A Statistical Look at the Gauss-Kuzmin Distribution

Steven Duff
Bucknell University
2011 Joint Mathematics Meetings
Special Session on Continued Fractions

Simple Continued Fractions

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \ddots}}}}$$

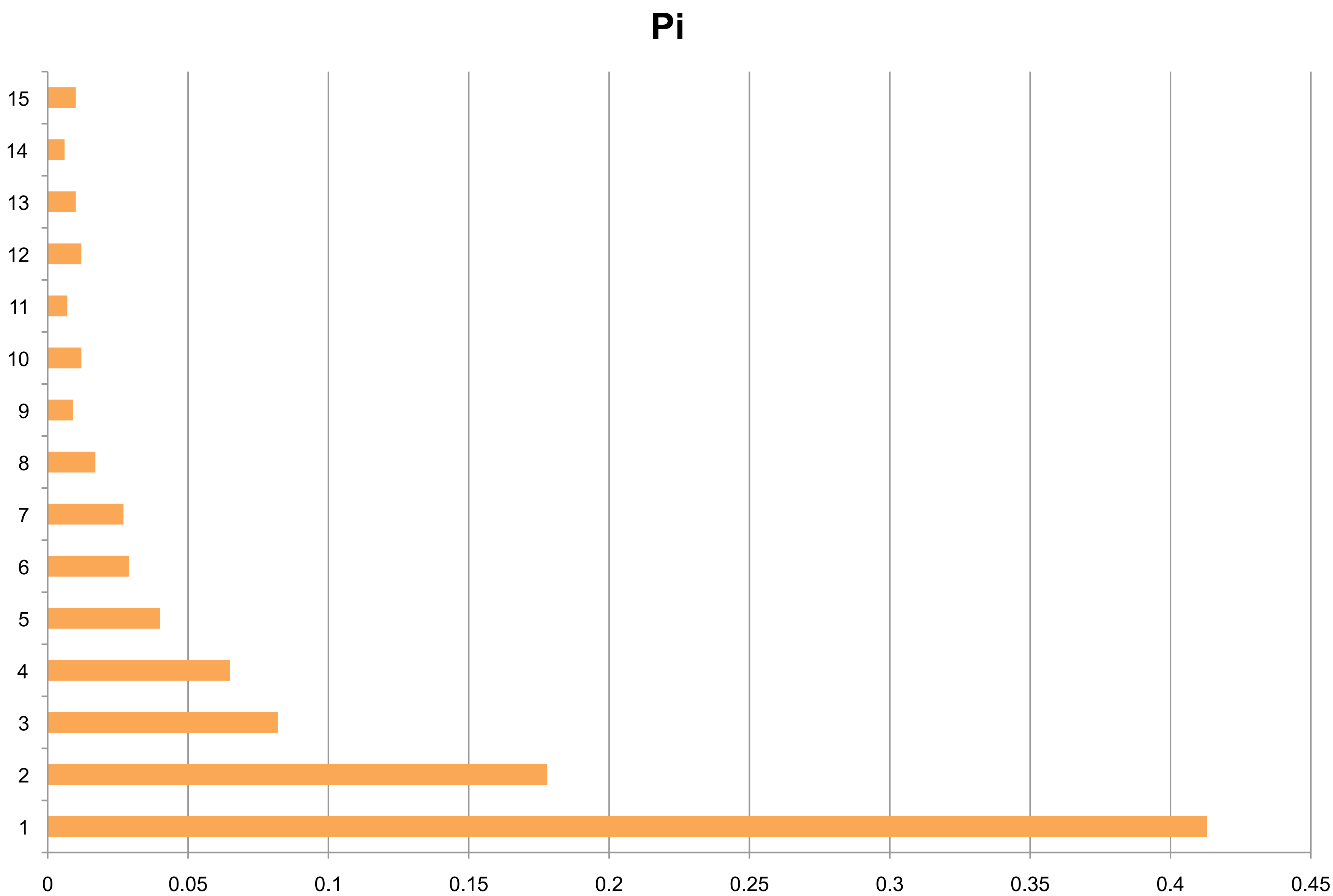
x real

Terms: $\{a_n\}$ integers

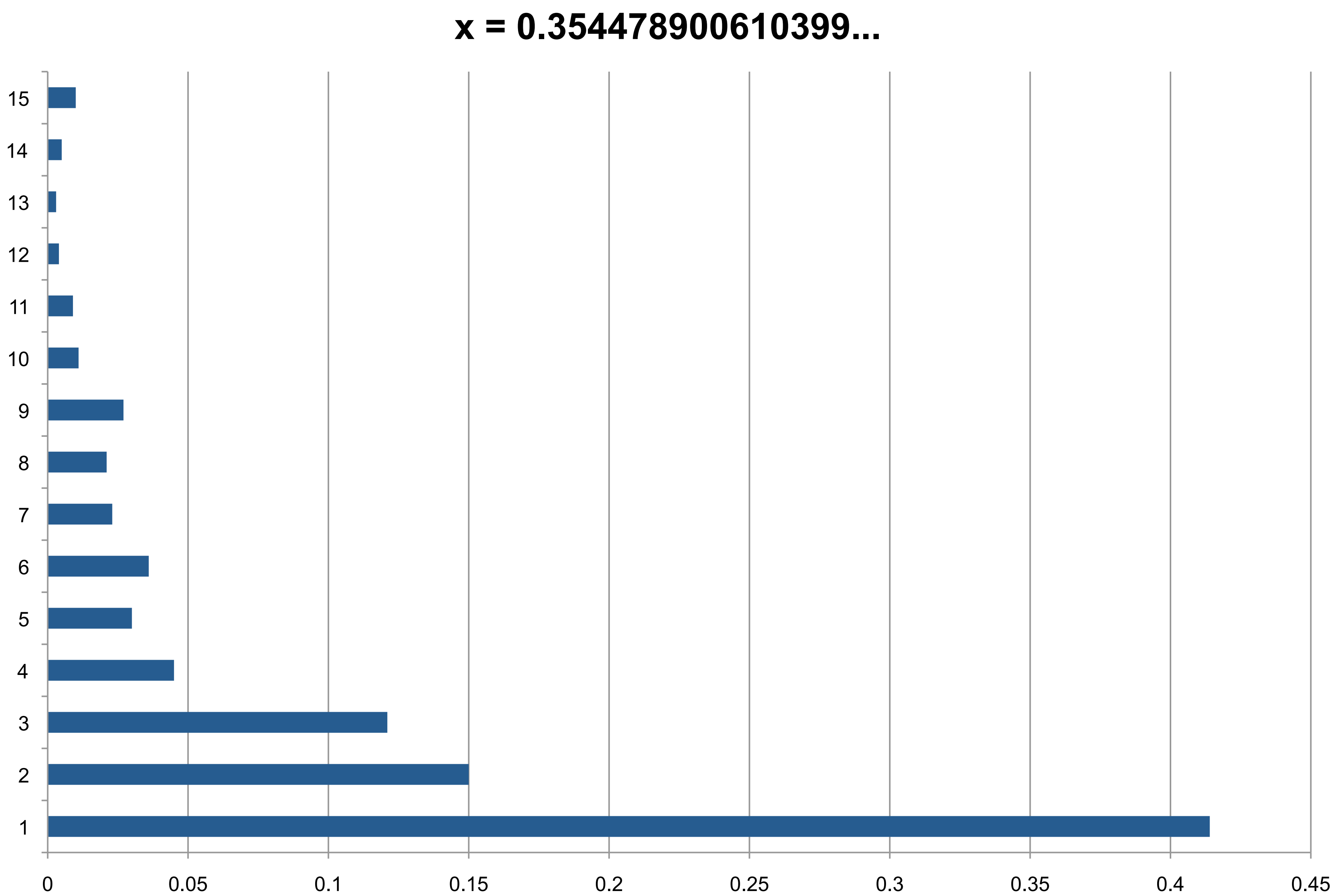
$a_k > 0$ for all $k > 0$

All reals have unique simple cf expansion

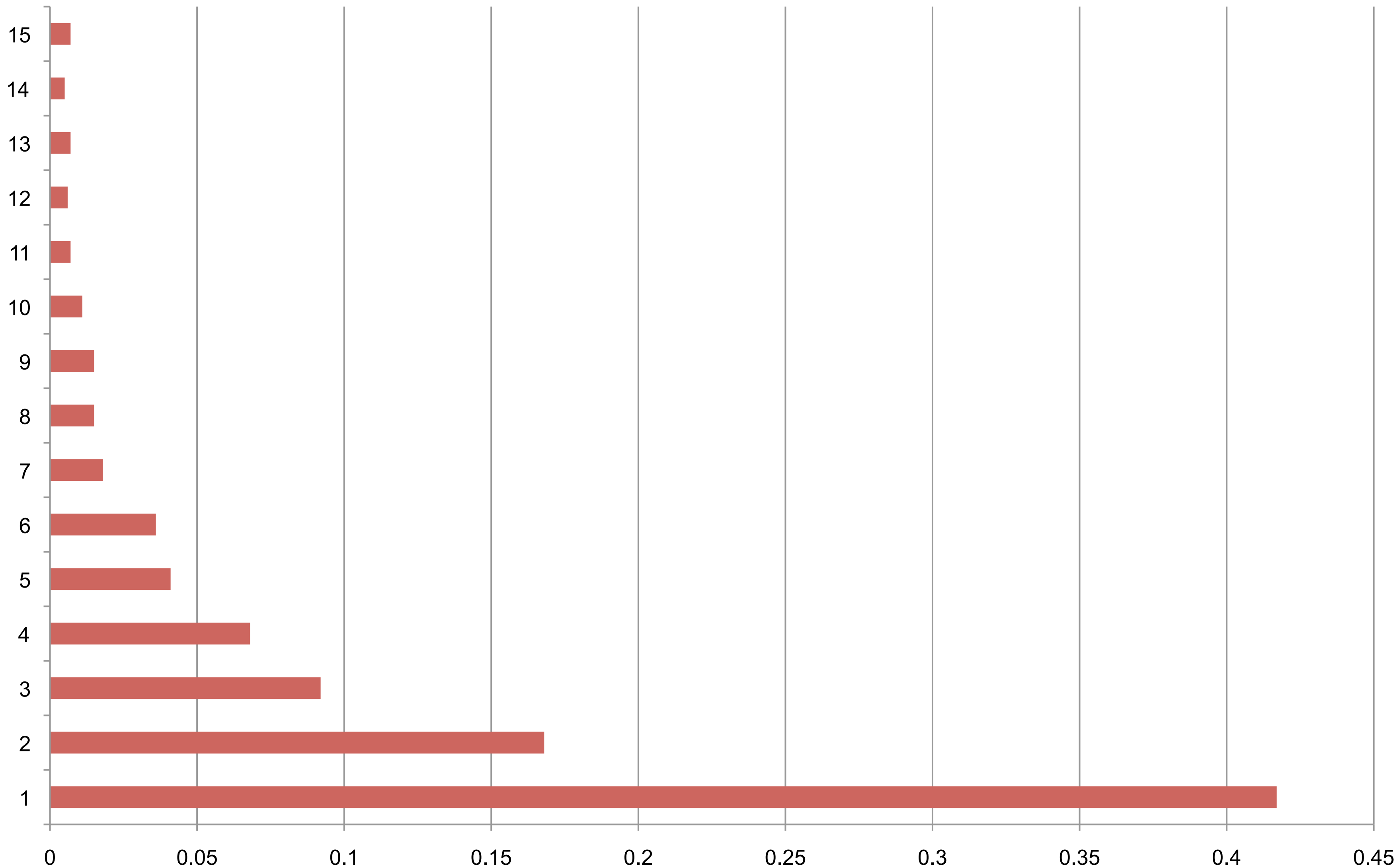
$$\pi = 3+\cfrac{1}{7+\cfrac{1}{15+\cfrac{1}{1+\cfrac{1}{292+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{3+\cfrac{1}{1+\cfrac{1}{14+\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\ddots}}}}}}}}}}}}}}}}}}}}$$

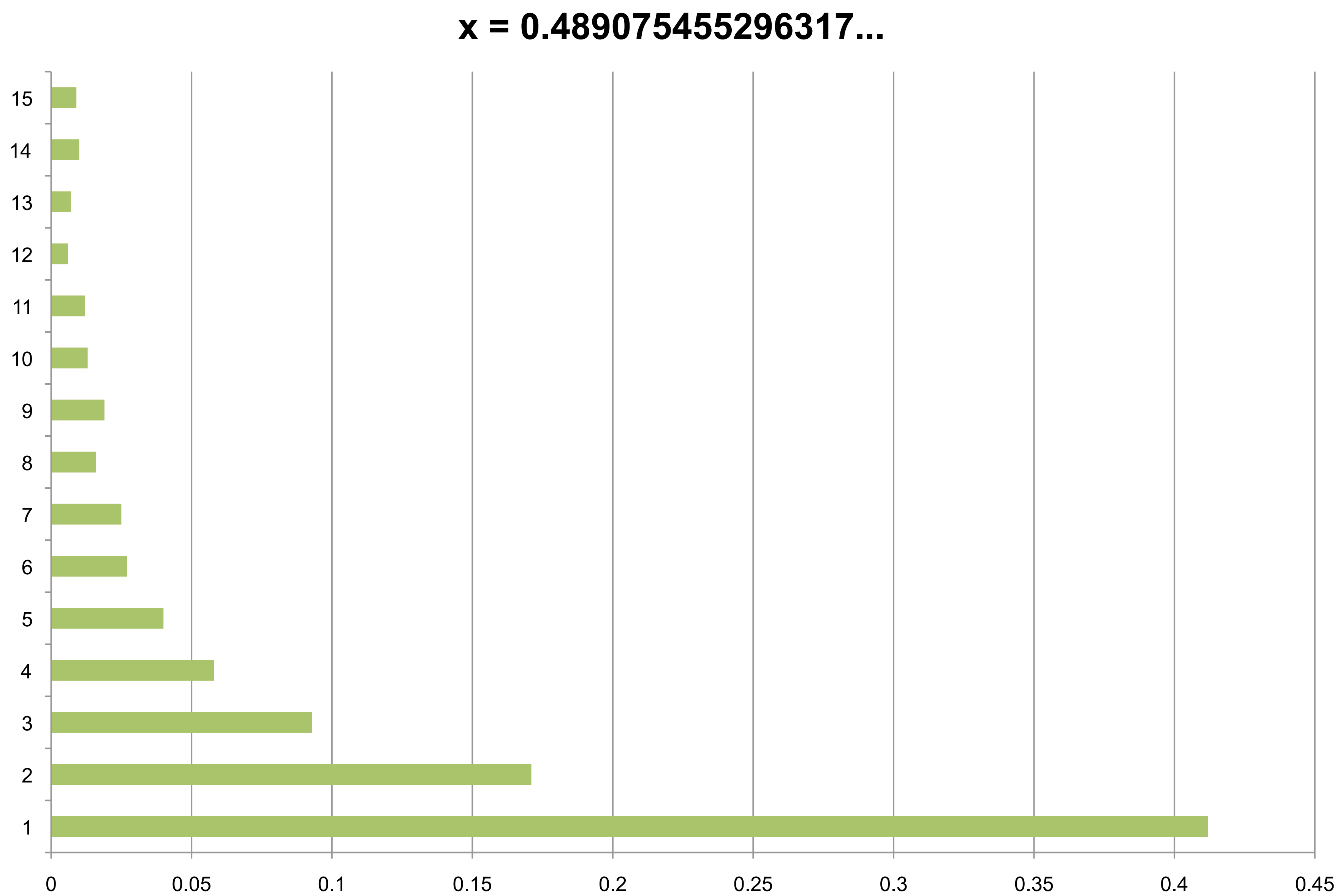


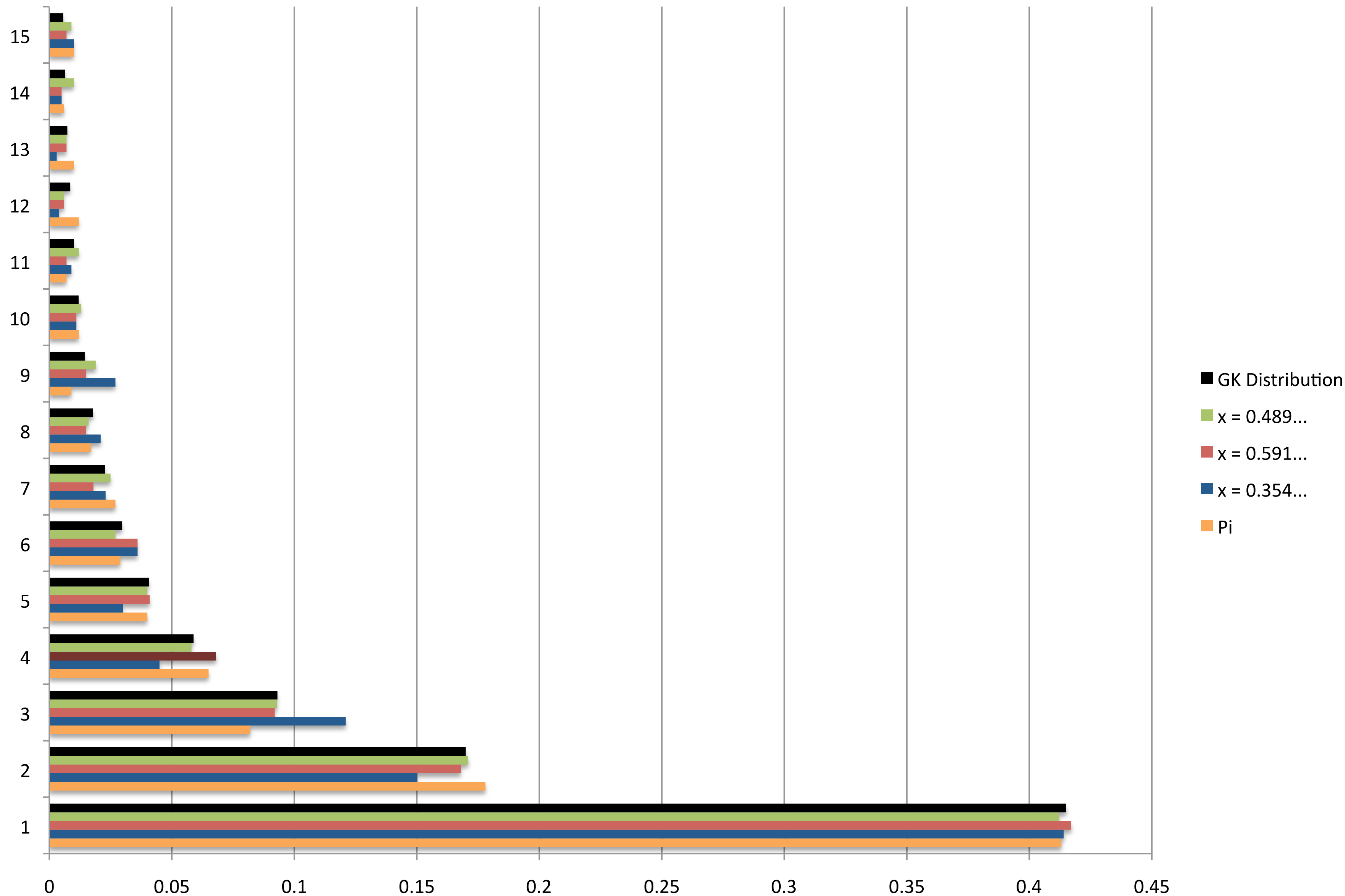
Pi
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45


x = 0.354478900610399...
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45

x = 0.591627180102407...




x = 0.489075455296317...
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45


15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
GK Distribution
x = 0.489...
x = 0.591...
x = 0.354...
Pi

Gauss-Kuzmin Distribution

$$\left| \mathsf{Prob}(a_n = k) - \log_2\left(1 + \frac{1}{k(k+2)}\right) \right| \leqslant \frac{A}{k(k+1)} e^{-B\sqrt{n-1}}$$

Holds for a set of full measure $K \subset \mathbb{R}$

Exceptional Set Z

$$\mathbb{Q} \subset Z$$

Bounded cf's $\subset Z$

Quadratic irrationals $\subset Z$

Linearly periodic cf's $\subset Z$

Linearly Periodic CF's

$$x = [a_0, \ldots, a_n, f_1(0), f_2(0), \ldots, f_k(0), f_1(1), \ldots, f_k(1), f_1(2), \ldots]$$

$f_1, f_2, \ldots, f_k$ linear functions

e.g.

$$e = [2; 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, \ldots]$$

for e, $f_1(n) = 1, f_2(n) = 2n, f_3(n) = 1$

Linearly Periodic CF's $\subset Z$ (Dimofte)

Randomly Generating CF's

Computing cf expansions

We want $\{r_n\}$, such that:

$$r_n = a_n + \cfrac{1}{a_{n+1} + \frac{1}{\ddots}} = a_n + \frac{1}{r_{n+1}}$$

1. Take random $x \in \mathbb{R}$
2. $r_0 = x$
3. $a_n = \lfloor r_n \rfloor$
4. $r_{n+1} = \dfrac{1}{(r_n - a_n)}$

Random CF Term by Term

Convergents

$$\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n]$$

$$\text{Put } \frac{p'_n}{q'_n} = [a_0, a_1, \ldots, a_n + 1]$$

1. Randomly choose $x \in \mathbb{R}$
2. $a_0 = \lfloor x \rfloor$
3. Randomly choose $x_n \in \left[\frac{p_n}{q_n}, \frac{p'_n}{q'_n}\right]$
4. $r_{n+1} = \frac{p_{n-1} - x_n \cdot q_{n-1}}{x_n \cdot q_n - p_n}$
5. $a_{n+1} = \lfloor r_{n+1} \rfloor$

Experiment

Looking at $x \in \mathbb{R}$

$$H_0 : x \in K$$

$$H_a : x \in Z$$

Chi Square Test

$O_i =$ observed number in i^{th} bin

$E_i =$ expected number in i^{th} bin

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Parameters

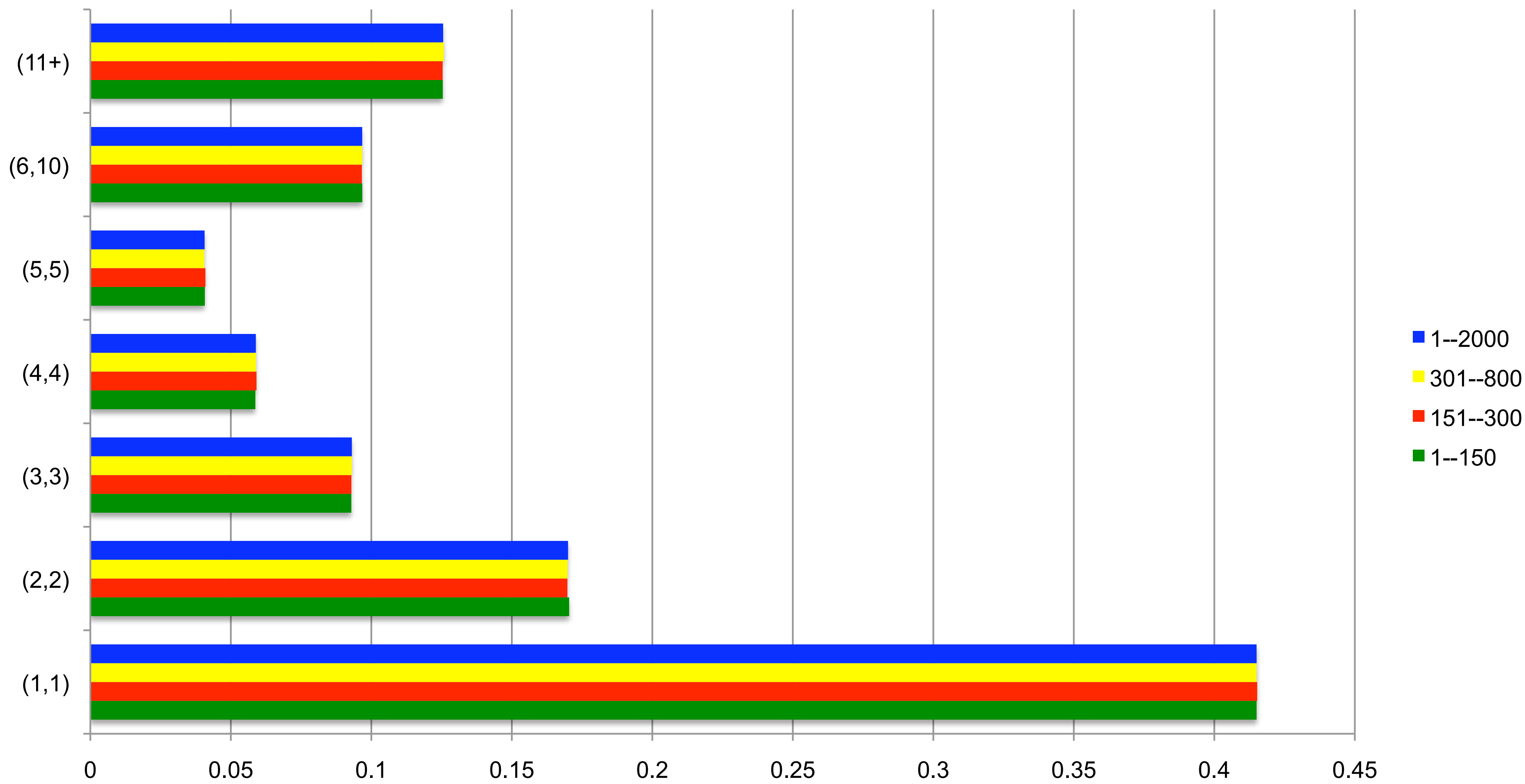
Bins

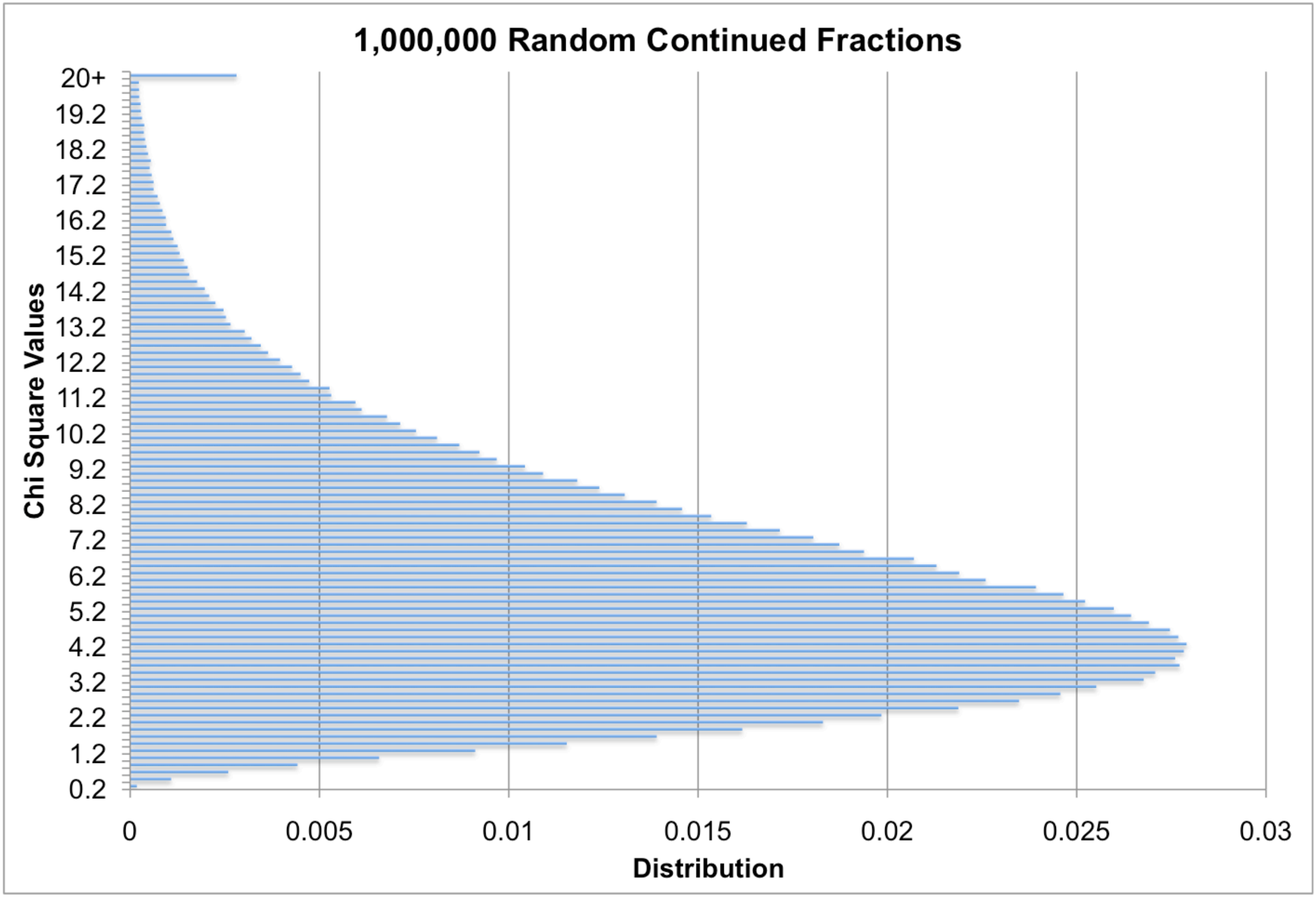
Terms

Chi Square Values of Different Bins

Bins	Mean	Standard Deviation
1, 2, 3, 4, 5, 6-10, 11+	5.901621415	3.350807445
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21+	20.03536582	6.719941612
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11+	9.97929274	4.524431181
1-2, 3-4, 5-6, 7-8, 9-10, 11+	5.037319422	3.349067771
1-2, 3-4, 5-6, 7-8, 9-10, 11-12, 13-14, 15-16, 17-18, 19-20, 21+	10.04140366	4.731914795
1-5, 6-10, 11-20, 21+	2.965167451	2.406082356
1-5, 6-10, 11-15, 16-20, 21+	3.993537546	3.005040382

Distributions from Different Terms




1,000,000 Random Continued Fractions
Chi Square Values
20+
19.2
18.2
17.2
16.2
15.2
14.2
13.2
12.2
11.2
10.2
9.2
8.2
7.2
6.2
5.2
4.2
3.2
2.2
1.2
0.2
0
0.005
0.01
0.015
0.02
0.025
0.03
Distribution

Does e Follow Gauss-Kuzmin?

$$H_0 : e \in K, \quad H_a : e \in Z$$

$$\chi^2 = 118.94$$

p-value$< .01$

What about e^n, for integers n?

	χ^2	p-value
e^2	292.385	$< .01$
e^3	1.200	.98
e^4	8.801	.17
e^5	3.541	.73
e^6	4.424	.60
e^7	2.716	.84
e^8	9.579	.13
e^9	4.398	.61
e^{10}	4.757	.56