Maths Inspiration

David Spiegelhalter

Statistical Laboratory and

MRC Biostatistics Unit, Cambridge

Fields Institute, May 27th 2009

With thanks to Mike Pearson, Ian Short , Yin-Lam Ng, Hauke Riesch, Owen Smith, Arciris Garay, etc etc

What is the Winton programme trying to do?

Improve the public handling of quantitative aspects of risk and uncertainty, through

- Educational lectures, workshops
- The 'Risk Roadshow'
- Website
- Engagement with media
- Working with people who want to communicate risk
- Inter-disciplinary research

www.understandingunceratinty.org

Search

Understanding Uncertainty

Home | Blog | For Educators | Fun Stuff | About Us

Contact us Log out

Home :: Featured Content

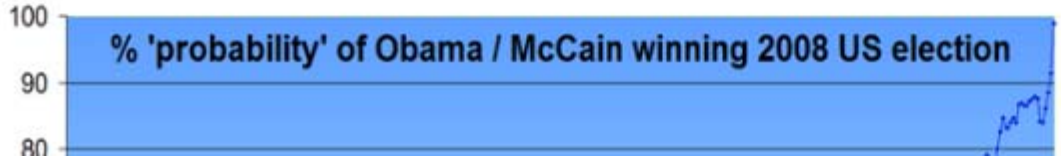
What was the probability that Barack Obama would win the US election?

View | Edit | Revisions | Workflow | Clone

Posted December 1st, 2008 by David Spiegelhalter in level 1, probability betting

On the face of it this seems an odd question. After all, he won. But before the election it was uncertain whether Obama would win, and probability is the way that uncertainty is quantified, so maybe it is reasonable to ask what that probability was.

We know that there were betting odds – a betting exchange such as Intrade allows people both to accept or make bets and so converges, at any point in time, to a certain set of odds at which people are willing to be either the better or the bookmaker. This prediction market provides a 'probability' on Obama winning that kept changing for the year before the election – this is shown in Figure 1 with some of the main events of the year marked in.



Featured Content

- One game to play!
- 2845 ways to spin the Risk
- A predictable pattern of murder?
- Nightingale's 'Coxcombs'
- What was the probability that Barack Obama would win the US election?
  - Laplace's law of succession
- Coincidences
- National Lottery
- Premier League
- What is Probability?
- Risk in the media
- How long are you going to live?

Pending Content

- Bayes' Theorem

“6/49” lottery


LOTTO 649
27-MAY/MAI-2009
$2.00
QUICK PICK/MISE-ÉCLAIR
10 12 17 25 30 42
ENCORE ENCORE ENCORE
NOT ENTERED/ NON-INSCRIT
Ticket No. Billet
11153-1954-2022-0618-56406
008901
400100
SIGN/SIGNEZ
SIGNATURE
WEDNESDAY´S LOTTO 6/49
OLG


City CONVENIENCE

Statistics : Lotto 6/49

Numeric Order

Appearance of Winning Numbers by Frequency Order

Past six months: **52 draws**
From 26/11/2008 to 23/05/2009

Numbers	Total	Bonus
34	16	2
11	14	3
48	14	2
20	13	1
39	12	1
4	11	2
8	11	1
37	11	0

Since the start: **2,644 draws**
From 12/06/1982 to 23/05/2009

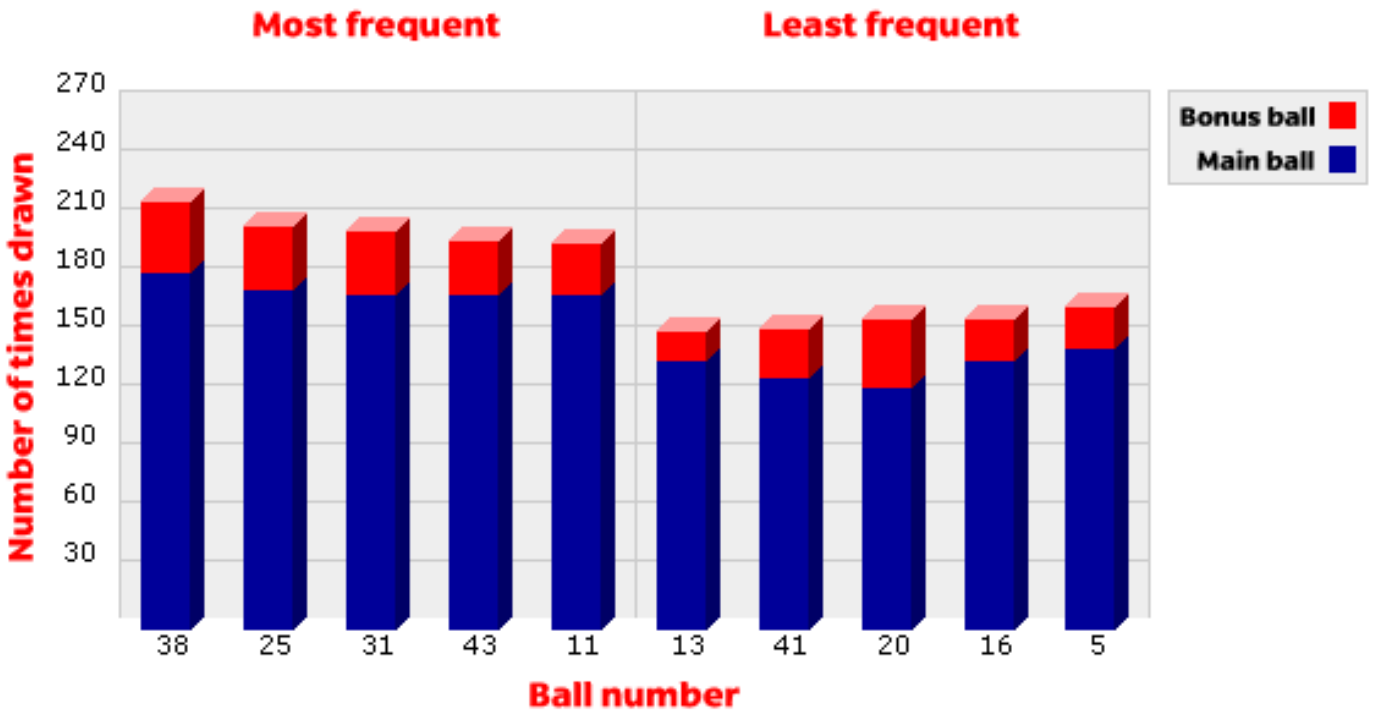
Numbers	Total	Bonus
34	426	61
31	421	55
43	419	58
47	415	59
45	408	55
46	403	55
40	400	46
27	399	57

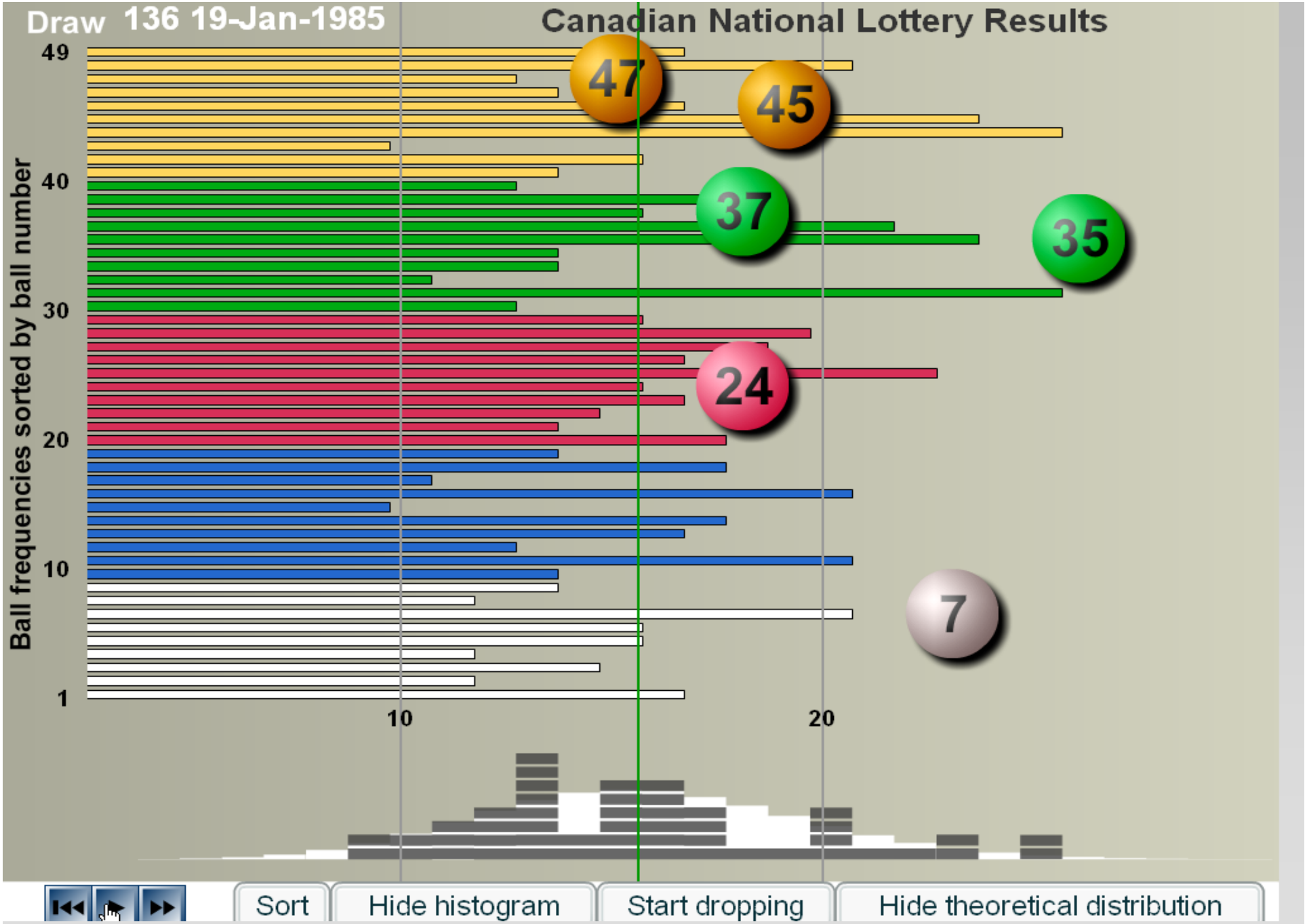
Find out which numbers have been drawn the most frequently, and which have been drawn the least. Despite the draws being totally random, some numbers have a habit of cropping up more than others, while others hardly appear at all! Please note, these results include the Lotto Bonus Draws held on 18th May 2002, 1st June 2002, 6th November 2004, and the £5 million jackpot-only draw held on 29th April 2006.

All numbers | Individual number

Game: Lotto | Time period: To Date | GO

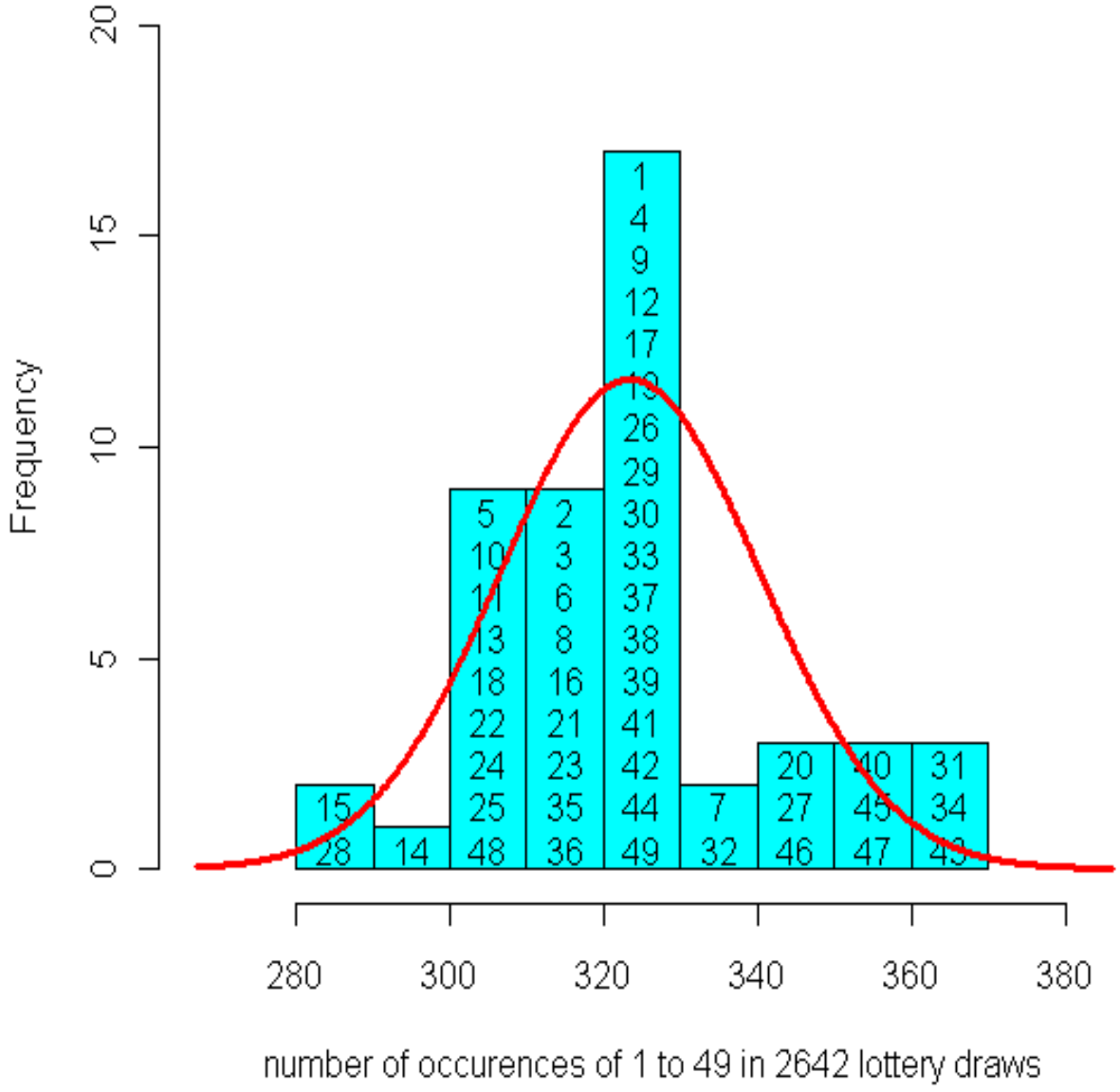
lotto Five most and least frequently drawn balls



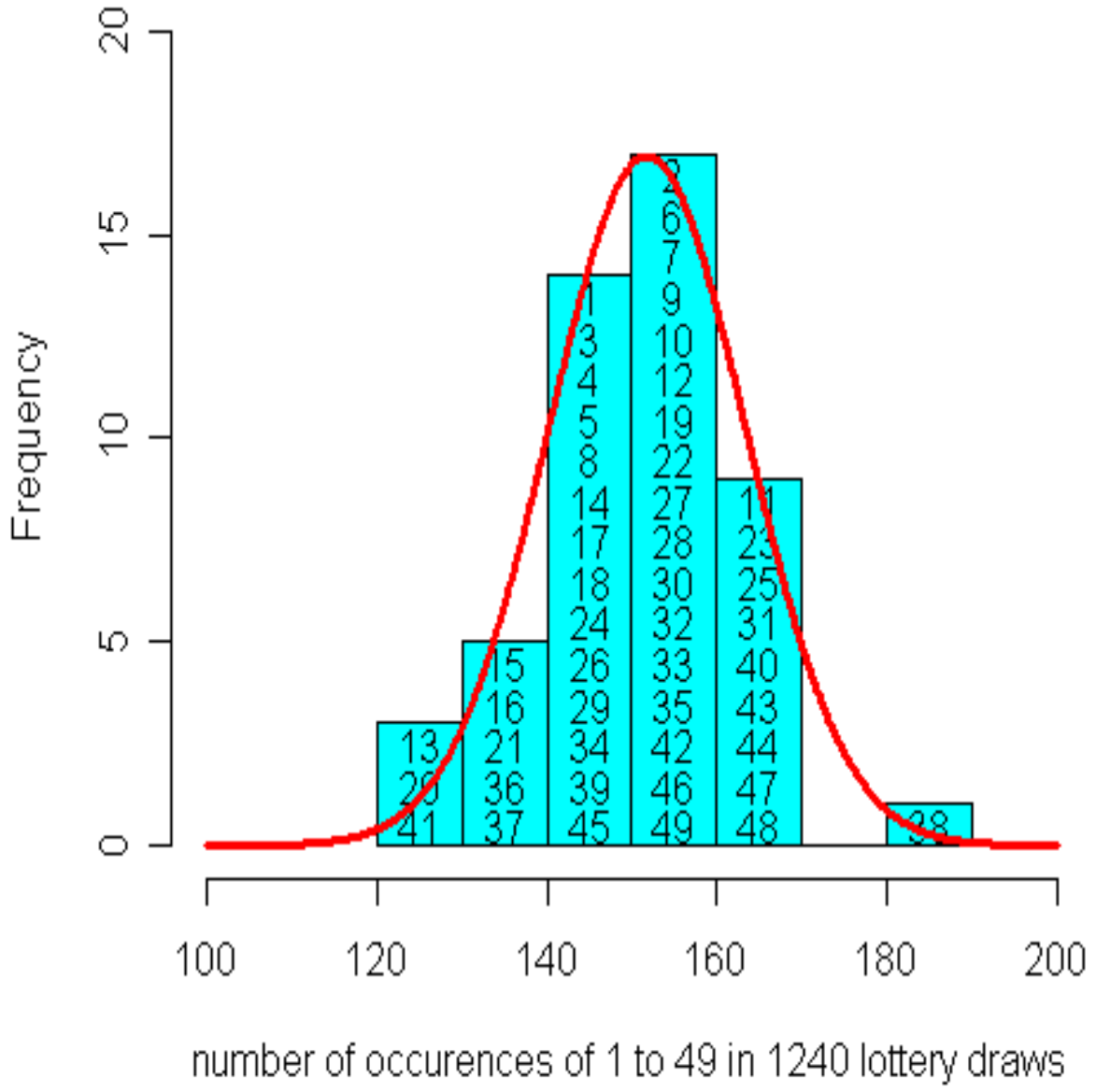


Lottery animation: www.understandinguncertainty.org/node/39

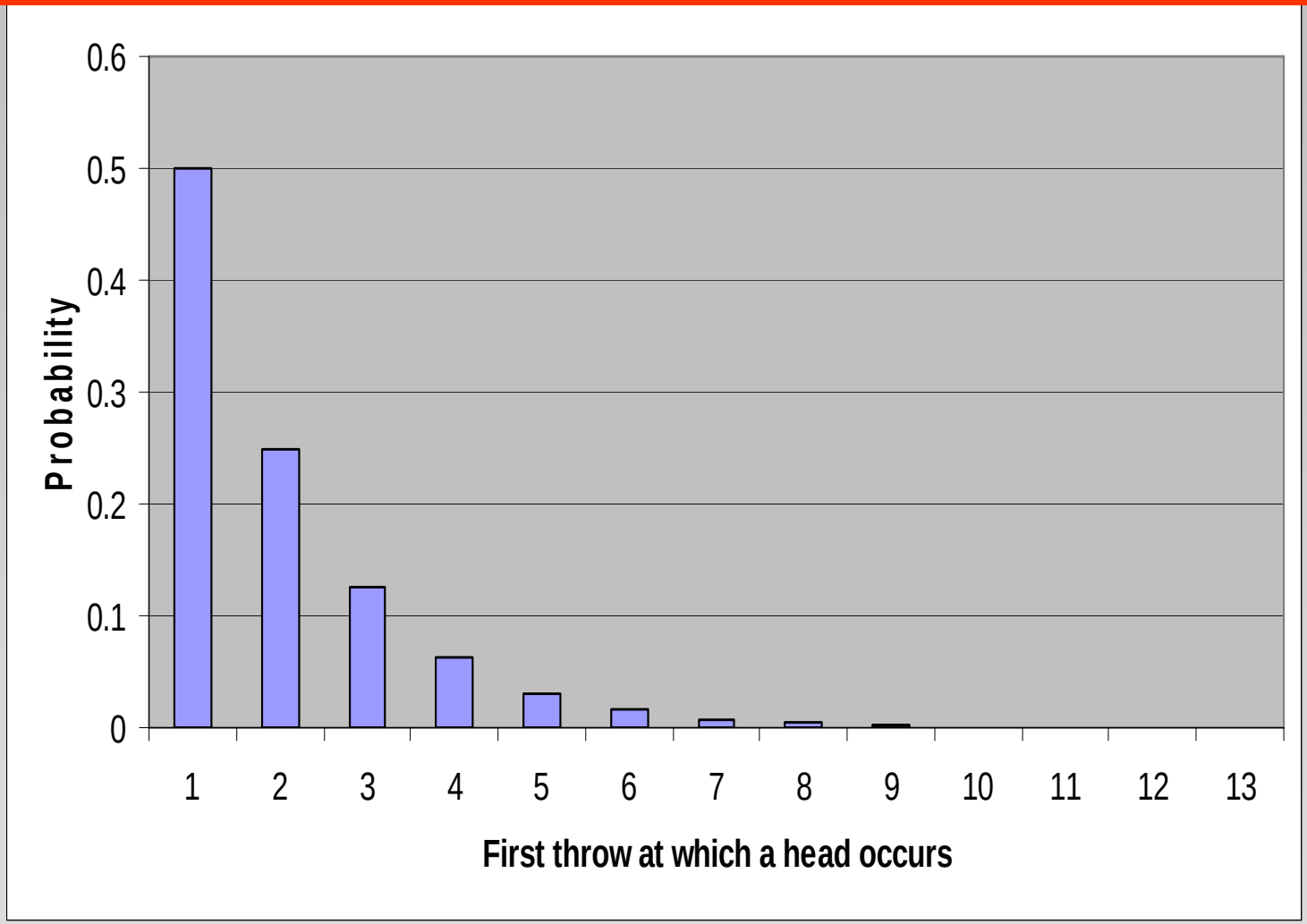
Counts in Canada 649, 1982-2009



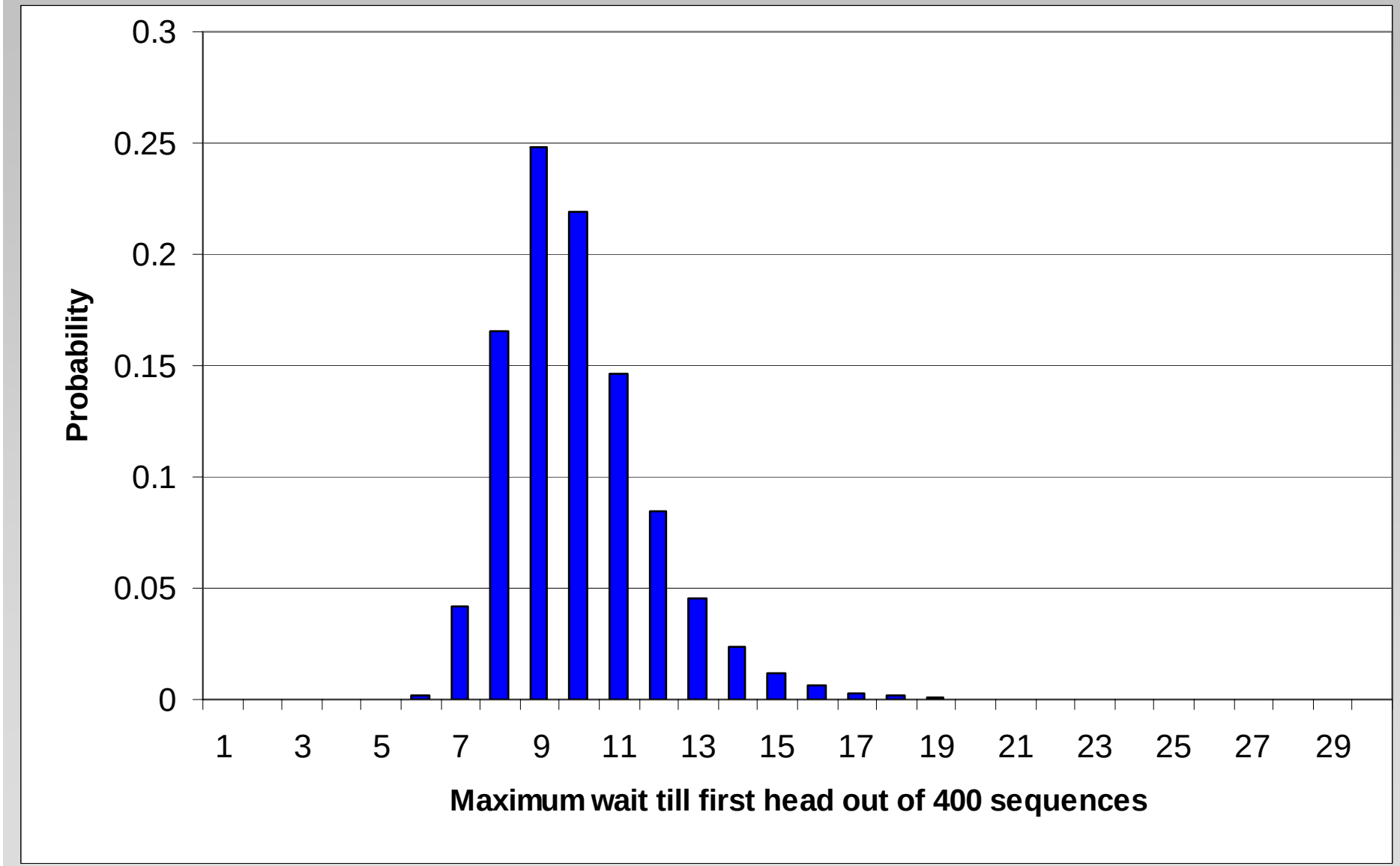
Counts (in the UK!) obey the rules of probability



A *geometric* distribution - probability of first head on n^{th} toss



(This is also the probability of having to wait *more* than n tosses until the first head)


0.3
0.25
0.2
0.15
0.1
0.05
0
Probability
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
Maximum wait till first head out of 400 sequences

- The chance of throwing a head for the first time on throw *x*

$$= P(X = x) = 1/2^x \quad (1/2 , 1/4 , 1/8 \ldots$$

- This is also the chance of waiting *longer* than x for the first head

$$= P(X > x) = 1/2^x \quad (1/2 , 1/4 , 1/8 \ldots$$

- Suppose *n* people flip coin until first head,

P(maximum wait $\leq x$) = P(*all* waits $\leq x$)

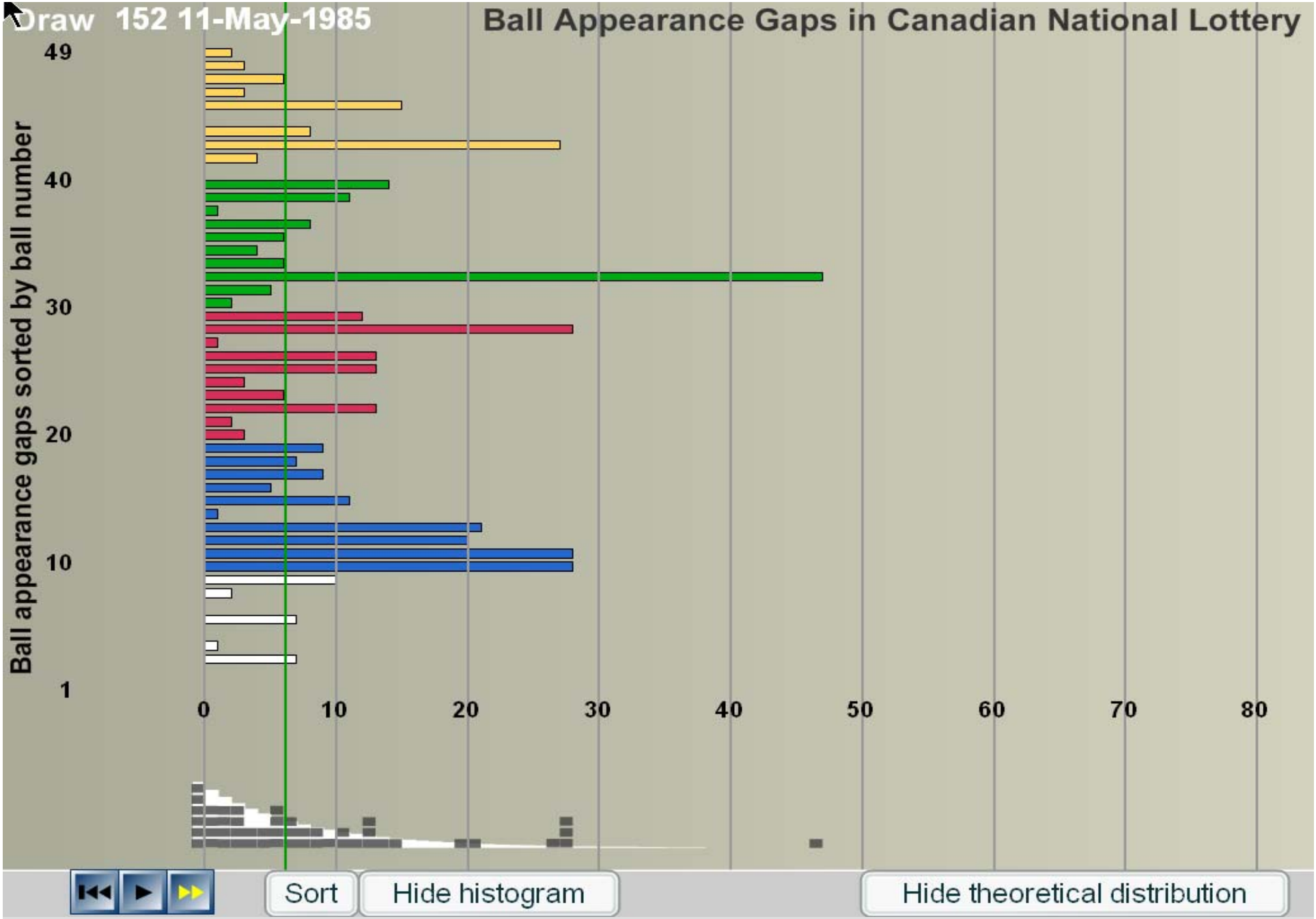
$$= P(X \leq x)^n$$

$$= [1 - P(X > x)]^n$$

$$= [1 - 1/2^x]^n$$

- *So* P(max-wait = x) $= P$(max-wait $\leq x$) - P(max-wait $\leq x\text{-}1$)

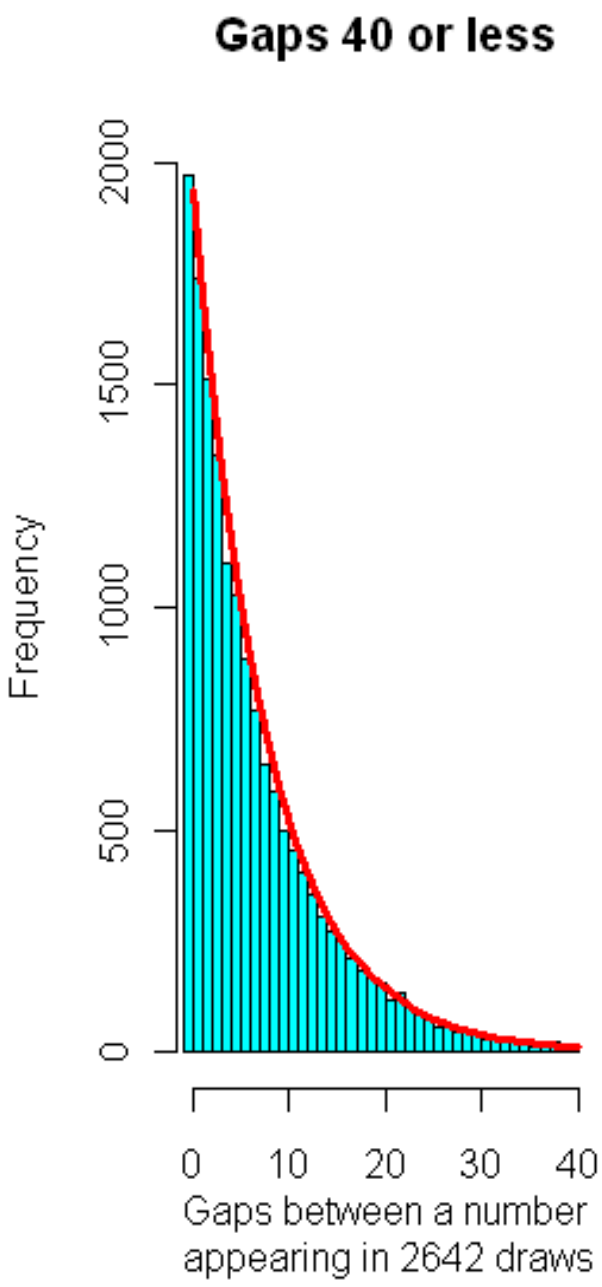
$$= [1 - 1/2^x]^n - [1 - 1/2^{x-1}]^n$$

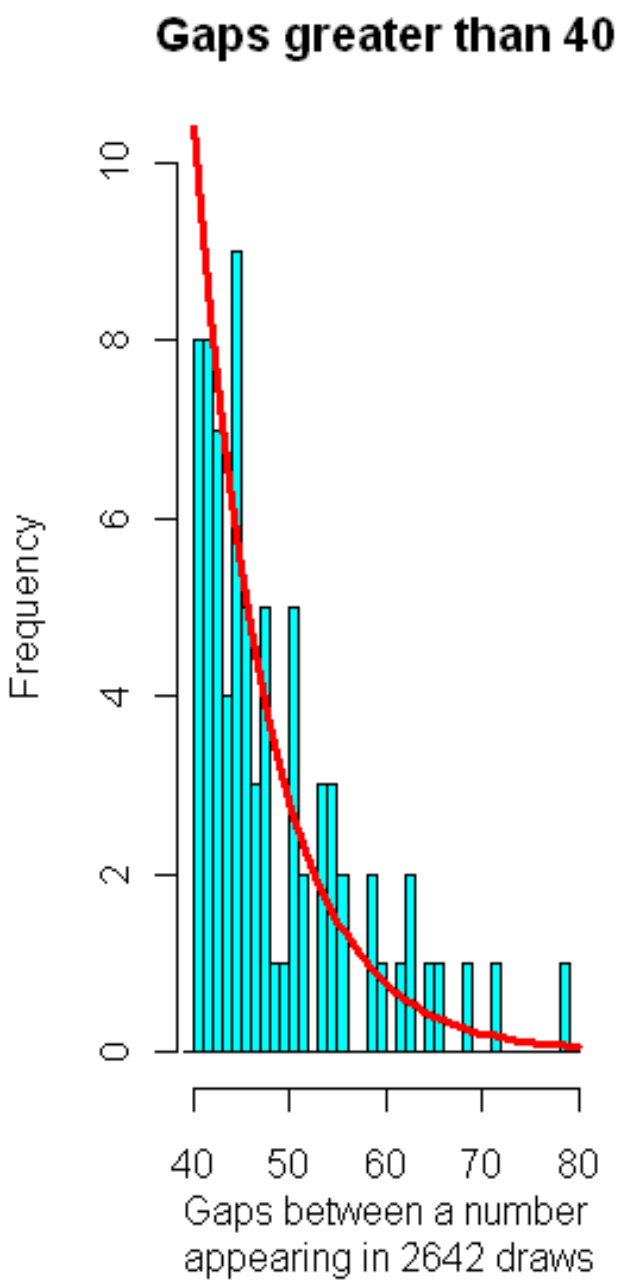


Lottery animation: www.understandinguncertainty.org/node/39

As expected, a geometric distribution of gaps

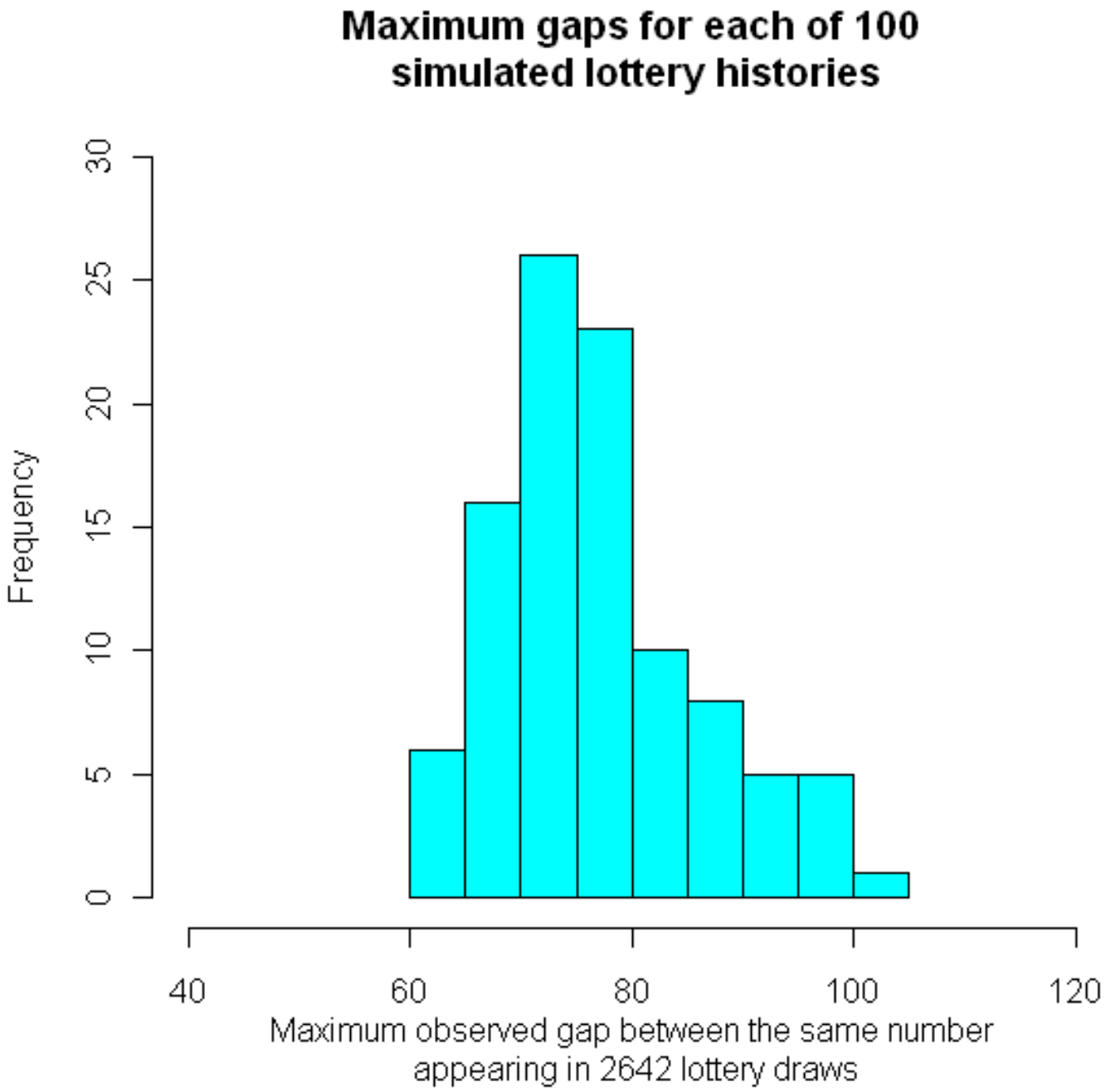
But is a maximum gap length of 79 surprising?





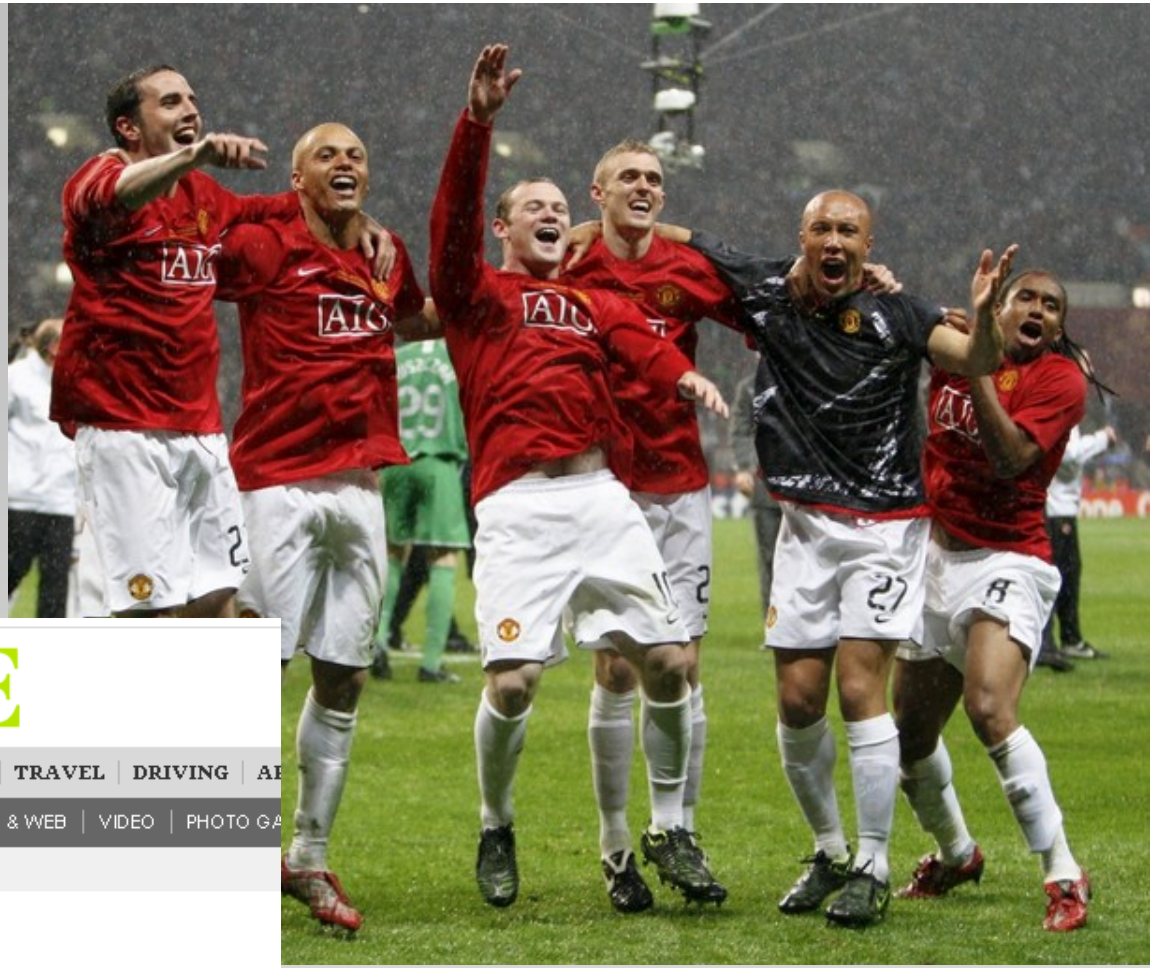
Simulate 100 full lottery histories

79 is almost exactly the expected maximum gap



- How much of the English Premier football league is due to chance?

- 22%



TIMESONLINE

NEWS | COMMENT | BUSINESS | MONEY | SPORT | LIFE & STYLE | TRAVEL | DRIVING | AR

UK NEWS | WORLD NEWS | POLITICS | ENVIRONMENT | WEATHER | TECH & WEB | VIDEO | PHOTO GA

Where am I? › Home › Sport › Football › Premier League › Newcastle

From The Times

May 23, 2009

Alan Shearer believes Newcastle will stay up

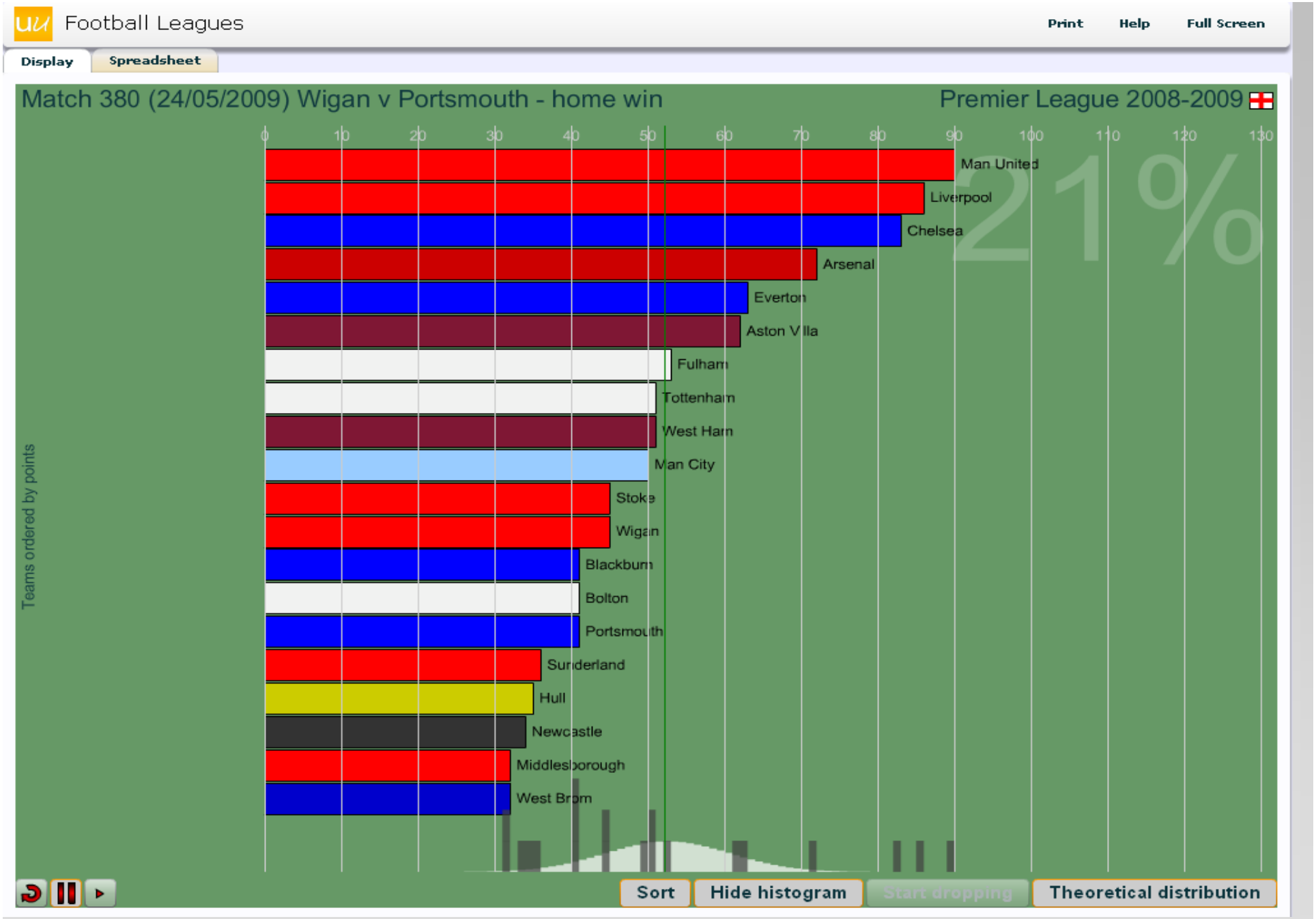
Manager expresses confidence that his club will escape relegation but warns of the implications of failure on Sunday

George Caulkin

Alan Shearer employed words such as "disaster" and "devastation" yesterday to describe the prospect of relegation, but while he spoke about the possibility of job losses at St James' Park, his is increasingly unlikely to be

CLUB DETAILS

Ground:	St James' Park
Capacity:	52,387



Football animation: www.understandinguncertainty.org/node/228

Team	Points	Goals for	'Attack strength'	Goals against	'Defence weakness'
Man United	87	67	1.46	24	0.52
Liverpool	83	74	1.61	26	0.57
Chelsea	80	65	1.41	22	0.48
Arsenal	69	64	1.39	36	0.78
Everton	60	53	1.15	37	0.80
Aston Villa	59	53	1.15	48	1.04
Fulham	53	39	0.85	32	0.70
Tottenham	51	44	0.96	42	0.91
West Ham	48	40	0.87	44	0.96
Man City	47	57	1.24	50	1.09
Stoke	45	37	0.80	51	1.11
Wigan	42	33	0.72	45	0.98
Bolton	41	41	0.89	52	1.13
Portsmouth	41	38	0.83	56	1.22
Blackburn	40	40	0.87	60	1.30
Sunderland	36	32	0.70	51	1.11
Hull	35	39	0.85	63	1.37
Newcastle	34	40	0.87	58	1.26
Middlesbrough	32	27	0.59	55	1.20
West Brom	31	36	0.78	67	1.46

Predicting results using simple independent Poisson model

Hull City vs Manchester United: expected goals

Hull: = home-average x attack strength x defence weakness of opposition

=1.06 x 0.85 x 0.52 = 0.60

Man U: = 1.36 x 1.46 x 1.37 = 2.12

Team	Expected goals	0	1	2	3	4	5
Hull City	0.60	55	33	10	2	0	0
Man U	2.12	12	25	27	19	10	4

Assume independent Poisson distributions to give probability of any result

Add to give win/draw/lose probabilities

Home	Away	Most likely	2nd most likely	3rd most likely	4th most likely
Arsenal	Stoke	2-0 (14%)	1-0 (13%)	2-1 (9%)	3-0 (9%)
Aston Villa	Newcastle	1-0 (10%)	2-0 (10%)	2-1 (10%)	1-1 (10%)
Blackburn	West Brom	1-1 (10%)	2-0 (10%)	2-1 (10%)	1-1 (10%)
Fulham	Everton	0-0 (19%)	1-0 (16%)	0-1 (14%)	1-1 (13%)
Hull	Man United	0-2 (14%)	0-1 (14%)	1-2 (9%)	1-1 (8%)
Liverpool	Tottenham	1-0 (16%)	2-0 (15%)	3-0 (10%)	2-1 (9%)
Man City	Bolton	2-1 (10%)	1-1 (10%)	1-0 (10%)	2-0 (10%)
Sunderland	Chelsea	0-1 (20%)	0-2 (15%)	0-0 (13%)	1-2 (8%)
West Ham	Middlesbrough	1-0 (19%)	0-0 (14%)	2-0 (13%)	1-1 (11%)
Wigan	Portsmouth	1-0 (17%)	2-0 (14%)	0-0 (11%)	1-1 (10%)

- Actual model used is Bivariate Poisson, allowing correlations (R function lm.bp)
- Found to best fit European league results
- Many more sophisticated models used

Page last updated at 14:11 GMT, Friday, 22 May 2009 15:11 UK

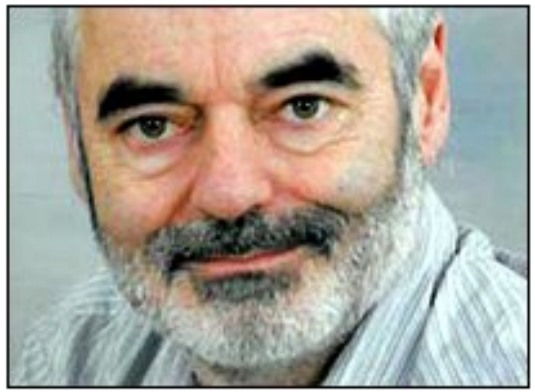
E-mail this to a friend Printable version

The professor's Premiership probabilities

By Professor David Spiegelhalter
Professor for the Understanding of Risk, Cambridge University

Why should anyone take any notice of what a Cambridge professor of statistics, who knows little about football and does not even support
bout this weekend's

in the increasingly
athematical models
sed by sports betting
t odds and identify
bets.


Professor David Spiegelhalter analyses the football table

So I would not recommend anyone using these odds for betting.

You have been warned.

Understanding Uncertainty: Animated Premier League Statistics

PREMIER LEAGUE PROBABILITIES

Read how the professor did

ARSENAL V STOKE

Home win: 72%

Draw: 19%

Away win: 10%

Verdict: 2-0 (14%)

ASTON VILLA V NEWCASTLE

Home win: 62%

Draw: 21%

Away win: 17%

Verdict: 1-0 (10%)

Coincidences

Coincidences – three children born on the same day?

- The MacKriell family in Gloucester: Robin 14, Rebecca 12, Ruby 0, all born on January 29th
- 1/365 x 1/365 chance (assuming uniform birth-dates)
  = 7.5 in 1,000,000
- But there are 1,000,000 families in the UK with 3 children
- So where are the other examples?

Coincidences –

- Joyce and Ron Pulsford of Pagham near Bognor Regis were both 80 on 08.08.08
- Are they unique in the country?
- DJS: "*As a rough assessment, I would say the odds on there being another couple are between 10 to 1 and 100 to 1 against."*
- *???*

Coincidences

Birthdays:

- 23 people: 51% chance that 2 share a birthday
- 35 people: 81% chance that 2 share a birthday
- 80 people: 99.99% chance that 2 share a birthday

Why does this happen?

- Imagine 35 people in a line
- First birthday can be anything
- 2nd birthday must be different from first: probability 364/365 = 0.997
- 3rd birthday must be different from 1st and 2nd: probability 363/365 = 0.995
- …..
- Probability that all 35 are different =

  0.997 x 0.995 x …. 0.907 = 0.19

How to amaze people!

- Get 20 of your friends together
- Ask them each to choose a number between 1 and 100
- Agree that if their numbers are *all different*, you will give them each a prize
- If there are any *matches*, they will each give you a prize
- Carry on playing until you are laden with prizes (and have no friends)

20 people choosing numbers between 1 and 100

- First number can be anything
- 2nd number must be different from first: probability 99/100 = 0.99
- 3rd number must be different from 1st and 2nd: probability 98/100 = 0.98
- …..
- Probability that all 20 are different =

  0.99 x 0.98 x …. 0.81 = 0.13

A neat trick

- Assume N people each choose a number between 1 and T
- Set $T = (N/2)^2$

- eg $N = 20$; $T = 100$
  $N = 40$; $T = 400$
  $N = 400$; $T = 40000$
- Then the probability that all choose different numbers ≈ 0.13

Coincidences –

- What's the chance *p* of the *specific* event?
- How many opportunities *N* are there for a 'similar' event to occur?
- Multiply to give expected number $E = Np$

Expected number of events	Chance no events occur	Chance at least one event occurs
1/2	61%	39%
1	37%	63%
2	13%	87%
3	5%	95%
4	2%	98%
5	1%	99%

What is the chance of winning?

- Imagine that the numbers on your lottery ticket were labelled as WIN
- Chance of picking first WIN ball = 6/49
- Chance of picking second WIN = 5/48
- Chance of picking all WIN balls =

6/49 x 5/48 x 4/47 x 3/46 x 2/45 x 1/44

= 1/13,983,816 !!

So why does anyone win the lottery?

- Each ticket has around 1/14,000,000 chance of winning
- They sell around 30,000,000 tickets
- So the expected number of Jackpot winners is around 2
- So the chance that nobody wins (a rollover) is around 0.13

Pick the same number?

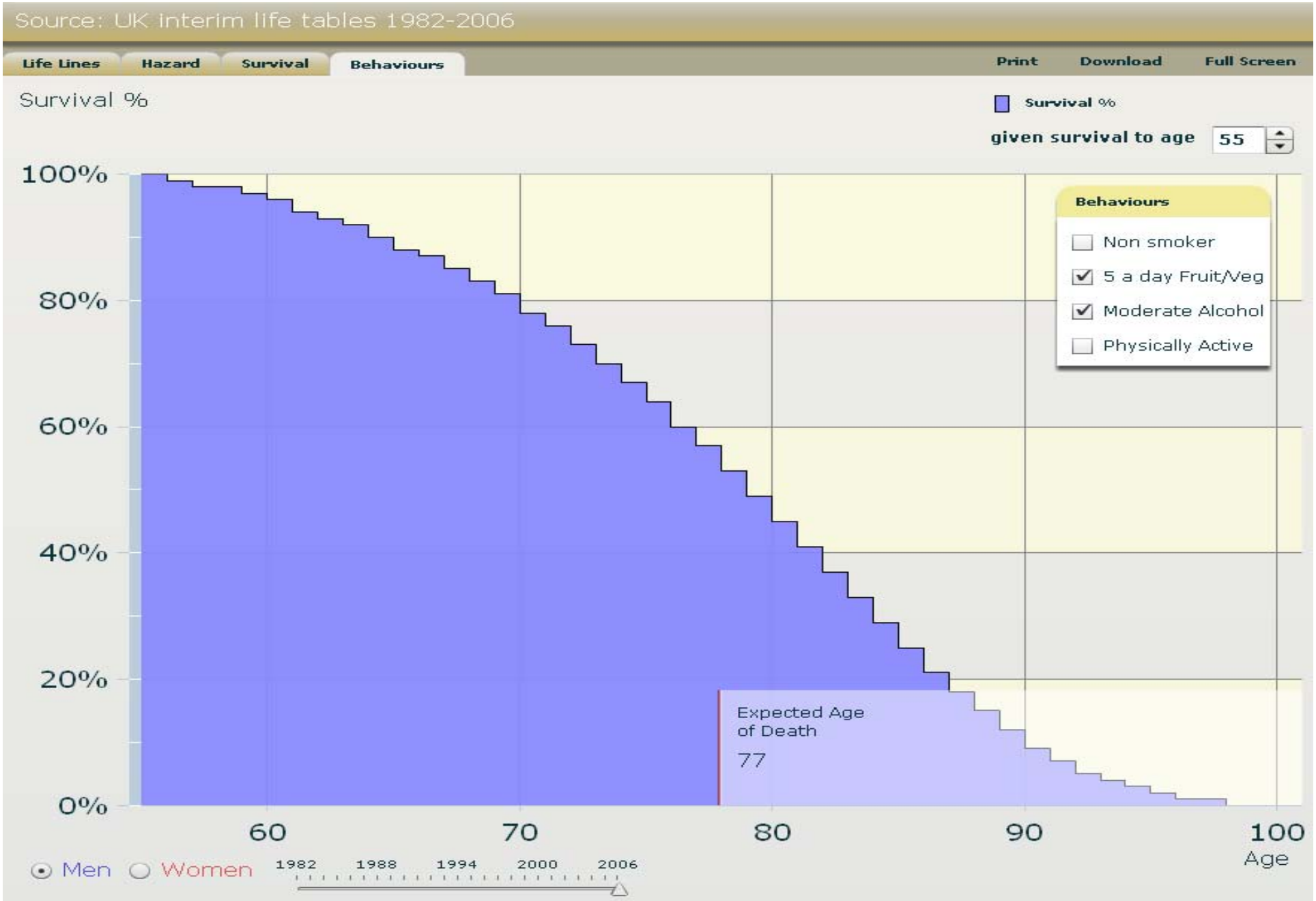
- Assume N people each choose a whole numbers between 1 and T
- Each pair has a $1/T$ chance of matching
- $N(N\text{-}1)/2 \approx N^2/2$ pairs of people
- So $E \approx N^2/(2T)$
- Prob "no match" $\approx \exp(-N^2/(2T))$
- So if $T = (N/2)^2$, then

  Prob "match" $\approx 1 - \exp(-2) \approx 0.87$

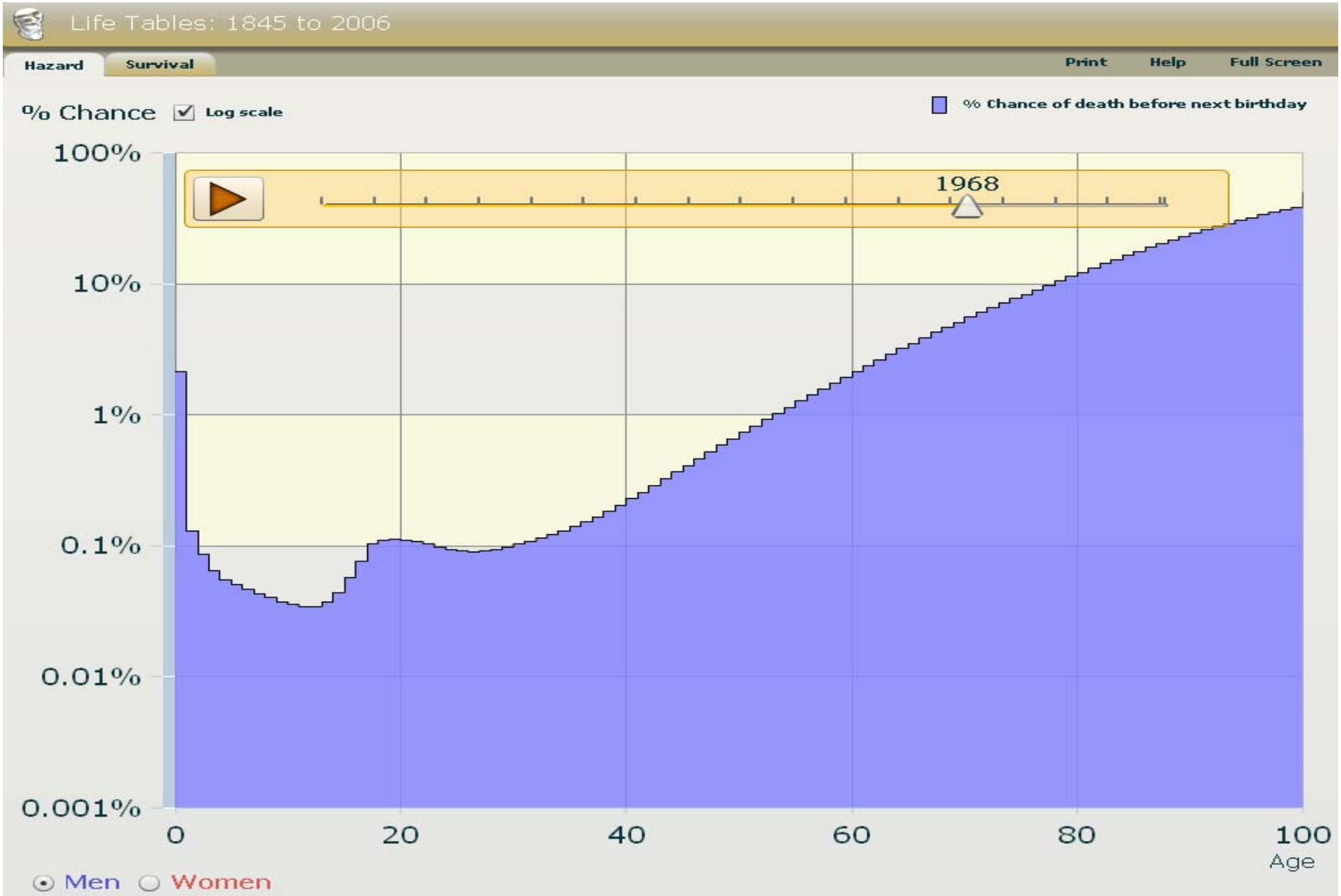
Choose how many people

- 50
- 100
- 200
- 300
- 400

How long are you going to live?



Survival: www.understandinguncertainty.org/node/76



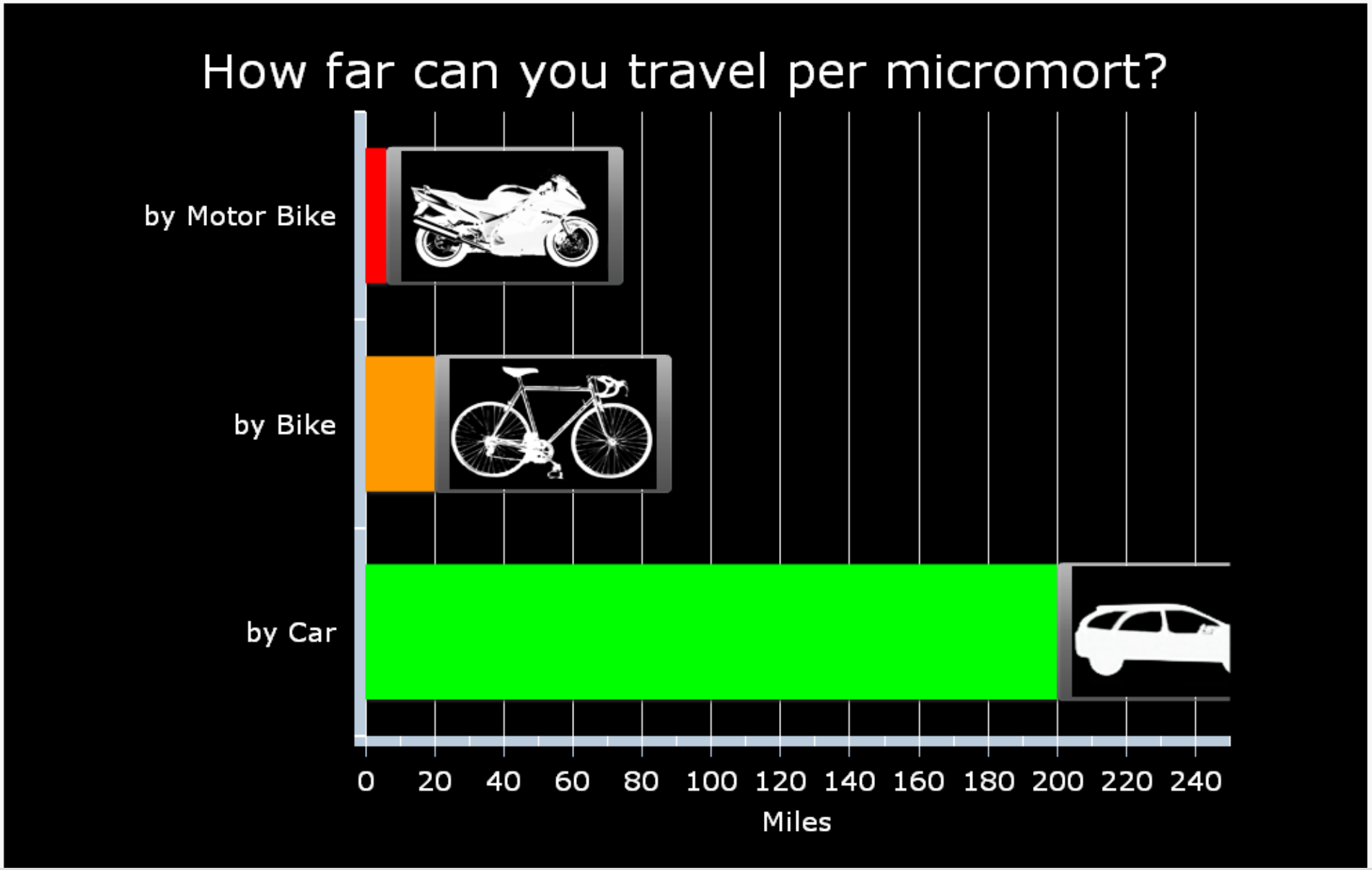
Past survival: www.understandinguncertainty.org/node/210

What are the risks?

- Need a friendly unit of deadly risk
- A ***Micromort*** is a 1-in-a-million chance of dying
- Each day 50 people are killed in England and Wales (about 50 million)
- So just living means that we experience a micromort every day (on average)

Risk Calculator

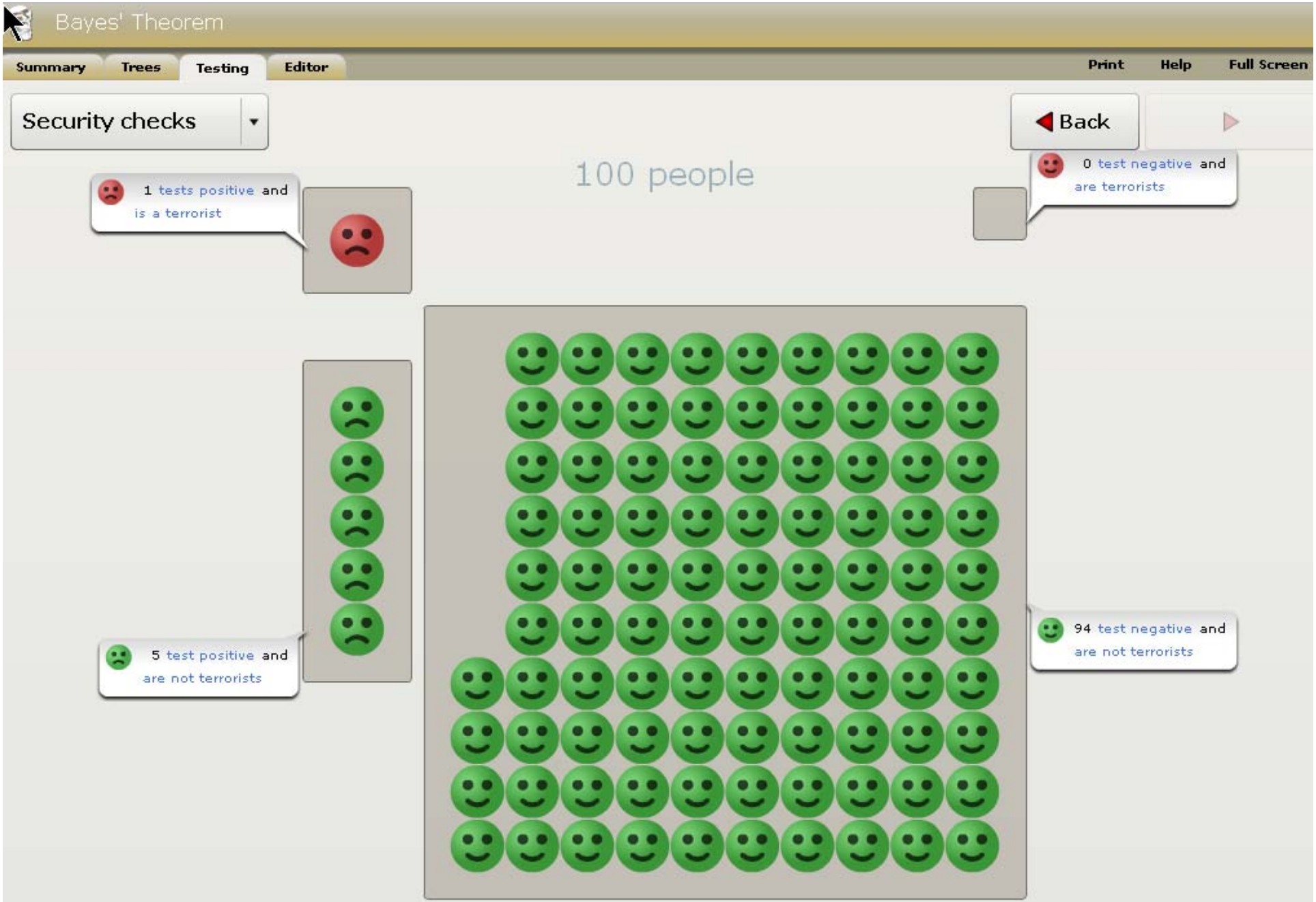
Transport | Embed | Print | Help | Full Screen



Micromort animation:

Lie detectors

- A terrorist hides in a room with 99 innocent people
- You have a lie detector that is 95% accurate
- You get people out one at a time and ask them if they are a terrorist
- They all say no
- Eventually the machine goes ‘ping!’
- What is the chance that you have caught the terrorist?
- (a) 95% (b) 84% (c) 50% (d) 16% (e) 5%



Bayes theorem: www.understandinguncertainty.org/node/238