

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Test of Hypothesis on Fair Die (Simulation of Chi Square Tests)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

paditz@informatik.htw-dresden.de

Abstract:

The considered experiments help our students better to understand the randomness and the statistic methods of every day life. At first we initialize the random number generator of our **CASIO FX 2.0 PLUS** in the RUN-menu by the help of **Ran# 0**. Let us begin with an experiment on a die which has been rolled **N = 100** times. Each face does not appear an equal number of times. Is there something wrong with the die? In **M = 250** of such experiments the chi square variable is computed, i.e. the die is rolled **N * M = 25000** times by the help of the **CASIO FX 2.0 PLUS**. We check up that indeed chi square variables are simulated (only in the case of a fair die). The chi square variable is a statistical measure on the difference between the expected outcome and the actual outcome. The probability theory tells us that we should expect each face of the die to appear **N / 6** times. But in actuality this usually does not happen. By the help of the **CASIO FX 2.0 PLUS** we simulate **M = 250** chi square variable to answer the question „What is the significance of the chi square test?“ and „How close to zero must the chi square variable be to conclude to have a fair die?“. Here we simulate an unfair die with the probability distribution $P(X = k) = 2 / 11$ for $k=1,2,3,4,5$ and $P(X = k) = 1 / 11$ for $k=6$.

Keywords: chi square goodness of fit test, random number generator, simulation and exploration of data

1. Discussion on the considered problem:

The **chi square goodness of fit test** computes the chi square variable C, which we have simulated **M** (= 250) times (**M** experiments), to decide the hypothesis on the fairness of the rolled die.

The **null hypothesis** is $P(X = k) = 1 / 6$ for all $k = 1, 2, 3, \dots, 6$.

(The **alternative** let be

$P(X = k) = 2 / 11$ for all $k = 1, 2, 3, 4, 5$, and $P(X = k) = 1 / 11$ for $k = 6$)

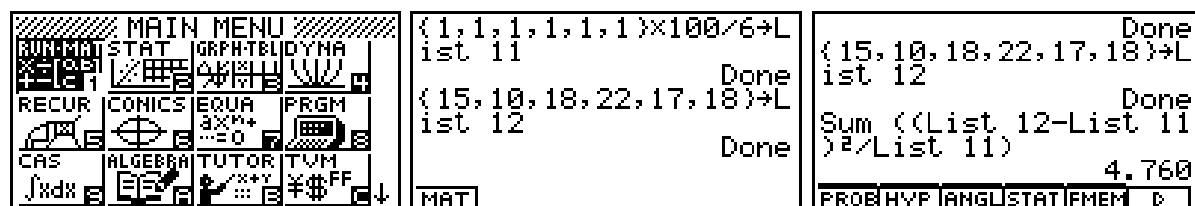
If in one experiment we roll the die **N** (=100) times, we have

the **expected frequencies** $\text{List 11} = \{N/6, N/6, N/6, N/6, N/6, N/6\}$ and

the **observed frequencies** $\text{List 12} = \{H_1, H_2, H_3, H_4, H_5, H_6\}$ with $H_1 + H_2 + \dots + H_6 = N$

E.g. let be $\text{List 12} = \{15, 10, 18, 22, 17, 18\}$

We compute in the RUN-menu $C = \text{chi square value} = \text{Sum}((\text{List 12} - \text{List 11})^2 / \text{List 11}) = 4.76$:



Practically by the help of **one** chi square value **C** we have to decide between the null hypothesis and the alternative.

What is with the error of first kind, if we decide against the null hypothesis and the null hypothesis was valid?

What is with the error of second kind, if we decide for the null hypothesis and the null hypothesis was false?

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

Remember:

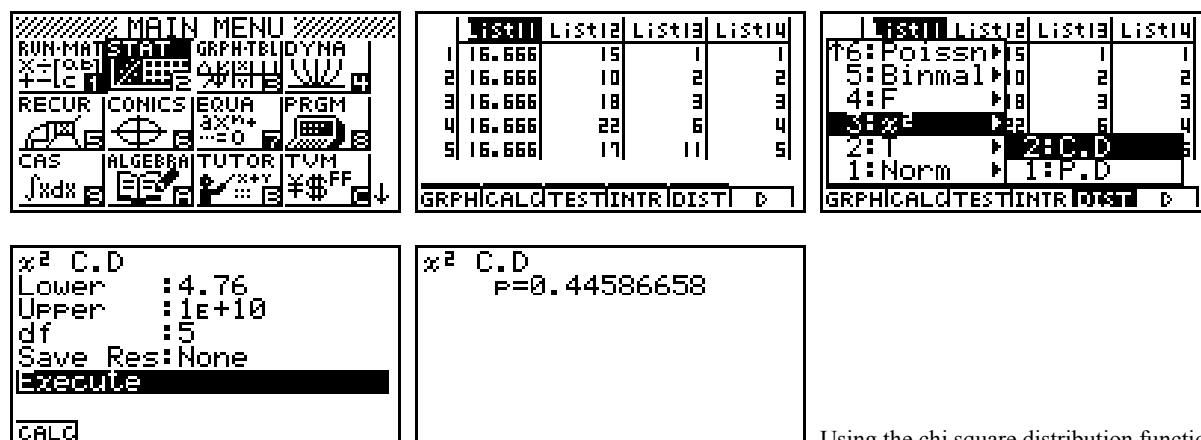
We consider the probability of the error of the first or of the second kind.

By the help of chi square distribution (5 degrees of freedom) we know:

$$P(C \geq 4.76 \mid \text{null hypothesis is valid}) = 1 - \text{Int}(\sqrt{X^3 * e^{-X}} / (18\pi), 0, 4.76, 10^{-6}) = 0.446 = 0.45 = \text{alpha}$$

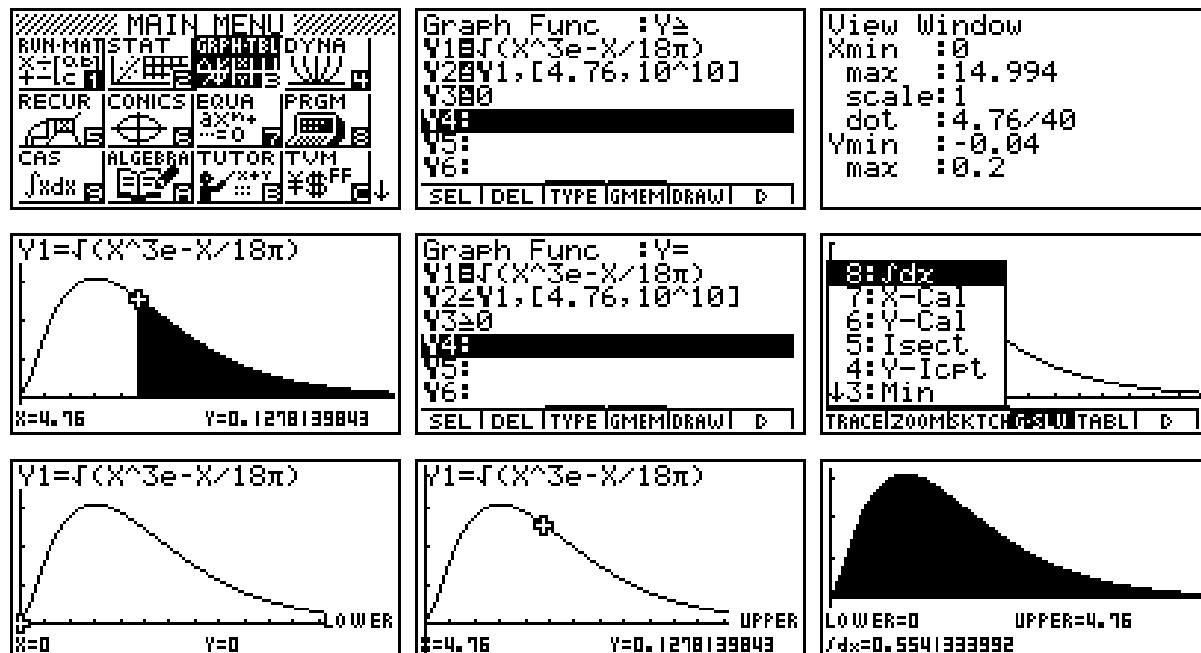
(Here the simulation shows $P(C \geq 4.76 \mid \text{null hypothesis is valid}) = 112/250 = 0.448 = 0.45$)

Thus alpha% = 45% of our experiments give a chi square value of a fair die, which is at least 4.76! If we decide against the null hypothesis, then the probability of the error of first kind is alpha% = 45%!

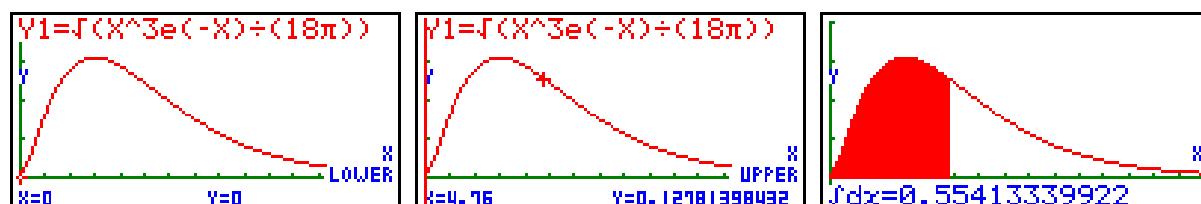


Using the chi square distribution function

It is possible to draw the chi square density function in the GRPH-TBL-menu and compute the probability $1-\text{alpha} = 1-0.446 = 0.554$ by the help of an integral (area):



Why is the whole area black? The same screenshots from the CASIO CFX 9850GB PLUS:



Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

Next, on the other hand by the help of our simulation,

$$P(C \leq 4.76 \mid \text{our alternative is valid}) = 45/250 = 0.18 = \text{beta}$$

Thus if we decide (because of $C \leq 4.76$) not against the null hypothesis, then the probability of the error of the second kind is beta% = 18% !

Finally the problems:

The chi square goodness of fit test is not contained in the STAT-menu (cp. F3: 3: chi square).

Practically the chi square goodness of fit test works with a significance of alpha% = 5%, i.e. we need the quantil $C_{0.95}$ with $P(C \geq C_{0.95}) = 0.05$ or $P(C \leq C_{0.95}) = 0.95$.

How to compute the quantil $C_{0.95}$ by the help of CASIO FX 2.0 PLUS? In the STAT-menu (cp. F5: 1:) we have have the Inverse of the normal distribution function but not in F5: 3: the Inverse of a chi square distribution!

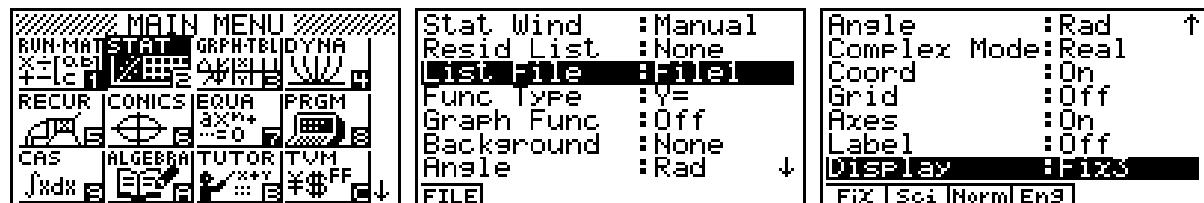
An other question ist the solution C_{α} of the Equation $\alpha = \text{beta}$, i.e.

$\alpha = P(C \geq C_{\alpha} \mid \text{null hypothesis is valid}) = P(C \leq C_{\alpha} \mid \text{our alternative is valid}) = \text{beta}$
and the computation of the error probability alpha (= beta) by the help of the CASIO FX 2.0 PLUS.

The CASIO FX 2.0 PLUS can not solve these problems in a direct manner (EQUA-menu or CAS-menu) but by the help of numerical integration and tabulation the (empirical) distribution functions of the simulated data in the RUN-menu. In the STAT-menu we can observe the functions in form of x-y-lines and search the solutions of the considered equations.

2. Screenshots on the simulation and statistical/graphical exploration of data:

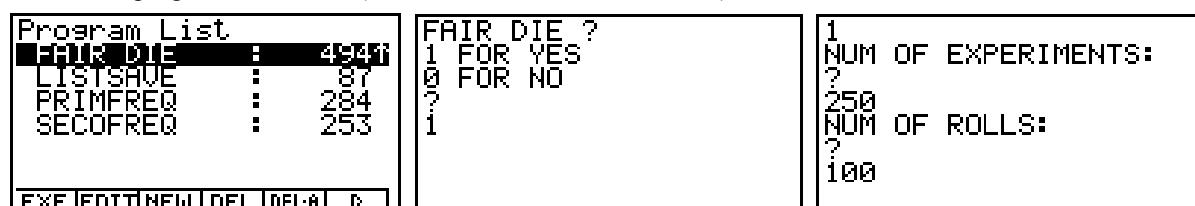
In the STAT-menu make the following SET UP: List File: File 1, Display: Fix3 :



Start of the random number generator after resetting: RUN-menu: Ran# 0 and using the generator Ran# 1



Start of the program FAIR DIE (Simulation of a fair die, code = 1)



The CASIO FX 2.0 PLUS needs approximately 30min to generate 250 chi square data. Some chi square data:

- Disp -		- Disp -		Done	
EXP. NO:	1.000	EXP. NO:	3.000	NUM OF EXPERIMENTS:	250.000
CHI^2=	5.480	CHI^2=	10.760	NUM OF ROLLS:	100.000

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

```
Program List
FAIR DIE : 494↑
LISTSAVE : 87
PRIMFREQ : 284
SECOFREQ : 253
EXE EDIT NEW DEL DEL-AT D
```

```
A MOMENT PLEASE, BUSY
...
- Disp -
```

```
Done
PRIMFREQ FINISHED
```

```
Program List
FAIR DIE : 494↑
LISTSAVE : 87
PRIMFREQ : 284
SECOFREQ : 253
EXE EDIT NEW DEL DEL-AT D
```

```
A MOMENT PLEASE, BUSY
...
- Disp -
```

```
Done
SECOFREQ FINISHED
```

```
Program List
FAIR DIE : 494↑
LISTSAVE : 87
PRIMFREQ : 284
SECOFREQ : 253
EXE EDIT NEW DEL DEL-AT D
```

```
LIST TRANSLATION
- Disp -
```

```
LIST TRANSLATION
Done
```

Start of the program **FAIR DIE** (Simulation of an unfair die, code = 0)
The CASIO FX 2.0 PLUS needs approximately 30min to generate 250 chi square data.

```
FAIR DIE ?
1 FOR YES
0 FOR NO
?
0
```

```
0
NUM OF EXPERIMENTS:
?
250
NUM OF ROLLS:
?
100
```

```
Done
NUM OF EXPERIMENTS:
250.000
NUM OF ROLLS:
100.000
CODE FAIRDIE:
0.000
```

```
Program List
FAIR DIE : 494↑
LISTSAVE : 87
PRIMFREQ : 284
SECOFREQ : 253
EXE EDIT NEW DEL DEL-AT D
```

```
Done
PRIMFREQ FINISHED
```

```
Program List
PRIMFREQ : 284↑
SECOFREQ : 253
EXE EDIT NEW DEL DEL-AT D
```

```
Done
SECOFREQ FINISHED
```

```
MAIN MENU
RUN-MAT STAT GRAPH-DYNA
X-Y-Z-EQUA PRGM
RECUR COMICS EQUA PRGM
CAS ALGEBRA TUTOR TVM
SolvE F2 F3 F4 F5 F6 F7 F8
```

Stat Wind	:	Manual
Resid List	:	None
List File	:	FILE
Func Type	:	Y=
Graph Func	:	On
Background	:	None
Angle	:	Rad
FILE		↓

Remember, using List File 1

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

	List 6	List 7	List 8	List 9	
1	0.2	0.2	1	0	
2	0.44	0.44	1	0.5	
3	0.8	0.8	1	1.5	
4	1.16	1.16	3	2.5	
5	1.16	1.28	5	3.5	
					0.2

	List 7	List 8	List 9	List 10	
1	0.2	1	0	3	
2	0.44	1	0.5	3	
3	0.8	1	1.5	29	
4	1.16	3	2.5	42	
5	1.28	5	3.5	33	
					0.2

Chi square data (fair die)

	List 1	List 2	List 3	List 4	
1	0.32	0.32	1	0	
2	0.92	0.92	1	0.5	
3	1.4	1.4	1	1.5	
4	1.52	1.52	1	2.5	
5	1.76	1.76	1	3.5	
					0.32

	List 2	List 3	List 4	List 5	
1	0.32	1	0	2	
2	0.92	1	0.5	2	
3	1.4	1	1.5	3	
4	1.52	1	2.5	10	
5	1.76	1	3.5	15	
					0.32

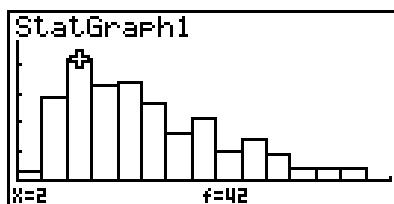
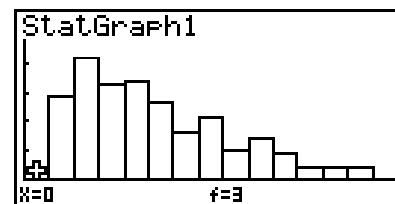
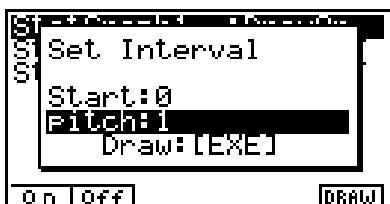
Chi square data (unfair die)

Now the definition and the histogram of the fair die chi square data follow:

	List 6	List 7	List 8	List 9	
1	0.2	0.2	1	0	
5: Set	44	1	0.5		
4: Select	8	1	1.5		
3: S-Gph3	16	3	2.5		
2: S-Gph2	28	5	3.5		
1: S-Gph1				0.2	

	StatGraph1	View Window
Graph Type	:Hist	Xmin :0
XList	:List7	max :15
Frequency	:List8	scale:1
		dot :0.11904761
i	[LIST]	Ymin : -8
		max :55
		INIT TRIG STD STO RCL

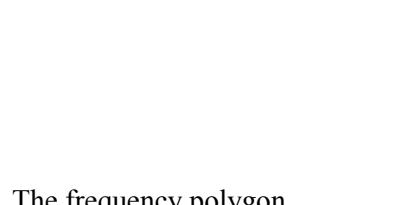
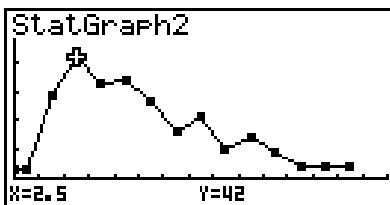
	StatGraph1	StatGraph2	StatGraph3
on/off		Set Interval	
		Start:0	
		Pitch:1	
		Draw:[EXE]	
DRAW			



	StatGraph1	StatGraph2	StatGraph3
Graph Type	:xyLine		
XList	:List9		
YList	:List10		
Frequency	:1		
Mark Type	:•		
	□ x ■		

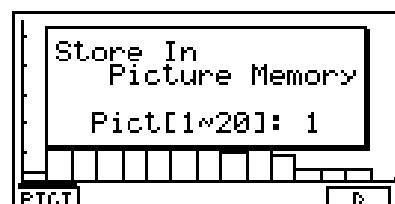
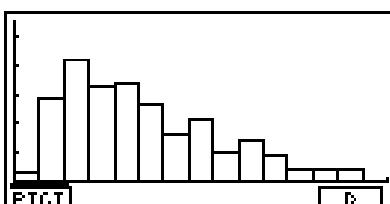
	List 1	List 2	List 3	List 4
1	0.32	0.32	1	0
5: Set	82	1	0.5	
4: Select	4	1	1.5	
3: S-Gph3	62	1	2.5	
2: S-Gph2	16	1	3.5	
1: S-Gph1				0.32

	StatGraph1	StatGraph2	StatGraph3
on/off			
		DrawOn	
		DrawOff	
		DrawOff	
DRAW			

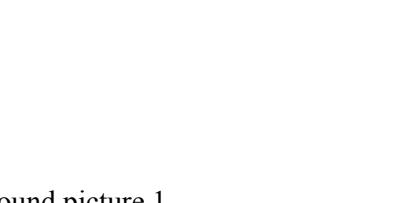
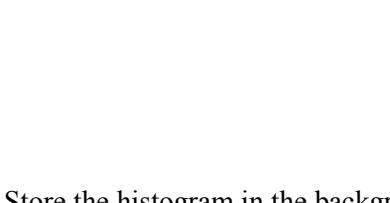


The frequency polygon

	StatGraph1	StatGraph2	StatGraph3
Graph Type	:Hist		
XList	:List7		
Frequency	:List8		
PICT			



	Stat_Wind	Resid_List	List_File	Func_Type	Graph_Func	Background	Hscale	NonePICT
	:Manual	:None	:File1	:y=	:On	:Pict1	:Rad	



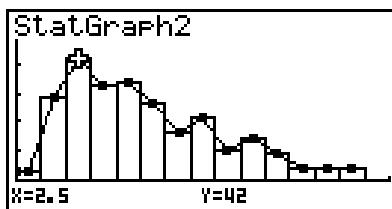
Store the histogram in the background picture 1

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

```
StatGraph1 :DrawOff
StatGraph2 :DrawOn
StatGraph3 :DrawOff
```

On Off [DRAW]

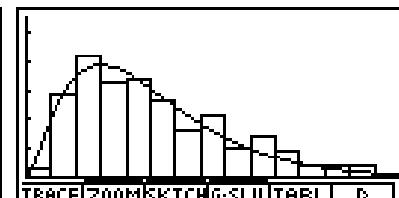


polygon and histogram together

In the GRPH-TBL menu we draw the chi square density function with the background picture 1:

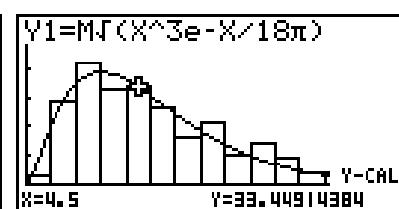
MAIN MENU
RUN-MAT STAT GRPH-TBL DYNA
RECUR CONICS EQUA PRGM
CAS ALGEBRA TUTOR TVM
Jdxh

Graph Func :Y=
Y1=MJ(X^3e-X/18π)
Y2:
Y3:
Y4:
Y5:
Y6:
SEL | DEL | TYPE | GMEM | DRAW |



t6:Y-Cal
5:Isect
4:Y-Icf
3:Min
2:Max
1:Root
TRACE | ZOOM | SKETCH | SOLVE | TABLE |

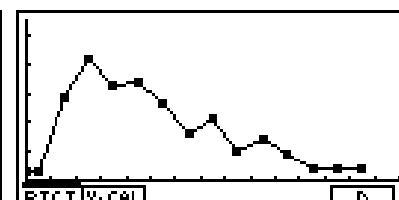
Enter X-Value
X:4.5



Now we store the background picture 2, the polygon:

MAIN MENU
RUN-MAT STAT GRPH-TBL DYNA
RECUR CONICS EQUA PRGM
CAS ALGEBRA TUTOR TVM
Jdxh

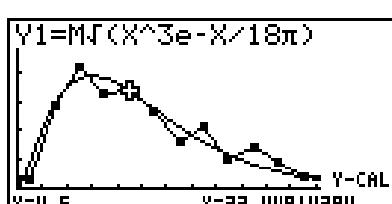
Stat Wind :Manual
Resid List :None
List File :File1
Func Type :Y=
Graph Func :On
Background :None
Hscale :Rad
None|PICT|



Store In Picture Memory
Pict[1~20]: 2
PICT|Y-CAL|

MAIN MENU
RUN-MAT STAT GRPH-TBL DYNA
RECUR CONICS EQUA PRGM
CAS ALGEBRA TUTOR TVM
Jdxh

Variable :Range
Draw Type :Connect
Graph Func :On
Dual Screen :Off
Simul Graph :Off
Derivative :Off
Background :Pict2
None|PICT|



The chi square density function together with the polygon (fair die)

3. Now some error conditions:

MAIN MENU
RUN-MAT STAT GRPH-TBL DYNA
RECUR CONICS EQUA PRGM
CAS ALGEBRA TUTOR TVM
Jdxh

	List 1	List 2	List 3	List 4
1:New	0.32	0.32	0	0
2:Set	02	02	0.5	0
4>Select	0.4	0.4	1.5	0
3:S-Gph3	62	62	2.5	0
2:S-Gph2	16	16	3.5	0
1:S-Gph1			0.32	0

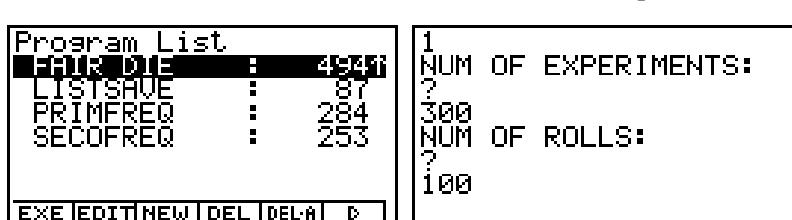
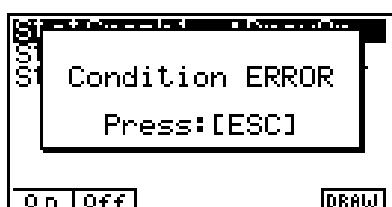
ABPCALC TEST INTR DISTI | D

```
StatGraph1 :DrawOn
StatGraph2 :DrawOn
StatGraph3 :DrawOff
```

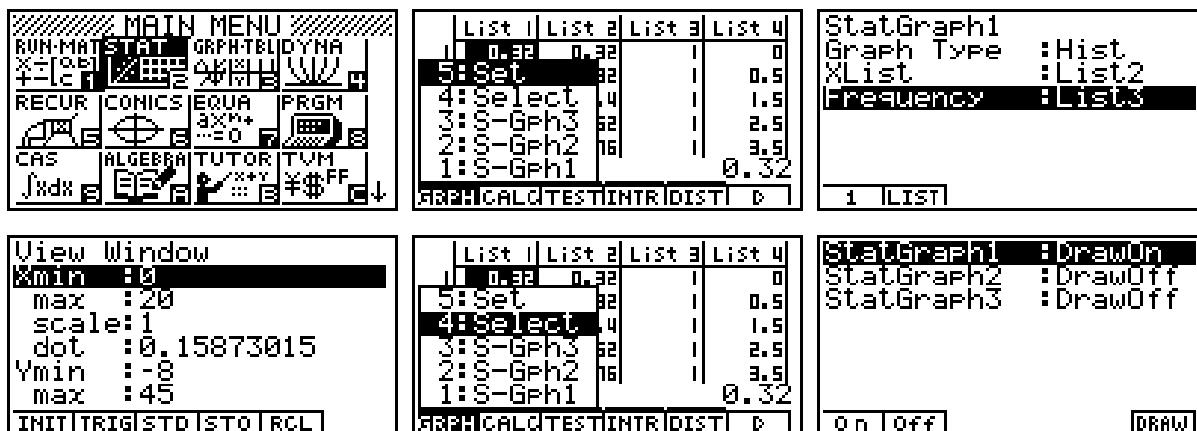
On Off [DRAW]

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

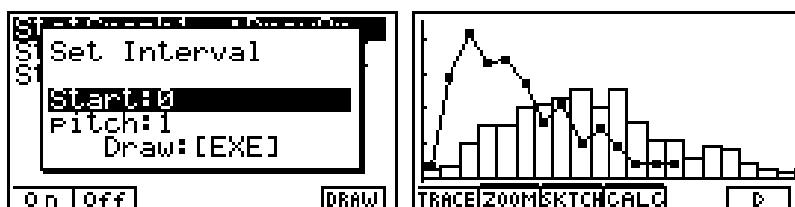
Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany



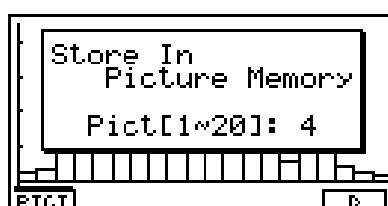
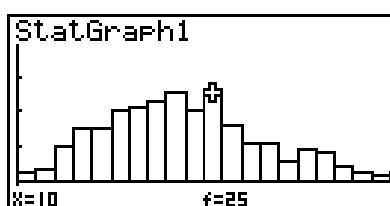
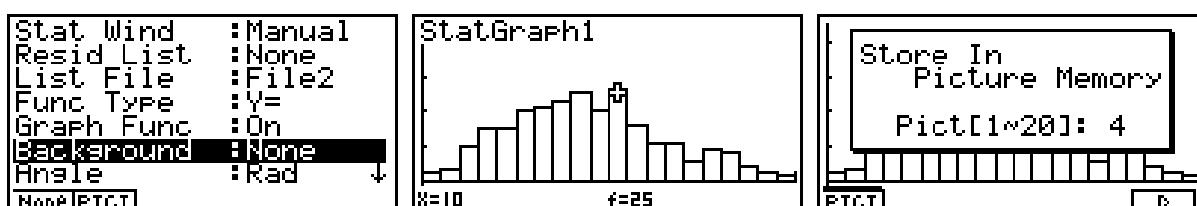
4. Screenshots of the chi square data of the unfair die:



The polygon (background picture 3 in the new view window, fair die) together with the histogram of the chi square data (unfair die):



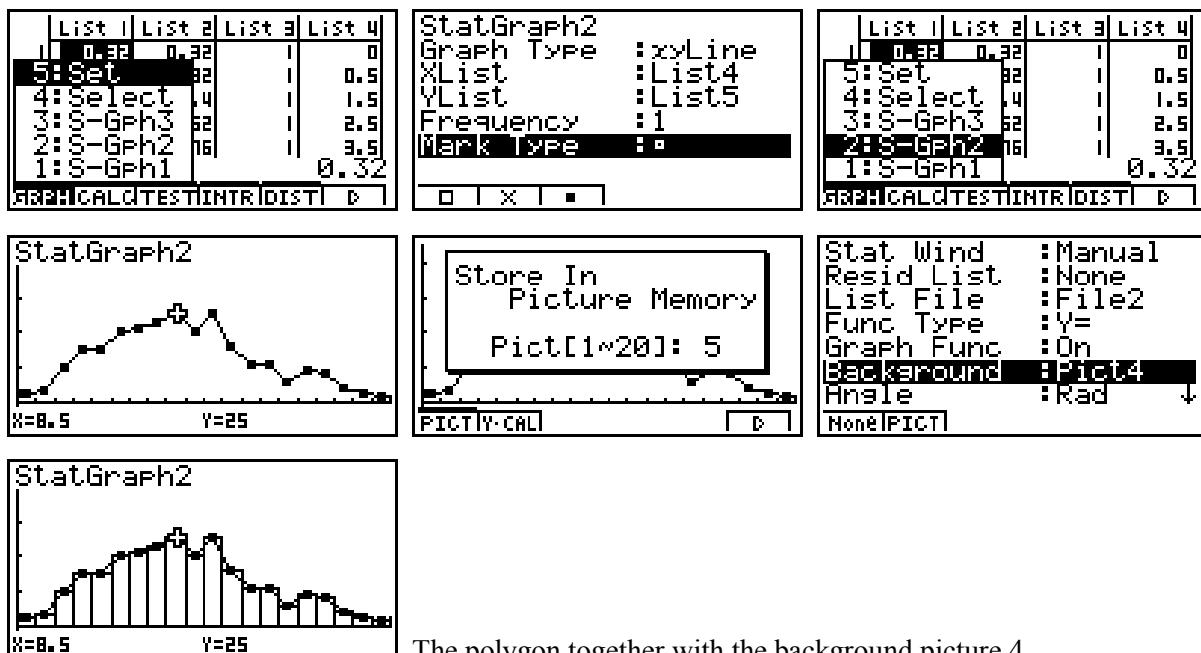
The histogram (unfair die) we store in picture 4:



Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

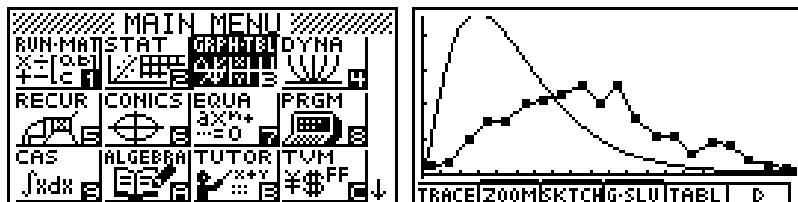
Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

The polygon (unfair die) we store in picture 5:



The polygon together with the background picture 4

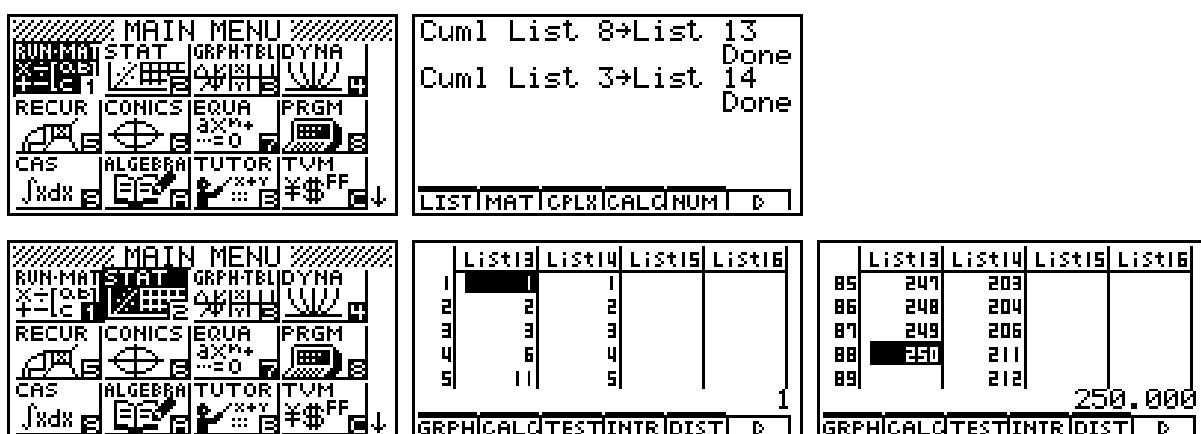
Finally we draw the chi square density function (fair die) and the polygon of the unfair die:



Now we can see, that the computed data of an unfair die are not chi square data!

5. Drawing the empirical distribution functions of the simulated data

To draw the empirical distribution functions of the simulated data we create the cumulative List 13 (fair die) and the cumulative List 14 (unfair die) respectively in the RUN-menu. In the STAT-menu we can observe these lists:



Next we create the table of values of the theoretical chi square distribution in the RUN-menu:

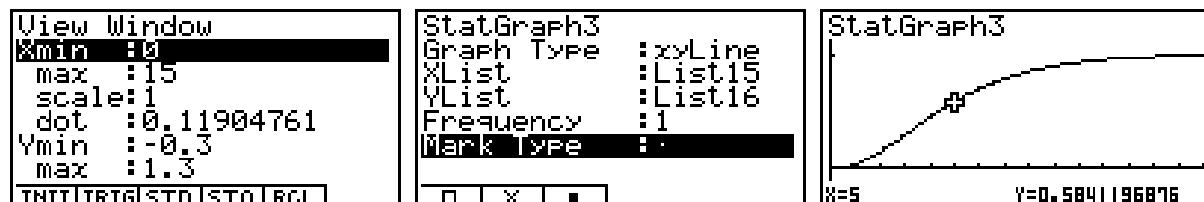
Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

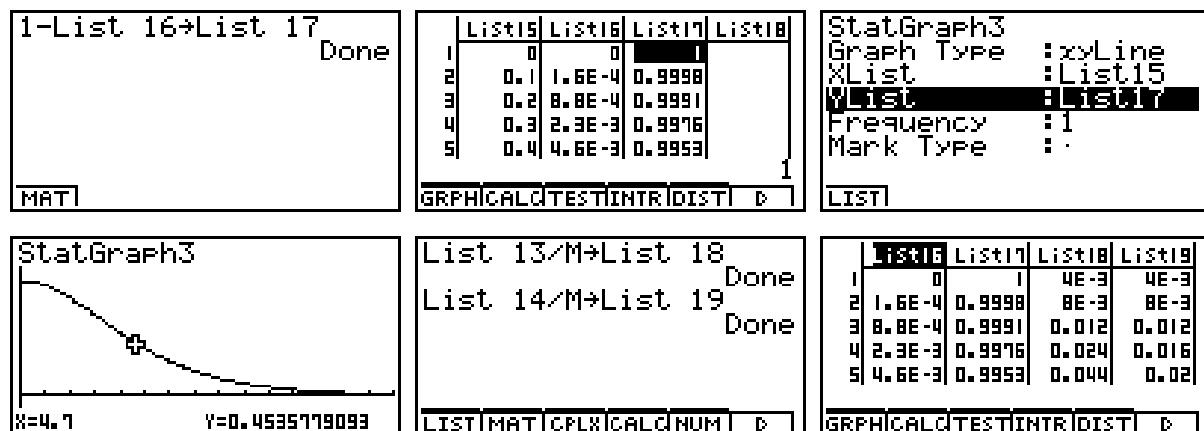
List 15 = { X | X = 0 (0.1) 15 }, List 16 = { Y | Y = $\int(X^{1.5} * e^{-X/2}) / \sqrt{18\pi}, 0, List 15, 0.001$ }

151→Dim List 15 Done	For 0→B To 15 Step 0. 1→ B→List_15[10×B+1]← $\int(X^{1.5} * e^{-X/2}) / \sqrt{18\pi}$, 0, B, .001)→List 16[1 0×B+1]← Next→ IF !FOR (WHILE CTRL LOGIC) B	1→ B→List_15[10×B+1]← $\int(X^{1.5} * e^{-X/2}) / \sqrt{18\pi}$, 0, B, .001)→List 16[1 0×B+1]← Next→ IF !FOR (WHILE CTRL LOGIC) B
-------------------------	--	---

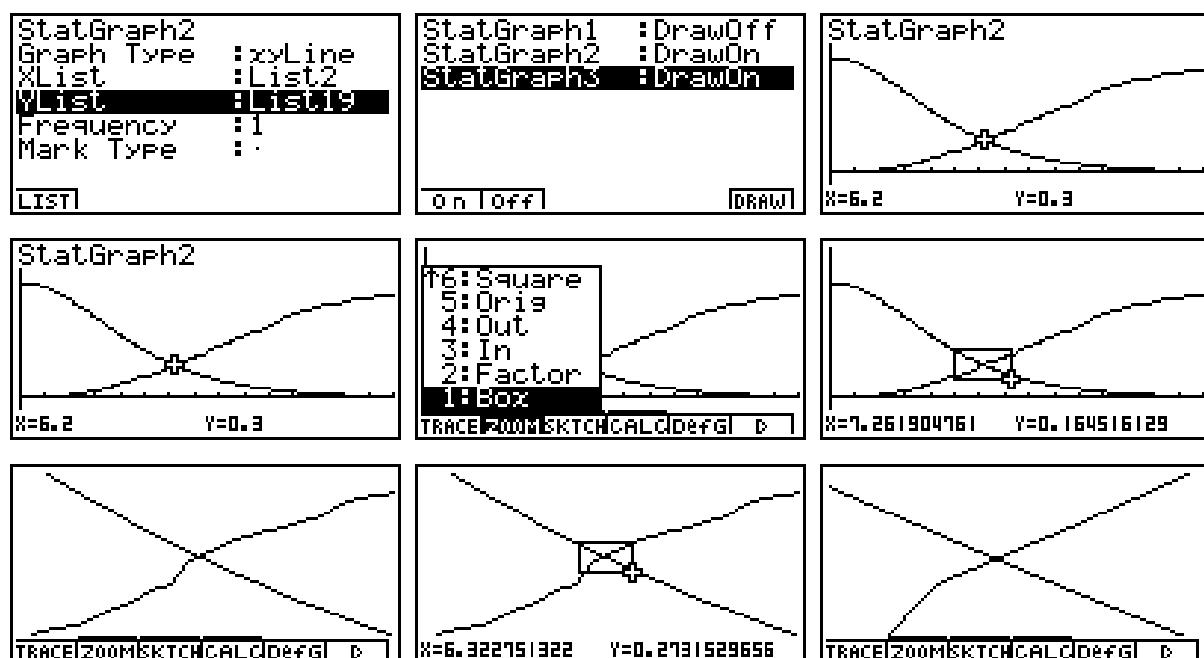
In the STAT-menu we draw the chi square distribution function (5 degrees of freedom):



In the RUN-menu we compute List 17 = 1 – List 16 to draw the function $Y = 1 - \text{distribution_function}$

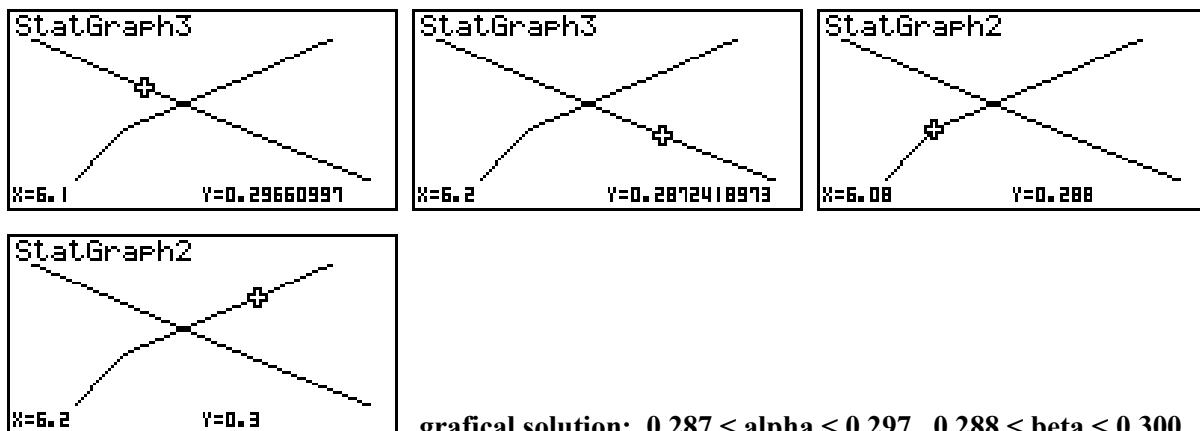


Finally we intersect the $Y = 1 - \text{distribution_function}$ (fair die) with the empirical distribution function (unfair die) to solve the equation $\alpha=\beta$:



Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

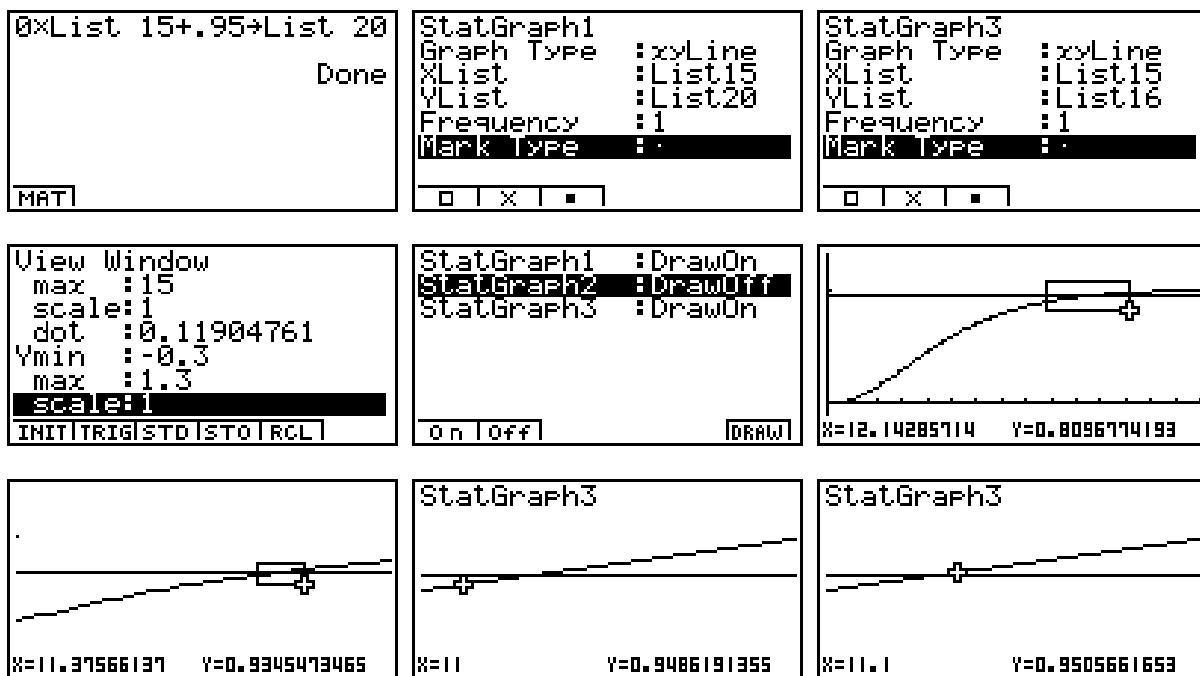


graphical solution: $0.287 < \alpha < 0.297$, $0.288 < \beta < 0.300$.

Thus, the solution of the equation $\alpha = \beta$ is approximately $\alpha = \beta = 0.29 = 29\%$ (say) and the quantile is approximately 6.15 (say), i.e. in our considered simulations 79 of 250 chi square data (fair die) were at least 6.15 ($\alpha = 79/250 = 31.6\%$, theoretically 29%) and 72 of 250 simulated data of the unfair die ($\beta = 72 / 250 = 29\%$) are smaller than 6.15.

If we decide for or against the fair die by the help of the critical value 6.15, then the probability of the error of second kind is not greater than the probability of the error of first kind.

To compute the quantile $C_{1-\alpha}$ for a given α we solve the problem by means of graphical methods in the STAT-menu: E.g. let be $\alpha = 0.05$ the given significance of a chi square goodness of fit test. We intersect the confidence $1-\alpha = 0.95$ (StatGraph1) with the chi square distribution function (StatGraph3): the interpolation gives $C_{1-\alpha} = 11 + 0.1 * 0.00138 / 0.00195 = 11.07$.



5. Table of the simulated data of $M = 250$ experiments, if in each experiment the die rolls 100 times.

The simulated data: A random sample of size 250 (the list of the chi square data, the raw data list), sort in ascending order (ListFile1 List6, fair die), the critical value let be 6.15 (blue ≤ 6.15 , red otherwise)

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

(Start of the random number generator after resetting: RUN-menu: Ran# 0 and using the generator Ran# 1)

0.20	2.60	3.80	5.36	7.76
0.44	2.60	3.92	5.36	7.88
0.80	2.72	3.92	5.36	7.88
1.16	2.72	3.92	5.48	7.88
1.16	2.72	3.92	5.48	7.88
1.16	2.72	3.92	5.60	8.00
1.28	2.72	3.92	5.60	8.12
1.28	2.84	4.04	5.60	8.24
1.28	2.84	4.04	5.60	8.36
1.28	2.84	4.04	5.60	8.36
1.28	2.84	4.04	5.72	8.36
1.40	2.84	4.16	5.72	8.48
1.40	2.84	4.16	5.84	8.60
1.40	2.84	4.16	5.84	8.60
1.40	2.84	4.16	5.96	8.84
1.40	2.84	4.16	5.96	9.08
1.40	2.84	4.28	5.96	9.20
1.52	2.84	4.28	5.96	9.32
1.52	2.84	4.28	6.08	9.32
1.52	2.96	4.28	6.08	9.32
1.52	2.96	4.28	6.08	9.44
1.64	2.96	4.28	6.20	9.44
1.64	2.96	4.28	6.20	9.56
1.64	2.96	4.28	6.20	9.56
1.64	3.08	4.40	6.32	9.80
1.76	3.08	4.40	6.32	9.92
1.76	3.20	4.40	6.44	9.92
1.76	3.20	4.40	6.44	9.92
1.76	3.32	4.40	6.44	9.92
1.88	3.32	4.52	6.56	10.04
1.88	3.32	4.52	6.56	10.28
1.88	3.32	4.52	6.56	10.40
2.00	3.32	4.52	6.80	10.52
2.00	3.44	4.64	6.80	10.52
2.12	3.44	4.64	7.04	10.64
2.12	3.44	4.64	7.04	10.76
2.12	3.44	4.64	7.04	10.76
2.12	3.56	4.64	7.04	10.76
2.24	3.56	4.76	7.16	11.24
2.24	3.68	4.88	7.28	11.72
2.24	3.68	4.88	7.40	11.84
2.24	3.68	5.00	7.40	11.84
2.24	3.68	5.12	7.52	12.08
2.36	3.68	5.12	7.52	12.44
2.36	3.68	5.12	7.64	12.44
2.48	3.68	5.12	7.64	12.56
2.48	3.68	5.12	7.76	13.04
2.60	3.80	5.24	7.76	13.16
2.60	3.80	5.24	7.76	13.28
2.60	3.80	5.36	7.76	13.88

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

Primary data and frequencies (ListFile1 List7 and ListFile1 List8, fair die), again the critical value let be 6.15 (blue \leq 6.15, red otherwise)

chi-square data	frequency	Cuml	chi-square date	frequency	Cuml
0.20	1	1	6.08	3	171
0.44	1	2	6.20	3	174
0.80	1	3	6.32	2	176
1.16	3	6	6.44	3	179
1.28	5	11	6.56	3	182
1.40	6	17	6.80	2	184
1.52	4	21	7.04	4	188
1.64	4	25	7.16	1	189
1.76	4	29	7.28	1	190
1.88	3	32	7.40	2	192
2.00	2	34	7.52	2	194
2.12	4	38	7.64	2	196
2.24	5	43	7.76	5	201
2.36	2	45	7.88	4	205
2.48	2	47	8.00	1	206
2.60	5	52	8.12	1	207
2.72	5	57	8.24	1	208
2.84	12	69	8.36	3	211
2.96	5	74	8.48	1	212
3.08	2	76	8.60	2	214
3.20	2	78	8.84	1	215
3.32	5	83	9.08	1	216
3.44	4	87	9.20	1	217
3.56	2	89	9.32	3	220
3.68	8	97	9.44	2	222
3.80	4	101	9.56	2	224
3.92	6	107	9.80	1	225
4.04	4	111	9.92	4	229
4.16	5	116	10.04	1	230
4.28	8	124	10.28	1	231
4.40	5	129	10.40	1	232
4.52	4	133	10.52	2	234
4.64	5	138	10.64	1	235
4.76	1	139	10.76	3	238
4.88	2	141	11.24	1	239
5.00	1	142	11.72	1	240
5.12	5	147	11.84	2	242
5.24	2	149	12.08	1	243
5.36	4	153	12.44	2	245
5.48	2	155	12.56	1	246
5.60	5	160	13.04	1	247
5.72	2	162	13.16	1	248
5.84	2	164	13.28	1	249
5.96	4	168	13.88	1	250

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

Secundary (grouped) data and frequencies

(ListFile1 List9 and ListFile1 List10, fair die, and ListFile1 List4 and ListFile1 List5, unfair die)

chi-square data (class limits $k \leq X < k+1$ class mark $k + 0.5$)	frequency (fair die)		chi-square data (class limits $k \leq X < k+1$ class mark $k + 0.5$)	frequency (unfair die)
(0)	(3)		(0)	(2)
0.5	3		0.5	2
1.5	29		1.5	3
2.5	42		2.5	10
3.5	33		3.5	15
4.5	34		4.5	15
5.5	27		5.5	20
6.5	16		6.5	21
7.5	21		7.5	23
8.5	10		8.5	25
9.5	14		9.5	20
10.5	9		10.5	25
11.5	4		11.5	16
12.5	4		12.5	11
13.5	4		13.5	11
			14.5	6
			15.5	9
			16.5	8
			17.5	4
			18.5	2
			19.5	1
			20.5	2
			21.5	1

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

A random sample of size 250 (the list of the chi square data, the raw data list), sort in ascending order
 (ListFile1 List1, **unfair** die), the critical value let be 6.15 (red \leq 6.15, blue otherwise)

0.32	5.36	7.64	9.92	12.32
0.92	5.48	7.64	9.92	12.44
1.40	5.48	7.64	9.92	12.68
1.52	5.48	7.76	9.92	12.80
1.76	5.60	7.76	10.04	12.92
2.00	5.60	7.76	10.04	12.92
2.00	5.60	7.76	10.16	13.04
2.36	5.72	7.88	10.16	13.04
2.36	5.72	7.88	10.16	13.04
2.48	5.84	8.00	10.28	13.04
2.60	5.84	8.12	10.28	13.04
2.72	5.84	8.12	10.28	13.16
2.72	5.84	8.12	10.40	13.28
2.72	5.96	8.12	10.40	13.28
2.84	5.96	8.12	10.40	13.28
3.08	6.08	8.24	10.52	13.64
3.08	6.08	8.36	10.52	13.88
3.20	6.08	8.36	10.52	14.00
3.32	6.08	8.36	10.52	14.24
3.44	6.08	8.48	10.52	14.24
3.56	6.08	8.48	10.52	14.36
3.68	6.08	8.48	10.64	14.36
3.68	6.20	8.48	10.64	14.96
3.68	6.20	8.60	10.64	15.08
3.68	6.20	8.60	10.64	15.08
3.68	6.32	8.60	10.76	15.20
3.68	6.32	8.60	10.76	15.32
3.80	6.32	8.72	10.88	15.32
3.92	6.56	8.72	10.88	15.44
3.92	6.56	8.72	11.00	15.68
4.04	6.56	8.72	11.12	15.68
4.04	6.68	8.84	11.24	15.80
4.16	6.80	8.96	11.24	16.04
4.16	6.80	8.96	11.24	16.04
4.28	6.92	9.08	11.24	16.04
4.52	6.92	9.08	11.36	16.04
4.52	7.04	9.20	11.36	16.04
4.52	7.04	9.20	11.36	16.28
4.52	7.04	9.32	11.48	16.76
4.64	7.04	9.32	11.48	16.76
4.64	7.16	9.32	11.60	17.24
4.76	7.16	9.32	11.84	17.48
4.76	7.28	9.32	11.84	17.48
4.76	7.40	9.32	11.96	17.60
4.76	7.52	9.44	11.96	18.56
5.00	7.52	9.44	12.08	18.56
5.12	7.52	9.68	12.08	19.40
5.12	7.52	9.80	12.20	20.00
5.12	7.52	9.80	12.20	20.72
5.12	7.52	9.80	12.20	21.20

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

Primary data and frequencies (ListFile1 List2 and ListFile1 List3, unfair die), again the critical value let be 6.15 (red \leq 6.15, blue otherwise)

chi-square variable	frequency	Cuml	chi-square variable	Frequency	Cuml
0.32	1	1	9.08	2	136
0.92	1	2	9.20	2	138
1.40	1	3	9.32	6	144
1.52	1	4	9.44	2	146
1.76	1	5	9.68	1	147
2.00	2	7	9.80	3	150
2.36	2	9	9.92	4	154
2.48	1	10	10.04	2	156
2.60	1	11	10.16	3	159
2.72	3	14	10.28	3	162
2.84	1	15	10.40	3	165
3.08	2	17	10.52	6	171
3.20	1	18	10.64	4	175
3.32	1	19	10.76	2	177
3.44	1	20	10.88	2	179
3.56	1	21	11.00	1	180
3.68	6	27	11.12	1	181
3.80	1	28	11.24	4	185
3.92	2	30	11.36	3	188
4.04	2	32	11.48	2	190
4.16	2	34	11.60	1	191
4.28	1	35	11.84	2	193
4.52	4	39	11.96	2	195
4.64	2	41	12.08	2	197
4.76	4	45	12.20	3	200
5.00	1	46	12.32	1	201
5.12	4	50	12.44	1	202
5.36	1	51	12.68	1	203
5.48	3	54	12.80	1	204
5.60	3	57	12.92	2	206
5.72	2	59	13.04	5	211
5.84	4	63	13.16	1	212
5.96	2	65	13.28	3	215
6.08	7	72	13.64	1	216
6.20	3	75	13.88	1	217
6.32	3	78	14.00	1	218
6.56	3	81	14.24	2	220
6.68	1	82	14.36	2	222
6.80	2	84	14.96	1	223
6.92	2	86	15.08	2	225
7.04	4	90	15.20	1	226
7.16	2	92	15.32	2	228
7.28	1	93	15.44	1	229
7.40	1	94	15.68	2	231
7.52	6	100	15.80	1	232
7.64	3	103	16.04	5	237
7.76	4	107	16.28	1	238
7.88	2	109	16.76	2	240
8.00	1	110	17.24	1	241
8.12	5	115	17.48	2	243
8.24	1	116	17.60	1	244
8.36	3	119	18.56	2	246
8.48	4	123	19.40	1	247
8.60	4	127	20.00	1	248
8.72	4	131	20.72	1	249
8.84	1	132	21.20	1	250
8.96	2	134			

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

6. Program List FAIR DIE

```
Filename:FAIR DIE
ClrText:ClrGraph
ClrList1:ClrList2
ClrList3:ClrList4
ClrList5
"FAIR DIE ?"↓
"1 FOR YES"↓
"0 FOR NO"↓
?→A↓
"NUM OF EXPERIMENTS:"↓
?→M↓
"NUM OF ROLLS:"↓
?→N↓
N÷6→E↓
M→Dim List 1↓
6→Dim List 2↓
0→K↓
ClrText↓
Lbl M↓
K+1→K↓
0×List 2→List 2↓
If A=1↓
Then For 1→I To N Step 1↓
Int (6×Ran#1+1)→D↓
List 2[D]+1→List 2[D]↓
Next↓
Else For 1→I To N Step 1↓
Int (5.5×Ran#1)+1→D↓
List 2[D]+1→List 2[D]↓
Next↓
IfEnd↓
Sum (((List 2-E)^2)÷E)→C↓
Rnd:Ans→C↓
Locate 1,2,"EXP. NO:"↓
Locate 10,3,K↓
Locate 1,4,"CHI^2="↓
Locate 10,5,C↓
C→List 1[K]↓
If K<M↓
Then Goto M↓
IfEnd↓
ClrText↓
Locate 1,2,"NUM OF EXPERIMENTS:"↓
Locate 10,3,M↓
Locate 1,4,"NUM OF ROLLS:"↓
Locate 10,5,N↓
Locate 1,6,"CODE FAIRDIE:"↓
Locate 10,7,A↓
SortA(List 1)↓
Stop↓
```

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

Program List **PRIMFREQ**

```
Filename: PRIMFREQ
ClrText↔
"A MOMENT PLEASE, BUSY ..."↔
0→L↔
M→Dim List 2↔
M→Dim List 3↔
Lbl S↔
List 1[1]→List 2[1]↔
1→List 3[1]↔
1→K↔
1→I↔
Lbl P↔
K+1→K↔
If List 1[K]=List 1[K-1]↔
Then Goto Q↔
IfEnd↔
I+1→I↔
List 1[K]→List 2[I]↔
1→List 3[I]↔
Goto R↔
Lbl Q↔
List 3[I]+1→List 3[I]↔
Lbl R↔
If K<M↔
Then Goto P↔
IfEnd↔
If L=1↔
Then Goto T↔
IfEnd↔
I→Dim List 2↔
I→Dim List 3↔
1→L↔
Goto S↔
Lbl T↔
ClrText↔
Locate 1,2,"PRIMFREQ FINISHED"↔
Stop↔
```

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

Program List SECOFREQ

```
Filename: SECOFREQ
"A MOMENT PLEASE, BUSY ..."↓
Dim List 2→A↓
Int (List 2[A]+2)→Dim List 4↓
Int (List 2[A]+2)→Dim List 5↓
Int (List 2[A]+1)→B↓
Seq(-.5+X,X,0,B,1)→List 4↓
0→List 4[1]↓
0×List 4→List 5↓
0→K↓
Lbl Q↓
K+1→K↓
Int (List 1[K])+2→L↓
List 5[L]+1→List 5[L]↓
If K<M↓
Then Goto Q↓
IfEnd↓
List 5[2]→List 5[1]↓
ClrText↓
Locate 1,2,"SECOFREQ FINISHED"↓
Stop↓
```

Program List LISTSAVE

```
Filename: LISTSAVE
"LIST TRANSLATION"↓
List 1→List 6↓
List 2→List 7↓
List 3→List 8↓
List 4→List 9↓
List 5→List 10↓
Stop↓
```

Remember:

Unfair die: List1, List2, List3, List4, List5, List 14 (Cum List 3), List 19 (List14/M)

Fair die: List6, List7, List8, List9, List10, List 13 (Cum List 8), List 18 (List13/M)

Example: List11, List12 (one experiment with 100 rolls)

Chi square distribution function: List15, List16, List17 (= 1-List16)

Confidence niveau 1-alpha: List 15 and List 20

Simulation and Statistical Exploration of Data (e.g. Fair Die or Unfair Die)

Ludwig Paditz, University of Applied Sciences Dresden (FH), Germany

Program files you get by the help of the CASIO Program-Link FA-123 (Software) here:

<http://www.informatik.htw-dresden.de/~paditz/FAIRDIE1.CAT>

7. References:

[1] Aulenbacher, Paditz, Wabel-Frenk: **Lehr- und Übungsbuch Mathematik Bd.3**
(Teil Stochastik: Beispiel 6.1 und Aufgabe 19.1),
Fachbuchverlag Leipzig 2001 (2.Aufl.) (ISBN 3-446-21682-0)

Internet:

<http://www.kno.de/kod-bin/isuche.exe?dbname=Buchkatalog&lang=deutsch&field1=SB&name1=3-446-21682-0&Aktion=Suchen>

[2] Edwards, C.C.: *Does a TI-8x Cast a Fair Die?*
in: **Eightysomething!** (<http://www.comcal.net/www.ti.com/calc/docs/80xthing.htm>), Vol. 6, No.3, 1997, p. 9-10.

Internet: <http://www.comcal.net/www.ti.com/calc/docs/act83stat.htm> and
<http://www.blackgold.ab.ca/curric/math/pdf/s9780som.pdf> (pdf-File) and
<http://www.comcal.net/www.ti.com/calc/docs/act/pdf/s9780so3.pdf> (pdf-file)

[3] Paditz, L.: *Der gezinkte Würfel – Workshop zu statistischen Datensimulationen und Untersuchungen zur Testgröße und zur Testentscheidung beim Test auf Gleichverteilung (Chancengleichheit aller Augenzahlen)*

in: **Praktische Anwendungsbeispiele zur Schulmathematik mit Graphiktaschenrechnern**
Ein Sammelband mathematischer Einzelbeiträge zum Schulunterricht mit dem CFX-9850GB Plus, Hrg. v. CASIO Computer Co. GmbH Deutschland, Norderstedt 2001 (1.Aufl.), S. 66–82

Internet: <http://www.informatik.htw-dresden.de/~paditz/images/Casio2.jpg> and
<http://www.informatik.htw-dresden.de/~paditz/images/Casio21.jpg>

[4] Shawer, Mayer Y.: *Understanding statistics through probability*
in: **International conference on mathematics education into the 21st century: New ideas in mathematics education.** Proceedings (Ed. Rogerson, Alan), Palm Cove 19-24 August 2001 (Australia), p. 229-233

Internet: <http://math.unipa.it/~grim/cairms> and <http://math.unipa.it/~grim/AShawer1> (pdf-file)

Note: The addresses of the internet sites mentioned in this paper were accessed in May of 2002

Contact: paditz@informatik.htw-dresden.de

(Update: 16 May 2002)