The **Law of Averages**

The Expected Value & The Standard Error

Where Are We Going?

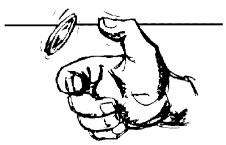
- Sums of random numbers
- The law of averages
- Box models for generating random numbers
- Sums of draws: the Expected Value
- Standard error of a sum: the square root law
- Normal curve approximation to sum

Part 1

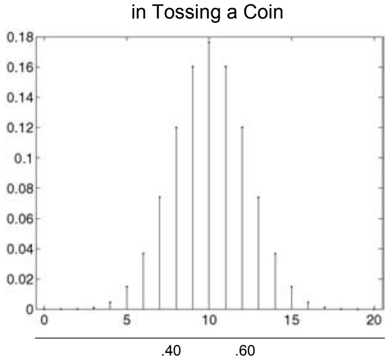
In which we meet the law of averages

Which bet would you choose?

- 10 heads on 20 tosses or 50 heads on 100 tosses?
- Between 8 and 12 heads in 20 tosses or between 48 and 52 on 100 tosses?
- Between 40-60% heads on 20 tosses or 40-60% heads on 100 tosses?

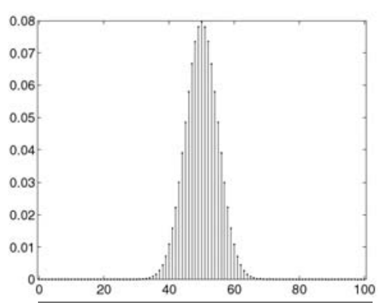


Binomial Probabilities: Number of Heads in Tossing a Coin

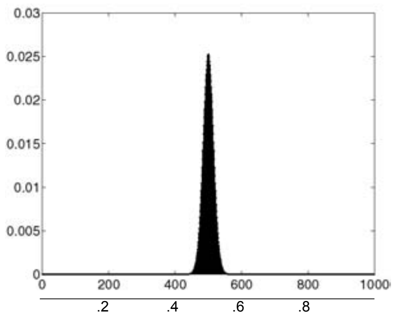


20 tries, p = .5

100 Tries, p = .5



1000 Tries, p = .5



The Law of Averages

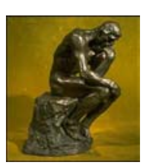
As the number of tosses increases, the fraction of heads tends to a constant.

Number of heads = half the number of tosses

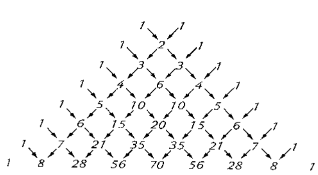
+ chance error

The chance error does not tend to zero, but the chance error divided by the number of tosses does.

If you throw a fair coin 5 times, *every* sequence is equally likely. HTHTH has the same chance as HHHHH. Same holds for 20, 100 tosses, etc. How then can the average be predictable?



When the number of tosses is large, most of these sequences have about half heads and half tails.

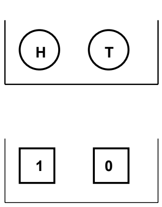


Choices: toss a fair coin 100 times or 1000 times?

- You win if there are more than 60% heads.
- You win if there are fewer than 55% heads.
- You win if get exactly half heads.
- You win if between 45% and 55% are heads.

Part 2
In which we are introduced to box and ticket models

Box Models for Chance Processes: Tossing a Coin



The number of heads equals the sum of the values of the tickets you draw from this box

Box Models for Chance Processes: Roulette

Wheel has 18 red, 18 black, and 2 green slots. You can bet on black or red, a specific number, or several other choices. When betting on black or red you either win or lose $1.

- 18 black tickets
- 18 red tickets
- 2 green tickets

$1	18 black tickets
-$1	18 red tickets
-$1	2 green tickets

Your winning is the value of the ticket you draw from the box.

If you play several times, your total winning is the sum of the values.

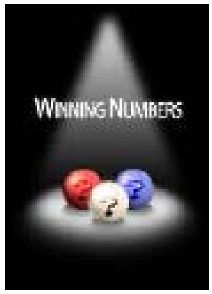
Bet on a Number: 00,0,1,2,...,36

$35 -$1 -$1 -$1 -$1

-$1 -$1 -$1 -$1 -$1

-$1 -$1 etc....

California Lottery: Win and Spin
135,000,000 tickets
Buy a ticket for $1



Prize	Number of Tickets
$1	10,800,000
$2	8,100,000
$5	3,240,000
$10	540,000
$50	54,000
$100	27,000
$500	6,073
$1000	1,350
$10,000	150

All the rest are tickets with -$1

Part 3
In which the Expected Value appears

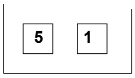
The Expected Value

1	1	1	5

What would you expect the sum to be if you drew 100 tickets from this box?

How many 5's would you expect?

How many 1's would you expect?

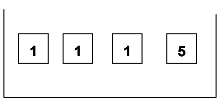


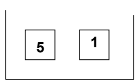
What do you expect the value of the sum of 100 draws to be?

How many 5's would you expect?

How many 1's would you expect?

Another Way



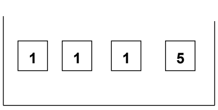


Average ticket value

EV for 100 draws

The EV of the sum of draws equals the number of draws times the box average

A More Mathematical Approach



This box-ticket model is a representation for random numbers, also called "random variables."

X is a random number. $P(X=1) = \frac{3}{4}$; $P(X=5) = \frac{1}{4}$

"Expected value of X" = E(X) = sum of all possible values of X, weighted by their probabilities:

$$E(X) = 1 \times P(X=1) + 5 \times P(X=5)$$
$$= 1 \times \frac{3}{4} + 5 \times \frac{1}{4} = 2$$

The expected value of the sum of n realizations of X is n E(X)

The Expected Value: Betting on Black 100 Times in Roulette

$1	18 black tickets
-$1	18 red tickets
-$1	2 green tickets

Bet on a Number: 00,0,1,2,...,36

$35 -$1 -$1 -$1 -$1

-$1 -$1 -$1 -$1 -$1

-$1 -$1 etc....

$35 -$1 -$1 -$1 -$1

-$1 -$1 -$1 -$1 -$1

-$1 -$1 etc....

Box average =

In 100 draws you expect to lose

Your *expected* losses are the same for both ways of playing.

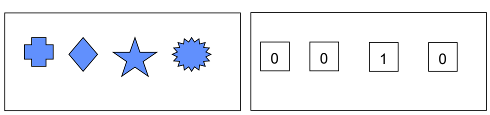
Prize	Number of Tickets
$1	10,800,000
$2	8,100,000
$5	3,240,000
$10	540,000
$50	54,000
$100	27,000
$500	6,073
$1000	1,350
$10,000	150

Your winning equals the prize value minus $1. So the values of the tickets equals the prize values minus $1. The rest of the 135,000,000 tickets have value -$1.

Box Average = -$.56

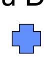
In 100 draws you would expect to lose $56

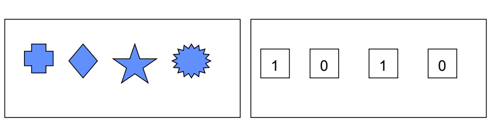
Box Models for Counts: The Number of Times Do You Draw a 



Box Average =

If you draw 20 times, you expect ________ 

Box Models for Counts: The Number of Times Do You Draw a  or a 



Box Average =

? A gambling house offers the following game. A letter is drawn at random from the sentence

WIN OR YOU PAY US

If the letter comes from the word "WIN" you win $1. If it comes from the word "PAY" you pay $1. Otherwise you pay nothing. How much money do you expect to have after playing 40 times?

Tickets and their values:

Box Average:

Expected value:

About how far off of this are you likely to be? That's the next question

Part 4
Wherein the Standard Error is introduced

Sum = Expected Value + Chance Error

How big do we expect the chance error to be? Will define the *Standard Error* (**SE**) of the sum.

How does its size depend on the values of the tickets in the box?

How does its size depend on the number of draws?

Box: 1 2 3

Average of Box = 2

Chance error = Sum - EV

4 Draws: EV = 8

Sums	9	10	6	7	10	10	8	8	10	9
Chance Errors	1	2	-2	-1	2	2	0	0	2	1

16 Draws: EV = 32

Sums	32	30	36	34	25	29	33	26	34	33
Chance Errors	0	-2	4	2	-7	-3	1	- 6	2	1

64 Draws: EV = 128

Sums	139	134	126	128	126	125	122	136	130	119
Chance Errors	11	6	-2	0	-2	-3	-6	8	2	-9

Sum = Expected Value + Chance Error

The **Standard Error (SE)** is a measure of how big the chance error is likely to be.

The **Square Root Law**: the standard error of the sum of draws is

$\sqrt{\text{number of draws}}$ x (SD of the box)

Box: 1 2 3

Average of Box = 2

SD of Box = .8

4 Draws (SE =

Sums	9	10	6	7	10	10	8	8	10	9
Chance Errors	1	2	-2	-1	2	2	0	0	2	1

16 Draws (SE =

Sums	32	30	36	34	25	29	33	26	34	33
Chance Errors	0	-2	4	2	-7	-3	1	- 6	2	1

64 Draws (SE =

Sums	139	134	126	128	126	125	122	136	130	119
Chance Errors	11	6	-2	0	-2	-3	-6	8	2	-9

Conditions for the Square Root Law to Hold

- The draws are all from the same box.
- The draws are independent (with replacement).

A gambling house offers the following game. A letter is drawn at random from the sentence

WIN OR YOU PAY US

If the letter comes from the word "WIN" you win $1. If it comes from the word "PAY" you pay $1. Otherwise you pay nothing. How much money do you expect to have after playing 40 times?

Tickets and their values:

Box Average:

Expected value:

About how far off of this are you likely to be?

Notation and Terminology

X: a random number, or "random variable"

x_1 x_2 ,…, x_n: a list of the values that X can take on. These are the numbers on the tickets in the box.

$p(x_1)$, $p(x_2)$,…, $p(x_n)$: the probabilities of taking on those values.

$E(X) = x_1\,p(x_1) + x_2\,p(x_2) + \ldots + x_n\,p(x_n)$

The expected value of X. Often denoted by μ

$E(X) = \mu$

$Var(X) = (x_1 - \mu)^2 p(x_1) + (x_2 - \mu)^2 p(x_2) + ... + (x_n - \mu)^2 p(x_n) = \sigma^2$

The variance of X

The standard deviation of X is the square root of the variance. Usually denoted by σ

Suppose you have n independent realizations of this random variable, like n draws from a box with replacement. Their sum is

$S = X_1 + X_2 + \ldots X_n$

S is also a random variable

The expected value of S is

$E(S) = n\,\mu$

The variance of S is

$Var(S) = n\,\sigma^2$

The SE (also called SD) of S is the square root of the variance

Viewed from this more mathematical perspective, the device of "tickets in a box" enables us to compute the expected value and standard deviation of a random number by representing it as the value of a ticket drawn from the box.

In a more traditional development, we would have defined the expectation, variance, and standard deviation of a "random variable" (a random number) and then prove facts about the sum of independent random variables.

Roulette: Betting on Black

Ticket	Count
$1	18 black tickets
-$1	18 red tickets
-$1	2 green tickets

Box Average = -.0526
SD of Box = .99

Draw 100 times

EV of sum =

SE of sum of =

Chance error = sum – (- $5.26)

Sums:	11	6	-2	0	-2	-3	-6	-9
Errors	16.26	11.26	3.26	5.26	3.26	2.26	-.74	-3.74

Roulette: Betting on a Number

Box contains 1 ticket worth $35 and 37 tickets worth -$1

Box Average = -.0526
Box SD = $5.75

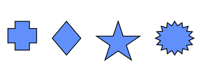
Draw 100 times

Expected value of sum =
SE of sum of =

Chance error = sum - expected value

Sums:	-64	8	-28	-100	-28
Errors:	-58.74	13.26	-22.74	-94.74	-22.74

SE's for Counts: The Number of Times You Draw a 



0 0 1 0

Box Average = 1/4

Special Formula for SD of a 0-1 box:

$$\sqrt{(\text{fraction of 1's}) \times (\text{fraction of 0's})}$$

For this box, SD = $\sqrt{(1/4)\text{x}(3/4)}$ = .87

0	0	1	0

Box Average = 1/4 Box SD = .87

Suppose you draw 100 times

How many ★ would you expect?

What is the SE of the number of ★?

Would you be surprised by 30? By 70?

Shortcut for a box with only two values: a and b

SD of box is

$$|a-b|\sqrt{(fraction\ of\ a)\times(fraction\ of\ b)}$$

Example

A multiple choice exam has 100 questions. Each question has 5 possible answers, one of which is correct. Four points are given for the right answer and a point is taken off for the wrong answer.

A student guesses randomly for each question. The student expects to score ______ give or take ____.

What's the box model:

4	-1	-1	-1	-1

Box average:

Box SD:

)

100 draws. A student guesses randomly for each question. The student expects to score ?? give or take ??

Investment Diversification

Option 1: Invest $1000 in each of ten companies. Lose $100 with chance .40 and gain $100 with chance .60.

Option 2: Invest $100 in each of 100 companies. For each one, lose $10 with chance .4 and gain $10 with chance .6. Returns are *independent.*

Which is better?

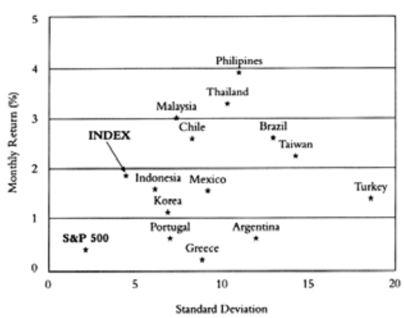
:

Option 1: Invest $1000 in each of 10 companies. For each one, lose $100 with chance .4 and gain $100 with chance .6. Returns are *independent.*

Consider 10 draws from a box with 40% of tickets worth -$100 and 60% of tickets worth +$100.

Option 2: Invest $100 in each of 100 companies. For each one, lose $10 with chance .4 and gain $10 with chance .6. Returns are *independent.*

Consider 100 draws from a box with 40% of tickets worth -$10 and 60% of tickets worth +$10.



The blessings of diversification. The track records of 13 emerging stock markets compared to the index (average of 13) and the S&P 500 from January 1992 through June 1994. The data are in percentages per month.

Box Model Demo

Part 5
We look back at the landscape we have traversed

Where Have We Been?

- *The law of averages*: As the number of draws from a box increases, their average value tends to the expected value (the box average).
- The *expected value* of the sum of the draws equals the number of draws times the box average.
- Sum = Expected Value + Chance Error
- The chance error of a sum does not tend to 0.

- The *standard error* is a measure of how big the chance error is likely to be.
- The *square root law*: the standard error of a sum is the square root of the number of draws times the SD of the box.
- Special rule for the SD of a box that only has two numbers in it:

$$(\text{big \# - small \#}) \times \sqrt{\text{frac big \# } \times \text{ frac small \#}}$$