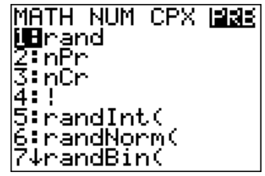
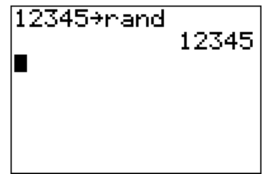
The Birthday Paradox

Dr. Schultz

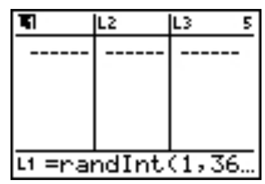
What is the probability that at least two people in a randomly selected group the size of this class share the same birthday? Traditionally, this problem is framed in terms of how large must a group be in order to ensure that there is at least a 50% probability that two or more people in the group share the same birthday. The "paradox" comes in because the group size is much smaller than you would expect. We are going to use our TI-83/84 calculators to run a Monte Carlo simulation to answer the question that I posed above.

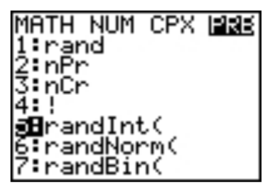


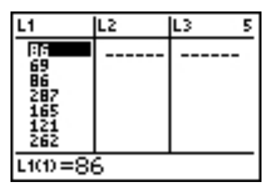


Your calculator cannot generate truly random numbers. Instead, it uses a algorithm to calculate pseudo-random numbers. To ensure that that we don't all get the same "random" numbers, you need to set the seed of your calculator's pseudo-random number generator. Type the last five digits of your Social Security number into your calculator and press the STO▸ button. Then, press MATH and scroll over to the PRB column and press 1 or ENTER. The five-digit number you entered will determine the sequence of "random" numbers generated by your calculator. You only need to set the seed once.

Here is the procedure that you will need to repeat <u>10 times</u>. Each of you will generate 10 trials, which we will then pool together to compute a relative frequency approximation of the probability that at least two people out of a group of 35 people share the same birthday.







1. Press STAT ENTER to access the stat editor. Highlight the L_1 list name and press CLEAR ENTER to clear out the contents of this list.
2. With the cursor still positioned on the L_1 list name, press the MATH key.
3. Scroll with the ▸ key to the PRB menu. Then, scroll down to `5:randInt(` and press ENTER.
4. You will find yourself back in the stat editor with the cursor blinking at the bottom of the display. Type `1,365,35)` and press ENTER. This command tells your calculator to generate 35 random numbers that are between 1 and 365, and to save them in L_1. Each number represents a day of the year (January 1 is "1," December 31 is "365," and all the other days fall in between.) To simplify the simulation, we will ignore the possibility of leap day birthdays.
5. L_1 now contains a list of 35 numbers between 1 and 365. These numbers are the simulated "birthdays" of 35 randomly selected "people." Determine whether any of these people share the same birthday (i.e., whether there are any repeated numbers in the list). It will be much easier to do so if you sort L_1 first. Press STAT, scroll down to `2:SortA(`, and press ENTER. Then, press 2nd 1) ENTER to sort L_1 in ascending order. Press STAT ENTER to return to the stat editor and inspect your list. If there are any repeated numbers, enter "YES" in Table 1 on the next page; otherwise, enter "NO" in the table.
6. Repeat steps 1-5 until you have completed a total of ten trials.

7. Tell me the number of "YES" trials you had so that I can record your data in the class spreadsheet. I'll also need to enter your birth date (month and day).
8. Consult the class data spreadsheet and complete Table 2 below.

Table 1. Individual Data

Trial	Shared Birthday? (YES or NO)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
Number of "YES" Trials:	

Table 2. Pooled Class Data

Total Number of "YES" Trials	Total Number of Trials (10 per person)	List any shared birthdays

☑ Using the pooled class data, compute a relative frequency approximation of the probability that a randomly selected group of 35 people will have at least two people who share the same birthday.

$$\frac{\text{Number of ``YES'' trials}}{\text{Total number of trials}} = \_\_\_\_\_\_\_\_\_\_\_\_$$

☑ How does this approximation compare to the actual probability of 0.814?

☑ How many students attended class today? __________

☑ Were there any shared birthdays among the students attending class today? __________

☑ What is the probability that there will be at least one shared birthday among the people in a randomly selected group this size? Before answering this question, consider the following:

- Let A = event that no two people attending class today have the same birthday
- Then, $\overline{A}$ = event that two or more people share the same birthday
- By the rule of complementary events, the probability that two or more people have the same birthday is given by $P(\overline{A}) = 1 - P(A)$.
- Now, let's build up the class starting with one person.
  - If there is only one person in the room, then the probability that no two people have the same birthday is 1.
  - Also, the probability is $\frac{365}{365} = 1$ that the person's birthday will be on one of the 365 days of the year.
- Let's add a second person to the class. The first person has taken one of the 365 days, so there are 364 possible birthdays that will not be the same as that of the first person. Hence, the probability that the first and second individuals do not share the same birthday is

$$\frac{365}{365} \cdot \frac{364}{365} = 0.997$$

- If there are three people, then the probability than nobody shares a birthday with anyone else is

$$\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} = 0.992$$

- Here's the general pattern: If there are n people in the room, then the probability that no two people have the same birthday is

$$P(A) = \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \ldots \cdot \frac{365-n+1}{365}$$

- By the rule of complementary events, the probability that there are two or more people out of n who do share the same birthday is

$$P(\overline{A}) = 1 - P(A)$$

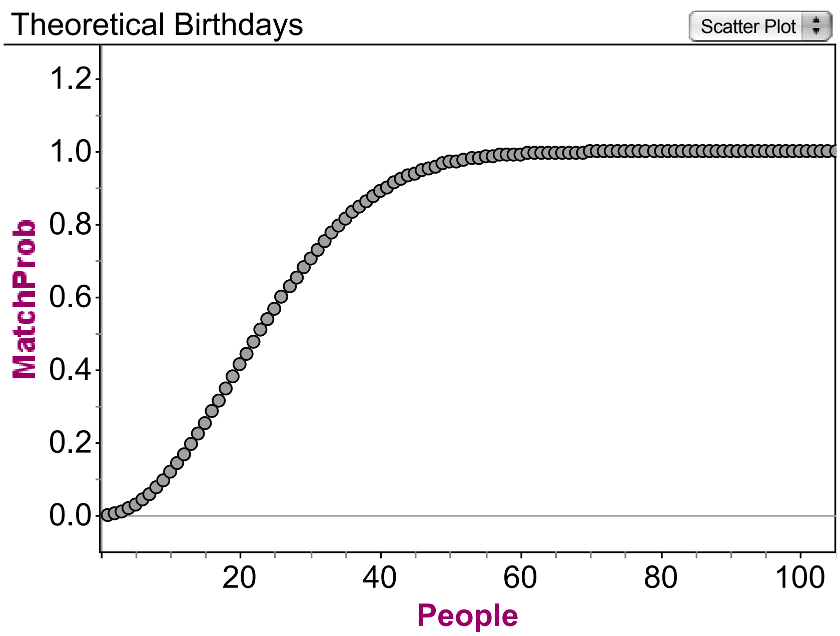
- For example, if n = 33, then

$$P(\overline{A}) = 1 - P(A) = 1 - \left[\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \frac{362}{365} \cdot \frac{361}{365} \cdot \frac{360}{365} \cdot \ldots \cdot \frac{333}{365}\right] = 0.775$$

- Hence, the theoretical probability that at least two people out of a group of 33 people have the same birthday is 0.775.

☑ I used Fathom and the formulas on the previous page to generate this table of the theoretical probabilities for groups of various sizes. Consult this table and determine the theoretical probability that at least two people out of ______ (number of students attending class today) share the same birthday.

☑ Consult this table again, and determine what size group is required for there to be at least a 50% probability that there are two or more people in the group who share the same birthday. (The graph below shows the probabilities for groups of up to 105 people.)



People	MatchProb
1	0
2	0.00273973
3	0.00820417
4	0.0163559
5	0.0271356
6	0.0404625
7	0.0562357
8	0.0743353
9	0.0946238
10	0.116948
11	0.141141
12	0.167025
13	0.19441
14	0.223103
15	0.252901
16	0.283604
17	0.315008
18	0.346911
19	0.379119
20	0.411438
21	0.443688
22	0.475695
23	0.507297
24	0.538344
25	0.5687
26	0.598241
27	0.626859
28	0.654461
29	0.680969
30	0.706316
31	0.730455
32	0.753348
33	0.774972
34	0.795317
35	0.814383
36	0.832182