Math/Stat 341: Probability
Second Lecture

Steven J Miller
Williams College

sjm1@williams.edu

http://www.williams.edu/Mathematics/sjmiller/public_html/341

Bronfman 105
Williams College, September 14, 2015

Homework

Homework

The day... https://www.youtube.com/watch?v=uAsV5-Hv-7U

Buddy Holly:

- https://www.youtube.com/watch?v=YwHrx0r0t2s
- https://www.youtube.com/watch?v=GMezwtB1oCU
- https://www.youtube.com/watch?v=ku5UeUT7yIQ

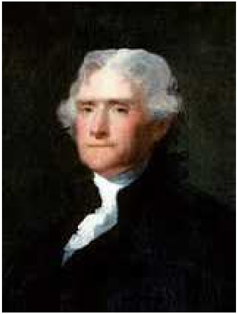
Ritchie Valens

- https://www.youtube.com/watch?v=Jp6j5HJ-Cok
- https://www.youtube.com/watch?v=-ziSLGVQOSg
- https://www.youtube.com/watch?v=HMcHbh6HBDk

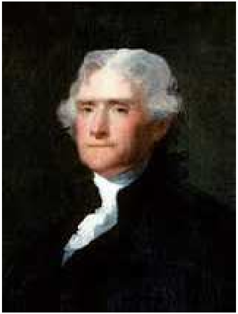
The Big Bopper

- https://www.youtube.com/watch?v=4b-by5e4saI
- https://www.youtube.com/watch?v=3NMklxiE6xw

Who are they and why are they being shown?


Jefferson, Adams and Monroe: July 4



Who are they and why are they being shown?

Truman and Ford, December 26

Clicker Problems

Birthday Problem I

Birthday Problem

How large must N be for there to be at least a 50% probability that two of the N people share a birthday?

Birthday Problem I

Birthday Problem

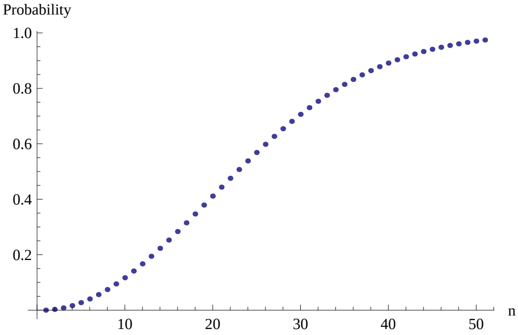
How large must N be for there to be at least a 50% probability that two of the N people share a birthday?

- (A) 11 people
- (B) 22 people
- (C) 33 people
- (D) 44 people
- (E) 90 people
- (F) 180 people
- (G) 365 people
- (H) 500 people.

Birthday Problem I

Birthday Problem

How large must N be for there to be at least a 50% probability that two of the N people share a birthday?



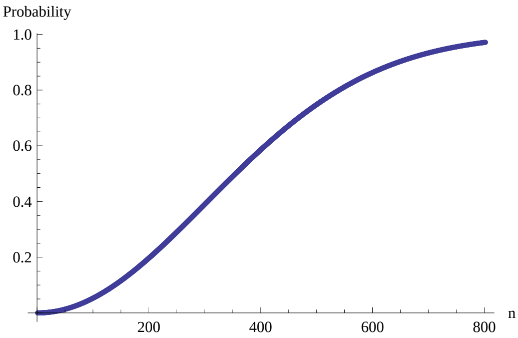
Birthday Problem II

How large must N be for there to be at least a 50% probability that two of N Plutonians share a birthday? ‘Recall’ one Plutonian year is about 248 Earth years (or 90,520 days).

- (A) 110 people
- (B) 220 people
- (C) 330 people
- (D) 440 people
- (E) 1,000 people
- (F) 5,000 people
- (G) 10,000 people
- (H) 20,000 people
- (I) more than 30,000 people.

Birthday Problem II

How large must *N* be for there to be at least a 50% probability that two of *N* Plutonians share a birthday? ‘Recall’ one Plutonian year is about 248 Earth years (or 90,520 days).



Probability Quantities

P_n: Probability at least two share birthday when n in room.

Q_n: Probability that no shared birthday with n in room.

Probability Quantities

P_n: Probability at least two share birthday when n in room.

Q_n: Probability that no shared birthday with n in room.

Note $Q_n = 1 - P_n$; oftentimes one is easier to compute than other.

What is the correct model?

Assumptions for Birthday Problem

Assume all days equally likely to be birthday, people independent, and the “no February 29th policy”.

Assumptions for Birthday Problem

Assume all days equally likely to be birthday, people independent, and the “no February 29th policy”.

Gladwell: Canadian Junior Hockey Championship: *March 11 starts around one side of the Tigers’ net, leaving the puck for his teammate January 4, who passes it to January 22, who flips it back to March 12, who shoots point-blank at the Tigers’ goalie, April 27. April 27 blocks the shot, but it’s rebounded by Vancouver’s March 6. He shoots! Medicine Hat defensemen February 9 and February 14 dive to block the puck while January 10 looks on helplessly. March 6 scores!*.”

Expansions: Q_n: no shared birthday n in room

$$Q_1 \;=\; 1 \text{ (not too surprising!).}$$

Expansions: Q_n: no shared birthday n in room

$Q_1 = 1$ (not too surprising!).

$$Q_1 = \frac{365}{365}$$

Expansions: Q_n: no shared birthday n in room

$Q_1 = 1$ (not too surprising!).

$$Q_2 = \frac{365}{365} \cdot \frac{364}{365}$$

Expansions: Q_n: no shared birthday n in room

$Q_1 = 1$ (not too surprising!).

$$Q_n = \frac{365}{365} \cdot \frac{364}{365} \cdots \frac{365-(n-1)}{365}.$$

Expansions: Q_n: no shared birthday n in room

$Q_1 = 1$ (not too surprising!).

$$Q_n = \frac{365}{365} \cdot \frac{364}{365} \cdots \frac{365-(n-1)}{365}.$$

Rewrite:

$$Q_n = \prod_{k=0}^{n-1} \left(1 - \frac{k}{365}\right).$$

Expansions: Q_n: no shared birthday n in room

$$Q_1 = 1 \text{ (not too surprising!).}$$

$$Q_n = \frac{365}{365} \cdot \frac{364}{365} \cdots \frac{365-(n-1)}{365}.$$

Rewrite:

$$Q_n = \prod_{k=0}^{n-1} \left(1 - \frac{k}{365}\right).$$

Pavlovian Response: See product,

Expansions: Q_n: no shared birthday n in room

$Q_1 = 1$ (not too surprising!).

$$Q_n = \frac{365}{365} \cdot \frac{364}{365} \cdots \frac{365-(n-1)}{365}.$$

Rewrite:

$$Q_n = \prod_{k=0}^{n-1}\left(1 - \frac{k}{365}\right).$$

Pavlovian Response: See product, take logarithm!

$$\log Q_n = \log\left[\prod_{k=0}^{n-1}\left(1-\frac{k}{365}\right)\right] = \sum_{k=0}^{n-1}\log\left(1-\frac{k}{365}\right).$$

Expansions: Q_n: no shared birthday n in room

$Q_1 \;=\; 1$ (not too surprising!).

$$Q_n \;=\; \frac{365}{365} \cdot \frac{364}{365} \cdots \frac{365-(n-1)}{365}.$$

Rewrite:

$$Q_n \;=\; \prod_{k=0}^{n-1}\left(1 - \frac{k}{365}\right).$$

More generally if D days in a year:

$$\log Q_n(D) \;=\; \sum_{k=0}^{n-1} \log\left(1 - \frac{k}{D}\right).$$

Why Calculus?

Calculus: Linearize the non-linear.
Taylor Expansion:

$$\log(1-u) = -u - \frac{u^2}{2} - \frac{u^3}{3} - \cdots \approx -u.$$

Why Calculus?

Calculus: Linearize the non-linear.
Taylor Expansion:

$$\log(1-u) \;=\; -u - \frac{u^2}{2} - \frac{u^3}{3} - \cdots \;\approx\; -u.$$

$$\log Q_n(D) \;=\; \sum_{k=0}^{n-1} \log\left(1 - \frac{k}{D}\right)$$

Why Calculus?

Calculus: Linearize the non-linear.
Taylor Expansion:

$$\log(1-u) = -u - \frac{u^2}{2} - \frac{u^3}{3} - \cdots \approx -u.$$

$$\begin{aligned}\log Q_n(D) &= \sum_{k=0}^{n-1} \log\left(1 - \frac{k}{D}\right) \\ &\approx \sum_{k=0}^{n-1} -\frac{k}{D} = -\frac{1}{D}\sum_{k=0}^{n-1} k\end{aligned}$$

Why Calculus?

Calculus: Linearize the non-linear.
Taylor Expansion:

$$\log(1-u) \;=\; -u - \frac{u^2}{2} - \frac{u^3}{3} - \cdots \;\approx\; -u.$$

$$\begin{aligned}
\log Q_n(D) &= \sum_{k=0}^{n-1} \log\left(1 - \frac{k}{D}\right) \\
&\approx \sum_{k=0}^{n-1} -\frac{k}{D} = -\frac{1}{D}\sum_{k=0}^{n-1} k \\
&= -\frac{1}{D}\frac{(n-1)n}{2} \;\text{ or }\; \approx -\frac{1}{D}\frac{(n-1/2)^2}{2}.
\end{aligned}$$

Estimating Solution

Let $n_{1/2;D}$ be value so that probability is 1/2. Then

$$\log Q_{n_{1/2;D}}(D) = \log \frac{1}{2} \approx -\frac{(n_{1/2;D} - 1/2)^2}{2D},$$

Estimating Solution

Let $n_{1/2;D}$ be value so that probability is 1/2. Then

$$\log Q_{n_{1/2;D}}(D) \;=\; \log\frac{1}{2} \;\approx\; -\frac{(n_{1/2;D}-1/2)^2}{2D},$$

which implies

$$n_{1/2;D} \;\approx\; \sqrt{2D\log 2}+\frac{1}{2} \;=\; \sqrt{D\log 4}+\frac{1}{2}.$$

Estimating Solution

Let $n_{1/2;D}$ be value so that probability is 1/2. Then

$$\log Q_{n_{1/2;D}}(D) \ = \ \log\frac{1}{2} \ \approx \ -\frac{(n_{1/2;D}-1/2)^2}{2D},$$

which implies

$$n_{1/2;D} \ \approx \ \sqrt{2D\log 2}+\frac{1}{2} \ = \ \sqrt{D\log 4}+\frac{1}{2}.$$

When $D = 365$ predict 22.99, stupendously close to 23!

Mathematica Code

```
(* Mathematica code to compute birthday probabilities *)
(* initialize list of probabilities of sharing and not *)
noshare = {{1, 1}};
share = {{1, 0}};
currentnoshare = 1; (* current probability don't share *)
For[n = 2, n <= 50, n++, (* will calculate first 50 *)
  {
   newfactor = (365 - (n-1))/365; (*next term in product*)
   (* update probability don't share *)
   currentnoshare = currentnoshare * newfactor;
   noshare = AppendTo[noshare, {n, 1.0 currentnoshare}];
   (* update probability share *)
   share = AppendTo[share, {n, 1.0 - currentnoshare}];
   }];
(* print probability share *)
Print[ListPlot[share, AxesLabel -> {"n", "Probability"}]]
```

Voting: Democratic Primaries

During the Democratic primaries in 2008, Clinton and Obama received exactly the same number of votes in Syracuse, NY. How probable was this?

Voting: Democratic Primaries

During the Democratic primaries in 2008, Clinton and Obama received exactly the same number of votes in Syracuse, NY. How probable was this? (Note: they each received 6001 votes.)

- (A) 1 / 10
- (B) 1 / 100
- (C) 1 / 1,000
- (D) 1 / 10,000
- (E) 1 / 100,000
- (F) 1 / 1,000,000 (one in a million)
- (G) 1 / 1,000,000,000 (one in a billion).

Voting: Democratic Primaries (continued)

Syracuse University mathematics Professor Hyune-Ju Kim said the result was less than one in a million, according to the Syracuse Post-Standard, which quoted the professor as saying, "It's almost impossible." Her comments were reprinted widely, as the Associated Press picked up the story. (Carl Bialik, WSJ, 2/12/08)

Voting: Democratic Primaries (continued)

Syracuse University mathematics Professor Hyune-Ju Kim said the result was less than one in a million, according to the Syracuse Post-Standard, which quoted the professor as saying, "It's almost impossible." Her comments were reprinted widely, as the Associated Press picked up the story. (Carl Bialik, WSJ, 2/12/08)

Far greater than 1/137! What's going on?

Voting: Democratic Primaries (continued)

Syracuse University mathematics Professor Hyune-Ju Kim said the result was less than one in a million, according to the Syracuse Post-Standard, which quoted the professor as saying, "It's almost impossible." Her comments were reprinted widely, as the Associated Press picked up the story. (Carl Bialik, WSJ, 2/12/08)

Far greater than 1/137! What's going on?

Prof. Kim's calculation ... was based on the assumption that Syracuse voters were likely to vote in equal proportions to the state as a whole, which went for Ms. Clinton, its junior senator, 57%-40%. .... Prof. Kim said she had little time to make the calculation, so she made the questionable assumption ... for simplicity.

From Shooting Hoops
to the Geometric Series Formula

Simpler Game: Hoops

Game of hoops: first basket wins, alternate shooting.

Simpler Game: Hoops: Mathematical Formulation

Bird and **Magic** (I'm old!) alternate shooting; first basket wins.

- **Bird** always gets basket with probability p.
- **Magic** always gets basket with probability q.

Let x be the probability **Bird** wins – what is x?

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

- **Bird** wins on 1^{st} shot: p.

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

- **Bird** wins on 1^{st} shot: p.
- **Bird** wins on 2^{nd} shot: $(1-p)(1-q)\cdot p$.

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

- **Bird** wins on 1st shot: p.
- **Bird** wins on 2nd shot: $(1-p)(1-q)\cdot p$.
- **Bird** wins on 3rd shot: $(1-p)(1-q)\cdot(1-p)(1-q)\cdot p$.

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

- **Bird** wins on 1st shot: p.
- **Bird** wins on 2nd shot: $(1-p)(1-q)\cdot p$.
- **Bird** wins on 3rd shot: $(1-p)(1-q)\cdot(1-p)(1-q)\cdot p$.
- **Bird** wins on nth shot:
  $(1-p)(1-q)\cdot(1-p)(1-q)\cdots(1-p)(1-q)\cdot p$.

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

- **Bird** wins on 1st shot: p.
- **Bird** wins on 2nd shot: $(1-p)(1-q)\cdot p$.
- **Bird** wins on 3rd shot: $(1-p)(1-q)\cdot(1-p)(1-q)\cdot p$.
- **Bird** wins on nth shot: $(1-p)(1-q)\cdot(1-p)(1-q)\cdots(1-p)(1-q)\cdot p$.

Let $r=(1-p)(1-q)$. Then

$$\begin{aligned} x &= \text{Prob}(\textbf{Bird}\text{ wins}) \\ &= p+rp+r^2p+r^3p+\cdots \\ &= p\left(1+r+r^2+r^3+\cdots\right), \end{aligned}$$

the geometric series.

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p +$$

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p + (1 - p)(1 - q)$$

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p + (1 - p)(1 - q)x$$

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p + (1-p)(1-q)x = p + rx.$$

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird}\text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird}\text{ wins}) = p + (1-p)(1-q)x = p + rx.$$

Thus

$$(1-r)x = p \quad \text{or} \quad x = \frac{p}{1-r}.$$

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird}\text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird}\text{ wins}) = p + (1-p)(1-q)x = p + rx.$$

Thus

$$(1-r)x = p \quad \text{or} \quad x = \frac{p}{1-r}.$$

As $x = p(1 + r + r^2 + r^3 + \cdots)$, find

$$1 + r + r^2 + r^3 + \cdots = \frac{1}{1-r}.$$

Lessons from Hoop Problem

- Power of Perspective: Memoryless process.
- Can circumvent algebra with deeper understanding! (Hard)
- Depth of a problem not always what expect.
- Importance of knowing more than the minimum: connections.
- Math is fun!