**Statistical Foundations of Data Science**
**Homework, due March 9**

---

The content of this assignment is based on the second lecture in the course titled "How to lie with statistics".

1. Famously, the Literary Digest correctly predicted the U.S. election in 1920, 1924, 1928 and 1932. In 1936 they pooled ten million of their own readers (with two million four hundered thousand responses) and predicted that Landon would beat FDR. Landon only won two states Maine and Vermont. Explain what the flaw in their survey was.

   Literary digest sampled people by telephone. In 1936 having a telephone put you (on average) in a more affluent than normal category. This slanted the survey.

2. Monty Hall decides to shake up his game show. He now has seven doors behind two of which are cars, one car is red and the other is blue and behind the other five doors are goats. The doors are labelled A,B,C, D, E, F and G. Our protagonist chooses a door. Monty (who knows where the cars are) will then open two doors that the player did not choose to reveal goats. The player is then offered a chance to switch to another door.

   (a) Find the probability of winning either car if the player does not switch.
   $\frac{2}{7} = 0.28571428..$

   (b) Find the probability of winning the red car if we does not switch.
   $\frac{1}{7} = 0.14285714..$

   (c) Find the probability of winning either car if the player switches.
   $\frac{2}{7} \cdot \frac{1}{4} + \frac{5}{7} \cdot \frac{2}{4} = \frac{3}{7} = 0.42857142...$

   (d) Find the probability of winning the red car the player switches.
   $\frac{1}{7} \cdot 0 + \frac{1}{7} \cdot \frac{1}{4} + \frac{5}{7} \cdot \frac{1}{4} = \frac{3}{17} = 0.21428571...$

3. Four friends play magic the gathering every week. Each week each friend randomly chooses one of the five colours used in the game (White, Green, Red, Blue and Black) and builds a deck around that colour. Each friend is equally likely to choose any colour each week and no friend influences any other friend's choice of colour.

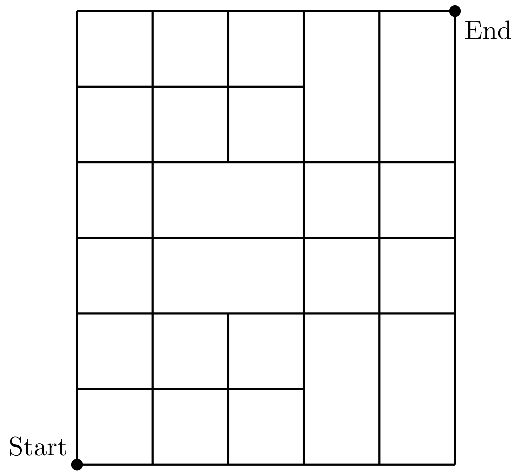
   (a) The probability that at least two of the friends build a deck of the same colour.
   $1 - \frac{4}{5} \cdot \frac{3}{5} \cdot \frac{2}{5} = \frac{101}{125} = 0.808$

   (b) The probability that all four choose the same colour.
   $\frac{5}{5^4} = \frac{1}{125} = 0.008$

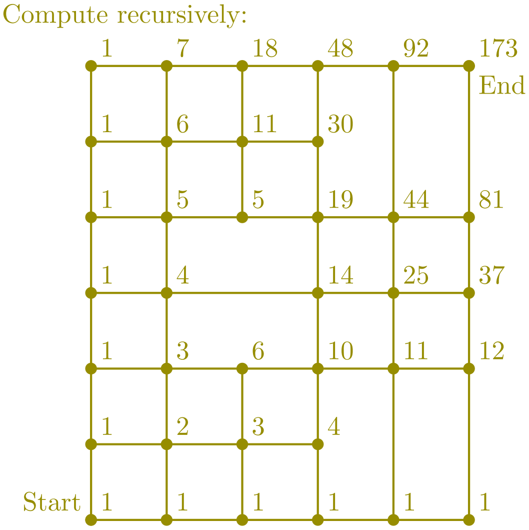
   (c) The probability that exactly three choose the same colour and that the fourth does not.
   $4 \cdot \frac{1}{5^2} \cdot \frac{4}{5} = \frac{16}{125} = 0.128$

   (d) The probability that two of the friends choose the same colour as each other and the other two friends choose the same colour as each other but not the same colour as the first two (the two pair situation). $3 \cdot \frac{1}{5} \cdot \frac{4}{5} \cdot \frac{1}{5} = \frac{12}{125} = 0.096$

4. How many ways can one go from the start to the end in the picture below if you are only allowed to take right and upwards steps?




Compute recursively:




5. (a) 180 students take a combinatorics course. Ribbons with the numbers 1, 2, 3 ,... ,10 are awarded to the top ten students. In how many ways can these awards be made, you may assume no ties are possible and leave factorials in your answer.
$180 \cdot 179 \cdot ... \cdot 171 = \frac{180!}{170!}$

(b) The next year 185 students take the course and once again ties are impossible. This year however five "top 5" ribbons are awarded as are five "top 10" ribbons (to places 6 through 10) and ten "top 20" ribbons (places 11 through 20). In how many ways can these be distributed. Again you may use factorials and/or binomial coefficents in your answer.
$\binom{180}{5} \cdot \binom{175}{5} \cdot \binom{170}{20} = \frac{180!}{175!5!} \cdot \frac{175!}{170!5!} \cdot \frac{170!}{160!10!} = \frac{180!}{5!5!10!160!}$

monty_hall.png