



International Electronic Journal of
Mathematics Education

Volume 4, Number 3, October 2009

www.iejme.com

Special issue on “Research and Developments in Probability Education”

Manfred Borovcnik & Ramesh Kapadia (Eds)

ANNEX: ALL APPENDICES TO
PARALLEL DISCUSSION OF CLASSICAL AND BAYESIAN WAYS
AS AN INTRODUCTION TO STATISTICAL INFERENCE

Ödön Vancsó

To article 

Glossary

[Glossary of terms](#)

Animation

[Animation](#) on the intransitivity of the favourable relation

Appendices – also contained in the body of the paper

[A. Some probability paradoxa](#)

[B. Examples of students’ work – in EXCEL](#)

Demonstration in Excel. To find the maximal number of balls in a lottery

[To the EXCEL sheet](#) 

[“Frozen” results](#) of the EXCEL demonstration

Classical methods

[Classical estimation](#) of the number of balls

[Classical confidence interval](#) for the number of balls

Bayesian methods

[Updating of prior distribution](#) on the maximal number by the result of one week

[Updating of prior distribution – Random draw – only in the EXCEL file](#)

[Cumulative – week by week – updating](#) of uniform prior

Data – draws of two Hungarian lotteries

[Six-numbers Lottery](#)

[Five-numbers Lottery](#)

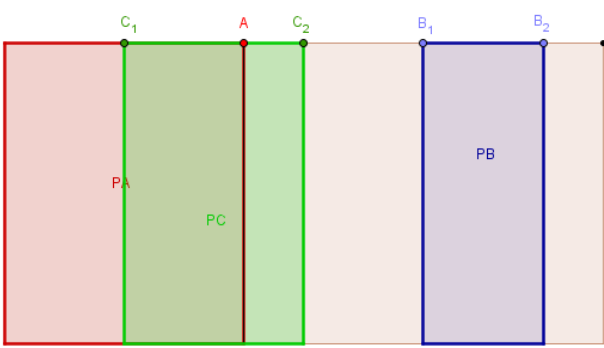
ANIMATION:

Animation on the intransitivity of the favourable relation

To article 

Top 

You can use the points to change the probabilities of the events A, B and C.
 You can see which relation is true for the events pair A, B; or B, C; or C and A.



$P(A)=0.4$
 $P(B)=0.2$
 $P(C)=0.3$

$P(BA)=0$	$P(B A)=0$	$A \searrow B \Leftrightarrow P(B A) < P(B)$
$P(CB)=0$	$P(C B)=0$	$B \searrow C \Leftrightarrow P(C B) < P(C)$
$P(CA)=0.2$	$P(C A)=0.5$	$A \nearrow C \Leftrightarrow P(C A) > P(C)$

The favourable relation is *not* transitive
 $A \uparrow B$ and $B \uparrow C$ does not imply $A \uparrow C$

$A \uparrow B$: A favours B , i. e. $P(B|A) > P(B)$.
 $A \hat{\uparrow} B$: A disfavors B , i. e. $P(B|A) < P(B)$
 $A \perp B$: A is neutral to B , i. e. $P(B|A) = P(B)$

$A \uparrow B$: A increases conditional probability of B

This seems paradox as one would guess that
 If A increases the (conditional) probability of B
 and B increases the probability of C – then
 A increases the conditional probability of C ??

We demonstrate that such a theorem cannot be true.
 Drag the points for the 3 events to see the effect
 on the favourable relation between them.

The favourable relation is *not* transitive

$A \uparrow B$ and $B \uparrow C$ does not imply $A \uparrow C$

$A \uparrow B$: A favours B , i. e. $P(B|A) > P(B)$.

$A \hat{\uparrow} B$: A disfavors B , i. e. $P(B|A) < P(B)$

$A \perp B$: A is neutral to B , i. e. $P(B|A) = P(B)$

$A \uparrow B$: A increases conditional probability of B .

This seems paradox as one would guess that

If A increases the (conditional) probability of B
 and B increases the probability of C – then
 A increases the conditional probability of C ??

We demonstrate that such a theorem cannot be true.

Drag the points for the 3 events to see the effect
 on the favourable relation between them.

APPENDIX A: SOME PROBABILITY PARADOXA

[To article](#) 

[Top](#) 

"The prisoner's dilemma"

This problem was originally formulated by Gardner (1959), and later taken up by several authors, e. g. Mosteller (1987). It was fiercely and widely discussed in the literature. For details, see the following [description](#) according to Weisstein (n. d.)

“In this problem, three prisoners A , B , and C with apparently equally good records have applied for parole, and the parole board has decided to release two of them, but not all three. A warden knows, which two are to be released, and one of the prisoners (A) asks the warden for the name of the one prisoner other than himself who is to be released. While his chances of being released before asking are $2/3$, he thinks his chances after asking and being told " B will be released" are reduced to $1/2$, since now either A and B or B and C are to be released. However, he is mistaken since his chances remain $2/3$.”

"Monty Hall problem"

This problem became famous as it was part of a popular TV show; the discussion to follow the solution by vos Savant (1990) was signified by remarkable wrong conceptions of the presented situation and a complete misunderstanding of the presented correct solution. It is amazing that a problem as simple as this one was so much disputed among mathematicians and statisticians, not to speak of other well-educated people or novices to the subject. More details may be found from [Monty Hall](#) (Wikipedia n. d.). The following [description](#) is according to Weisstein (n. d.):

“This problem is named for its similarity to the *Let's Make a Deal* television game show hosted by Monty Hall. The problem is stated as follows: Assume that a room is equipped with three doors. Behind two are goats, and behind the third is a shiny new car. You are asked to pick a door, and will win whatever is behind it. Let's say you pick door 1. Before the door is opened, however, someone who knows what is behind the doors (Monty Hall) opens one *of the other* two doors, revealing a goat, and asks you if you wish to change your selection to the third door (i.e., the door which neither you picked nor he opened). The Monty Hall problem is deciding what to do: change your choice or retain it.”

"Three discs problem"

Varga (1976) proposed a nice variant of the elder problem of Bertrand's drawers. Interesting details about the history, or the solution, may be found in [Bertrand's box paradox](#) (n. d.), [Darling, D.](#) (n. d.), or from [Everything2](#) (n. d.). The advantage of Varga's discs lies in the circumstance that it may easily be performed as an experiment in class. There are three discs marked as in Figure 1.

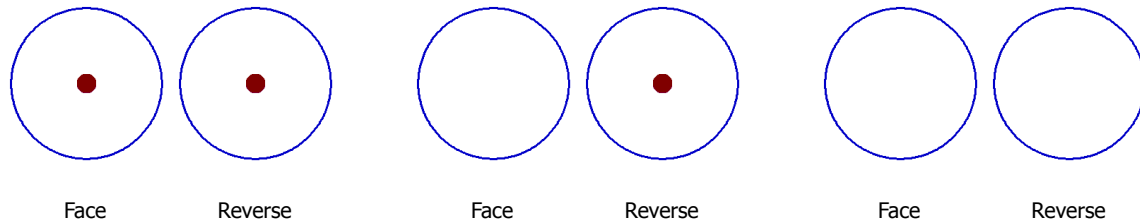


Figure 1. Varga's discs.

One of these discs is held up to the children; only one side is shown to them and they are asked to guess what is on the reverse 'spot or blank' (We used two different colours in our experiments). After a series of random guessing and getting the other side of the disc shown to see whether they had made the right guess, the children were asked to devise and write down a strategy for guessing, which they would apply each time subsequently.

For illustrative purpose, one class experiment with this game is reported:

A teacher played this game with 10–11 years old children. He summarised his observations briefly. "Some tried to repeat the last result in their prediction each time, others used blank and spot alternately for predicting the next result. None chose the best strategy: whatever is on the face is most likely to be also on the on the reverse side). He then let one child use this strategy and the results showed that he consistently scored best over a range of fifty trials. The children began to think and to suggest reasons as to why this might be. Their thinking was intuitively supported; no one came up with a numerical solution but their answers reflect that they had started to grasp some of the relevant ideas inherent in probability.

[To article](#) 

[Top](#) 

APPENDIX B: EXAMPLES OF STUDENTS' WORK – IN EXCEL

[To article](#) 

[Top](#) 

The following examples illustrate students' work. As is typical for the application of Bayesian methods, we had to use software; VisualBayes from Wickmann (2006), or EXCEL. In what follows, we present some graphs together with the problems and methods we used in the projects. Of course, for the paper, the layout has been enhanced.

The problem and various classical and Bayesian methods to deal with it

In a lottery n out of N , the number N of balls is assumed to be unknown. We draw $n = 6$ balls without replacement from the “urn”, the Lotto numbers; the numbers ordered are:

$$x_1 < x_2 < \dots < x_6$$

The problem is how to extract information on the unknown number N of balls from the numbers drawn?

- Classical methods for finding the maximal number of balls
 - Estimation of the number of balls
 - Confidence interval for the number of balls
- Bayesian methods for finding the maximal number of balls
 - Updating of prior distribution on the maximal number by the result of *one* week
 - Cumulative – week by week – updating of uniform prior

The reader will find more details in the [annex](#) or in the [EXCEL file](#) where it is also possible to simulate the results of the week and see the influence on the Bayesian result. Here, we will show only a few graphs to illustrate the difference in results between the classical and the Bayesian approach.

Classical estimation of the number N of balls

There are several estimators of the unknown number, all with different properties. We refer only to a few:

Median	Mean	MLE	Extreme gaps	Mean gap
$2 \cdot \text{med}(x_1, \dots, x_6)$	$2 \cdot \text{mean}(x_1, \dots, x_6) - 1$	$\max(x_1, \dots, x_6)$	$\max + \min - 1$	$\max + \frac{\max - 6}{6}$

Table 3. Various estimators for the number N of balls.

The data of the lottery since its start are analyzed to show the behaviour of these estimators. For classical estimators, two properties are most relevant: Whether the estimator is unbiased (or correctly centred), and whether it has a small variance, which means that in repetitions of the situation the new estimate would not differ too much from the first estimate.

From the graphs one may see the following. The median estimator has a great variance but is centred correctly; the extreme gaps estimator is better than the median estimator but with respect to the maximum likelihood estimator, it is worse. However, the MLE estimator on the other hand is not unbiased (it is only asymptotically unbiased, which means that the systematic error converges to 0 as the sample size increases to infinity).

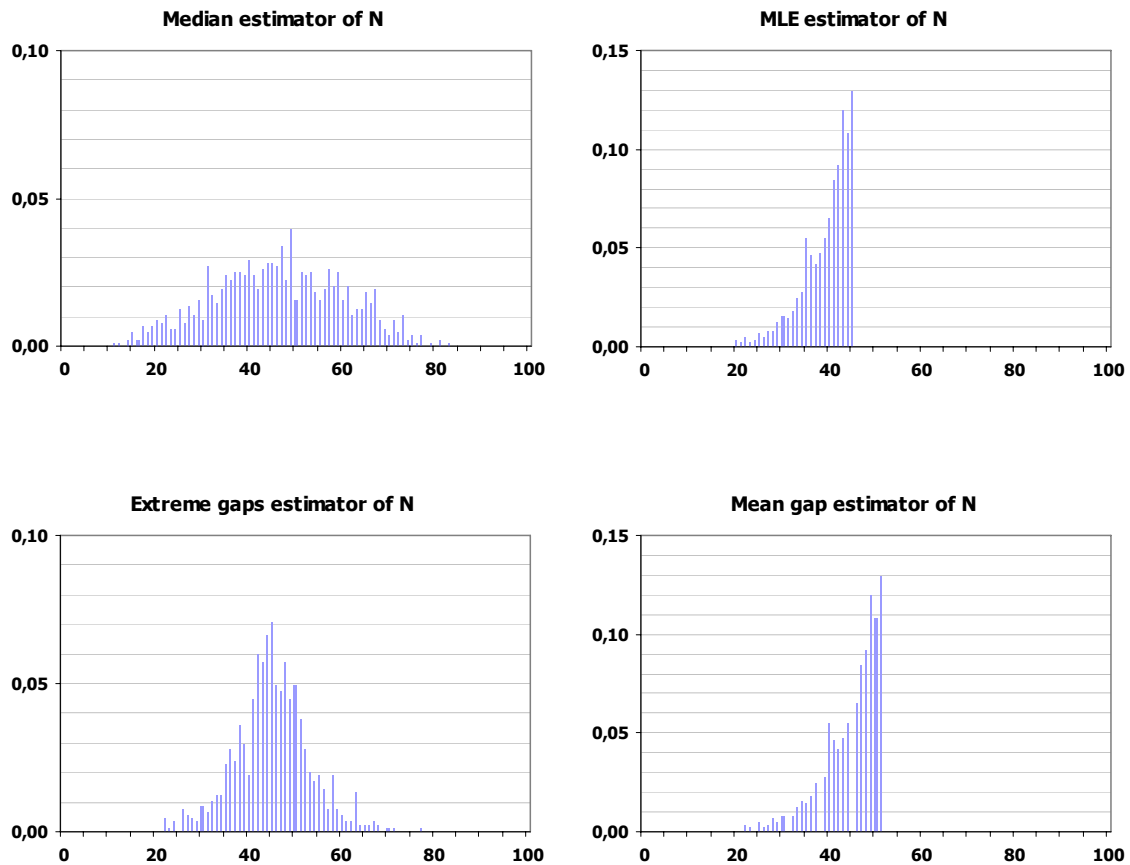


Figure 3. Repeated estimations of the unknown number of balls by various estimators.

Classical confidence intervals

This method yields intervals, which cover the unknown parameter (here the number N of balls) with a pre-assigned probability – supposed that it is applied in repeated cases under the *same* conditions. The graph shows the intervals for N week by week, calculated from the week's drawn lotto numbers. The global coverage rate is 95.7%. A disadvantage with classical confidence intervals is that it is not easy to integrate the data cumulatively from the past to give one summary confidence interval for N .

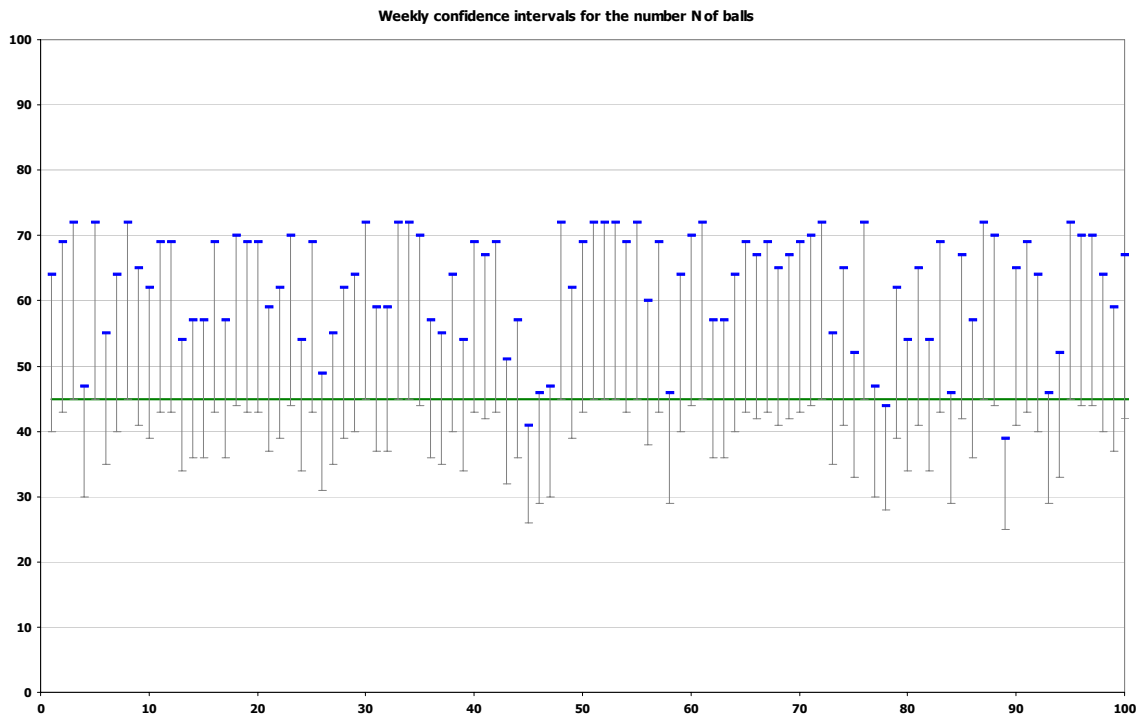


Figure 4. Classical confidence intervals on a weekly basis – the line in the middle represents the "true" value.

Bayesian methods for finding the maximal number of balls

For a Bayesian solution, it is necessary to model the prior knowledge by a distribution. Here, "complete" ignorance of this number N will be modelled by a uniform distribution on the interval $[31, 80]$. This prior is updated by the results of one week to a new posterior distribution on N reflecting the information of the data of one week. This new status of knowledge on the maximal numbers of balls is calculated and graphically presented.

Updating of uniform prior in dependence of week's maximum -detail

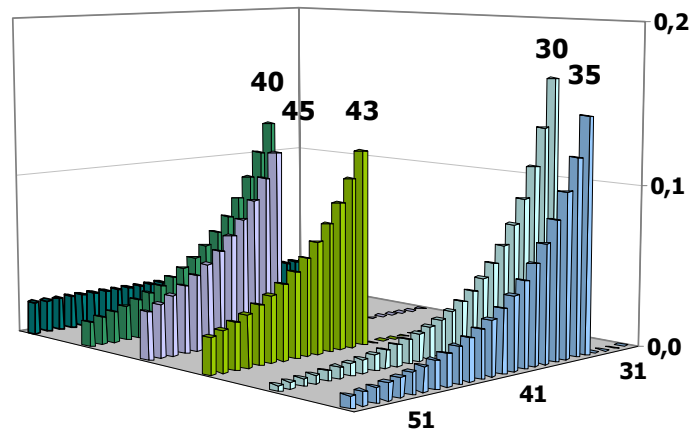


Figure 5a. The posterior distribution of the number balls is dependent on the week's results.

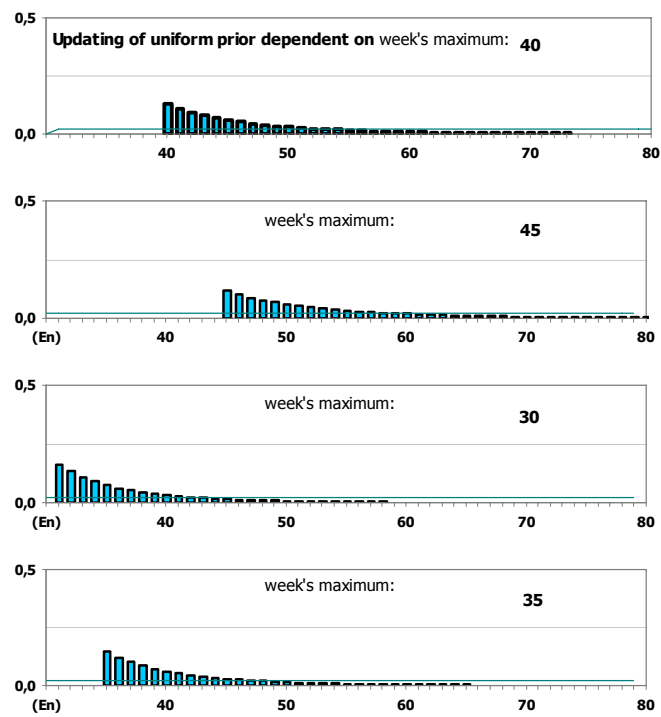


Figure 5a. Another representation of the posterior distribution of the number balls.

Within the Bayesian framework, the posterior distribution of the number of balls after one week may serve as prior distribution for the next week. Thus, a continuous process of updating may be applied. The following graphs show the resulting learning process: after 9 weeks, the initial ignorance on the interval $[31, 80]$ – modelled as uniform distribution – has been changed to a distribution, which is reduced to the values between 45 and 48, all other values have already a negligible probability at that stage. After 30 weeks, this posterior distribution expresses a status of knowledge that 45 has a high probability, while the slight “risk” still to have 46 or more balls is virtually negligible.

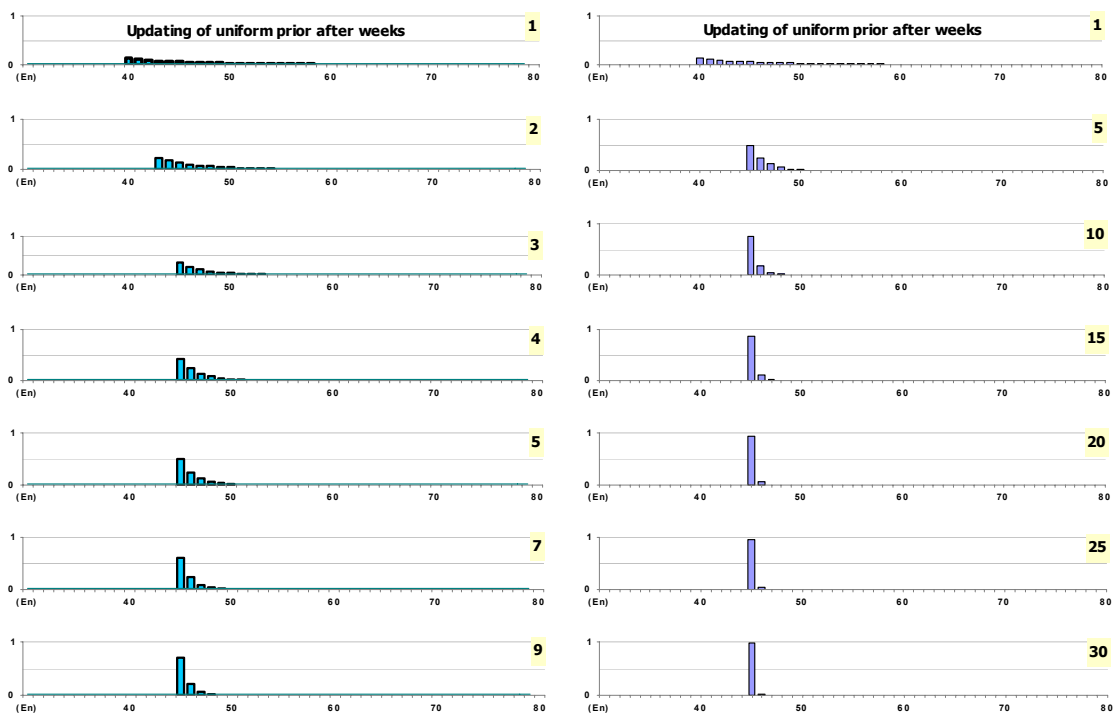


Fig. 6 a. After 9 weeks,
only 45–48 bear a non-negligible probability.

Figure 6b. After 30 weeks,
only 45 is relevant; there is a small probability for 46.

The result in Figure 6a and b clearly shows the convergence of the posterior probability distribution with time. In fact, the true number N equals 45; we deal here with the 6 out of 45 Hungarian lotto. The repeated updating accumulates all information from the past and yields a present status of information about the unknown parameter.

Ordered numbers of successive draws

Nr	x_1	x_2	x_3	x_4	x_5	x_6
1	6	15	20	24	38	40
2	9	14	23	32	40	43
3	1	7	14	43	44	45
4	6	9	11	14	18	30
5	12	13	19	30	34	45
6	16	17	19	20	34	35
7	4	23	26	35	39	40
8	11	32	40	41	42	45
9	1	5	11	25	40	41
10	3	27	33	34	35	39
11	7	12	20	32	38	43
12	5	21	29	33	39	43
887	9	10	11	14	25	28
888	3	6	10	17	38	43
889	6	9	10	24	37	38
890	3	16	24	26	36	40

Table 4. Data of the Hungarian six-numbers lotto since its start

The whole data on the lottery since its beginning is contained in the EXCEL file where the reader may find also data on the 5 out of 90 lotto in Hungary. Furthermore, the file contains a detailed analysis of the problem by classical and Bayesian methods done as part of the project by the students. Here, we focused on a presentation of the methods used to solve the problem.

[To article](#) 

[Top](#) 

DEMONSTRATION IN EXCEL

[To article](#) 
[Top](#) 

Classical estimation of the number of balls

Lottery n out of N: Number N of balls is unknown

We draw n = 6 without replacement: the Lotto numbers; ordered: $x_1 < x_2 < \dots < x_6$

There are different possibilities to estimate N

They have different properties

[Contents](#)

[Estimators](#)

[Distribution of estimators](#)

[Calculation of estimators](#)

	Median	Mean	MLE	Extreme gaps	Mean gap
Formula	$2 \times \text{med}(x_1, \dots, x_6)$	$2 \times \text{mean}(x_1, \dots, x_6) - 1$	$\max = x_6$	$\max + \min - 1$	$\max + (\max - 6)/6$
Reason	med "=" $N/2$	mean = $(N+1)/2$		$\min - 1 = N - \max$	$N - \max$ "=" mean gap

From the 890 draws of the six numbers lottery, the various estimators are calculated below. These estimates are investigated here further:

Properties

Mean	45,942	45,229	39,419	45,126	44,989	
Stand. Dev.	13,952	9,990	5,150	7,987	6,009	
True value	45	45	45	45	45	
Unbiased	yes	yes	no	yes	yes	
Variance	5	4	1	3	2	Rank in the sense "better"

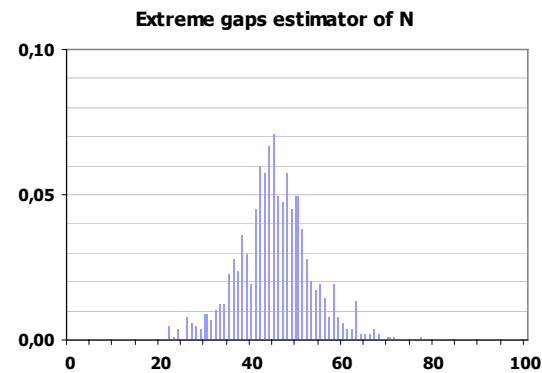
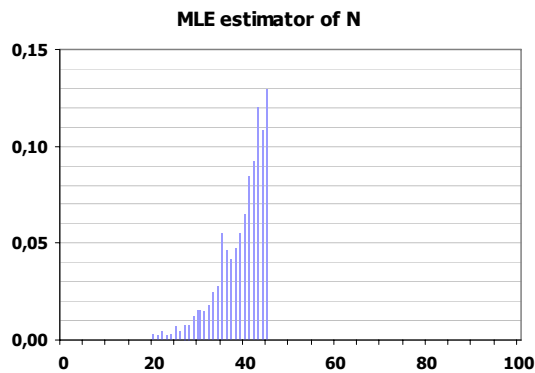
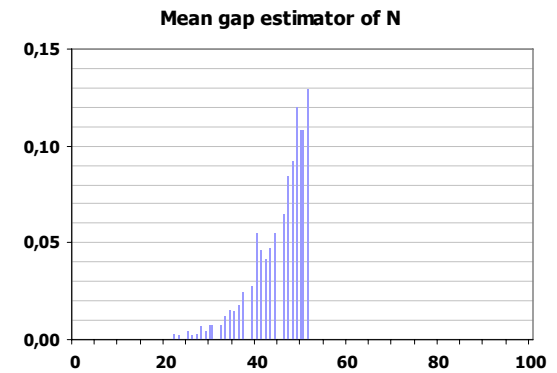
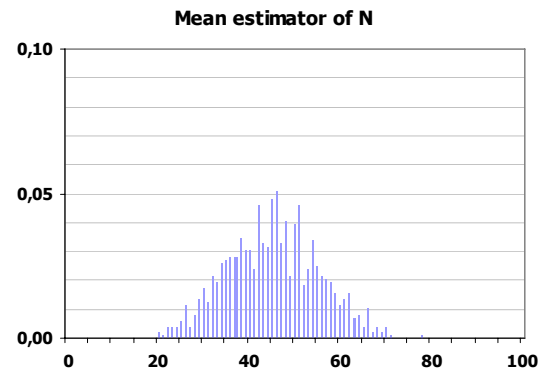
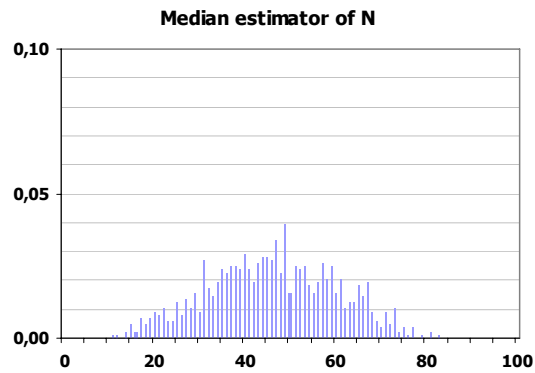
Graphs of the distribution of the estimators

[Contents](#)[Estimators](#)[Distribution of estimators](#)[Calculation of estimators](#)

From the graphs one may see:

Median and mean estimators have great variance but are centred correctly

Mean gap, extreme gaps, and MLE estimator have a small variance but the MLE is biased - it yields too small estimates

[To article](#) [Top](#) 

Calculation of the estimators

[Contents](#)
[Estimators](#)
[Distribution of estimators](#)
[Calculation of estimators](#)

Results of the draws are in a separate sheet

Draw Nr.	Median	Mean	MLE	Extreme gaps	Mean gap	
1	44	46,7	40,0	45,0	45,7	
2	55	52,7	43,0	51,0	49,2	
3	57	50,3	45,0	45,0	51,5	lines
889	34	40,3	38,0	43,0	43,3	hidden
890	50	47,3	40,0	42,0	45,7	
min	11	19,7	20,0	22,0	22,3	
max	83	78,3	45,0	77,0	51,5	

Frequencies	Median	Mean	MLE	Extreme gaps	Mean gap	
-0,5	0	0	0	0	0	-1
0,5	0	0	0	0	0	lines 0
99,5	0	0	0	0	0	99
100,5	0	0	0	0	0	100
sum	890	890	890	890	890	

Relative frequencies

-0,5	0,000	0,000	0,000	0,000	0,000	-1
0,5	0,000	0,000	0,000	0,000	0,000	0
1,5	0,000	0,000	0,000	0,000	0,000	lines 1
44,5	0,028	0,031	0,108	0,066	0,055	hidden 44
45,5	0,028	0,048	0,129	0,071	0,000	45
100,5	0,000	0,000	0,000	0,000	0,000	100
sum	1,000	1,000	1,000	1,000	1,000	

[To article](#) 
[Top](#) 

Classical confidence intervals for the number of balls

Lottery n out of N: Number N of balls is unknown

We draw $n = 6$ without replacement: the Lotto numbers; ordered: $x_1 < x_2 < \dots < x_6$

Task: to derive a **confidence interval for the unknown number N** of balls

[Contents](#)

[Confidence intervals](#)

[Graphs of conf. inter. 1](#)

[Graphs of conf. inter. 2](#)

[Calculation of confidence intervals](#)

Confidence interval for N at level $1 - \alpha$:

$[k, J_0 - 1]$

$k =$ **observed max** of drawn numbers

$J_0 =$ **smallest number that fullfils an inequality (*)**

Justification:

J = candidate for entering the confidence interval for N at level $1 - \alpha$:

J cannot enter the confidence interval if

$P(\text{all numbers} \leq k \mid J) \leq \alpha$

Let $J_0 :=$ smallest J that fullfils the inequality then the confidence interval is: $[k, J_0 - 1]$

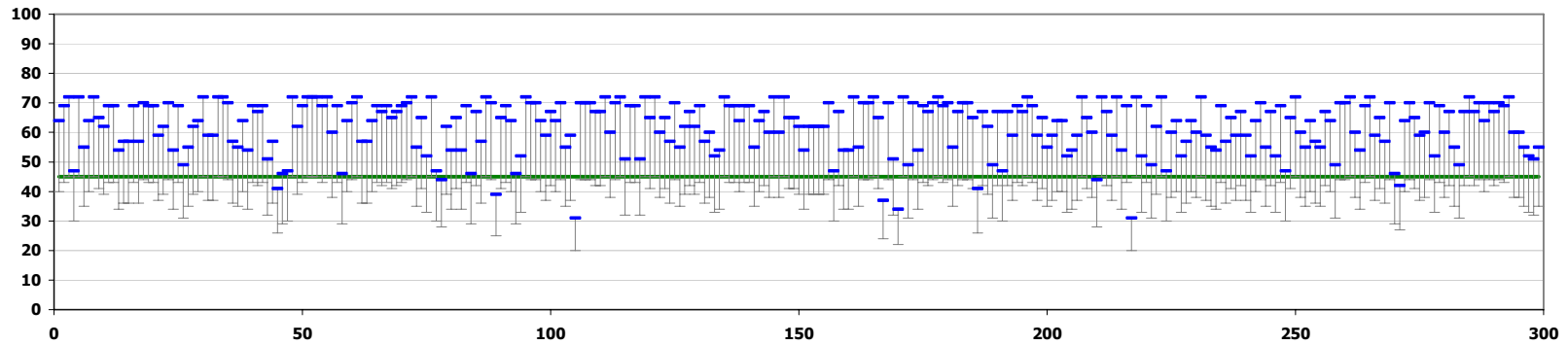
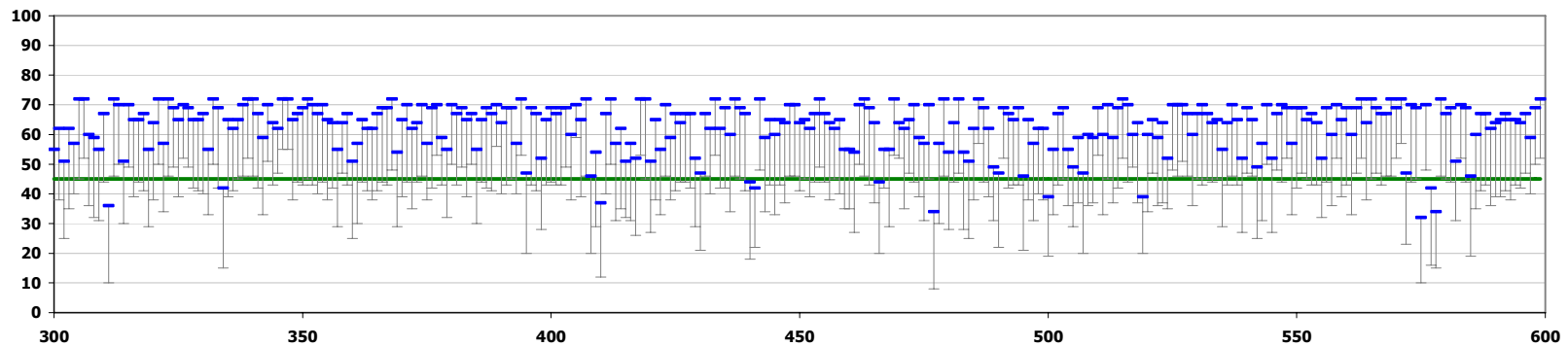
Detail of justification:

$P(\text{all numbers} \leq k \mid J) = k/J \times (k-1)/(J-1) \times \dots \times (k-5)/(J-5)$

smallest J:	$k/J \times (k-1)/(J-1) \times \dots \times (k-5)/(J-5)$	\leq	specific value of α
(*)	$20 \times k \times (k-1) \times \dots \times (k-5)$	\leq	$J \times (J-1) \times \dots \times (J-5)$

[To article](#) 

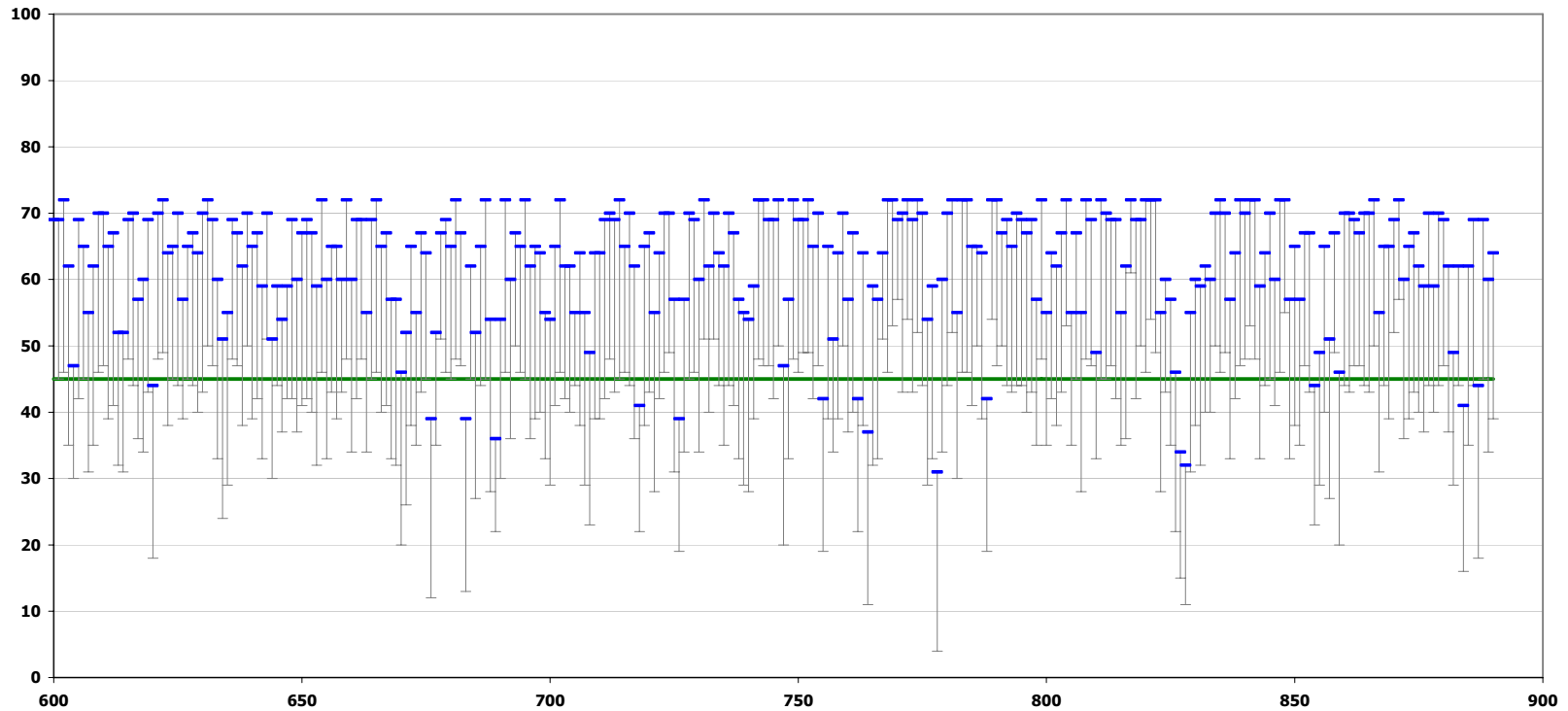
[Top](#) 

Graphs of the confidence intervals for all draws[Contents](#)[Confidence intervals](#)[Graphs of conf. inter. 1](#)[Graphs of conf. inter. 2](#)[Calculation of confidence intervals](#)**Weekly confidence intervals for the number N of balls****Weekly confidence intervals for the number N of balls**[To article](#) [Top](#) 

Graphs of the confidence intervals for all draws - continued

[Contents](#) [Confidence intervals](#) [Graphs of conf. inter. 1](#) [Graphs of conf. inter. 2](#) [Calculation of confidence intervals](#)

Weekly confidence intervals for the number N of balls - last period enlarged



[To article](#) 

[Top](#) 

Calculation of confidence intervals for all draws

[Contents](#)
 [Confidence intervals](#)
 [Graphs of conf. inter. 1](#)
 [Graphs of conf. inter. 2](#)
 [Calculation of confidence intervals](#)

Results of the draws are in a separate sheet

Draw Nr.	lower	upper	coverage	true number	
0					
1	40	64	1	45	24
2	43	69	1	45	26
3	45	72	1	45	27
889	38	60	1	45	22
890	40	64	1	45	24
min	20	31	852	count	
max	45	72	95,7	%	

Calculation of J_0

k = max	$20 \times k \times (k-1) \times \dots \times (k-5)$	smallest J	$J \times (J-1) \times \dots \times (J-5)$	indicating that (*) is fulfilled	
6	14.400,00	12	665.280,00	1	
7	100.800,00	10	151.200,00	1	lines
44	101.650.348.800,00	71	103.117.679.280,00	1	hidden
45	117.288.864.000,00	73	122.565.925.440,00	1	

To article 

Top 

Updating of prior distribution on the maximal number by the results of one week

Lottery n out of N: Number N of balls is unknown

"Complete" ignorance on this number N will be modelled by a uniform distribution on [31, 80]

The prior is updated by the results of **one** week to a new **posterior distribution reflecting the information of the data of one week**

This new status of knowledge on the maximal numbers of balls is calculated and graphically presented.

This posterior is clearly dependent on the week's results, especially on the maximum of the drawn numbers.

[Contents](#)

[Graphs of posterior 1](#)

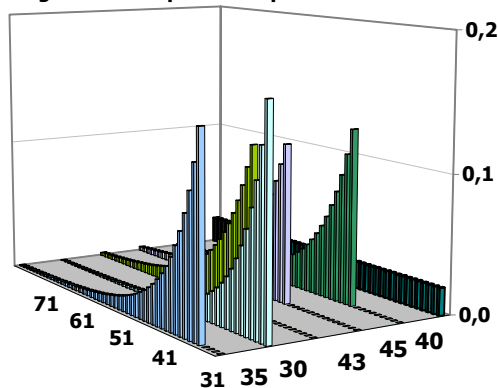
[Graphs of posterior 2](#)

[Calculation of posterior](#)

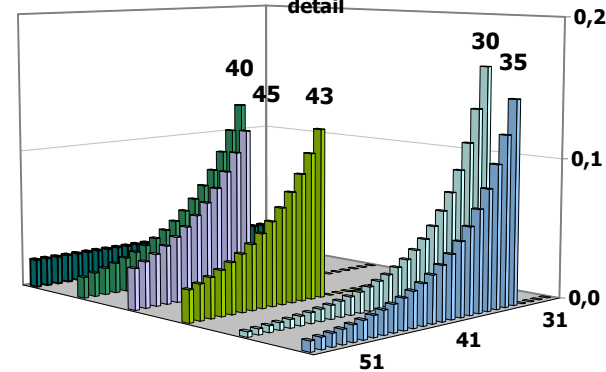
Graphical presentation how the posterior depends on the week's maximum

Nr	x_1	x_2	x_3	x_4	x_5	x_6	max
1	6	15	20	24	38	40	40
3	1	7	14	43	44	45	45
2	9	14	23	32	40	43	43

Updating of uniform prior in dependence of week's maximum



Updating of uniform prior in dependence of week's maximum - detail



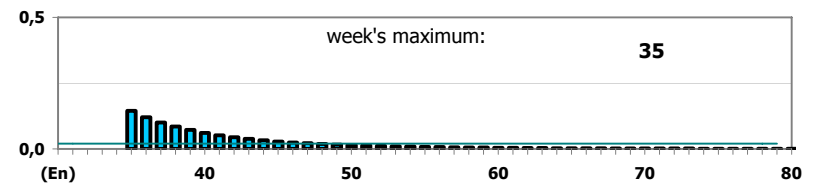
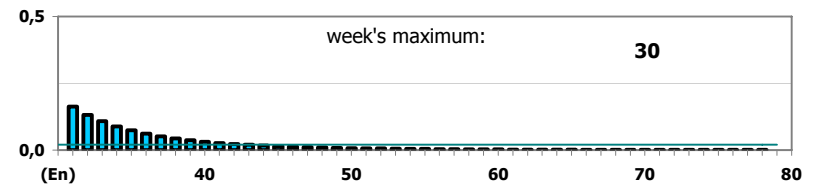
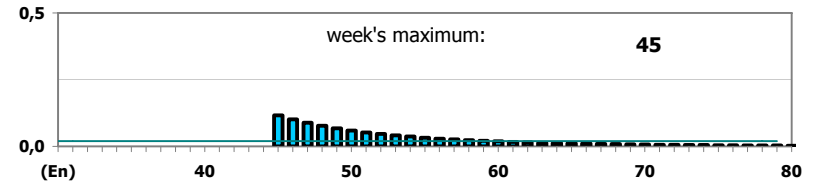
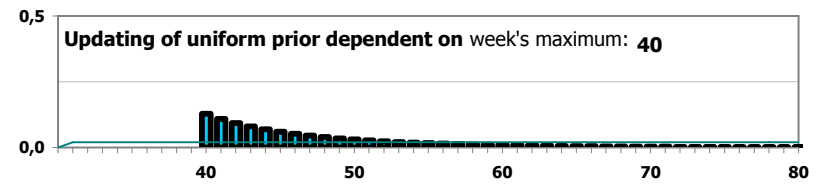
[To article](#) 

[Top](#) 

An other representation of the posterior distributions

[Contents](#)
[Graphs of posterior 1](#)
[Graphs of posterior 2](#)
[Calculation of posterior](#)

Nr	x_1	x_2	x_3	x_4	x_5	x_6	max
1	6	15	20	24	38	40	40
3	1	7	14	43	44	45	45
2	9	14	23	32	40	43	43


[To article](#)
[Top](#)

Calculation of the posterior distributions including only the actual week's results and the uniform prior

Contents	Graphs of posterior 1						Graphs of posterior 2						Calculation of posterior					
Nr	2	3	4	5	6	7	k			31	32	33	34	35	79	80		
								prior	$P(E_n)$	0,0200	0,0200	0,0200	0,0200	0,0200	0,0200	0,0200	1	
1	6	15	20	24	38	40	40	posterior	$P(E_n k)$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0018	0,0016	1	
2	9	14	23	32	40	43	43	posterior	$P(E_n k)$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0026	0,0024	1	
3	1	7	14	43	44	45	45	posterior	$P(E_n k)$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0034	0,0032	1	
4	6	9	11	14	18	30	30	posterior	$P(E_n k)$	0,1623	0,1318	0,1079	0,0888	0,0736	0,0004	0,0004	1	
5	12	13	19	30	34	45	45	posterior	$P(E_n k)$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0034	0,0032	1	
6	16	17	19	20	34	35	35	posterior	$P(E_n k)$	0,0000	0,0000	0,0000	0,0000	0,1445	0,0008	0,0008	1	
7	4	23	26	35	39	40	40	posterior	$P(E_n k)$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0018	0,0016	1	
8	11	32	40	41	42	45	45	posterior	$P(E_n k)$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0034	0,0032	1	
9	1	5	11	25	40	41	41	posterior	$P(E_n k)$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0020	0,0019	1	
10	3	27	33	34	35	39	39	posterior	$P(E_n k)$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0015	0,0014	1	

Cumulative - week by week - updating of uniform prior

Lottery n out of N: Number N of balls is unknown

The number of balls "is" unknown. "Complete" ignorance on this number N will be modelled by a uniform distribution on [31, 80]

The prior is updated weekly by the results of this week to a new posterior distribution

This posterior is then the prior distribution for the next week

The topical status of the information on the maximal number of balls N "converges" rapidly to the true maximal number, which is 45.

The process of updating week by week takes all draws of the past into account

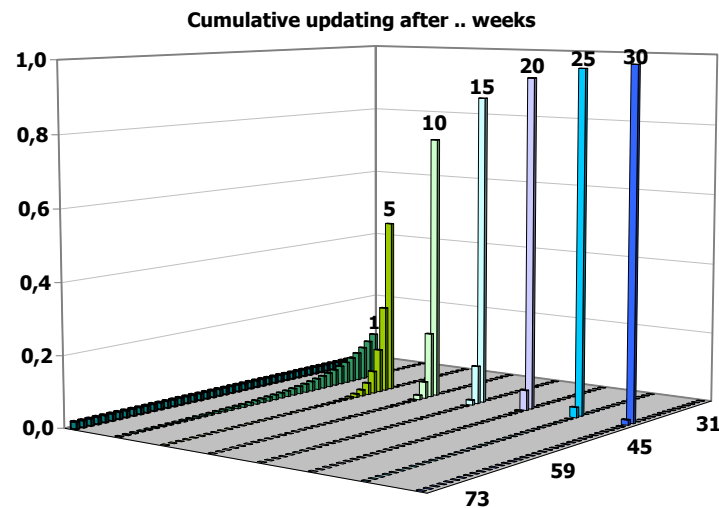
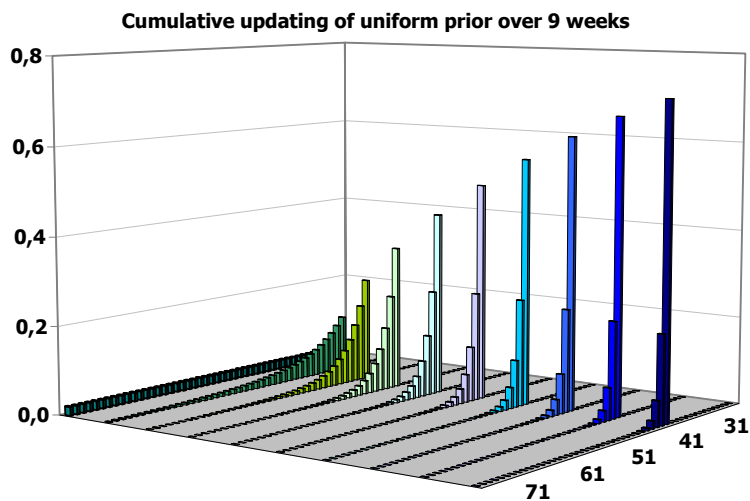
[Contents](#)

[Graphs of posterior 1](#)

[Graphs of posterior 2](#)

[Calculation of cumulative posterior](#)

Graphical presentation of the development of the posterior



[To article](#)

[Top](#)

An other representation of the posterior distributions

[Contents](#)

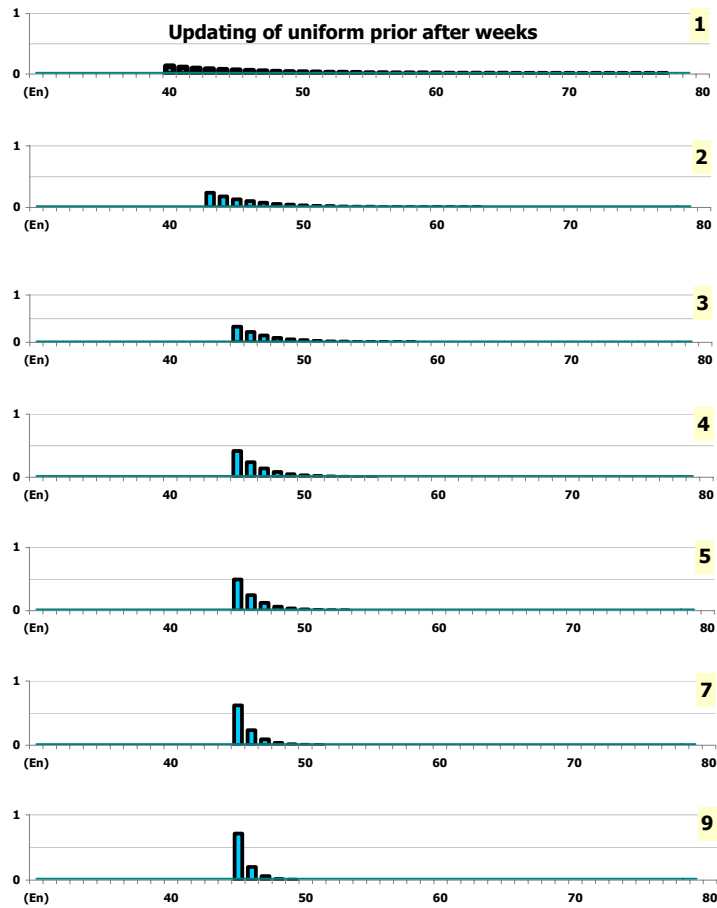
[Graphs of posterior 1](#)

[Graphs of posterior 2](#)

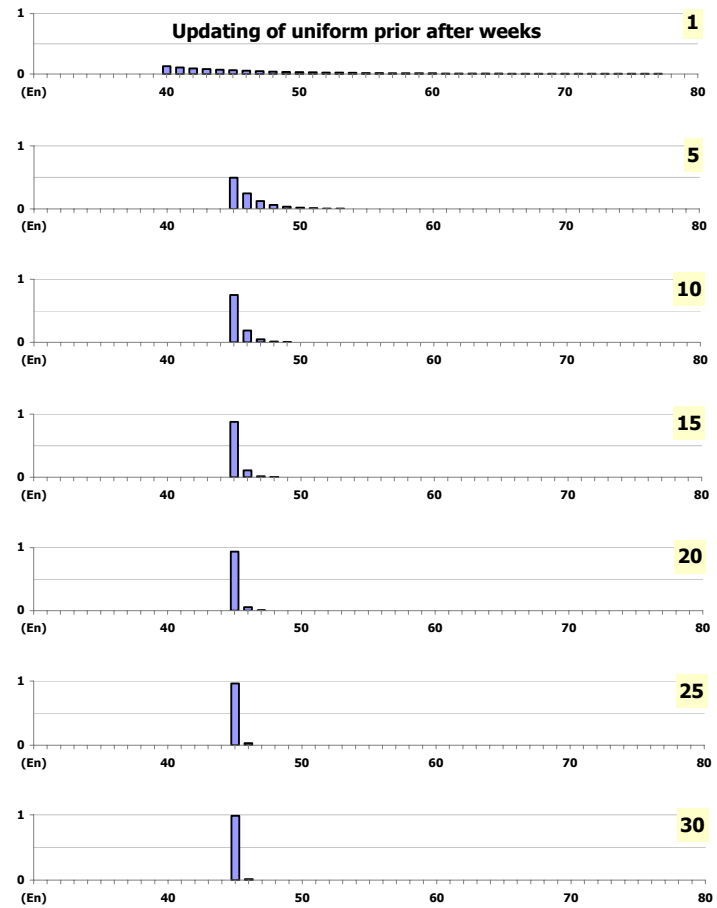
[Calculation of cumulative posterior](#)

[To article](#) 

[Top](#) 



After 9 weeks, only 45 - 48 bear a non-negligible probability.



After 30 weeks, only 45 is relevant; there is a small probability for 46.

Hungarian six-numbers lottery - results since its start

[Contents](#)

Ordered numbers of successive draws

Nr	x_1	x_2	x_3	x_4	x_5	x_6	
1	6	15	20	24	38	40	
2	9	14	23	32	40	43	
3	1	7	14	43	44	45	
4	6	9	11	14	18	30	
5	12	13	19	30	34	45	
6	16	17	19	20	34	35	
7	4	23	26	35	39	40	
8	11	32	40	41	42	45	
9	1	5	11	25	40	41	
10	3	27	33	34	35	39	
11	7	12	20	32	38	43	
12	5	21	29	33	39	43	
							lines
887	9	10	11	14	25	28	hidden
888	3	6	10	17	38	43	
889	6	9	10	24	37	38	
890	3	16	24	26	36	40	

[To article](#) 

[Top](#) 

Hungarian five-numbers lottery - results since its start

Contents

Ordered numbers of successive draws

Nr	x_1	x_2	x_3	x_4	x_5
1	10	16	29	34	42
2	20	43	53	54	58
3	17	21	57	62	71
4	4	7	31	54	57
5	7	11	15	66	79
6	7	11	15	66	79
7	14	32	53	70	86
8	20	44	52	74	85
9	26	31	33	36	37
10	11	25	57	79	86
11	7	22	79	86	88
12	5	19	26	62	64
lines					
2710	3	33	46	59	75
hidden					
2711	16	38	40	65	73
2712	20	55	64	73	75
2713	7	52	58	62	67
2714	5	27	45	46	48
2715	12	19	35	42	64
2716	3	5	19	20	48
2717	7	12	24	30	75
2718	22	26	22	27	66

Author : Ödön Vancsó
 E-mail : vancso@ludens.elte.hu
 Address : Institute for Mathematics, Eötvös Lóránd University, Budapest, Hungary

To article 
 Top 