












Random Processes



Random Processes

-  Over view of discrete probability (including hints how to treat continuous probability)
-  Probability distributions
-  How to generate random numbers
-  See how Monte Carlo can be viewed as a sampling technique
-  See how matlab can be used to generate random processes
-  See applications of Monte Carlo methods
-  See how random walks can be used to simulate experiments that typically can't be done in any other way.

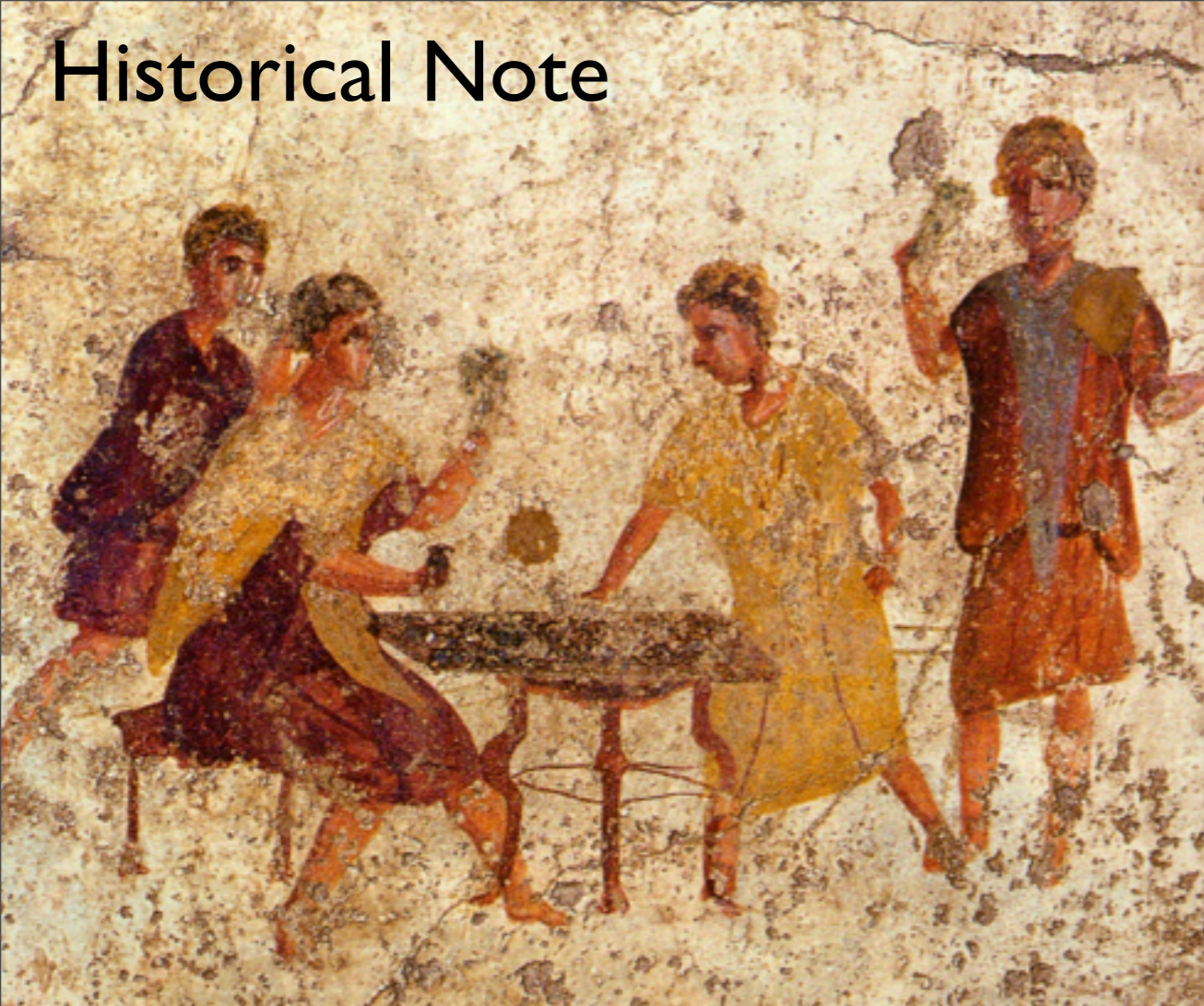
Random Processes

-  Random processes are important because most real world problems exhibit random variation. (stochastic processes)
-  Monte Carlo method is based on principles of probability and statistics

Continuous versus discrete probabilities

-  Discrete probability: if we have a finite number of possible outcomes (example: toss a die)
-  Continuous probability: we consider experiments where the random variable can take any value in a range, for example using a spinner on a table to get random coordinates.

Historical Note



Archaeologists have found evidence of games of chance on prehistoric digs, showing that gaming and gambling have been a major pastime for different peoples since the dawn of civilization. Given the Greek, Egyptian, Chinese, and Indian dynasties' other great mathematical discoveries (many of which predated the more often quoted European works) and the propensity of people to gamble, one would expect the mathematics of chance to have been one of the earliest developed.

Historical Note

Surprisingly, it wasn't until the 17th century that a rigorous mathematics of probability was developed by French mathematicians

Pierre de Fermat (Aug 1601 or 1607/8 – 12 Jan 1665)

and Blaise Pascal (19 June 1623 – 19 Aug 1662).

[Galileo Galileo wrote a treatise on chance in 1620, and the first text book on probability was written 1718 by Abraham de Moivre (26 May 1667 – 27 November 1754)]



Blaise Pascal



Pierre de Fermat

[Information taken from MathForum, <http://mathforum.org/isaac/problems/probl.html>]



Michelangelo Merisi da Caravaggio (Milan, 28 September 1571 - Porto Ercole, 18 July 1610) was an Italian artist active in Rome, Naples, Malta and Sicily between 1593 and 1610. He is commonly placed in the Baroque school, of which he is considered the first great representative.

The problem that inspired the development of mathematical probability in Renaissance Europe was the problem of points.

It can be stated this way: Two equally skilled players are interrupted while playing a game of chance for a certain amount of money. Given the score of the game at that point, how should the stakes be divided?

In this case "equally skilled" indicates that each player started the game with an equal chance of winning, for whatever reason. For the sake of illustration, imagine the following scenario. Pascal and Fermat are sitting in a cafe in Paris and decide, after many arduous hours discussing more difficult scenarios, to play the simplest of all games, flipping a coin.

If the coin comes up heads, Fermat gets a point. If it comes up tails, Pascal gets a point. The first to get 10 points wins.

Knowing that they'll just end up taking each other out to dinner anyway, they each ante up a generous 50 Francs, making the total pot worth 100. They are, of course, playing 'winner takes all'.

But then a strange thing happens. **Fermat is winning, 8 points to 7**, when he receives an urgent message that a friend is sick, and he must rush to his home town of Toulouse. The carriage man who has delivered the message offers to take him, but only if they leave immediately. Of course Pascal understands, but later, in correspondence, the problem arises: **how should the 100 Francs be divided?** How would you divide the 100 Francs? After we discuss probability you should be able to divide the money and argue that it is a fair division.

Probabilities and Events

- Probability measures the chance of a given event occurring. It is a number between 0 and 1 .
- The complement of an event is everything that is not in the event; for example the complement of 3 or less hurricanes hitting Florida in 2013 is 4 or more hurricanes hitting Florida.
- If the probability of an event is 1 then it is certain to occur.
- A probability of 0 represents an event that cannot occur, i.e., is impossible.
- Probabilities in between 0 and 1 are events that may occur and the bigger the number the more likely that the event will occur.
- "Rain tomorrow", "crop yield", "number of hurricanes to hit Florida's coast in 2013", "the number of times the sum of two dice thrown is even" are all examples of events for which we may be interested in determining their probabilities.

Probabilities and Events

- Let X denote the random variable, for example the sum of two dice when they are thrown.
- The set of all possible outcomes is denoted Ω . For example, if we roll one die then the possible outcomes are $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- We are interested in determining the probability of an event and will denote it $p(\mathcal{E})$. The sum of all probabilities for all possible outcomes is 1 .
- We will denote an event as ε . So for example, if we roll a single die and want to determine the likelihood of the result being an even number, then $\varepsilon = \{2, 4, 6\}$; note that $\varepsilon \subset \Omega$.
- The probability of an event will be the sum of the probabilities of each outcome in the event. The probability of the complement of an event ε will be $1 - p(\mathcal{E})$.

Probabilities and Events

Suppose the weather forecast is that the probability of rain tomorrow is $1/4=0.25$.

We say the likelihood of rain is 25%. What does this mean?

The complement of this event is that there is NO rain tomorrow so the probability of no rain is $1 - 0.25 = 0.75$

So there is a 75% chance of no rain tomorrow and thus it is 3 times more likely that there is no rain tomorrow than there is rain.

Unions and Intersections

Consider the following two outcomes:

 $A = \{ 3 \text{ or less hurricanes to hit Florida in } 2013 \}$

 $B = \{ 4 \text{ or } 5 \text{ hurricanes to hit Florida in } 2013 \}$

Then $A \cup B$ is $\{ 5 \text{ or less hurricanes to hit Florida in } 2013 \}$.

The probability of $A \cup B$ is $p(A \cup B) = p(A) + p(B)$ because the events are mutually exclusive.

Then $A \cap B$ is zero because both events can't be satisfied.

Unions and Intersections

Now consider two events which are **not mutually exclusive**, because the events can both happen at in the **same 'experiment'** (in contrast to the Hurricane example on the previous page.

 $C = \{ \text{rain tomorrow} \}$

 $D = \{ \text{rain the day after tomorrow} \}$

Then $C \cup D = \{ \text{rain in the next two days} \}$ (i.e., rain tomorrow or rain the next day or both days).

The intersection of C and D is everything that is in event C and in event D so

$C \cap D = \{ \text{rain tomorrow and rain the next day} \}$

What is the probability of $C \cup D$?

We can't simply sum the probabilities. For example if there was a probability of 1/4 each day (i.e., 25% chance) then we can't say there is a 50% chance that it will rain in the next two days.

In the case where the events are not mutually exclusive we have

$$p(C \cup D) = p(C) + p(D) - p(C \cap D)$$

If this is mysterious consider all the possibilities!



$C = \{ \text{rain tomorrow} \}$



$D = \{ \text{rain the day after tomorrow} \}$

$$p(C) = 1/4$$

$$p(D) = 1/4$$

$$p(\text{not } C) = 1 - p(C) = 3/4$$

$$p(\text{not } D) = 1 - p(D) = 3/4$$

we can now combine all possible events and finally calculate $p(C \cup D)$

$$p(C \cap D) = p(C) \times p(D) = 1/4 \times 1/4 = 1/16$$

$$p(\text{not } C \cap D) = p(\text{not } C) \times p(D) = 3/4 \times 1/4 = 3/16$$

$$p(C \cap \text{not } D) = p(C) \times p(\text{not } D) = 1/4 \times 3/4 = 3/16$$

$$p(\text{not } C \cap \text{not } D) = p(\text{not } C) \times p(\text{not } D) = 3/4 \times 3/4 = 9/16$$

$$p(C \cup D) = p(C \cap D) + p(\text{not } C \cap D) + p(C \cap \text{not } D) = 1/16 + 3/16 + 3/16 = 7/16$$

$$= 1 - p(\text{not } C \cap \text{not } D) = 1 - 9/16 = 7/16$$

$$= p(C) + p(D) - p(C \cap D) = 1/4 + 1/4 - 1/16 = 7/16$$

Conditional Probability

- If we know that one event has occurred it may change our view of the probability of another event. For example, let

$$A = \{\text{rain today}\}, \quad B = \{\text{rain tomorrow}\}, \quad C = \{\text{rain in 90 days from now}\}$$

Assume that knowledge that A has occurred will change your view of the probability that B will occur, but not of the probability that C will occur.

- We write $p(B|A)$ to denote the conditional probability of B , given A .
- With our assumptions, $p(C|A) = p(C)$ but $p(B|A) \neq p(B)$.
- We say that A and C are independent, but A and B are not.
- For independent events A and C , the probability of $p(A \cap C) = p(A)p(C)$.

Example If we roll a single die (with sides 1, 2, 3, 4, 5, 6)
what is the probability that you will roll the number 4?
If we roll two dice, what is the probability that they will both be 4?

Example If we roll a single die (with sides 1, 2, 3, 4, 5, 6) what is the probability that you will roll the number 4?
If we roll two dice, what is the probability that they will both be 4?

Clearly all numbers are equally probable and since the probabilities have to sum to 1 and there are 6 possible outcomes, $p(X = 4) = 1/6$.

To determine the probability that a 4 occurs on both dice, we note that they are independent events so the probability is $1/6 * 1/6 = 1/36$.

Example If we roll a single die (with sides 1, 2, 3, 4, 5, 6) what is the probability that you will roll the number 4?
If we roll two dice, what is the probability that they will both be 4?

Clearly all numbers are equally probable and since the probabilities have to sum to 1 and there are 6 possible outcomes, $p(X = 4) = 1/6$.

To determine the probability that a 4 occurs on both dice, we note that they are independent events so the probability is $1/6 * 1/6 = 1/36$.

Example Let's return to the Fermat/Pascal problem of flipping a coin. We know that heads and tails are each equally likely to occur. If heads occurs then Fermat gets a point, otherwise Pascal gets one. When the game is interrupted the score is Fermat 8 and Pascal 7. There is 100 francs at stake so how do you fairly divide the winnings? We first ask ourselves what is the maximum number of flips of the coin that guarantee someone will win? Clearly Fermat needs 2 points and Pascal 3. The worse case scenario would be for Fermat to get 1 and Pascal to get 3 so in 4 more flips we are guaranteed a winner.

If we determine the possible outcomes in these 4 flips then we can assign a winner to each. Then Fermat's probability of winning will be the number of outcomes he would win over the total number of outcomes; similarly for Pascal. Then we will know how to divide the money.

Here are the 16 possible outcomes for 4 flips and we have indicated the winner in each case remembering that Fermat needs 2 heads to win and Pascal 3 tails.

was our calculation on the last page correct?



Pierre de Fermat

HHHH (F) HTHH (F) TTTT (P) THTT (P)
HHHT (F) HTTH (F) TTHT (P) THHH (F)
HHTH (F) HTHT (F) TTTH (P) THTH (F)
HHTT (F) HTTT (P) TTHH (F) THHT (F)



Blaise Pascal

was our calculation on the last page correct?



Pierre de Fermat

HHHH (F) HTHH (F) TTTT (P) THTT (P)
HHHT (F) HTTH (F) TTHT (P) THHH (F)
HHTH (F) HTHT (F) TTTH (P) THTH (F)
HHTT (F) HTTT (P) TTHH (F) THHT (F)

So Fermat will win 11 out of 16 times so his probability is $11/16$. Pascal will win 5 out of 16 times so his probability is $5/16$. Fermat should get $11/16$ of the money and Pascal the rest.



Blaise Pascal

was our calculation on the last page correct?



Pierre de Fermat

HHHH (F) HTHH (F) TTTT (P) THTT (P)
HHHT (F) HTTH (F) TTHT (P) THHH (F)
HHTH (F) HTHT (F) TTTH (P) THTH (F)
HHTT (F) HTTT (P) TTHH (F) THHT (F)

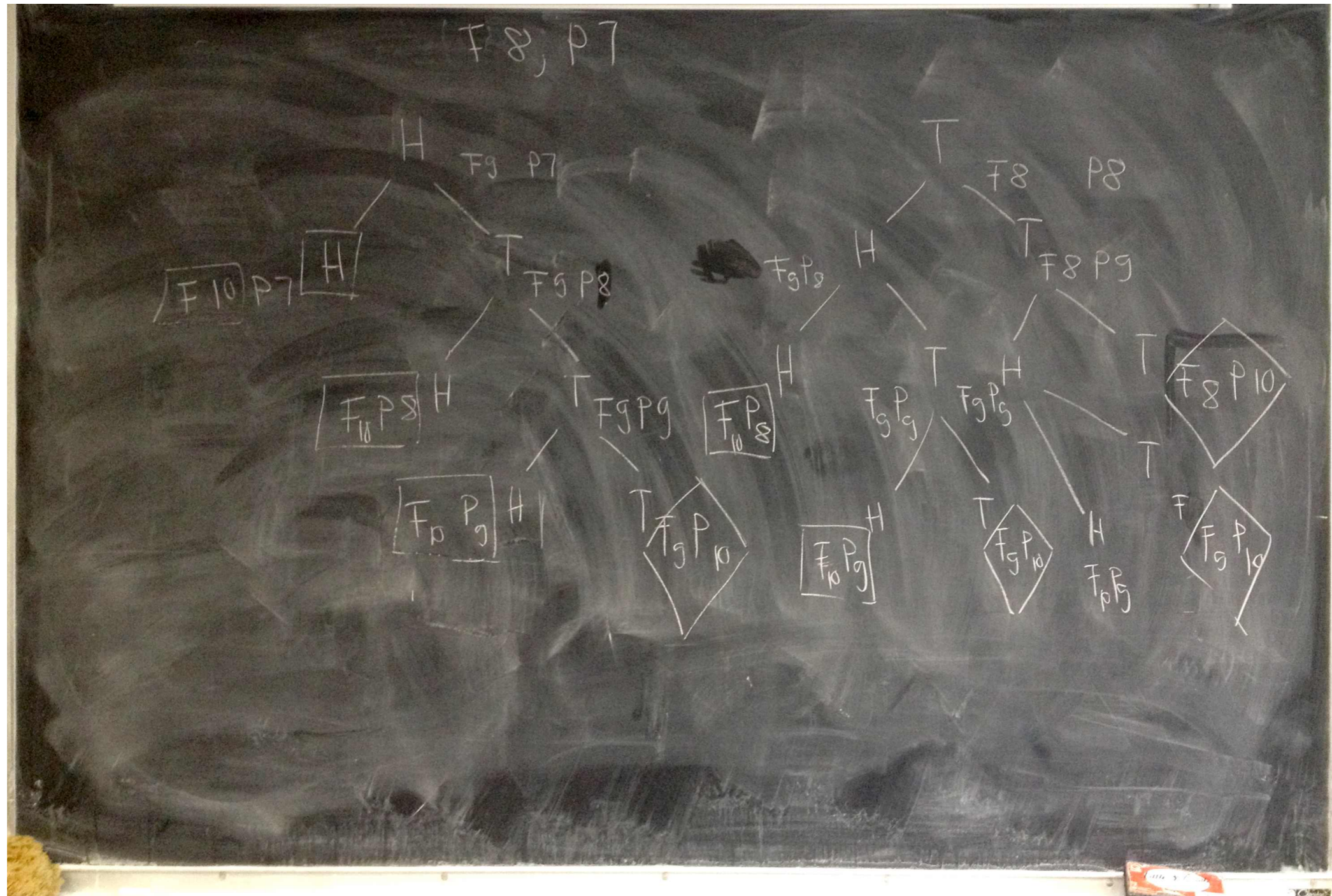
So Fermat will win 11 out of 16 times so his probability is $11/16$. Pascal will win 5 out of 16 times so his probability is $5/16$. Fermat should get $11/16$ of the money and Pascal the rest.



Blaise Pascal

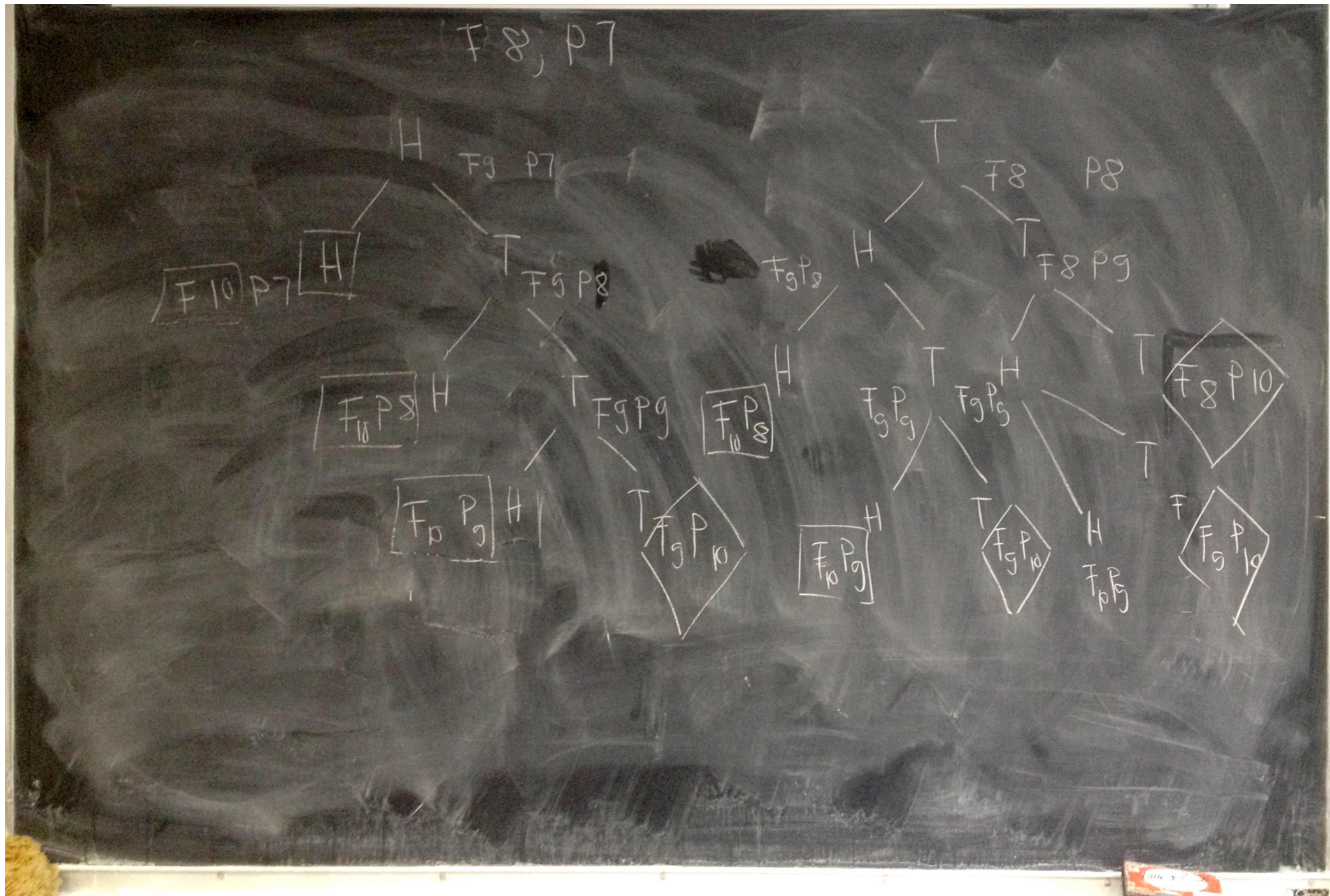
was our calculation on the last page correct?

was our calculation on the last page correct?



HH (F), HTH (F), HTTH (F), HTTT (P),
THH (F), THTH (F), THTT (P), TTTH (F), TTHT (P), TTT (P)

was our calculation on the last page correct?



HH (F), HTH (F), HTTH (F), HTTT (P),
THH (F), THTH (F), THTT (P), TTTH (F), TTHT (P), TTT (P)

Result: Fermat wins 6, Pascal wins 4 times out of 10.

Random numbers

Physical random number generators

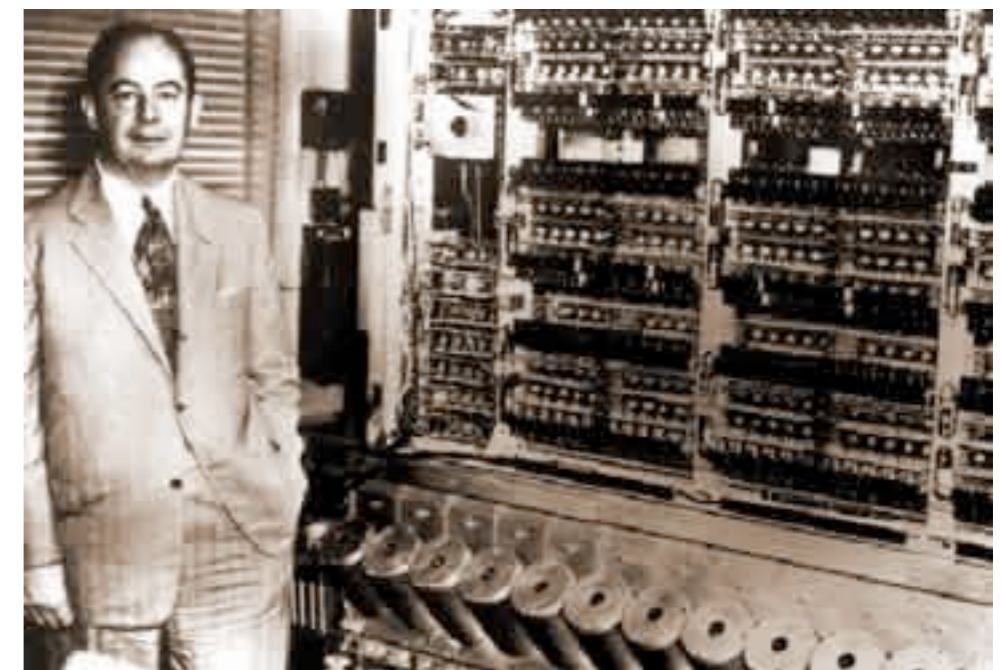
Lava-lamps

Geiger counter

Devices that combine
multiple events `/dev/random`

Pseudorandom numbers

John von Neumann proposed using the following method as one of the first random number generators. Suppose we want to create eight-digit numbers. Begin with an eight-digit number X_0 , which we call the *seed*, and create the next integer in the sequence by removing the middle eight digits from X_0^2 .

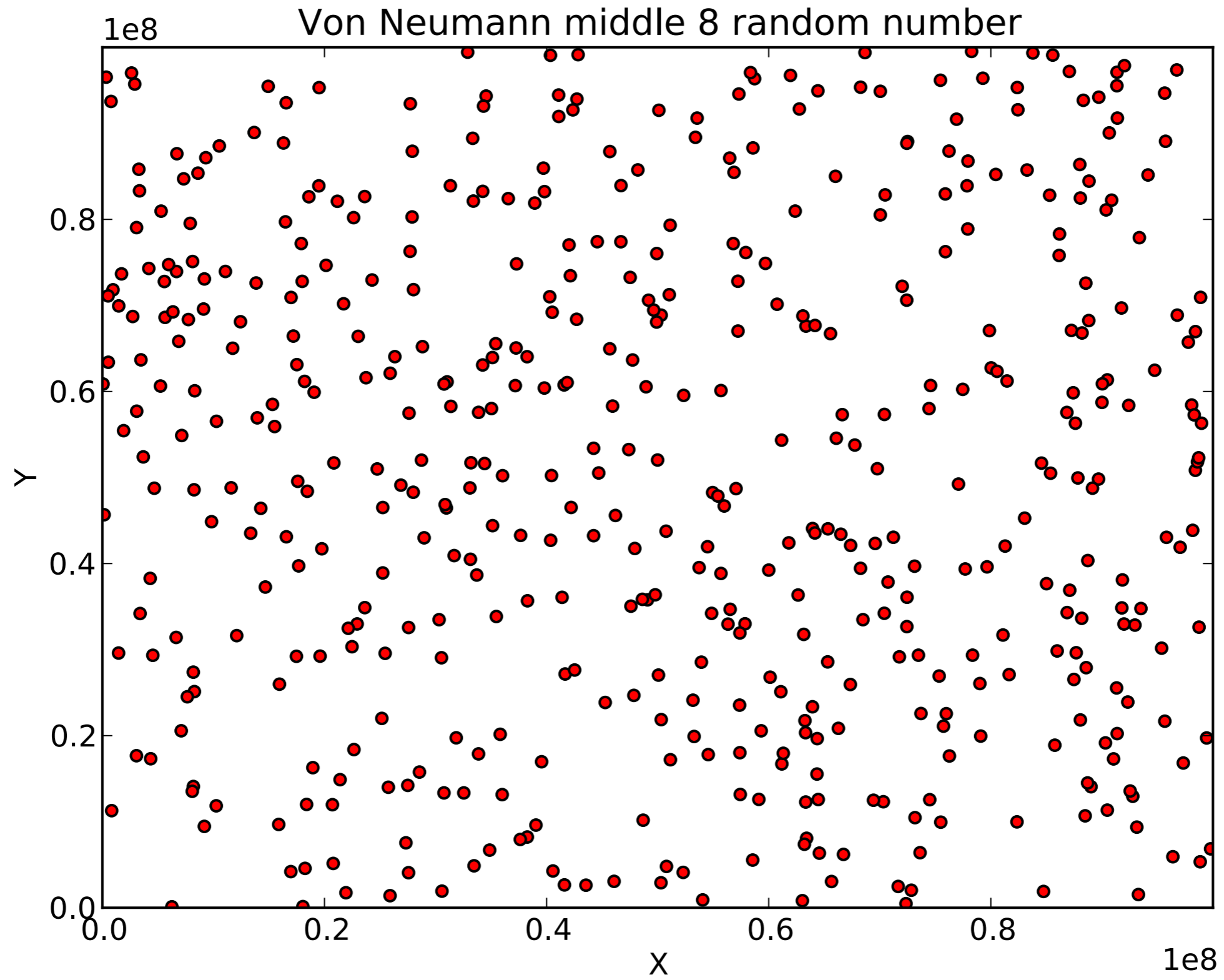


Pseudocode for von Neuman random number generator

```
initialize vector x
seed = input(a number with 8 digits)
for i in 1,n:
    x0 = seed * seed
    seed = take middle 8 digits of x0
    x[i] = seed
```

perhaps better and more informative

```
initialize vector x
seed = input(a number with 8 digits)
for i in 1,n:
    x0 = seed * seed
    s = convert_to_string(x0)
    middle = length(s)/2
    s = substring(s,middle-4,middle+4)
    seed = convert_to_integer(s)
    x[i] = seed
```

Simulating random processes using Matlab's rand command

- Random events are easily simulated in matlab with the function *rand*. This is a pseudorandom number that appears random for most purposes, for encryption one should use truly random numbers to guarantee secure data.

The matlab command $x=rand()$ can be used to get a random number between 0 and 1. It can also generate row or column vectors. For example, $rand(1,5)$ returns a row vector of five random numbers (1 row, 5 columns) such as

0.2416 0.9644 0.4519 0.3278 0.8913

```
a = rand ( )           <-- a scalar value;  
b = rand ( 5, 1 )     <-- a column vector;  
c = rand ( 1, 5 )     <-- a row vector;  
d = rand ( 3, 4 )     <-- a matrix;  
e = rand ( 5 )        <-- a 5x5 matrix;
```

- If we want to generate n integers between say 0 and k then we first generate the numbers between 0 and 1 using $x=rand(1,n)$ and then use either of

the following commands:

```
x=floor ( (k+1)*x ) )  
x=ceil ( k*x )
```

Pay attention: these two methods do not deliver the same boundaries:
floor results in a range: 0 .. k
ceil results in a range: 1 .. k

For example, if we generate 5 random numbers

.8147, .9058 .1770 .9134 .6324

and we want integers ≤ 10 then `x=floor (11*x))` gives {8, 9, 1, 10, 6}
and `x=ceil (10*x)` gives {9, 10, 2, 10, 7}.

- The only thing that may be disconcerting is that if we run our program which is based on generating random numbers two times in a row, we will get different numbers! It is sometimes useful (such as in debugging) to be able to reproduce your results exactly. If you continually generate random numbers during the same Matlab session, you will get a different sequence each time, as you would expect.
- However, each time you start a session, the random number sequence begins at the same place (0.9501) and continues in the same way. This is not true to life, as every gambler knows.

- Clearly we don't want to terminate our session every time we want to reproduce results so we need a way to "fix" this. To do this, we provide a *seed* for our random number generator. Just like Matlab has a seed at the beginning of a session we will give it a seed to start with at the beginning of our code.

Seeding rand

- The random number generator rand can be seeded with the statement

rand('state',n)

where n is any integer. By default, n is set to 0 when a Matlab session starts.

- Note that this statement does not generate any random numbers; it only initializes the generator.
- Sometimes people will use the system clock to make sure the seed is being reset

rand('state',sum(100*clock))

- Try setting the seed for rand, generating a random number, then resetting the seed and generate another random number; see that you get the same

random number two times in a row; e.g.,

```
rand('state', 12345)    rand(1)
```

Exercise Use two calls to Matlab's `rand` command to generate two random row vectors of length 4. Now repeat the exercise where you reset the seed before each call to `rand`. You should get the same vector each time in this latter case.