The Birthday Paradox

Birthday Paradox

Let $m < n$. We figure out m, n later.
We will put m balls into n boxes uniformly at random.
What is prob that some box has ≥ 2 balls?

Birthday Paradox

Let $m < n$. We figure out m, n later.
We will put m balls into n boxes uniformly at random.
What is prob that some box has ≥ 2 balls?

We ask opp: What is prob that NO box has ≥ 2 balls?

Birthday Paradox

Let $m < n$. We figure out m, n later.
We will put m balls into n boxes uniformly at random.
What is prob that some box has ≥ 2 balls?

We ask opp: What is prob that NO box has ≥ 2 balls?

Birthday Paradox

Let $m < n$. We figure out m, n later.
We will put m balls into n boxes uniformly at random.
What is prob that some box has ≥ 2 balls?

We ask opp: What is prob that NO box has ≥ 2 balls?

- Number of ways to put balls into boxes: n^m

Birthday Paradox

Let $m < n$. We figure out m, n later.
We will put m balls into n boxes uniformly at random.
What is prob that some box has ≥ 2 balls?

We ask opp: What is prob that NO box has ≥ 2 balls?

- Number of ways to put balls into boxes: n^m
- Number of ways to put balls into boxes: so that no box has ≥ 2 balls: $n(n-1)\cdots(n-m+1)$

Birthday Paradox

Let $m < n$. We figure out m, n later.
We will put m balls into n boxes uniformly at random.
What is prob that some box has ≥ 2 balls?

We ask opp: What is prob that NO box has ≥ 2 balls?

- Number of ways to put balls into boxes: n^m
- Number of ways to put balls into boxes: so that no box has ≥ 2 balls: $n(n-1)\cdots(n-m+1)$

Hence we seek

$$\frac{n(n-1)(n-2)\cdots(n-m+1)}{n^m}$$

Approx

$$\frac{n(n-1)(n-2)\cdots(n-m+1)}{n^m}$$

$$= \frac{n}{n} \times \frac{n-1}{n} \times \frac{n-2}{n} \times \cdots \times \frac{n-m+1}{n}$$

$$= 1 \times \left(1 - \frac{1}{n}\right) \times \left(1 - \frac{2}{n}\right) \times \cdots \times \left(1 - \frac{m-1}{n}\right)$$

Approx

$$\frac{n(n-1)(n-2)\cdots(n-m+1)}{n^m}$$

$$= \frac{n}{n} \times \frac{n-1}{n} \times \frac{n-2}{n} \times \cdots \times \frac{n-m+1}{n}$$

$$= 1 \times \left(1 - \frac{1}{n}\right) \times \left(1 - \frac{2}{n}\right) \times \cdots \times \left(1 - \frac{m-1}{n}\right)$$

Recall: $e^{-x} \sim 1 - x$ for x small. So we have

$$\sim e^{-1/n} \times e^{-2/n} \times \cdots e^{-(m-1)/n} = e^{-(1/n)(1+2+\cdots+(m-1))}$$

$$\sim e^{-m^2/2n}$$

Real Numbers!

If $m < n$ and you put m balls in n boxes at random then prob that ≥ 2 balls in same box is approx:

$$1 - e^{-m^2/2n}$$

To get this $> \frac{1}{2}$ need $1 - e^{-m^2/2n} > \frac{1}{2}$

$$e^{-m^2/2n} < \frac{1}{2}$$

$$-\frac{m^2}{2n} < \ln(0.5) \sim -0.7$$

$$\frac{m^2}{2n} > 0.7$$

$$m^2 > 1.4n$$

$$m > \sqrt{1.4n}$$

Real Numbers!

If $m > \sqrt{1.4n}$ and you put m balls in n boxes at random then prob that ≥ 2 balls in same box is over $\frac{1}{2}$.

$n = 365.$
$m = \left\lceil 1.4\sqrt{n} \right\rceil = 23$

Birthday Paradox: If there are 23 people in a room then prob two have the same birthday is $> \frac{1}{2}$.

Alternative Proof

Prob balls i, j in same box is $\frac{n}{n^2} = \frac{1}{n}$.
Prob balls i, j NOT in same box is $\frac{n}{n^2} = 1 - \frac{1}{n}$.

Alternative Proof

Prob balls i, j in same box is $\frac{n}{n^2} = \frac{1}{n}$.
Prob balls i, j NOT in same box is $\frac{n}{n^2} = 1 - \frac{1}{n}$.

Prob NO pair is in same box: Want to say $(1 - \frac{1}{n})^{\binom{m}{2}}$.

Alternative Proof

Prob balls i, j in same box is $\frac{n}{n^2} = \frac{1}{n}$.
Prob balls i, j NOT in same box is $\frac{n}{n^2} = 1 - \frac{1}{n}$.

Prob NO pair is in same box: Want to say $(1 - \frac{1}{n})^{\binom{m}{2}}$.

Not quite. That would be true if they are all ind. But this is good approx.

Alternative Proof

Prob balls i, j in same box is $\frac{n}{n^2} = \frac{1}{n}$.
Prob balls i, j NOT in same box is $\frac{n}{n^2} = 1 - \frac{1}{n}$.

Prob NO pair is in same box: Want to say $(1 - \frac{1}{n})^{\binom{m}{2}}$.

Not quite. That would be true if they are all ind. But this is good approx.

Prob NO pair is in same box $< (1 - \frac{1}{n})^{\binom{m}{2}} \sim e^{-m^2/2n}$.
Prob SOME pair is in same box $> 1 - e^{-m^2/2n}$.
Same as before.

Three Balls in a Box

Prob balls i, j, k in same box is $\frac{n}{n^3} = \frac{1}{n^2}$.
Prob balls i, j, k NOT in same box is $1 - \frac{1}{n^2}$.

Three Balls in a Box

Prob balls i, j, k in same box is $\frac{n}{n^3} = \frac{1}{n^2}$.
Prob balls i, j, k NOT in same box is $1 - \frac{1}{n^2}$.

Prob NO triple is in same box: APPROX $(1 - \frac{1}{n^2})^{\binom{m}{3}} \sim e^{-m^3/6n^2}$
Prob SOME triple is in same box: APPROX $1 - e^{-m^3/6n^2}$

Real Numbers!

If $m < n$ and you put m balls in n boxes at random then prob that ≥ 3 balls in same box is approx:

$$1 - e^{-m^3/6n^2}$$

Real Numbers!

If $m < n$ and you put m balls in n boxes at random then prob that ≥ 3 balls in same box is approx:

$$1 - e^{-m^3/6n^2}$$

To get this $> \frac{1}{2}$ need $1 - e^{-m^3/6n^2} > \frac{1}{2}$

$$e^{-m^3/6n^2} < \frac{1}{2}$$

$$-\frac{m^3}{6n^2} < \ln(0.5) \sim -0.7$$

Real Numbers!

If $m < n$ and you put m balls in n boxes at random then prob that ≥ 3 balls in same box is approx:

$$1 - e^{-m^3/6n^2}$$

To get this $> \frac{1}{2}$ need $1 - e^{-m^3/6n^2} > \frac{1}{2}$

$$e^{-m^3/6n^2} < \frac{1}{2}$$

$$-\frac{m^3}{6n^2} < \ln(0.5) \sim -0.7$$

Continued on Next Slide.

Real Numbers! (Cont)

$$0.7 < \frac{m^3}{6n^2}$$

Real Numbers! (Cont)

$$0.7 < \frac{m^3}{6n^2}$$

$$4.2n^2 < m^3$$

Real Numbers! (Cont)

$$0.7 < \frac{m^3}{6n^2}$$

$$4.2n^2 < m^3$$

$$m > (4.2)^{1/3} n^{2/3} \sim 1.61 n^{2/3}$$

Real Numbers! (Cont)

$$0.7 < \frac{m^3}{6n^2}$$

$$4.2n^2 < m^3$$

$$m > (4.2)^{1/3} n^{2/3} \sim 1.61 n^{2/3}$$

Birthday: $n = 365$ then need

$$m \geq (1.61)(365)^{2/3} \sim 82.$$

SO if 82 people in a room prob is $> \frac{1}{2}$ that three have same bday!

k Balls in a Box, $k \ll m$

Prob balls $i_1, \ldots, i_k$ in same box is $\frac{n}{n^k} = \frac{1}{n^{k-1}}$.
Prob balls $i_1, \ldots, i_m$ NOT in same box is $1 - \frac{1}{n^{k-1}}$.

k Balls in a Box, $k \ll m$

Prob balls $i_1, \ldots, i_k$ in same box is $\frac{n}{n^k} = \frac{1}{n^{k-1}}$.
Prob balls $i_1, \ldots, i_m$ NOT in same box is $1 - \frac{1}{n^{k-1}}$.

Prob NO k balls is in same box: APPROX

$$(1 - \frac{1}{n^{k-1}})^{\binom{m}{k}} \sim e^{-m^k/k!n^{k-1}}$$

Prob SOME triple is in same box: APPROX

$$1 - e^{-m^k/k!n^{k-1}}$$

Real Numbers!

If $m < n$ and you put m balls in n boxes at random then prob that $\geq k$ balls in same box is approx:

$$1 - e^{-m^k/k!n^{k-1}}$$

Real Numbers!

If $m < n$ and you put m balls in n boxes at random then prob that $\geq k$ balls in same box is approx:

$$1 - e^{-m^k/k!n^{k-1}}$$

To get this $> \frac{1}{2}$ need

$$1 - e^{-m^k/k!n^{k-1}} > \frac{1}{2}$$

$$e^{-m^k/k!n^{k-1}} < \frac{1}{2}$$

$$-\frac{m^k}{k!n^{k-1}} < \ln(0.5) \sim -0.7$$

Real Numbers!

If $m < n$ and you put m balls in n boxes at random then prob that $\geq k$ balls in same box is approx:

$$1 - e^{-m^k/k!n^{k-1}}$$

To get this $> \frac{1}{2}$ need

$$1 - e^{-m^k/k!n^{k-1}} > \frac{1}{2}$$

$$e^{-m^k/k!n^{k-1}} < \frac{1}{2}$$

$$-\frac{m^k}{k!n^{k-1}} < \ln(0.5) \sim -0.7$$

Continued on Next Slide.

Real Numbers! (Cont)

$$0.7 < \frac{m^k}{k!n^{k-1}}$$

Real Numbers! (Cont)

$$0.7 < \frac{m^k}{k!n^{k-1}}$$

$$0.7k!n^{k-1} < m^k$$

Real Numbers! (Cont)

$$0.7 < \frac{m^k}{k!n^{k-1}}$$

$$0.7k!n^{k-1} < m^k$$

$$m > (0.7k!)^{1/k} n^{k-1/k}$$