P_{h_2} = P_W \times P_L \times P_W = 0.6 \times 0.4 \times 0.6 = 0.144$

#5 $P_{h_3} = P_L \times P_W \times P_W = 0.4 \times 0.6 \times 0.6 = 0.144$

So from our intuitive derivation before:

$P_h(2,3)$ = 3 possible combinations × 0.144 probability of occurrence per combination.

$= 3 \times 0.144 = 0.432$

or using the mathematical description of our logic (Bernoulli equation, we will get the same thing)

$$P_h(2,3) = \binom{n}{k} \times P_W^k P_L^{n-k} = 3 \times 0.6^2 \times 0.4 = 0.432 \checkmark$$

neat!