

Simulation of probability experiments with the TI-83 Plus

Koen Stulens
Limburgs Universitair Centrum
Diepenbeek, Belgium



1. Introduction

The **TI-83 Plus** is a perfect piece of technology to demonstrate the early ideas of probability and link them with the mathematical theory of probability.

The probability theory has its roots in the world of gambling. Perhaps we can say that the first steps in the development of the probability theory were taken in search for all possible solutions of several dice games.

One of the first dissertations about probability, *Liber de Ludo Aleae*, dates from 1525 and is written by Girolamo Cardano (1501-1576). It is a book about throwing dice.



Cardano
1501-1576

Mathematical discussions about probability arise in the seventeenth century. The correspondence between Pascal and Fermat, which consists of five letters (± 1624), forms the fundamentals of probability. Based on this Christiaan Huygens (1629-1695) wrote the first printed work on probability, *De Ratiociniis in Ludo Aleae* in 1657.



Blaise Pascal
1623-1662

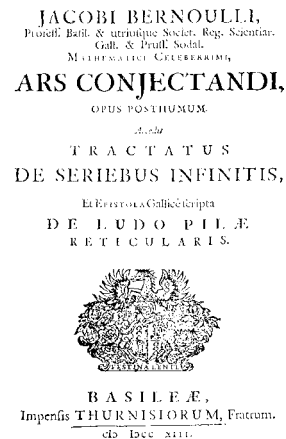


Christiaan Huygens
1629-1695

The work *Ars Conjectandi* of Jacob (James) Bernoulli is a masterpiece in the history of probability. It was published posthumously by his nephew Nikolaus Bernoulli in 1713.



Jacob Bernoulli
1654-1705



It was in this masterpiece the law of the large numbers appears for the first time. James formulates and proves that the relative frequency of an event approximates its probability in the long run.

It is this idea that will be used in this article to get information from the simulations.

Another important moment in the development of probability as a modern science is the axiomatic approach of probability by Kolmogorov (1903-1987) around 1933.

2. Coin tosses

Let us simulate a toss of a fair coin. Denote tails by 0 and heads by 1. The command `randInt(0,1)` generates at random the integer 0 or 1.

Although both heads and tails have the same chance to appear the picture alongside shows that the result of the first four tosses is heads every time.

```
randInt(0,1)
1
1
1
1
0
0
```

To simulate a hundred successive tosses you have to use the command `randInt(0,1,100)`. We count the number of heads and divide this result by a hundred.

```
randInt(0,1,100)
→L1:=sum(L1)/100
.49
.54
.51
.48
.47
.56
```

```
.44
.6
.48
.43
.47
.47
.52
```

```
.52
.57
.48
.57
.53
.5
.4
```

The result above shows that the relative frequency of the event heads is in the long run $\frac{1}{2}$, which is the theoretical probability of the event heads.

To visualize the previous result we define the following lists :

$L_1 = \text{seq}(X, X, 1, 100)$

$L_2 = \text{randInt}(0, 1, 100)$

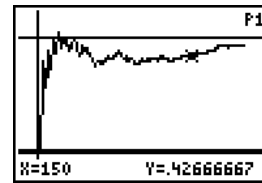
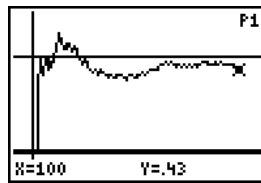
$L_3 = \text{cumSum}(L_2)$

$L_4 = L_3 / L_2$

```
seq(X,X,1,100)→L
1:randInt(0,1,10
0)→L2:cumSum(L2)
→L3:L3/L1→L4
```

The statistical plot, XY-line, of the lists L_1 and L_2 shows what will happen in the long run. In the pictures below you see the result of 100 and 200 tosses. Notice that more tosses don't automatically cause a better approximation. If we continue tossing the relative frequency will get as close to $\frac{1}{2}$ as we want. But no one can tell us how many tosses we have to do.

```
Plot2 Plot3
Off
Type: L1
Xlist: L1
Ylist: L4
Mark: +
```

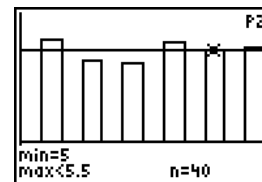


3. Throwing dice

$\text{randInt}(1, 6)$ simulates the throwing of a die. Several throws can be simulated for example with $\text{randInt}(1, 6, 5)$. A histogram of the simulated throws gives us a visualization of Bernoulli's law of large numbers.

```
randInt(1,6,5)→L
1:sum(L1=6)/5
0
0
.4
0
.2
```

```
randInt(1,6,240)
→L1:sum(L1=6)/24
0
.175
.1666666667
.1458333333
.175
```



A simulation of throwing two dice and counting the points can be executed with $\text{randInt}(1, 6) + \text{randInt}(1, 6)$. With the TEST-menu you can carry out several tests on the data created by randInt .

Dice 1	Dice 2	Dice 1	Dice 2	Dice 1	Dice 2	Dice 1	Dice 2	Dice 1	Dice 2	Dice 1	Dice 2
1	1	2	1	3	1	4	1	5	1	6	1
1	2	2	2	3	2	4	2	5	2	6	2
1	3	2	3	3	3	4	3	5	3	6	3
1	4	2	4	3	4	4	4	5	4	6	4
1	5	2	5	3	5	4	5	5	5	6	5
1	6	2	6	3	6	4	6	5	6	6	6

```
randInt(1,6,200)
+randInt(1,6,200)
)→L1
(7 5 7 6 4 6 6 ...
sum(L1<7)/200
.475
```

```
randInt(1,6,200)
+randInt(1,6,200)
)→L1
(8 4 6 7 9 3 11...
sum(L1=7)/200
.15
```

```
randInt(1,6,200)
+randInt(1,6,200)
)→L1
(5 6 9 9 2 10 6...
sum(L1>7)/200
.41
```

4. Law of large numbers : the modern version

We'll explain the (weak) law of numbers by throwing a die. This experiment has a uniform probability distribution. Each event has a probability to appear of $1/6$.

Let A be the event of throwing six points and consider the following random variables :

$X_n = 1$ if the event A occurs in the n^{th} throw and

$X_n = 0$ if the event A doesn't occur in the n^{th} throw.

The random variable $S_n = X_1 + \dots + X_n$ counts the number of throws of six points in n repetitions. S_n has a binomial distribution with parameters n en $p = \frac{1}{6}$. $E(S_n) = \frac{n}{6}$

The relative frequency of the event A can be seen as the random variable $\frac{S_n}{n}$.

Notice that $E\left(\frac{S_n}{n}\right) = \frac{1}{6}$. The (weak) law of large numbers tells us that for each $\varepsilon > 0$ and n

large enough the probability of getting a value for the random variable $\frac{S_n}{n}$ in the interval

$\left(\frac{1}{6} - \varepsilon, \frac{1}{6} + \varepsilon\right)$ will be very big (≈ 1) : $P\left[\left|\frac{S_n}{n} - \frac{1}{6}\right| < \varepsilon\right] \xrightarrow{n \rightarrow \infty} 1$.

So the relative frequency of each event gets close to its theoretical probability as more trials are made. The **STAT-editor** below shows the results of throwing a die 500 times. The list **L4** contains the relative frequency of each possible outcome.

X = number of points						
x	1	2	3	4	5	6
$P(X=x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

L1	L3	L4	Z
1	82	.164	
2	74	.148	
3	78	.156	
4	93	.186	
5	95	.19	
6	78	.156	
L2 = (1, 2, 3, 4, 5, 6)			

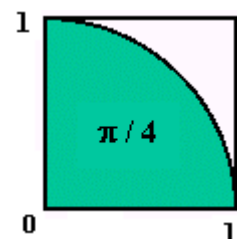
5. An approximation of π

The command **rand** generates at random a number between 0 and 1. **rand(5)** generates a list of five such numbers.

Consider a square with sides of length 1. In this square we draw a quarter of a circle with radius 1 and midpoint one of the vertices of the square.

If we choose at random a point in the square the probability that the point lies in the coloured part of the disk bounded by the circle is $\pi / 4$.

```
rand
.9819128515
.8187750612
.1795892308
.4194649398
rand(5)
(.9172672624 .4...
```



`{rand, rand}` generates at random a point in the square that's shown on the previous page. And with the test $\text{rand}^2 + \text{rand}^2 \leq 1$ you can measure whether a point lies in the coloured part of the disk or not.

If we execute the test mentioned above a hundred times, calculate the mean of these data and multiply it by four we get an approximation of π .

```
sum(seq(rand^2+rand^2<=1,X,1,400))/100
3.18
3.14
3.14
3.2
```

The same strategy can be used to find an approximation of the integral $\int_0^1 e^{-x^2} dx$.

6. Drawing balls

A lot of problems in probability theory can be transformed in terms of drawing balls out of an urn. There are two sorts of drawings : with or without replacement.

6.1 Drawing with replacement

A drawing with replacement is easy to simulate.

Suppose there are 10 white balls and 5 black balls in an urn. Denote the white balls by the numbers 1 to 10 and the black ones by 11 to 15.

The command `randInt(1, 15, 3)` simulates a drawing of three balls out of the urn.

To check how many balls are black you use the command :

$$\text{sum}(\text{randInt}(1, 15, 3) > 10).$$

The probability to get three black balls by a drawing with replacement of three balls is :

$$P(3 \times \text{black}) = P(\text{black}) \cdot P(\text{black}) \cdot P(\text{black}) = \left(\frac{5}{15}\right)^3 \approx 0.037.$$

```
seq(sum(randInt(1,15,3)>10),X,1,100)/100
.03
.03
.05
```

```
.04
.01
.03
.04
.01
.03
.05
```

6.2 Drawing without replacement

To simulate a drawing without replacement we have to refine the previous procedure. The short program below generates a drawing without replacement of three numbers out of 15 numbers.

```
Repeat sum(ΔList(L2)=∅)=∅
randInt(1, 15, 3)→L1
L1→L2
SortA(L2)
End
```

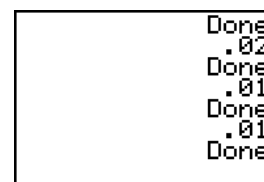
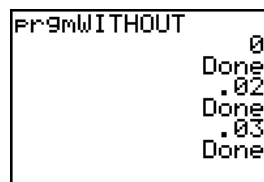
The drawing (`randInt(1,15,3)→L1`) continues until the condition of the **Repeat**-loop (`sum(ΔList(L2)=0)=0`) is true. This means that when the **Repeat**-loop ends all the elements of the list `ΔList(L2)=0` are zero. So none of the elements of `ΔList(L2)` are equal to zero and all the elements of `L1` are different.

To simulate in this case the probability of three black balls you have to put the commands above in a loop, for example a **For**-loop.

```

ClearList L1,L2,L3
For (I,1,1000)
Repeat sum(ΔList(L2)=0)=0
randInt(1,15,3)→L1
L1→L2
SortA(L2)
End
sum(L1>10)→L3(I)
End
sum(L3=3)/1000→N
Disp N

```



The theoretical probability of getting three times black is given by :

$$P(3 \times \text{black}) = \frac{5}{15} \frac{4}{14} \frac{3}{13} = \frac{2}{91} \approx 0.022$$

7. Random samples from a population

Consider the following heights of 200 children (age : 10 years) as a population (`LK200`).

120	123	151	142	137	128	133	142	136	137
144	126	135	142	135	134	140	149	137	129
128	140	129	137	141	137	135	129	133	139
132	125	124	132	129	139	132	145	140	138
137	133	137	138	131	137	131	127	134	134
150	140	144	137	133	139	130	141	136	124
130	135	124	122	136	132	133	133	142	127
142	130	135	125	136	132	153	145	131	131
134	145	139	132	136	143	138	141	141	141
136	148	128	137	134	138	130	145	135	141
131	143	146	132	127	129	133	142	157	133
139	128	123	140	140	152	136	125	130	153
130	126	129	157	144	142	128	138	142	135
141	139	132	135	145	134	140	136	138	143
122	141	122	132	136	129	138	130	129	135
134	141	133	128	121	131	137	140	133	135
138	132	140	145	128	140	134	128	146	132
131	142	133	137	126	128	129	124	137	127
139	141	157	146	128	136	130	141	129	143
137	143	139	141	121	131	128	133	136	146

1-Var Stats LK200 calculates some statistics of our population. You'll find the result below.

```
1-Var Stats
x̄=135.575
Σx=27115
Σx²=3686399
Sx=7.188375247
σx=7.170381789
↓n=200
```

```
1-Var Stats
fn=200
minX=120
Q1=130
Med=135.5
Q3=140
maxX=157
```

Generating a sample of size 4 out of our population is possible with the following instruction : `seq(LK200(randInt(1,200)),X,1,4)`. `randInt(1,200)` generates the index of an element of the list `LK200`.

```
seq(LK200(randInt(1,200)),X,1,4)
→L1:=mean(L1)
130.5
137.25
135.75
139
```

```
137.5
134
137
136.75
134.5
133.25
137.75
```

$$\frac{130.5+137.25+\dots+137.75}{11} = 135.93$$

The calculations above form a simulation of the fact that the sample mean is an unbiased estimator of the population mean ($E(\bar{X}) = \mu$).

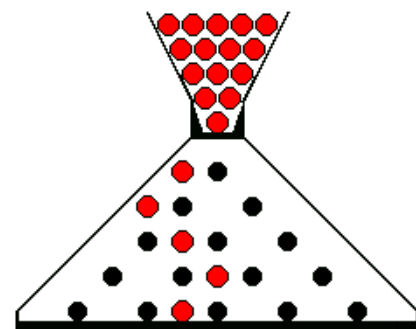
Notice that the sampling above is with replacement. If you want to create a sample without replacement you first have to generate a list of different indices (see paragraph 6).

When you deal with real datasets the size of a sample is usually much smaller than the size of the population. To estimate for example the mean of the population in such a situation there is practically no difference in sampling with or without replacement.

8. Galton's pinball machine

Galton's pinball machine is a device that allows dropping a ball through an array of pins on a board. The pins are attached to the board as the picture alongside shows.

When the ball hits a pin the probability it falls to the left or to the right is equal. After passing through all rows the ball is collected in one of the slots

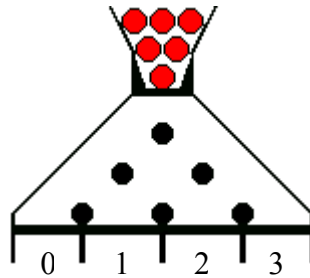


The questions we will ask about this device are :

- How many ways can a ball follow to reach a specific slot ?
- What's the probability that a ball falls into a specific slot ?

8.1 The probability distribution of Galton's pinball machine

We will start with a device with three rows of pins.



To reach the extreme left hand (right hand) slot the ball must fall to the left (right) every time. The way that leads to slot 0 we denote by LLL and the way to slot 3 by RRR.

There are three ways to reach slot 1 : namely the ways LLR, LRL and RLL. The ways RRL, RLR and LRR lead to slot 2.

Let X be the random variable that indicates the number of the slot the ball falls into. The probability distribution of X is given by (the probability of each route is $1/8$) :

<u>$X = \text{the number of the slot}$</u>				
x	0	1	2	3
possible ways to reach x	1	3	3	1
$P(X=x)$	$1/8$	$3/8$	$3/8$	$1/8$

Notice that a ball that falls into slot 0 never falls to the right, one that falls in slot 1 falls to the right once, one that falls in slot 2 twice and one that falls in slot 3 three times.

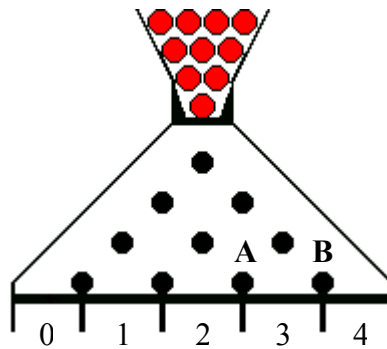
So the number of ways to reach slot k is the same as the number of combinations of k objects from among 3 objects. This number we denote by $\binom{3}{k}$.

With this notation we see that the random variable has the following binomial distribution.

$$X \sim B\left(3, \frac{1}{2}\right)$$

x	0	1	2	3
$P(X=x)$	$\binom{3}{0}\left(\frac{1}{2}\right)^3$	$\binom{3}{1}\left(\frac{1}{2}\right)^3$	$\binom{3}{2}\left(\frac{1}{2}\right)^3$	$\binom{3}{3}\left(\frac{1}{2}\right)^3$

Now we will extend our pinball machine with one row.



There is still one way to reach each of the extreme slots. How many ways are there to reach slot 3 ?

The number of ways to reach slot 3 is the sum of the number of ways to reach A and the number of ways to reach B. So the total number is $1 + 3 = 4$.

The probability that a ball falls into slot 3 is the sum of the probability to reach A and fall to the right and the probability to reach B and fall to the left.

The probability to fall in slot 1 is : $P(X = x) = \frac{1}{8} \cdot \frac{1}{2} + \frac{3}{8} \cdot \frac{1}{2} = \frac{1}{16} + \frac{3}{16} = \frac{4}{16} = \frac{1}{4}$

Counting also the ways to the other slots leads to the following probability distribution.

$X = \text{the number of the slot}$

x	0	1	2	3	4
possible ways to reach x	1	4	6	4	1
$P(X=x)$	1/16	4/16	6/16	4/16	1/16

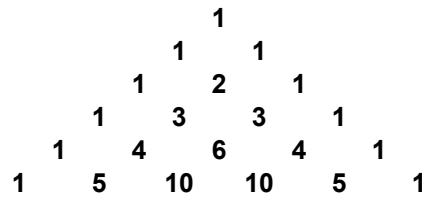
Again X has a binomial distribution with parameters $n = 4$ and $p = \frac{1}{2}$.

$$X \sim B(4, \frac{1}{2})$$

x	0	1	2	3	4
$P(X = x)$	$\binom{4}{0} \left(\frac{1}{2}\right)^4$	$\binom{4}{1} \left(\frac{1}{2}\right)^4$	$\binom{4}{2} \left(\frac{1}{2}\right)^4$	$\binom{4}{3} \left(\frac{1}{2}\right)^4$	$\binom{4}{4} \left(\frac{1}{2}\right)^4$

In general, for a pinball machine with n rows, the probability that a ball falls in slot x is given by the formula : $P(X = x) = \binom{n}{x} \left(\frac{1}{2}\right)^n \quad x = 0, \dots, n$.

If we extend our pinball machine repeatedly with one row of pins and count each time all the possible ways a ball can roll we construct Pascal's triangle.



8.2 A simulation of Galton's pinball machine

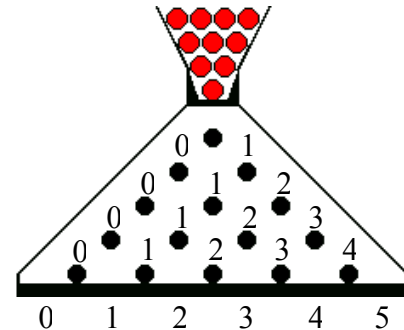
The following is a simulation of a pinball machine with 5 rows of pins.

We denote a fall to the right with 1 and a fall to the left with 0. The next program simulates the rolling of a ball.

```

Ø→P
For (J, 1, 5)
P+randInt(Ø, 1)→P
End
  
```

The value of the variable **P** determines in which slot the ball lands. The **For**-loop can be replaced with the command **randBin(5, Ø.5)**.



If you want to simulate more than one ball you have to put the program mentioned above in another **For**-loop. Each time you have to put the value of the variable **P** into a list.

Below you see some outputs of the program **GALTON**. The source code you can find in the appendix.

NUMBER OF BALLS	2: GRAPH	3: FREQUENCY	4: REL FREQUENCY
5		<pre> (Ø 0) (1 1) (2 2) (3 1) (4 0) (5 1) </pre>	<pre> (Ø 0 .03125) (1 .2 .15625) (2 .4 .3125) (3 .2 .3125) (4 0 .15625) (5 .2 .03125) </pre>
20		<pre> (Ø 0) (1 3) (2 4) (3 7) (4 6) (5 0) </pre>	<pre> (Ø 0 .03125) (1 .15 .15625) (2 .2 .3125) (3 .35 .3125) (4 .3 .15625) (5 0 .03125) </pre>
100		<pre> (Ø 3) (1 14) (2 28) (3 33) (4 19) (5 3) </pre>	<pre> (Ø .03 .03125) (1 .14 .15625) (2 .28 .3125) (3 .33 .3125) (4 .19 .15625) (5 .03 .03125) </pre>

9. Appendix : TI-83 Plus programs

9.1 Galton's pinball machine

```
ClrList L1,L2,L3,L4,L5
Input "NUMBER OF ROWS  ",R
Input "NUMBER OF BALLS  ",N
seq(X,X,0,R)→L1
For(I,1,R+1)
0→L2(I)
End
N→M
Lbl H
For(I,1,N)
0→P
For(J,1,R)
P+randInt(0,1)→P
End
P→L5(I)
L2(P+1)+1→L2(P+1)
End
L2/M→L3
binompdf(R,.5)→L4
Lbl G
ClrHome
Menu("GALTON","EXTRA BALLS",A,"GRAPH",B,"FREQUENCY",C,"REL
FREQUENCY",D,"TOTAL OF BALLS",E,"END",F)
Lbl A
Input "NUMBER OF BALLS  ",N
ClrList L5
M+N→M
Goto H
Lbl B
PlotsOff
FnOff
Plot1(Scatter,L1,L3,□)
Plot2(Scatter,L1,L4,*)
AxesOff
ZoomStat
DispGraph
Pause
Goto G
Lbl C
For(K,1,R+1)
Disp {L1(K),L2(K)}
Pause
End
Goto G
Lbl D
For(K,1,R+1)
Disp {L1(K),L3(K),L4(K)}
Pause
End
Goto G
Lbl E
Disp "NUMBER OF BALLS",M
Pause
Goto G
Lbl F
Stop
```

9.2 Drawing with or without replacement

```
Menu("REPLACEMENT", "WITHOUT", A, "WITH", B)
Lbl A
Disp "DRAW"
Input R
Disp "OUT OF"
Input N
Repeat sum( $\Delta$ List(L2)= $\emptyset$ )= $\emptyset$ 
randInt(1,N,R) $\rightarrow$ L1
L1 $\rightarrow$ L2
SortA(L2)
End
Disp L1
Stop
Lbl B
Disp "DRAW"
Input R
Disp "OUT OF"
Input N
randInt(1,N,R) $\rightarrow$ L1
Disp L1
Stop
```

References

1. G. Blom, *Probability and Statistics – Theory and applications*, Springer-Verlag New York, 1989.
2. G. Herweyers, K. Stulens, *Statistiek met een grafisch rekentoestel*, Acco Leuven, 2000.
3. D.S.Yates, D.S. Moore, G.P. McCabe, *The Practice of Statistics, TI-83 Graphing Calculator Enhanced*, W.H. Freeman and Company, New York, 1999.