UNIFORM PROBABILITY SPACES – Equally Likely Outcomes

Let S be a finite sample space.

S is called a uniform probability space if each outcome has the same probability.

Let $\omega \in S$, then $Pr(\omega) = \frac{1}{|S|}$.

Observe that if (S, Pr) is a uniform probability space, then for any event $A \subseteq S$,

$$Pr(A) = \sum_{\omega \in A} Pr(\omega) = \frac{|A|}{|S|}.$$

Example 1: Roll a fair die, observe the outcome.

$S = \{1, 2, 3, 4, 5, 6\}$

Each outcome is equally likely and hence Probability for each of them is $1/6$

Let A be the event of obtaining a number divisible by 2.

i.e. $A = \{2, 4, 6\}$

Thus $Pr(A) = \frac{|A|}{|S|} = \frac{3}{6} = \frac{1}{2}$.

Example 2: Probability of obtaining a full house

52 cards; subdivided in

4 suits (Hearts, Spades, Clubs, Diamonds)
13 ranks (A, K, Q, J, 10, 9, 8, 7, 6, 5, 4, 3, 2)

A hand of 5 cards is called a Full House if

- 3 cards are of same rank
- remaining 2 cards are also of same rank, but different than the first 3.

e.g of full house includes

→ {Q-S, Q-H, Q-D, 3-H, 3-C} → {10S, 10D, 10C, 5C, 5H}

→ {5S, 5D, 5C, 3H, 3D}

Assume each hand of 5 cards is equally likely,

What is the probability of getting a full house?

Note that there are $\binom{52}{5}$ possible hands of 5 cards, and each one is equally likely.

Thus $|S| = \binom{52}{5}$ and probability of obtaining a particular hand is $\dfrac{1}{\binom{52}{5}}$.

Define event A to be set of all full houses in S.

Then
$$Pr(A) = \frac{|A|}{|S|}.$$

Question: what is $|A|$, i.e, how many full houses are there?

We will use product rule to determine the number of full houses.

Procedure: choose a full house.

1st Task: choose the rank of 3 cards in the full house. $N_1 = 13$ as there are 13 possible ranks.

2nd Task: choose the suits of these 3 cards. $N_2 = \binom{4}{3}$ as we need to choose 3 suits out of 4.

3rd Task: Choose the rank of remaining 2 cards. $N_3 = 12$ as there are 12 ways to choose the rank. Only restriction is that we can't choose rank to be what we chose in the 1st task.

4th Task: choose the suits of these 2 cards
$N_4 = \binom{4}{2}$

Thus $|A| = \#$ ways to execute the procedure

$$= N_1 N_2 N_3 N_4 = 13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2} = 3{,}744.$$

Thus $Pr(A) = \frac{3744}{\binom{52}{5}} = \frac{3{,}744}{2{,}598{,}960} \approx 0.00144$ ∎

BIRTHDAY PARADOX

Consider a set of $n \geq 2$ people $P_1, P_2, \ldots, P_n$.

Let d be the number of days in a year and label the days as $1, 2, 3, \ldots, d$.

Consider the sample space

$$S_n = \{ (b_1, b_2, \ldots, b_n) \mid b_i \in \{1, 2, 3, \ldots, d\} \text{ for } 1 \leq i \leq n, \text{ where } b_i \text{ represents Date of Brith of } P_i \}$$

$|S_n| = d^n$.

Assume that anybody can be born on anyday, then each outcome $(b_1, b_2, \ldots, b_n)$ is equally likely.

Thus $\Pr(b_1, b_2, \ldots, b_n) = \dfrac{1}{|S_n|} = \dfrac{1}{d^n}$.

Consider the event $A_n \subseteq S_n$ where at least two people have birthday on the same day, i.e., $(b_1, b_2, \ldots, b_n)$, contains duplicates.

Thus $A_n = \{ (b_1, b_2, \ldots, b_n) \in S_n : b_1, b_2, \ldots, b_n \text{ contains duplicates} \}$.

Define $p_n = \Pr(A_n)$.

Note that if $n > d$, then $p_n = 1$ as we have more people than the # days in a year, and hence by pigeon hole principle, at least two people will have same birthday.

Assume $n \le d$.

* Evaluate $p_2 = Pr(A_2) = \dfrac{|A_2|}{|S_2|}$

$|S_2| = d^2$

The event $A_2 = \{(1,1), (2,2), (3,3), \text{----}, (d,d)\} \subseteq S_2$.

Thus $|A_2| = d$.

and $p_2 = \dfrac{|A_2|}{|S_2|} = \dfrac{d}{d^2} = \dfrac{1}{d}$.

* Let us evaluate p_n for a bit larger values of n.

Consider the complementary event

$$\bar{A}_n = \{(b_1, b_2, \dots, b_n) \in S_n \mid b_1, b_2, \dots, b_n \text{ are pairwise distinct}\}.$$

Therefore, $\bar{A}_n$ is set of all ordered sequences consisting of n pairwise distinct elements $\{1, 2, \dots, d\}$.

$$|\bar{A}_n| = \frac{d!}{(d-n)!}$$

$\Rightarrow$ For any n, $2 \le n \le d$,

$$p_n = Pr(A_n) = 1 - Pr(\bar{A}_n) = 1 - \frac{|\bar{A}_n|}{|S_n|}$$

$$= 1 - \frac{d!}{(d-n)!\, d^n}$$

Set $d = 365$ (# days in a non-leap year)

$p_{22} \equiv 0.476$

$p_{23} \equiv 0.507$ ← There is 50% chance that in a random collection of 23 people, two people will be born on same day.

$p_{40} \equiv 0.891$

$p_{100} \equiv 0.9999997$ ← In a random collection of 100 people, it is extremely likely that two will have same birthday.

—×—

BALLS INTO BINS

Consider d-bins: $B_1, B_2, B_3, B_4, \ldots, B_d$, where d is a large integer.

Experiment: Throw n balls one by one, where each ball lands in a random bin (with equal probability).

Let $p_n \equiv$ Probability that at least one bin contains at least two or more balls.

From the birthday paradox (Bins are days and balls are people) we know that

$$\boxed{p_n = 1 - \frac{d!}{(d-n)!\, d^n}}$$

RECALL: Recall from somewhere that for any real number x,

$$1 - x \leq e^{-x}.$$

Define $q_n = 1 - p_n \equiv \dfrac{d!}{(d-n)!\, d^n}$

$$= \frac{d\,(d-1)\,(d-2)\cdots(d-(n-1))}{d^n} \; \frac{(d-n)!}{(d-n)!}$$

$$= \frac{d}{d} \cdot \frac{d-1}{d} \cdot \frac{d-2}{d} \cdots\cdots \frac{d-(n-1)}{d}$$

$$q_n = \left(1-\frac{1}{d}\right)\left(1-\frac{2}{d}\right) \cdots\cdots \left(1-\frac{n-1}{d}\right)$$

$$\leq e^{-\frac{1}{d}} e^{-\frac{2}{d}} e^{-\frac{3}{d}} \cdots\cdots e^{-\frac{n-1}{d}}$$

$$= e^{-\frac{(1+2+3+\cdots+n-1)}{d}}$$

$$= e^{-\frac{n(n-1)}{2d}} \quad \text{as} \quad 1+2+3+\cdots+n-1 = \frac{n(n-1)}{2}.$$

$$\Rightarrow \quad q_n \leq e^{-\frac{n(n-1)}{2d}}$$

or $p_n = 1 - q_n \geq 1 - e^{-\frac{n(n-1)}{2d}}$

If n is large than $n(n-1) \sim n^2$

Thus $p_n \geq 1 - e^{-n^2/2d}$

Choose $n = \sqrt{2d}$,

$$p_n \gtrsim 1 - e^{-1} = 0.632$$

Conclusions: For large values of d, if we throw $\sqrt{2d}$ balls into d bins, than with $pr \geq 1 - \frac{1}{e}$, there is a bin that contains ~~moreet~~ more than one ball.

MONTY HALL PROBLEM.

Input: 3 closed doors,
A big prize behind one door.

No prizes behind the other two doors.

Problem: You are asked to choose a door, but it is not opened yet.

The host opens one of the remaining doors and reveals that that door has no prize

~~Que~~ Then the host asks, do you want to switch your door?

Question: Do you stay with your first choice, or should you switch?

We will show that probability of success is ~~at least~~ 2/3 if you switch.

Pf: Probability that car is behind any of the doors is 1/3.

Suppose you choose the first door.

Note that sample space consists of

{PNN, NPN, NNP}

prize no prize no prize

Host opens one of the doors among doors 2 or 3 which does not contain a prize

	Door 1	Door 2	Door 3	
PNN	P	N	$\boxed{N}$	— I
NPN	N	P	$\boxed{N}$	— II
NNP	N	$\boxed{N}$	P	— III

Your pick

Depending on the three members of the sample space, Host opens either Door 2 or Door 3.

Notice that if you switch from Door 1 to Door 2 in I you loose, but in II the switch helps you to win, and also in case III you win.

Thus there is $2/3$ chance that you will win by switching.

Basically by switching you are choosing 2 doors in place of one, and hence Pr of success $\equiv 2/3$.

A more convincing version.

Suppose we do the same game with 10 doors, and prize behind one of the doors.

Suppose you choose a door. (lets say door 1)

Host reveals 8 doors among doors 2 - 10 that do not contain prizes.

Should you switch?

Answer is Yes, as there is only $1/10$th chance that the prize is at door 1 (your initial choice) v/s $9/10$th chance that prize is among Doors 2 - 10.

Thus, you should switch.

BIG BOX PROBLEM

Input: Two boxes, one containing \$$x$ and other one containing \$$y$, where

$0 \le x < y \le 100$, but we do not know which one contains x and which one contains y.

Also we do not know x or y.

Let box containing x\$ be small box.

Let box containing y\$ be large box.

GOAL: Find the big box by performing the following operations

1. Choose a box, open it and count the amount of money it has
2. Final Decision: keep the open box or take the other box.

To Show: We can find the big box with $Pr \ge 0.505$.

Idea: Suppose we know z, such that

$x < z < y$.

After Step 1, if money in open box is $> z$, we keep it, otherwise we switch to the other box.

$\boxed{\text{Given } z, \text{ we can always find the big box.}}$

BUT we do not know z, neither do we know x or y.

TRICK: Choose z at random, i.e choose z from the set $B = \{1/2, 3/2, 5/2, \ldots, 100 - 1/2\}$, where each choice is equally likely.

(By choosing z to be fraction, $z \neq x$ or $z \neq y$).

ALGORITHM FINDBIGBOX

Step 1: Choose one of the two boxes uniformly at Random.

Open it and count money in it.

Let this amount be a.

Step 2: Choose z uniformly at random from set B.

Step 3:

If $a \geq z$, keep the box chosen in Step 1.

Otherwise, choose the other box.

Sample space

$S = \{(a, z) ; a \in \{x, y\} \text{ and } z \in B\}$

$|S| = 2 \times 100 = 200$ as $|\{x,y\}| = 2$ & $|B| = 100$.

Define $W \subseteq S$ be the event that FINDBIGBOX is successful.

$W = \{(a,z) \in S \mid \text{FIND BIGBOX is succesful}\}$

First consider the case when $a = x$.

There are two possibilities for z

- If $x = a > z$, then algo keeps small box and fails.
- If $x = a < z$, then algo takes the big box and succeds.

Define $W_x = \{(x, z) \mid z \in \{x + 1/2, x + 3/2, \dots, 100 - 1/2\}\}$

$|W_x| = 100 - x$

Now consider the case when $a = y$.

- If $y = a > z$, algo succeds.
- If $y = a < z$, algo fails.

Define
$$W_y = \{ (y, z) \mid z \in \{ 1/2, 3/2, \dots, y - 1/2 \} \}.$$

$$|W_y| = y.$$

Note that $W = W_x \cup W_y$

Since $W_x \cap W_y = \phi$,

$$\Pr(W) = \Pr(W_x) + \Pr(W_y)$$
$$= \frac{|W_x|}{|S|} + \frac{|W_y|}{|S|}$$
$$= \frac{100 - x}{200} + \frac{y}{200}$$
$$= \frac{100}{200} + \frac{y - x}{200}$$

$$\Pr(W) \geq \frac{1}{2} + \frac{1}{200}$$ Since $y > x$ or $y - x \geq 1$

$$= 0.505 \quad \blacksquare$$