Statistical Literacy: Coincidence

MILO SCHIELD,
Augsburg College
Director, W. M. Keck Statistical Literacy Project
US Rep, International Statistical Literacy Project
Member, International Statistical Institute

National Numeracy Network Workshop
Oct 11, 2014.
www.StatLit.org/pdf/2014-Schield-NNN1-Slides.pdf

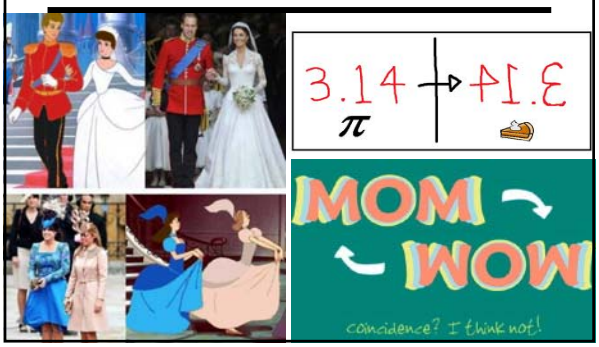
Law of Very-Large Numbers

Not the same as Law of Large Numbers!!!

Unlikely is almost certain given enough tries.

Given an event: one chance in N.
In N tries, one event is 'expected';
* More likely than not. Schield (2012)

Coincidence?



The "Birthday" Problem: Chance of a matching birthday

Richard von Mises (1938)
In a group of 28 people,
a birthday match is *expected*.
The trick is to show it,
– not just to prove

Try this Excel den
www.StatLit.org/Excel/2012Schield-Bday.xls

The "Birthday" Problem Math Answer

If the chance of an rare event is p and $p = 1/k$, then this event is "expected" in k trials.

In a group of size N, there are (N-1)(N/2) pairs.

Solve for N(k). k = (N-1)(N/2) = (N^2-N)/2

Quadratic: $N^2 - N - 2k = 0$

Estimate: $N^2 \sim 2/p$.

Trial and error: $27^2 \sim 2*364$

Q. Are students convinced? No!!!

49 Connections: Quadrant 1

Schield (2011) RICHARD VON MISES' BIRTHDAY PROBLEM 28 People

		Month	10	11	11	9	4	7	6		
		Day	16	18	8	9	13	25	24		
Month	Day									Month	Day
8	20							1		7	25
10	29									8	16
4	11									11	6
3	3									11	29
1	3									8	3
3	30									3	24
10	28									1	15
		Month	5	2	6	2	1	7	5		
		Day	28	8	6	12	14	1	25		

2014 Coincidence NNN1 7

49 Connections: Quadrant 2

Schield (2011) RICHARD VON MISES' BIRTHDAY PROBLEM 28 People

		Month	8	12	7	11	6	4	2		
		Day	28	2	15	15	5	24	2		
Month	Day									Month	Day
10	8									2	5
5	17									2	17
9	13									12	26
11	18									3	6
12	21							2		4	20
2	28									10	2
10	11									3	23
		Month	10	7	4	12	8	4	8		
		Day	22	22	10	6	4	20	21		

2014 Coincidence NNN1 8

49 Connections: Quadrant 3

Schield (2011) RICHARD VON MISES' BIRTHDAY PROBLEM 28 People

		Month	3	8	7	5	6	8	11		
		Day	4	5	25	27	19	4	26		
Month	Day									Month	Day
7	15									12	13
4	31									7	30
11	3									2	1
8	15									4	14
3	28									10	25
3	18									1	18
2	26		3							12	23
		Month	2	3	2	4	6	11	9		
		Day	26	26	23	6	30	11	8		

2014 Coincidence NNN1 9

49 Connections: Quadrant 4

Schield (2011) RICHARD VON MISES' BIRTHDAY PROBLEM 28 People

		Month	11	11	3	5	1	5	2		
		Day	5	27	17	3	5	19	4		
Month	Day									Month	Day
11	5		4							11	12
11	17									8	24
8	2									5	1
4	26									3	28
4	22									10	13
10	8									4	4
12	22									8	11
		Month	1	7	5	5	12	10	5		
		Day	2	1	23	7	20	14	14		

2014 Coincidence NNN1 10

49 Connections: Side-To-Side

Schield (2011) RICHARD VON MISES' BIRTHDAY PROBLEM 28 People

		Month	2	3	10	6	6	9	6		
		Day	14	3	13	27	13	7	24		
Month	Day									Month	Day
1	24									1	31
9	8	E								6	28
12	6									12	24
12	28									10	1
10	27									11	19
9	18								W	9	8
4	12									4	16
		Month	8	8	6	5	7	4	7		
		Day	13	3	19	3	30	9	18		

2014 Coincidence NNN1 11

49 Connections: Top-to-Bottom

Schield (2011) RICHARD VON MISES' BIRTHDAY PROBLEM 28 People

		Month	11	8	10	10	8	10	3		
		Day	19	3	28	17	27	29	5		
Month	Day					S				Month	Day
5	23									1	12
1	1									11	17
9	6									12	3
10	13									7	29
7	14									2	17
8	30									4	2
1	8									8	17
					N						
		Month	12	3	10	9	12	9	5		
		Day	24	6	17	19	1	20	29		

2014 Coincidence NNN1 12

21 Connections: Same-Side

Schield (2011) RICHARD VON MISES' BIRTHDAY PROBLEM 28 People

		Month	3	2	2	3	9	3	5		
		Day	4	5	9	29	20	5	20		
Month	Day									Month	Day
6	22								E	4	1
10	8									7	10
5	5									3	26
11	23									3	10
3	27								E	4	1
10	2									9	8
2	21									5	7
		Month	8	1	10	12	9	5	5		
		Day	18	6	11	9	3	26	19		

Connections and Chance

Pairs	GROUP	Details
196	Quadrants 1-4	49 pairs each
49	Side-to-Side	
49	Top-to-Bottom	
84	Within each side	21 pairs each
378	TOTAL	

A "birthday" match has one chance in 365.

In a group of 28, we have 378 pairs: (N-1)(N/2).

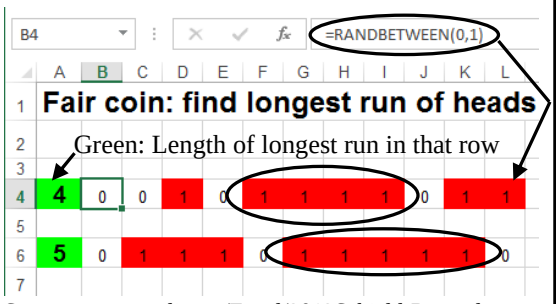
A match is expected: Match is more likely than not.

Runs: Flipping Coins

Law of Very-Large Numbers (Qualitative):
The very unlikely is almost certain given enough tries

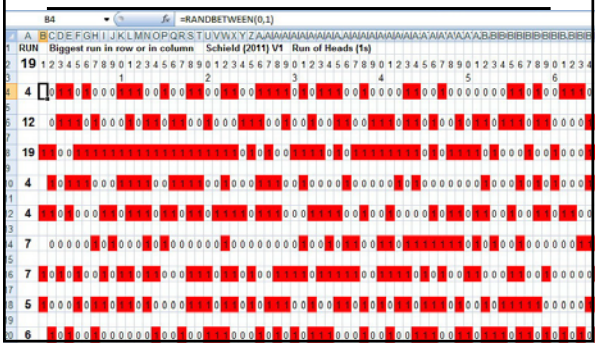
Law of Expected Values:
Events with 1 chance in k are "expected" in k tries.

Flip coins in rows. 1=Heads (Red fill) Adjacent Red cells is a Run of heads.



Source: www.statlit.org/Excel/2012Schield-Runs.xls

Chance of a run of 19 heads: One chance in 2^19 = 1 in 524,288



Consider a run of 10 heads? What is the chance of that?

Question is ambiguous! Doesn't state context!

1. Chance of 10 heads on **the next 10 flips**?
   p = 1/2; k = 10.
   P = p^k = (1/2)^10 = one chance in 1,024

2. What is the chance of *at least* one set of 10 heads [**somewhere**] when flipping 1,024 sets of 10 coins each? At least 50%.*

* Schield (2012)

Runs in Flipping a Fair Coin

1) Unlikely is expected given enough tries.

2) Unlikely (1 chance in k) is *expected* in k tries

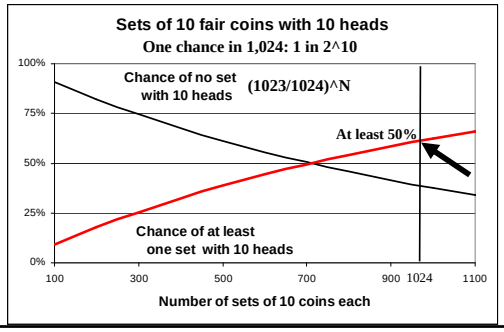
Run of 6 is expected in 64 tries: 2^6 = 64.

Run of 7 is expected in 128 tries: 2^7 = 128

Run of 8 is expected in 256 tries: 2^8 = 256

k tries = k flips of a coin

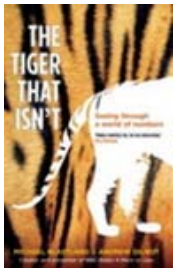
Coincidence increases as data size increases



Michael Blastland's *The Tiger that Isn't*

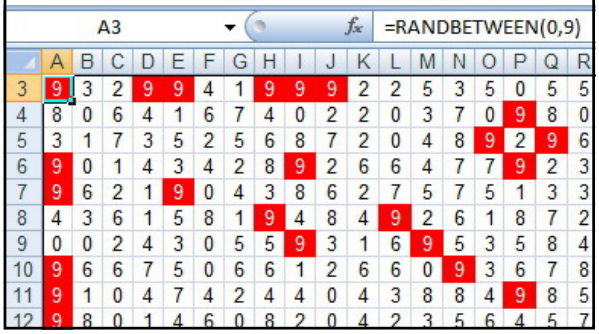
With rice scattered in two dimensions, people can often see memorable shapes.

After this webinar, check out this Excel scattered-rice demo with 1 chance in 100 per cell:



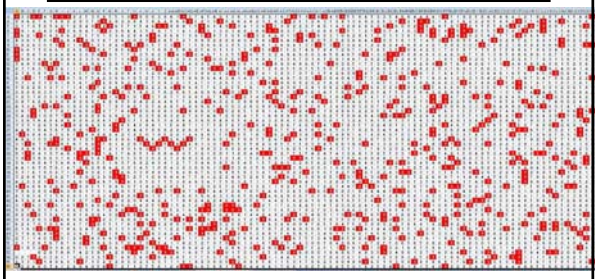
www.StatLit.org/Excel/2012Schield-Rice.xls

Patterns in Rice: # Touching
2:1/100; 4:1/10,000; 6: 1/1,000,000



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
3	9	3	2	9	9	4	1	9	9	9	2	2	5	3	5	0	5	5
4	8	0	6	4	1	6	7	4	0	2	2	0	3	7	0	9	8	0
5	3	1	7	3	5	2	5	6	8	7	2	0	4	8	9	2	9	6
6	9	0	1	4	3	4	2	8	9	2	6	6	4	7	7	9	2	3
7	9	6	2	1	9	0	4	3	8	6	2	7	5	7	5	1	3	3
8	4	3	6	1	5	8	1	9	4	8	4	9	2	6	1	8	7	2
9	0	0	2	4	3	0	5	5	9	3	1	6	9	5	3	5	8	4
10	9	6	6	7	5	0	6	6	1	2	6	6	0	9	3	6	7	8
11	9	1	0	4	7	4	2	4	4	0	4	3	8	8	4	9	8	5
12	9	8	0	1	4	6	0	8	2	0	4	2	3	5	6	4	5	7

3 touching: 1in 1,000
6 touching: 1 in a million



Coincidence Outcomes

Students must "see" that coincidence

- may be more common than expected
- depends on the context
- may be totally spurious
- may be a sign of causation

References

Papers:

Schield (2012). Coincidence in Runs and Clusters
www.statlit.org/pdf/2012Schield-MAA.pdf

Schield (2014). Two Big Ideas for Teaching Big Data
www.statlit.org/pdf/2014-Schield-ECOTS.pdf

Downloadable spreadsheets:

- Birthdays: www.statlit.org/Excel/2012Schield-Bday.xls
- Runs of Coins: www.statlit.org/Excel/2012Schield-Runs.xls

Statistical Literacy: Coincidence

MILO SCHIELD,

Augsburg College

Director, W. M. Keck Statistical Literacy Project

US Rep, International Statistical Literacy Project

Member, International Statistical Institute

National Numeracy Network Workshop

Oct 11, 2014.

www.StatLit.org/pdf/2014-Schield-NNN1-Slides.pdf

Law of Very-Large Numbers

Not the same as Law
of Large Numbers!!!

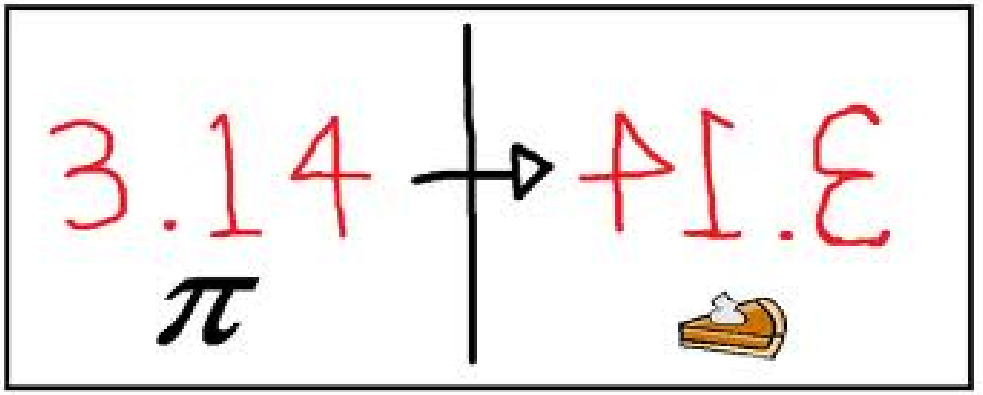
Unlikely is almost certain
given enough tries.

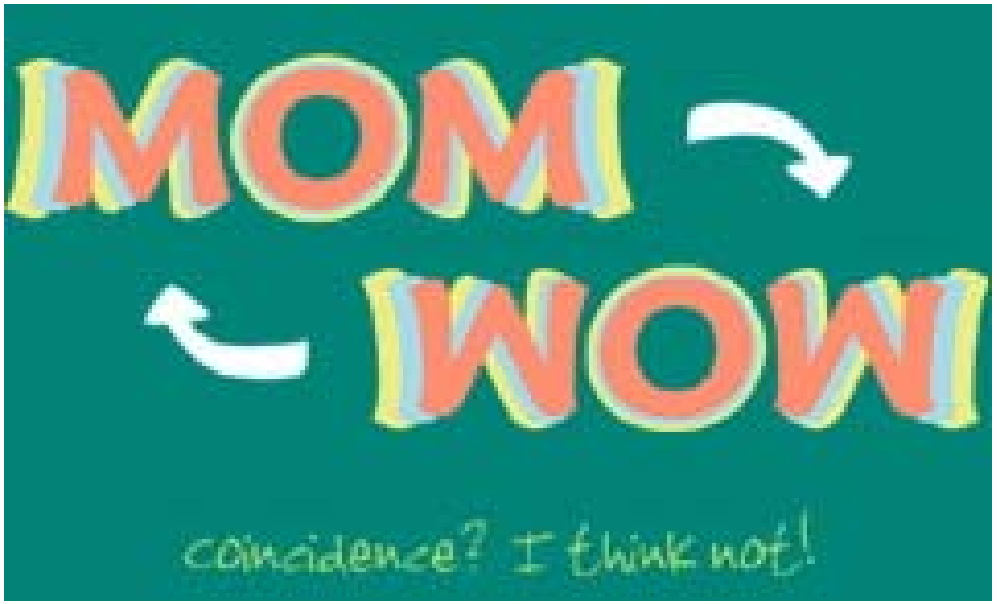
Given an event: one chance in N.

In N tries, one event is ‘expected’;

* More likely than not. Schield (2012)

Coincidence?





The "Birthday" Problem: Chance of a matching birthday

Richard von Mises (1938)

In a group of 28 people,
a birthday match is *expected*.

The trick is to show it,
– not just to prove

Try this Excel den

www.StatLit.org/Excel/2012Schield-Bday.xls

The "Birthday" Problem
Math Answer

If the chance of an rare event is p and $p = 1/k$, then this event is "expected" in k trials.

In a group of size N, there are (N-1)(N/2) pairs.

Solve for N(k). k = (N-1)(N/2) = (N^2-N)/2

Quadratic: $N^2 - N - 2k = 0$

Estimate: $N^2 \sim 2/p$.

Trial and error: $27^2 \sim 2*364$

Q. Are students convinced? No!!!

49 Connections: Quadrant 1

Schield (2011)			RICHARD VON MISES' BIRTHDAY PROBLEM							28 People	
		Month	10	11	11	9	4	7	6		
		Day	16	18	8	9	13	25	24		
Month	Day									Month	Day
8	20							1		7	25
10	29									8	16
4	11									11	6
3	3									11	29
1	3									8	3
3	30									3	24
10	28									1	15
		Month	5	2	6	2	1	7	5		
		Day	28	8	6	12	14	1	25		

49 Connections: Quadrant 2

Schield (2011)			RICHARD VON MISES' BIRTHDAY PROBLEM							28 People	
		Month	8	12	7	11	6	4	2		
		Day	28	2	15	15	5	24	2		
Month	Day									Month	Day
10	8									2	5
5	17									2	17
9	13									12	26
11	18									3	6
12	21							2		4	20
2	28									10	2
10	11									3	23
		Month	10	7	4	12	8	4	8		
		Day	22	22	10	6	4	20	21		

49 Connections: Quadrant 3

Schield (2011)			RICHARD VON MISES' BIRTHDAY PROBLEM							28 People	
		Month	3	8	7	5	6	8	11		
		Day	4	5	25	27	19	4	26		
Month	Day									Month	Day
7	15									12	13
4	31									7	30
11	3									2	1
8	15									4	14
3	28									10	25
3	18									1	18
2	26		3							12	23
		Month	2	3	2	4	6	11	9		
		Day	26	26	23	6	30	11	8		

49 Connections: Quadrant 4

Schield (2011)			RICHARD VON MISES' BIRTHDAY PROBLEM							28 People	
		Month	11	11	3	5	1	5	2		
		Day	5	27	17	3	5	19	4		
Month	Day									Month	Day
11	5		4							11	12
11	17									8	24
8	2									5	1
4	26									3	28
4	22									10	13
10	8									4	4
12	22									8	11
		Month	1	7	5	5	12	10	5		
		Day	2	1	23	7	20	14	14		

49 Connections: Side-To-Side

Schield (2011)			RICHARD VON MISES' BIRTHDAY PROBLEM								28 People	
		Month	2	3	10	6	6	9	6			
		Day	14	3	13	27	13	7	24			
Month	Day										Month	Day
1	24										1	31
9	8	E									6	28
12	6										12	24
12	28										10	1
10	27										11	19
9	18									W	9	8
4	12										4	16
		Month	8	8	6	5	7	4	7			
		Day	13	3	19	3	30	9	18			

49 Connections: Top-to-Bottom

Schield (2011)			RICHARD VON MISES' BIRTHDAY PROBLEM							28 People	
		Month	11	8	10	10	8	10	3		
		Day	19	3	28	17	27	29	5		
Month	Day					S				Month	Day
5	23									1	12
1	1									11	17
9	6									12	3
10	13									7	29
7	14									2	17
8	30									4	2
1	8									8	17
					N						
		Month	12	3	10	9	12	9	5		
		Day	24	6	17	19	1	20	29		

21 Connections: Same-Side

Schield (2011)			RICHARD VON MISES' BIRTHDAY PROBLEM								28 People	
		Month	3	2	2	3	9	3	5			
		Day	4	5	9	29	20	5	20			
Month	Day										Month	Day
6	22									E	4	1
10	8										7	10
5	5										3	26
11	23										3	10
3	27									E	4	1
10	2										9	8
2	21										5	7
		Month	8	1	10	12	9	5	5			
		Day	18	6	11	9	3	26	19			

Connections and Chance

Pairs	GROUP	Details
196	Quadrants 1-4	49 pairs each
49	Side-to-Side	
49	Top-to-Bottom	
84	Within each side	21 pairs each
378	TOTAL	

A “birthday” match has one chance in 365.

In a group of 28, we have 378 pairs: (N-1)(N/2).

A match is expected: Match is more likely than not.

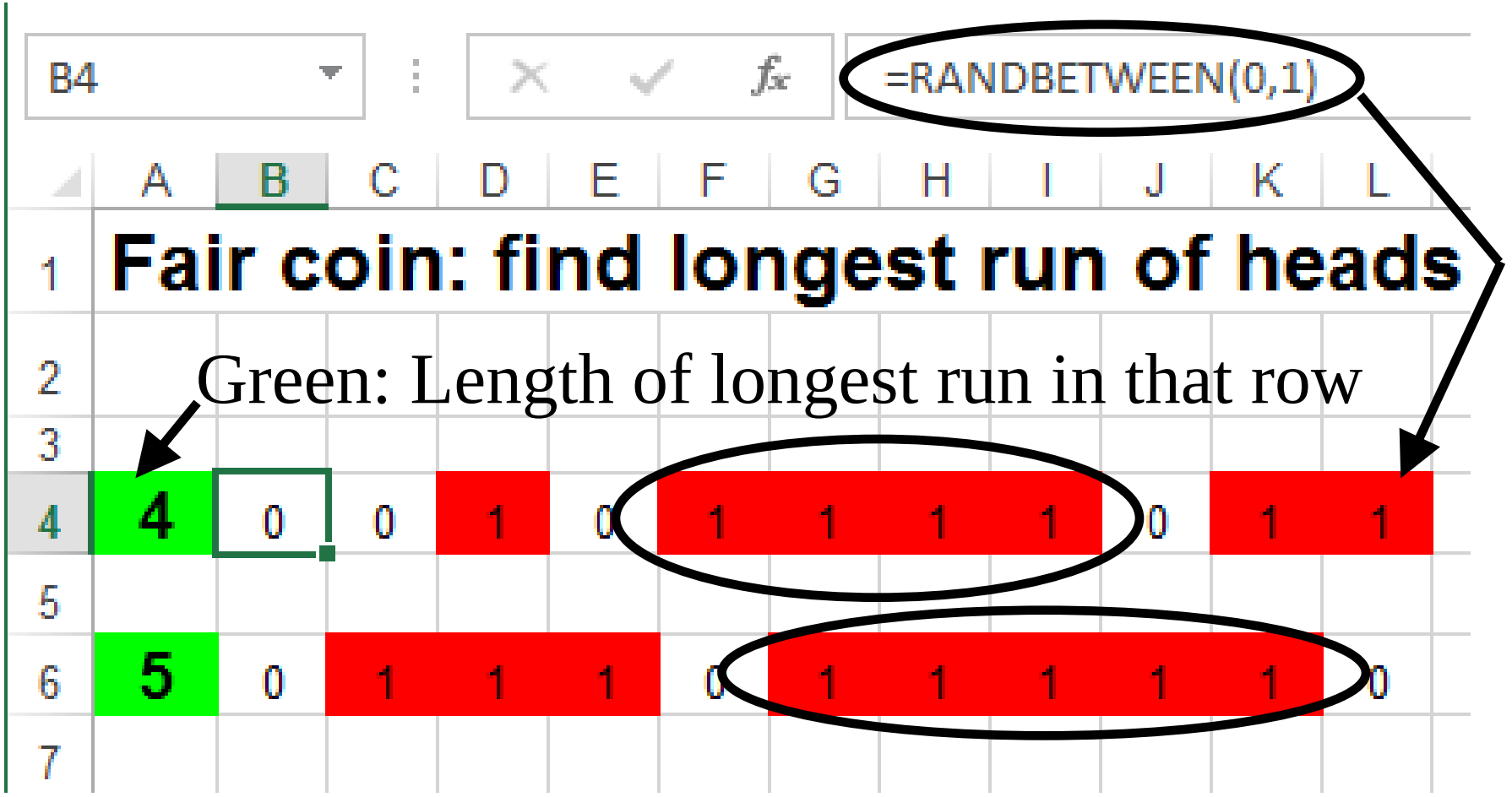
Runs: Flipping Coins

Law of Very-Large Numbers (Qualitative):
The very unlikely is almost certain given enough tries

Law of Expected Values:

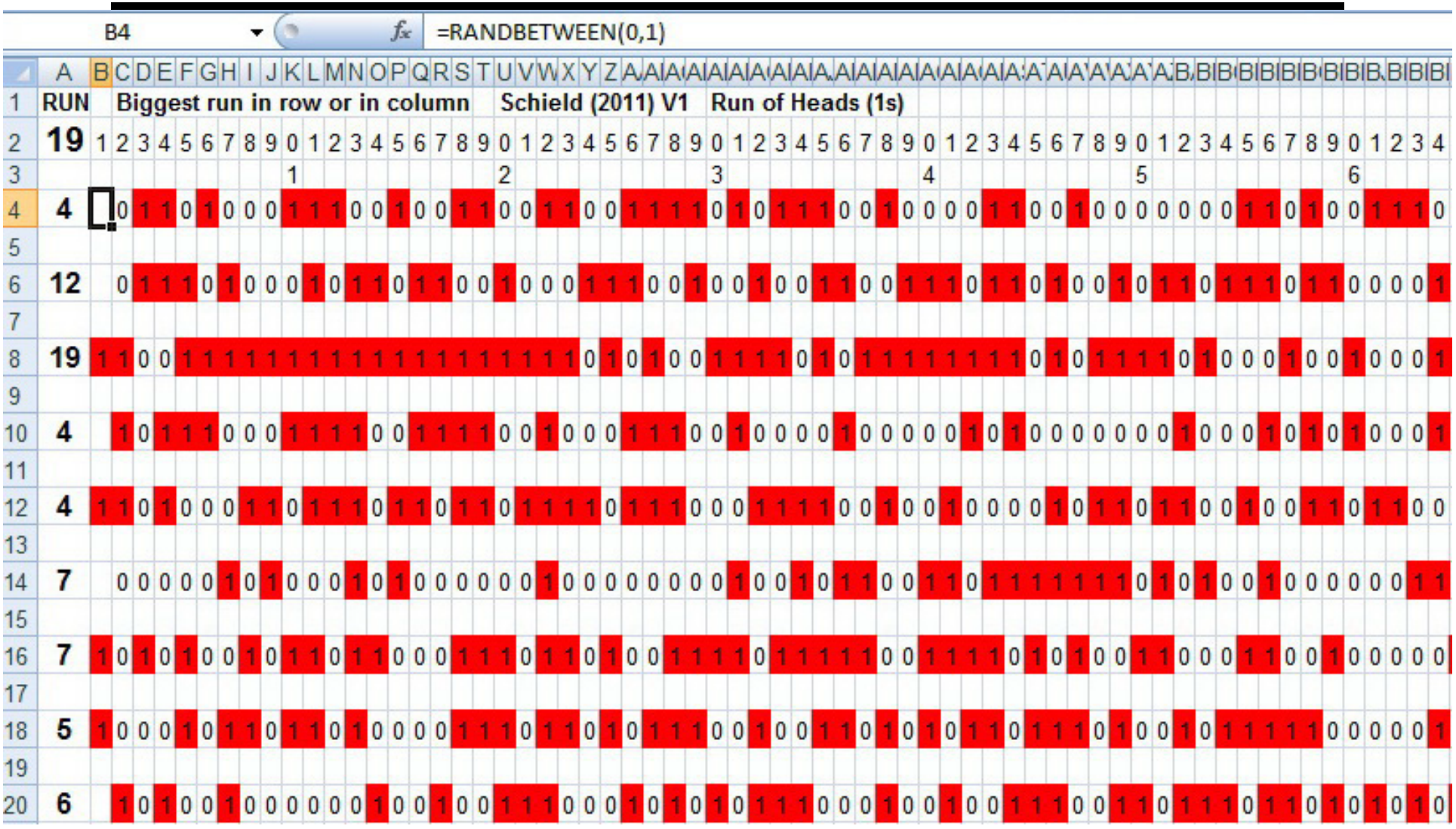
Events with 1 chance in k
are “expected” in k tries.

Flip coins in rows. 1=Heads (Red fill)
Adjacent Red cells is a Run of heads.



Source: www.statlit.org/Excel/2012Schield-Runs.xls

Chance of a run of 19 heads:
One chance in 2^19 = 1 in 524,288



Consider a run of 10 heads? What is the chance of that?

Question is ambiguous! Doesn't state context!

1. Chance of 10 heads on **the next 10 flips**?
   p = 1/2; k = 10.
   P = p^k = (1/2)^10 = one chance in 1,024

2. What is the chance of *at least* one set of 10 heads [**somewhere**] when flipping 1,024 sets of 10 coins each? At least 50%.*

* Schield (2012)

Runs in Flipping a Fair Coin

1) Unlikely is expected given enough tries.

2) Unlikely (1 chance in k) is *expected* in k tries

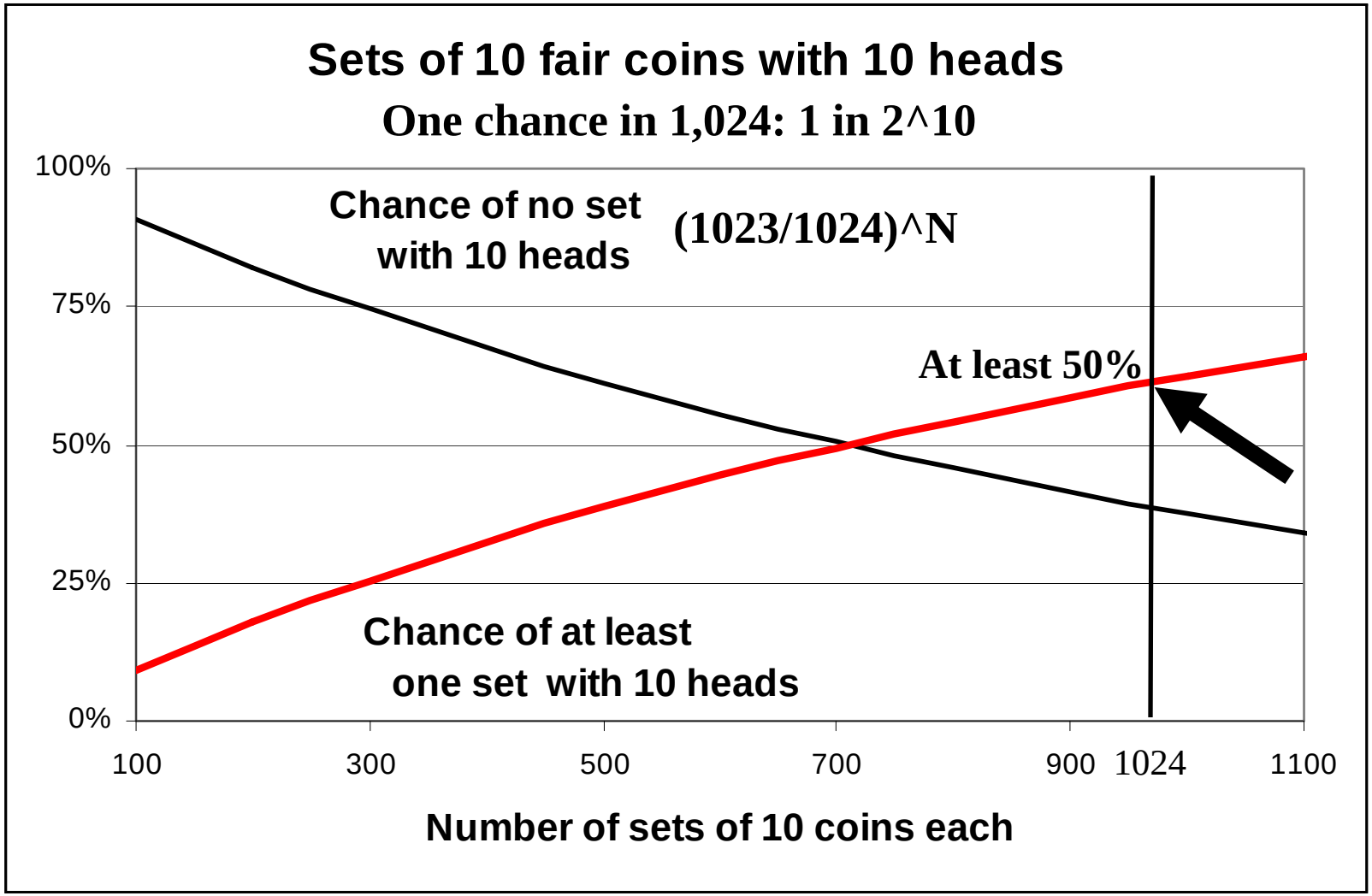
Run of 6 is expected in 64 tries: 2^6 = 64.

Run of 7 is expected in 128 tries: 2^7 = 128

Run of 8 is expected in 256 tries: 2^8 = 256

k tries = k flips of a coin

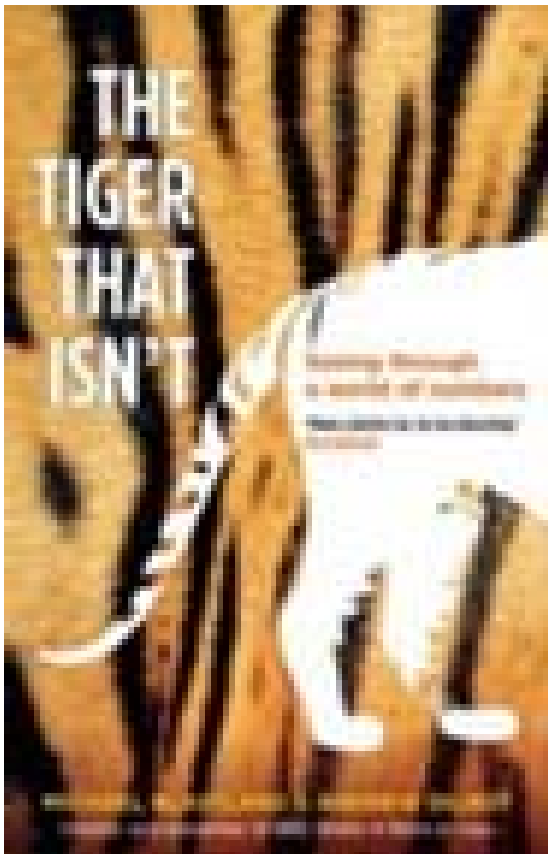
Coincidence increases as data size increases



Michael Blastland's *The Tiger that Isn't*

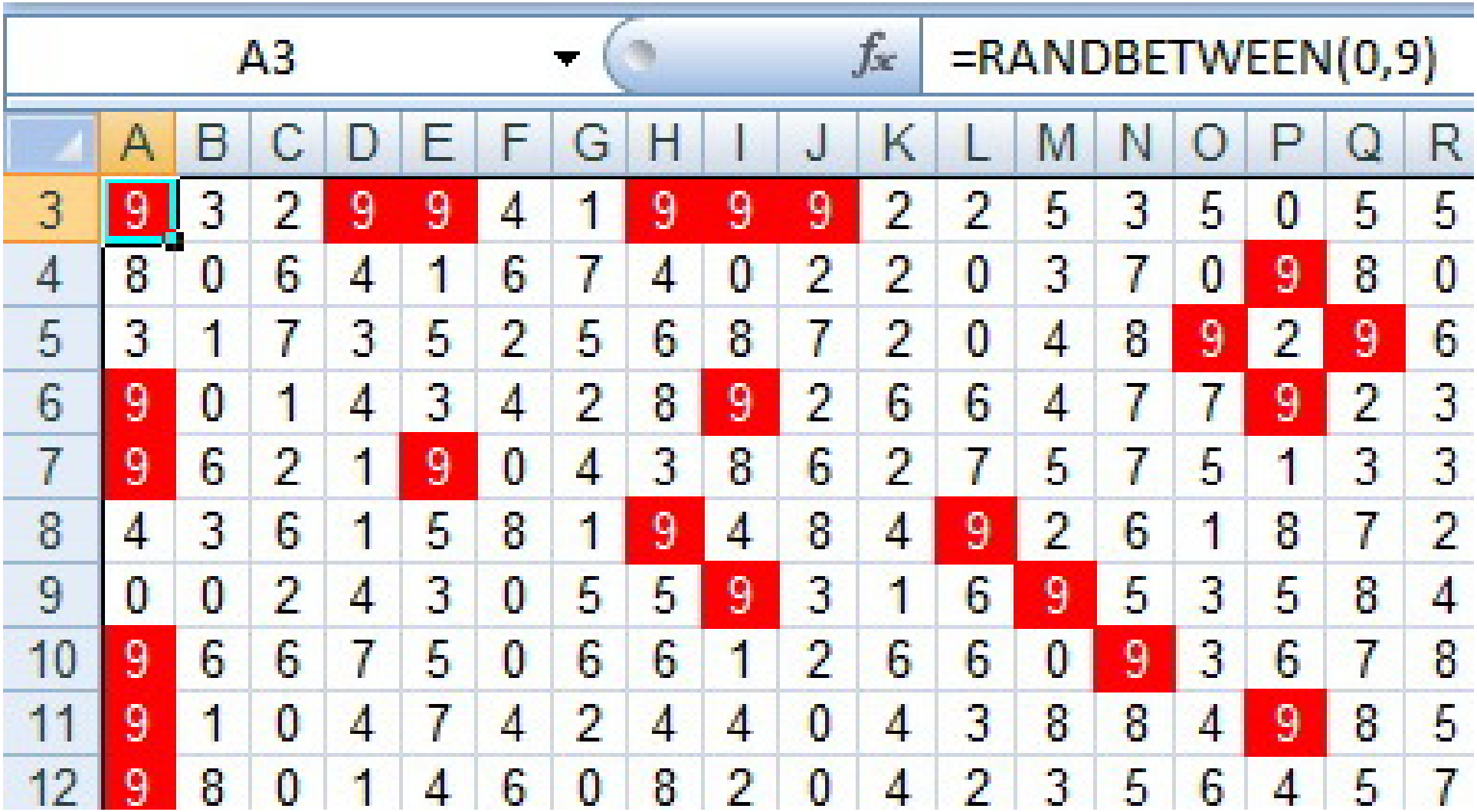
With rice scattered in two dimensions, people can often see memorable shapes.

After this webinar, check out this Excel scattered-rice demo with 1 chance in 100 per cell:



www.StatLit.org/Excel/2012Schield-Rice.xls

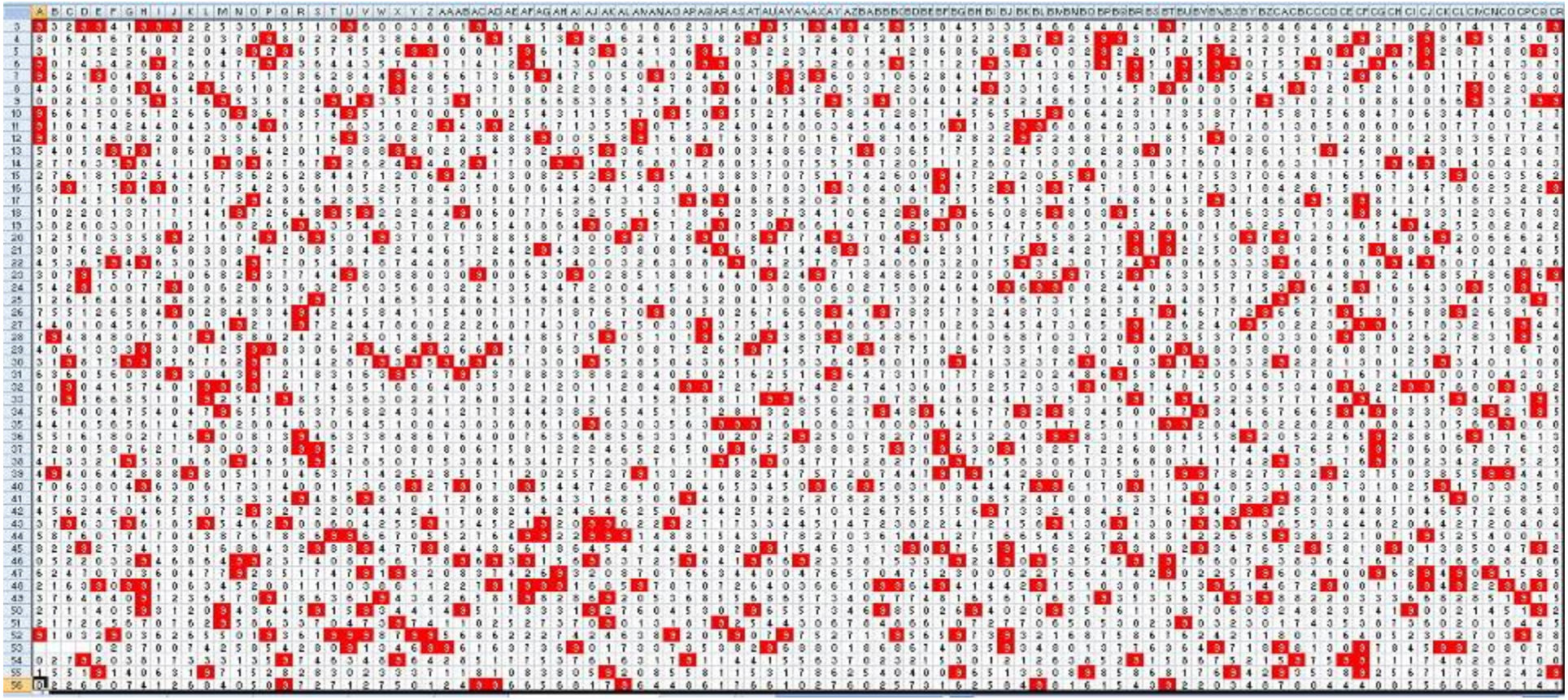
Patterns in Rice: # Touching
2:1/100; 4:1/10,000; 6: 1/1,000,000



A3 =RANDBETWEEN(0,9)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
3	9	3	2	9	9	4	1	9	9	9	2	2	5	3	5	0	5	5
4	8	0	6	4	1	6	7	4	0	2	2	0	3	7	0	9	8	0
5	3	1	7	3	5	2	5	6	8	7	2	0	4	8	9	2	9	6
6	9	0	1	4	3	4	2	8	9	2	6	6	4	7	7	9	2	3
7	9	6	2	1	9	0	4	3	8	6	2	7	5	7	5	1	3	3
8	4	3	6	1	5	8	1	9	4	8	4	9	2	6	1	8	7	2
9	0	0	2	4	3	0	5	5	9	3	1	6	9	5	3	5	8	4
10	9	6	6	7	5	0	6	6	1	2	6	6	0	9	3	6	7	8
11	9	1	0	4	7	4	2	4	4	0	4	3	8	8	4	9	8	5
12	9	8	0	1	4	6	0	8	2	0	4	2	3	5	6	4	5	7

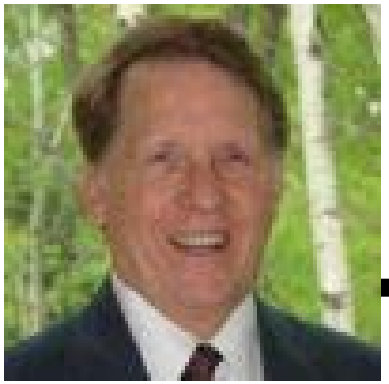
3 touching: 1in 1,000
6 touching: 1 in a million



Coincidence Outcomes

Students must "see" that coincidence

- may be more common than expected
- depends on the context
- may be totally spurious
- may be a sign of causation



References

Papers:

Schield (2012). Coincidence in Runs and Clusters
www.statlit.org/pdf/2012Schield-MAA.pdf

Schield (2014). Two Big Ideas for Teaching Big Data
www.statlit.org/pdf/2014-Schield-ECOTS.pdf

Downloadable spreadsheets:

- Birthdays: www.statlit.org/Excel/2012Schield-Bday.xls
- Runs of Coins: www.statlit.org/Excel/2012Schield-Runs.xls