May 2012 1

Coincidence in Runs and Clusters

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

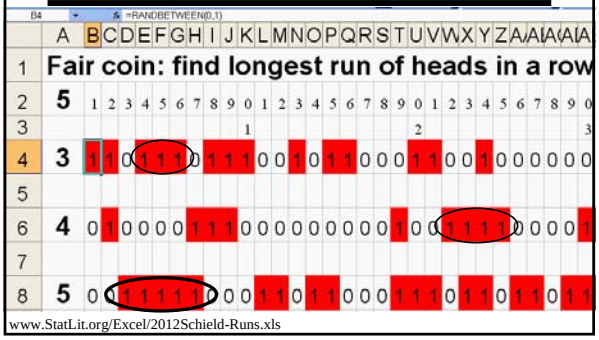
May, 2012
Slides at www.StatLit.org/pdf/2012Schield-eCOTS-6up.pdf

May 2012 2

Coincidence?

May 2012

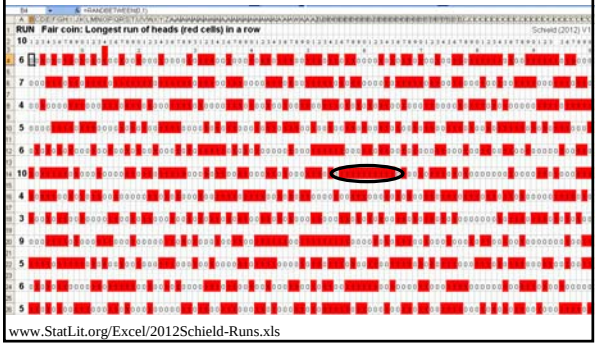
Run of Heads (Red Cells): Chance of 5 Touching: 1 in 32 (2^5)



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

May 2012

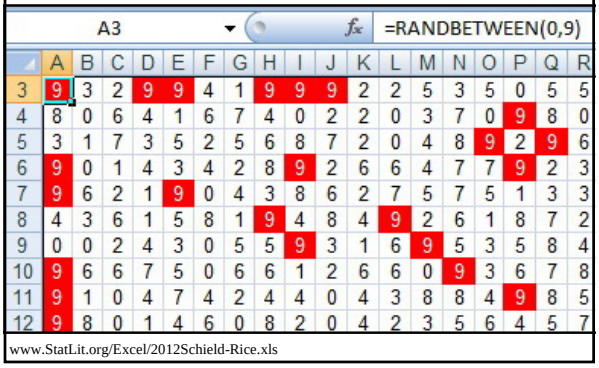
Run of 10 heads (Ones): Chance = 1 in 1,024 (2^10)



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

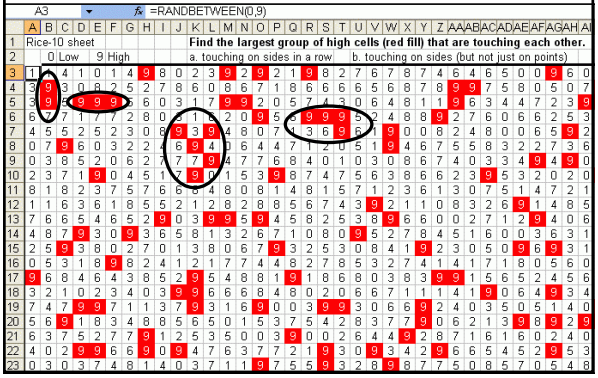
May 2012

Clusters of Rice Chance of Red: One in 10



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

Chance of a Red Cell: 1 in 10 Chance Two Reds Touch: 1 in 100



May 2012 7

Runs and clusters are much more likely than expected!

When students press F9, they often get:

RUNS	CLUSTERS
a run of 10 heads:	a cluster of six squares:
one chance in 2^10	one chance in 10^6
a "thousand-year flood" every year	a "million-year flood" every year

They get unlikely results every time!

May 2012 8

Runs and clusters are much more likely than expected!!!

WHY?

May 2012 9

Explanation #1

What is the chance of that?

I. The question is ambiguous?

- Before or after the fact (ex post vs. ex ante)
- Specific place or anywhere
- Painting the target before or after the shooting

May 2012 10

Explanation #2

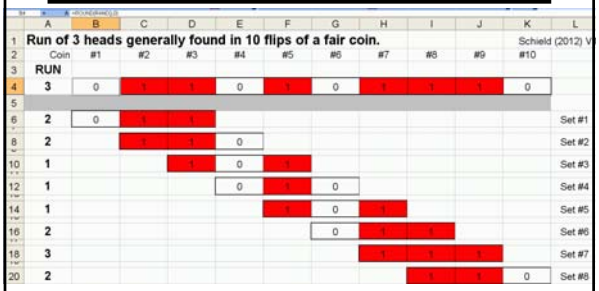
Law of Very-Large Numbers:

'Impossible': almost certain given enough tries

The Law of Very-Large Numbers isn't covered in most intro-stats courses where the focus is on sampling error in small samples.

May 2012 11

#3 Compression: Flipping 10 coins. Almost the same as 24: 8 sets of 3



May 2012 12

Students must "see" that coincidences:

are much more common than expected

1. **involve an ambiguity (ex-ante/ex-post)**
2. **involve the Law of Very-Large Numbers**
3. **involve compression**
4. **may still be signs of causation.**

Example: Cholera outbreak in London around a particular pump.