Birthday Problem

Why Important ?

Analogy to!
- Encryption → Security
- Queues → Process threading
- Hashing → Dictionaries Key/Value
- Filtering → Data Compression
- Games → Collisions

Concepts

Counting

$$ {}_N P_m^r = N^m \quad , \quad {}_N P_m = \frac{N!}{(N-m)!} $$

$$ \sum_{i=1}^{N} i = \frac{N(N+1)}{2} $$

Approximations

$$ e^x = 1 + x/1! + x^2/2! + x^3/3! \cdots . $$

$$ e^x \approx (1+x) + O(x^2) \leftarrow \text{H.O.T} $$

Functions

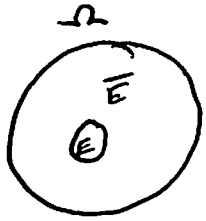
$$ a = D^x \qquad b = D^y $$

$$ a \cdot b = D^{x+y} $$

$$ x + y = \log_D a \cdot b $$

A product is a sum with logarithms

Probabilities



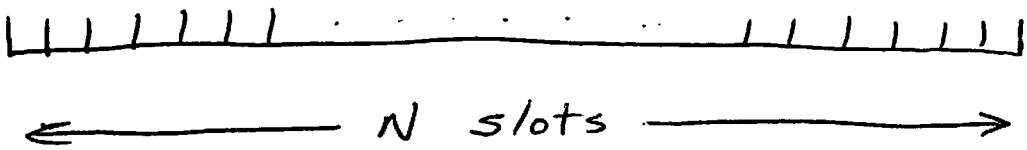
$$ E, \bar{E} \subseteq \Omega \qquad 0 \le P(E) = \frac{\eta(E)}{\eta(\Omega)} \le 1 $$

$$ E \cup \bar{E} = \Omega $$

$$ E \cap \bar{E} = 0 \qquad P(E) + P(\bar{E}) = 1 $$

MEL

Setup



N slots

m balls randomly dropped into slots

By analogy N days in a year → 365

m students in CSC 7

Question: How many students have the same birthday?

Define Events

$C \rightarrow$ Collision, Students have same BD

$\bar{C} \rightarrow$ No Collision, No Student same BD

$\Omega \rightarrow \bar{C} \cup C$, All possible scenarios

Calculations

$$n(\Omega) = {}_N P_m^r = N^m$$

$$n(\bar{C}) = {}_N P_m = N!/(N-m)!$$

$$P(\bar{C}) = (N!/(N-m)!)/N^m = \frac{n(\bar{C})}{n(\Omega)}$$

MEK

Simplification

$$P(\bar{C}) = \frac{N}{N} \cdot \frac{(N-1)}{N} \cdot \frac{(N-2)}{N} \cdot \ldots \cdot \left(\frac{N-m+1}{N}\right)$$

$$= 1 \cdot (1-1/N) \cdot (1-2/N) \cdot \cdots \left(1 - \frac{m-1}{N}\right)$$

Using $e^{-i/N} \times (1 - i/N) + H.O.T.$

$$\approx e^{-0/N} \cdot e^{-1/N} \cdot e^{-2/N} \cdots e^{-(m-1)/N}$$

$$= \prod_{i=0}^{m-1} e^{-i/N}$$

Using products of exponents are sums

$$= e^{-1/N \sum_{i=0}^{m-1} i}$$

$$= e^{-(m-1)m/2N} \approx e^{-m^2/2N}$$

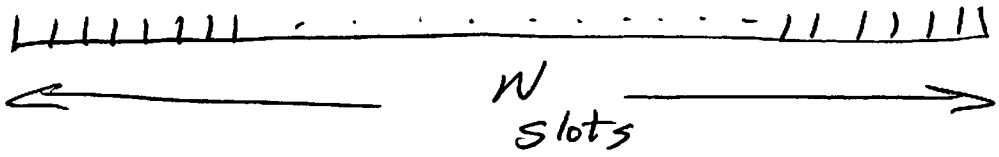
Where $P(\bar{C}) = P(C) = 1/2$

$$1/2 \approx e^{-m^2/2N}$$

$$m = \sqrt{2N \log_e 2}$$

MEL

Same setup



m balls randomly choosen

Calculated the sum!

$$E(\Sigma_m) = \sum_{i=1}^{m} x_i P(x_i)$$

$$= \sum_{i=1}^{m-1} x_i P(x_i) + X_m P(X_m)$$

$$= E(\Sigma_{m-1}) + X_m P(X_m)$$

$$= E(\Sigma_{m-1}) + 1 \cdot \frac{N - E(\Sigma_{m-1})}{N}$$

$$= \frac{(N-1)}{N} E(\Sigma_{m-1}) + 1$$

Recursive Sequence

Non-Recursive closed-form solution

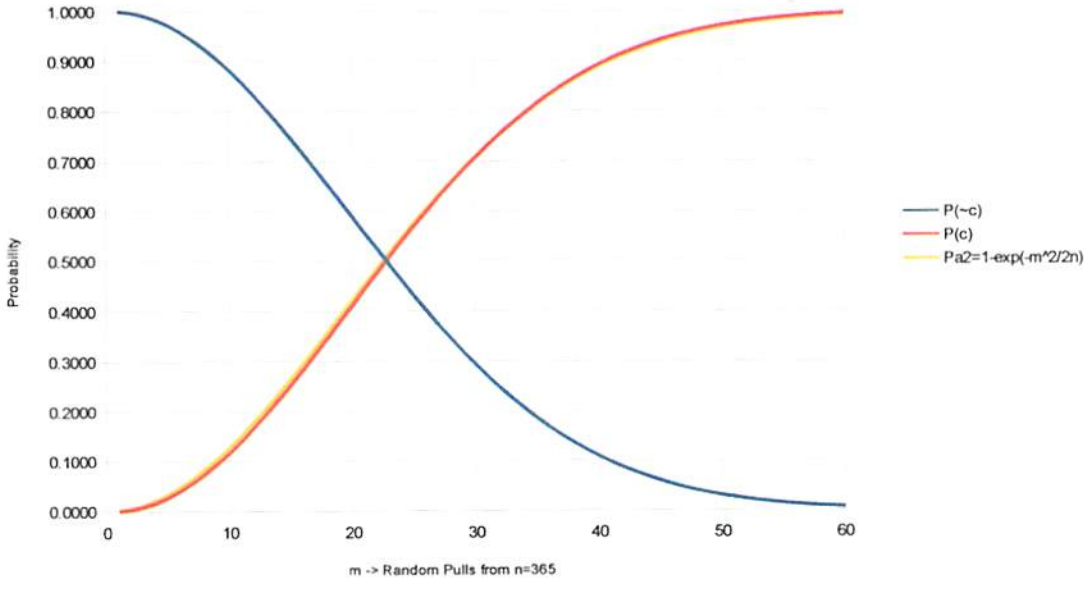
$$E(0) = 0$$
$$E(1) = 1$$
$$E(2) = (2N-1)/N$$
$$E(3) = (3N^2 - 3N + 1)/N^2$$
$$E(4) = (4N^3 - 6N^2 + 4N - 1)/N^3$$
$$E(5) = (5N^4 - 10N^3 + 10N^2 - 5N + 1)/N^4$$
$$\vdots$$
$$E(m) = \left(\sum_{i=1}^{m} {}_N C_i \, N^{m-i} (-1)^{i+1} \right) / N^{m-1}$$

or

$$E(m) = \left(\sum_{i=1}^{m} \binom{N}{i} N^{m-i} (-1)^{i+1} \right) / N^{m-1}$$

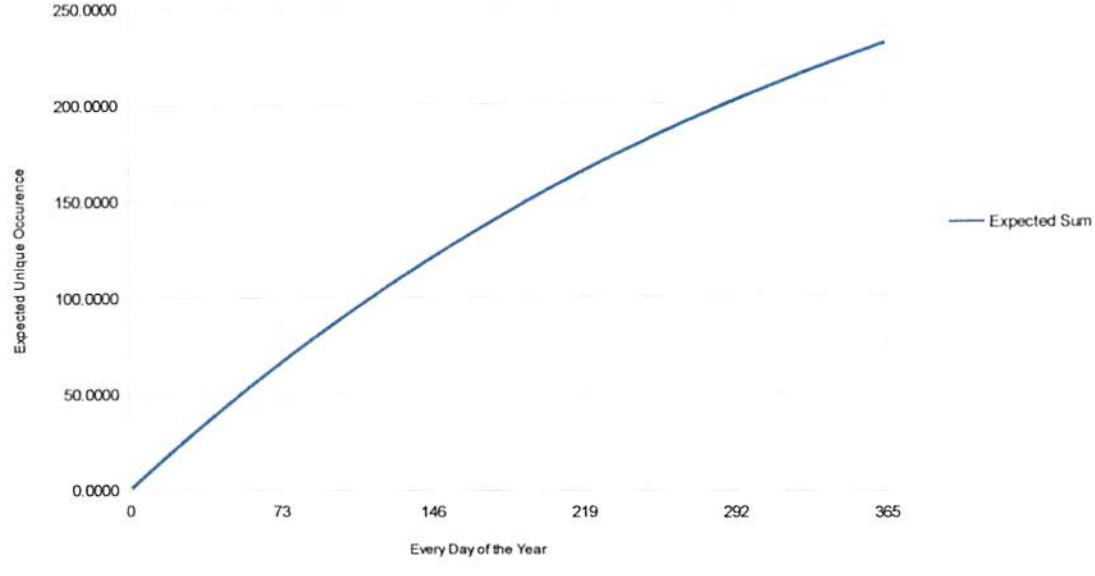
Birthday Problem or

Collisions are Bound to Occur



Number of Unique Birthdays

Up to the days of the year



The Birthday Problem

Days/Slots of the Year n-m	Conditional Probability of no collision delta P	Probability of no collision P(~c)	Probability of a collision P(c)	Expected Sum E(sum)	Approx Probability of no collision Pa_1=exp(-m*(m-1)/2n)	Further Approximation Probability of a C Pa_2=1-exp(-m^2/2n)
1	1.0000	1.0000	0.0000	1.0000	1.0000	0.0014
2	0.9973	0.9973	0.0027	1.9973	0.9973	0.0055
3	0.9945	0.9918	0.0082	2.9918	0.9918	0.0123
4	0.9918	0.9836	0.0164	3.9836	0.9837	0.0217
5	0.9890	0.9729	0.0271	4.9727	0.9730	0.0337
6	0.9863	0.9595	0.0405	5.9591	0.9597	0.0481
7	0.9836	0.9438	0.0562	6.9427	0.9441	0.0649
8	0.9808	0.9257	0.0743	7.9237	0.9262	0.0839
9	0.9781	0.9054	0.0946	8.9020	0.9061	0.1050
10	0.9753	0.8831	0.1169	9.8776	0.8840	0.1280
11	0.9726	0.8589	0.1411	10.8505	0.8601	0.1527
12	0.9699	0.8330	0.1670	11.8208	0.8346	0.1790
13	0.9671	0.8056	0.1944	12.7884	0.8076	0.2067
14	0.9644	0.7769	0.2231	13.7534	0.7793	0.2355
15	0.9616	0.7471	0.2529	14.7157	0.7500	0.2652
16	0.9589	0.7164	0.2836	15.6754	0.7198	0.2958
17	0.9562	0.6850	0.3150	16.6325	0.6889	0.3269
18	0.9534	0.6531	0.3469	17.5869	0.6576	0.3584
19	0.9507	0.6209	0.3791	18.5387	0.6259	0.3901
20	0.9479	0.5886	0.4114	19.4879	0.5942	0.4219
21	0.9452	0.5563	0.4437	20.4345	0.5625	0.4534
22	0.9425	0.5243	0.4757	21.3785	0.5311	0.4847
23	0.9397	0.4927	0.5073	22.3200	0.5000	0.5155
24	0.9370	0.4617	0.5383	23.2588	0.4695	0.5457
25	0.9342	0.4313	0.5687	24.1951	0.4396	0.5752
26	0.9315	0.4018	0.5982	25.1288	0.4105	0.6039
27	0.9288	0.3731	0.6269	26.0600	0.3823	0.6316
28	0.9260	0.3455	0.6545	26.9886	0.3550	0.6584
29	0.9233	0.3190	0.6810	27.9146	0.3288	0.6840
30	0.9205	0.2937	0.7063	28.8381	0.3037	0.7085
31	0.9178	0.2695	0.7305	29.7591	0.2797	0.7319
32	0.9151	0.2467	0.7533	30.6776	0.2569	0.7541
33	0.9123	0.2250	0.7750	31.5935	0.2354	0.7750
34	0.9096	0.2047	0.7953	32.5070	0.2150	0.7948
35	0.9068	0.1856	0.8144	33.4179	0.1959	0.8133
36	0.9041	0.1678	0.8322	34.3264	0.1780	0.8306
37	0.9014	0.1513	0.8487	35.2323	0.1613	0.8467
38	0.8986	0.1359	0.8641	36.1358	0.1457	0.8617
39	0.8959	0.1218	0.8782	37.0368	0.1313	0.8755
40	0.8932	0.1088	0.8912	37.9353	0.1180	0.8883
41	0.8904	0.0968	0.9032	38.8314	0.1058	0.9000
42	0.8877	0.0860	0.9140	39.7250	0.0945	0.9108
43	0.8849	0.0761	0.9239	40.6162	0.0842	0.9206
44	0.8822	0.0671	0.9329	41.5049	0.0749	0.9295
45	0.8795	0.0590	0.9410	42.3912	0.0664	0.9376
46	0.8767	0.0517	0.9483	43.2750	0.0587	0.9449
47	0.8740	0.0452	0.9548	44.1565	0.0517	0.9515
48	0.8712	0.0394	0.9606	45.0355	0.0455	0.9574
49	0.8685	0.0342	0.9658	45.9121	0.0399	0.9627
50	0.8658	0.0296	0.9704	46.7863	0.0349	0.9674
51	0.8630	0.0256	0.9744	47.6582	0.0304	0.9716
52	0.8603	0.0220	0.9780	48.5276	0.0264	0.9754
53	0.8575	0.0189	0.9811	49.3946	0.0229	0.9787
54	0.8548	0.0161	0.9839	50.2593	0.0198	0.9816
55	0.8521	0.0137	0.9863	51.1216	0.0171	0.9841
56	0.8493	0.0117	0.9883	51.9815	0.0147	0.9864
57	0.8466	0.0099	0.9901	52.8391	0.0126	0.9883
58	0.8438	0.0083	0.9917	53.6944	0.0108	0.9900
59	0.8411	0.0070	0.9930	54.5473	0.0092	0.9915
60	0.8384	0.0059	0.9941	55.3978	0.0078	0.9928
61	0.8356	0.0049	0.9951	56.2460	0.0066	0.9939
62	0.8329	0.0041	0.9959	57.0919	0.0056	0.9948
63	0.8301	0.0034	0.9966	57.9355	0.0047	0.9956
64	0.8274	0.0028	0.9972	58.7768	0.0040	0.9963
65	0.8247	0.0023	0.9977	59.6158	0.0034	0.9969
66	0.8219	0.0019	0.9981	60.4524	0.0028	0.9974
67	0.8192	0.0016	0.9984	61.2868	0.0023	0.9979
68	0.8164	0.0013	0.9987	62.1189	0.0019	0.9982
69	0.8137	0.0010	0.9990	62.9487	0.0016	0.9985
70	0.8110	0.0008	0.9992	63.7763	0.0013	0.9988
71	0.8082	0.0007	0.9993	64.6015	0.0011	0.9990
72	0.8055	0.0005	0.9995	65.4245	0.0009	0.9992