

# Wahrscheinlichkeitsrechnung und Statistik

Otto Wurnig

Der folgende Kommentar stellt eine Zusammenfassung der Erfahrungen dar, die in der Oberstufe der AHS mit dem Einsatz des TI-92 in den genannten Stoffgebieten gemacht wurden.

Da der Einsatz des TI-92 in der Unterstufe noch nicht allgemein akzeptiert wird, soll kurz erläutert werden, an welchen im Lehrplan der 5. und 6. Klasse vorgesehenen Kapiteln wichtige TI-92 Befehle erarbeitet werden, die dann in der 7. und 8. Klasse bei der Behandlung der Wahrscheinlichkeitsrechnung und Statistik zur Verfügung stehen.

In der **5. Klasse AHS** ist das Kapitel „*Darstellen und Analysieren von Daten- und Beziehungsstrukturen*“ vorgesehen.

R. FISCHER schreibt 1991 im offiziellen Lehrplankommentar (S.78f), dass es dabei im ersten Bereich um die klassische **beschreibende Statistik** gehe.

Es geht um die Bearbeitung der **Werte einer Variablen**:

- um *graphische Darstellungen* wie die Angabe von Häufigkeitsverteilungen in Tabellen, Histogrammen, Kreisdiagrammen oder Kastenschaubildern,
- um die Beschreibung von Datenmengen durch *Kennzahlen*, wie das arithmetische Mittel, Median, Modus sowie Spannweite, Quartile, mittlere lineare Abweichung oder empirische Standardabweichung.

Es geht um die Bearbeitung der **Werte zweier Variablen**, es geht oft darum, einen Zusammenhang zwischen diesen Variablen herzustellen, d.h.:

- eine Darstellung in einem zweidimensionalen Koordinatensystem vorzunehmen und
- gegebenenfalls eine *Trendkurve* für diese Punktwolke zu finden.

Für die *Bearbeitung solcher Listen* stellt der TI-92 eigene Befehle zur Verfügung, die *im HOME-Fenster* oder professioneller **im Daten-Editor** verwendet werden.

(vgl. REICHEL-MÜLLER 1997, S 54ff)

In der **6. Klasse AHS** werden im Kapitel *Funktionen - Analysis* viele Grundbegriffe erarbeitet, die gerade mit dem TI-92 gut behandelt und dargestellt werden können.

Große Bedeutung hat

- die *Darstellung und Behandlung von reellen Funktionen* im **Home-, Tabellen- und Graphikfenster**,
- die *Darstellung und Behandlung von Zahlenfolgen* als besondere Funktionen im *Home-Fenster* und im **SEQUENCE-Mode** im *Tabellen- und Graphik-Fenster*.

Für die Darstellung von Zahlenfolgen steht im Home-Fenster ein eigener **SEQUENCE-Befehl** zur Verfügung:

Die Zahlenfolge {1, 4, 9, 16} kann man durch den Befehl **seq( $n^2$ , n, 1, 4)** erhalten werden.

# Stochastik mit dem TI-92

## Übersicht

- 1) Zufallsgenerator - Simulation des Würfels
- 2) Berechnung der Binomialkoeffizienten durch Simulation des Münzwurfs
- 3) Simulation der Binomialverteilung
- 4) Graphische Darstellung der Binomialverteilung
- 5) Ableitung von  $\mu$  und  $\sigma$  einer Binomialverteilung bei gegebenem  $n$
- 6) Eine Musteraufgabe zur Binomialverteilung
- 7) Die Poissonverteilung als Sonderfall der Binomialverteilung
- 8) Darstellung der Binomialverteilung als Histogramm und Wahrscheinlichkeitsdichtepolygon
- 9) Der Zugang zur Normalverteilung über die Binomialverteilung
- 10) Diskussion der GAUSSschen Glockenkurve
- 11) Eine Aufgabe – Vier Lösungsvarianten
- 12) Der Zugang zur Normalverteilung über die Augensumme dreier Würfel
- 13) Eine Musteraufgabe zur Normalverteilung
- 14) Beispiele zur Normalverteilung (zusammengestellt von Alfred Eisler)

Literatur

# 1. Der TI-92 Zufallsgenerator

Um zum Modell der Laplace-Münze (L-Münze) und des L-Würfels zu kommen, sind zum besseren Verständnis zuerst eigene Versuche und anschließend Simulationen von großer Bedeutung. Dazu steht beim TI-92 ein eigener Zufallsgenerator zur Verfügung.

**TI-92 Befehle:** (TI-92 Handbuch 1996, S.432)

<b>RandSeed</b>	<i>Zahl</i>
<i>Zahl = 0</i>	setzt die Ausgangsbasis („seed“) für den Zufallsgenerator auf die Werkseinstellung zurück.
<i>Zahl ≠ 0</i>	erzeugt 2 Basen, die in den Systemzahlen seed1 u. seed2 gespeichert werden.
<b>rand(n)</b>	n ist eine ganze Zahl ≠ 0
rand()	gibt die nächste Zufallszahl aus einer Zufallszahlenfolge zwischen 0 und 1 aus.
rand(6)	gibt eine ganzzahlige Zufallszahl $1 \leq x \leq 6$ aus.
rand(-6)	gibt eine ganzzahlige Zufallszahl $-6 \leq x \leq -1$ aus.

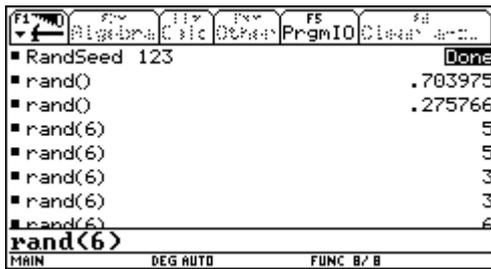


Fig. 1

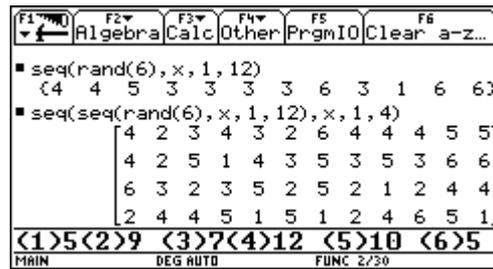


Fig. 2

Es kann somit problemlos im HOME-Fenster „gewürfelt“ werden (Fig. 1). Für eine Folge von Zufallszahlen wird der TI-Befehl SEQ herangezogen. Da aber im Fenster nur wenige Zufallszahlen sichtbar sind, schlagen REICHEL/MÜLLER (S.159f) zwei verschachtelte SEQ-Befehle vor (Fig. 2). Durch Abzählen kann festgestellt werden, dass die Augenzahl 5 10mal vorkommt.

Da aber beim L-Würfel alle sechs Augenzahlen mit gleicher Wahrscheinlichkeit vorkommen und dies bei den 48 Würfeln des TI-Würfels in Fig. 2 auch nicht angenähert der Fall ist, sollte die Qualität des TI-Zufallsgenerators mit einer größeren Folge von Zufallszahlen (z.B. 960) getestet werden. Dafür ist die Verwendung des Matrix-Editors bestens geeignet.

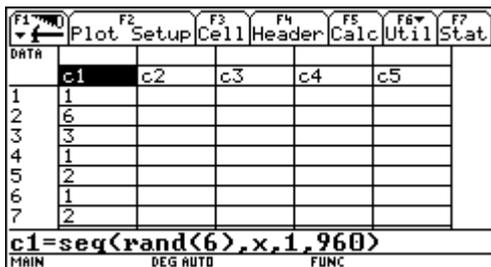


Fig. 3

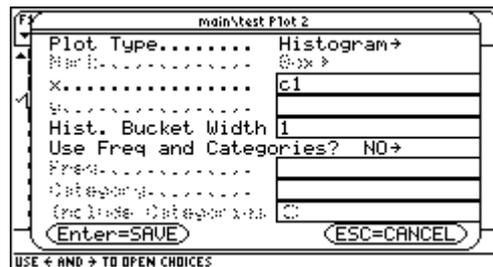


Fig. 4

Um die absoluten Häufigkeiten der Augenzahlen des TI-Würfels ohne eigenes Abzählen feststellen zu können, kann man zu Spalte c1 ein Histogramm (Fig. 6) herstellen lassen.

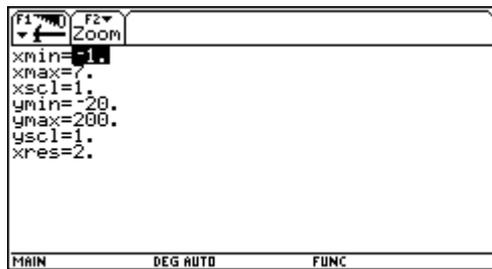


Fig. 5

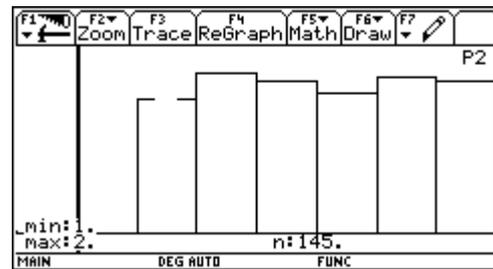


Fig. 6

Der TI kann im TRACE-Modus selbst die absoluten Häufigkeiten der Augenzahlen feststellen. In Fig. 6 kommt die Augenzahl 1 bei 960 Würfeln 145mal vor. Die relative Häufigkeit für die Augenzahl 1 ist somit  $145/960$  oder in gekürzter Form  $29/192$ . Gibt man der Reihe nach die absoluten Häufigkeiten für 1,2,3,4,5,6 mit der Hand im Daten-Editor ein, so erhält man Fig. 7.

	c1	c2	c3
1	1	29/192	
2	2	57/320	
3	3	163/960	
4	4	151/960	
5	5	167/960	
6	6	163/960	
7			

c3=

Fig. 7

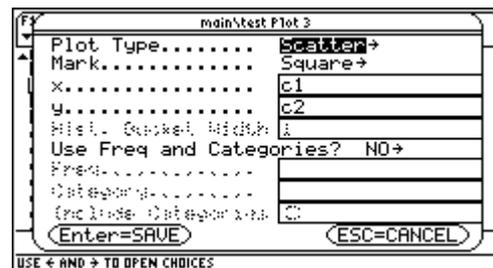


Fig. 8

Die sechs Zahlenpaare  $[1, 29/192], \dots, [6, 163/960]$  lassen sich mit Hilfe des Darstellungstyps SCATTER leicht im Graphik-Fenster als Punkte (Fig. 10) darstellen. Abschließend kann durch Einzeichnen der Geraden  $y = 1/6$  anschaulich überprüft werden, ob die sechs Punkte nahe bei dieser Geraden liegen.

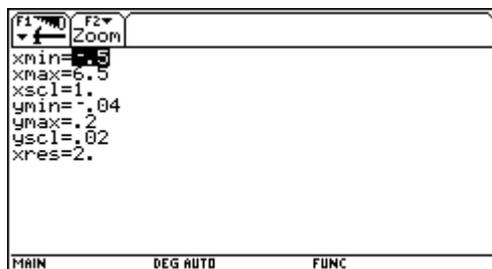


Fig. 9

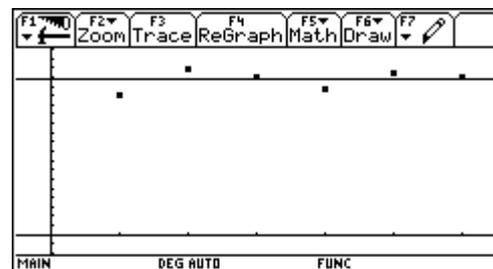


Fig. 10

Da im dargestellten simulierten Experiment alle sechs Punkte nahe bei der Geraden  $y = 1/6$  liegen, hat der TI-Würfel den L-Würfel, d.h. eine Gleichverteilung, gut simuliert. A. ENGEL (1973, S. 111f) nennt einen Würfel dann „gut“, wenn der  $\chi^2$ -Anpassungstest einen Wert kleiner als 8,16 ergibt. Im dargestellten Fall liefert der  $\chi^2$ -Anpassungstest den sehr guten Wert 3,09.

## 2. Berechnung der Binomialkoeffizienten durch Simulation des Münzwurfs

Aufgaben zur Berechnung von Wahrscheinlichkeiten waren früher vielfach mit Kombinatorik verknüpft. Der Lehrplan 1989 sieht dieses Kapitel aber nicht mehr ausdrücklich vor. Daher ist der Vorschlag von R. FRASER (GB), die Binomialkoeffizienten durch ein stochastisches Experiment zu erarbeiten, von besonderer Aktualität. Es werden  $n$  Münzen (z.B. 1, 2, 3, 4) sehr oft aufgeworfen (günstig ist z.B. 64, 128, 256 mal) und dann ausgewertet, wie oft **Zahl** (1) und **Wappen** (0) vorkommt. Der Münzwurf läßt sich sehr schön mit dem TI-92 durch den Befehl **RAND(2)-1** simulieren und kann sofort im Daten-Editor ausgeführt werden.

Im anschließenden Beispiel wurde in den Spalten  $c_1, c_2, c_3, c_4$  je 512mal der Münzwurf simuliert: **seq(rand(2)-1,x,1,512)** Anschließend wurde in  $c_5, c_6, c_7$  der Reihe nach  $c_1+c_2, c_1+c_2+c_3, c_1+c_2+c_3+c_4$  berechnet und dadurch beim Wurf mit zwei, drei, vier Münzen festgestellt, ob Zahl jeweils 0, 1, 2, 3, 4 mal vorkommt.

Um sich das direkte Abzählen der Häufigkeiten zu ersparen, kann man mit dem TI-92 das Histogramm bilden und im Graphik-Fenster die absoluten Häufigkeiten im TRACE-Mode ablesen.

	c1	c2	c3	c4	c5	c6	c7
1	1	0	1	1	1	2	3
2	0	1	1	0	1	2	2
3	0	0	1	0	0	1	1
4	0	1	0	0	1	1	1
5	1	1	0	0	2	2	2
6	1	1	1	0	2	3	3
7	1	0	0	0	1	1	1

**c7=c1+c2+c3+c4**

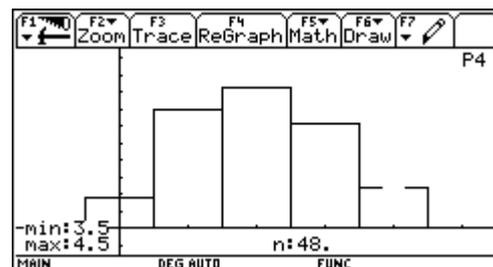
```

main/muenze Plot 4
Plot Type..... Histogram->
Mark:.....
X:..... c7
Y:.....
Hist. Bucket Width 1
Use Freq and Categories? NO->
Freq:.....
Category:.....
(Include Categories) C:
(Enter=SAVE) (ESC=CANCEL)
    
```

$c_1, c_2, c_3, c_4$  ... Wurf mit je einer Münze;  $c_5=c_1+c_2, c_6=c_1+c_2+c_3, c_7=c_1+c_2+c_3+c_4$

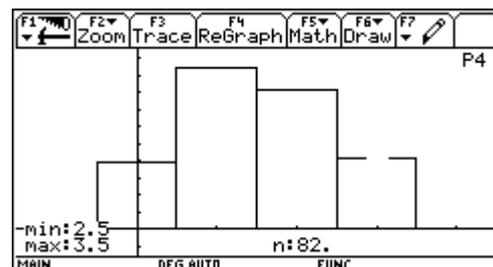
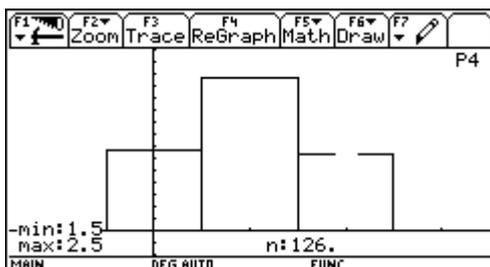
```

F1 Zoom
xmin=1.5
xmax=5.5
xsc1=1.
ymin=-30.
ymax=210.
ysc1=20.
xres=2.
    
```



Fenstereinstellung für  $n=4$

Histogramm für  $n=4, k=0, 1, 2, 3, 4$



Histogramm für  $n=2, k=0, 1, 2$

Histogramm für  $n=3, k=0, 1, 2, 3$

### Auswertung der Histogramme

In der Auswertung der Histogramme soll das jeweilige Verhältnis der Säulen (der absoluten Häufigkeiten) zueinander festgestellt werden. Dieses Verhältnis errechnet man dadurch, dass die Anzahl der Versuche (also 512) für  $n=1,2,3,4$  der Reihe nach durch 256,128,64,32 dividiert werden. Als Ergebnis erhält man die ersten Zeilen des Pascalschen Dreieckes.

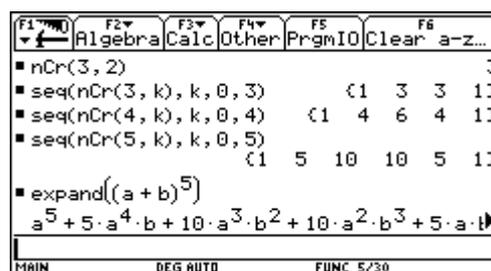
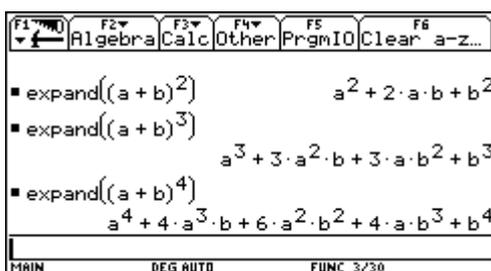
<b>n = 1</b>			
k	0	1	$\Sigma$
absolute Häufigkeit	269	243	512
Division	269 : 256	243 : 256	512 : 256
(gerundetes) Ergebnis	<b>1</b>	<b>1</b>	<b>2</b>

<b>n = 2</b>				
k	0	1	2	$\Sigma$
absol. Häufigkeit	133	253	126	512
Division	133 : 128	253 : 128	126 : 128	512 : 128
(gerund.) Ergeb.	<b>1</b>	<b>2</b>	<b>1</b>	<b>4</b>

<b>n = 3</b>					
k	0	1	2	3	$\Sigma$
abs. Häufig.	79	189	162	82	512
Division	79 : 64	189 : 64	162 : 64	82 : 64	512 : 64
(ger.) Ergeb.	<b>1</b>	<b>3</b>	<b>3</b>	<b>1</b>	<b>8</b>

<b>n = 4</b>						
k	0	1	2			$\Sigma$
abs. Häufig.	36	140	166	122	48	512
Division	36 : 32	140 : 32	166 : 32	122 : 32	48 : 32	512 : 32
(ger.) Erg.	<b>1</b>	<b>4</b>	<b>6</b>	<b>4</b>	<b>1</b>	<b>16</b>

Diese erhaltenen Zahlen heißen **Binomialkoeffizienten**, da sie beim Potenzieren des Binoms  $(a+b)^n$  mit  $n \in \mathbb{N}$ , auftreten. Sie können mit dem TI-92 mit  $nCr(n,k)$  berechnet werden.



### 3. Simulation der Binomialverteilung

REICHEL/MÜLLER (1997, S.159f) simulieren den Reißnagelwurf ( $P(\text{„Spitze unten“}) = 0,7$  und  $P(\text{„Spitze oben“}) = 0,3$ ) mit dem SEQ-Befehl in Verbindung mit einer WENN-DANN-SONST - Abfrage (Fig.1). Zum Erzeugen einer Versuchsreihe verwenden sie zwei verschachtelte SEQ-Befehle (Fig.2).

```

F1 Algebra F2 Calc F3 Other F4 PrgmIO F5 Clear a-z... F6
▀ seq(when(rand(10)<=7,1,0),i,1,12)
  1
  0,else
  seq(1,rand(10)<=7,i,1,12)
  0,else
  1 1 1 1 1 1 1 1 1 1 0
*seq(when(rand(10)<=7,1,0),i,1,12)
MAIN DEG AUTO FUNC 2/30
  
```

Fig. 1

```

F1 Algebra F2 Calc F3 Other F4 PrgmIO F5 Clear a-z... F6
▀ seq(seq(1,rand(10)<=7,i,1,12),j,1,5)
  0 1 1 1 1 0 1 1 1 1 1
  1 0 1 0 1 1 1 1 1 1 1
  0 1 1 0 1 1 1 1 0 1 1
  1 0 1 1 0 1 0 1 0 1 1
  0 1 1 1 1 1 1 0 1 1 0
*seq(seq(when(rand(10)<=7,1,0),i,1,12),j,1,5)
MAIN DEG AUTO FUNC 3/30
  
```

Fig. 2

G. MALLE betont im Kommentar zur AHS-Oberstufe 1991 (S.163), dass zum Verständnis der **Binomialverteilung** es zunächst notwendig ist, herauszuarbeiten, „daß die **absolute Häufigkeit  $H$**  des Eintretens eines Ereignisses in einer Versuchsserie als eine **Zufallsvariable** aufgefaßt werden kann und eine gewisse Wahrscheinlichkeit besitzt“. Da in Fig. 2 genau eine solche Versuchsserie dargestellt ist, sind alle Voraussetzungen vorhanden, um mit dem TI-92 Aufgabenstellungen zur Binomialverteilung zu simulieren.

Beispiel:

Peter und Klaus spielen gegeneinander Tennis. Peter gewinnt i. a. 7 von 10 Sätzen. Wie groß ist die Wahrscheinlichkeit, dass Peter bei 10 gespielten Sätzen mindestens 6 Sätze gewinnt?

a) Es sollen nun **1280 Sätze durch den TI-92 simuliert** werden und dann **in 10er-Gruppen** auf dem Gerät **ausgewertet** werden.

Die erzeugte Versuchsreihe wird als **Matrix** mit 10 Zeilen und 128 Spalten am Home-Fenster ausgegeben und unter den Namen **mat1** gespeichert (Fig. 3). Um die absolute Häufigkeit der Siege (der Einsen) bei 10 Sätzen zu erhalten, benötigt man die **Spaltensumme**. Der TI-Befehl **sum(mat1)** gibt die Summe der Elemente aus den Spalten der Matrix mat1 als Zeilenvektor aus (in Fig. 4 [6, 6, 7, ...]). Da aber im Daten-Editor zur weiteren Verarbeitung eine Liste benötigt wird, wird der Vektor mit dem TI-Befehl **mat▶list([ ... ])** in eine Liste {...} verwandelt.

```

F1 Algebra F2 Calc F3 Other F4 PrgmIO F5 Clear Up F6
▀ seq(seq(1,rand(10)<=7,i,1,128),j,1,10)
  1 1 0 0 1 1 0 0 0 1 1 0 1 0
  0 0 1 1 0 0 1 0 0 1 1 1 1 1
  1 0 1 0 1 0 1 0 1 1 0 1 1 0
  0 0 0 1 1 1 1 1 0 1 1 1 1 1
  1 1 1 1 1 1 1 1 1 1 0 1 1
...7,1,0),i,1,128),j,1,10)▶mat1
MAIN RAD AUTO FUNC 30/30
  
```

Fig. 3

```

F1 Algebra F2 Calc F3 Other F4 PrgmIO F5 Clear Up F6
▀ sum(mat1)
  6 6 7 7 8 7 7 5 7 7 5 8 8 7
▀ mat▶list([6 6 7 7 8 7 7 5 7 7 5 8 8 7
  6 6 7 7 8 7 7 5 7 7 5 8 8 7
...8,9,6,6,10,7,9,8,9,5,7,6]])
MAIN RAD AUTO FUNC 25/30
  
```

Fig. 4

Diese Liste der absoluten Häufigkeit der Siege bei je 10 gespielten Sätzen wird mit dem Befehl COPY kopiert und mit dem Befehl PASTE in die Spalte c1 des Daten-Editors geholt (Fig. 5). Anschließend wird von der Spalte c1 das Histogramm im Graphik-Fenster dargestellt (Fig. 8).

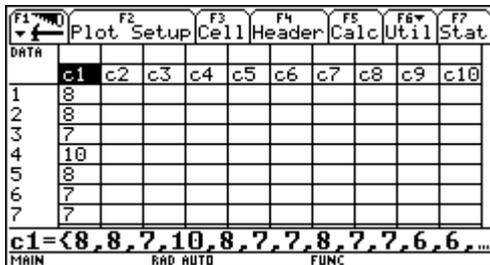


Fig. 5

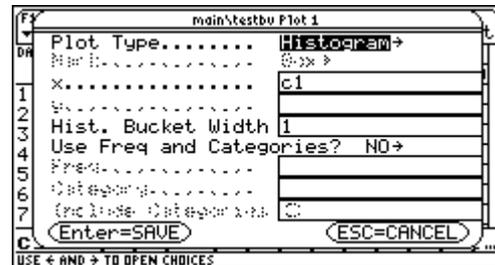


Fig. 6

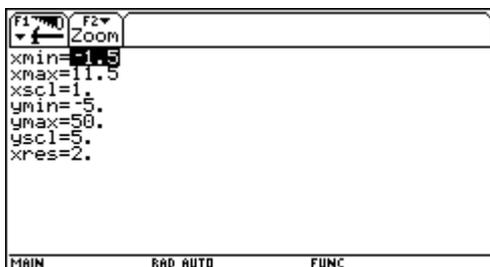


Fig. 7

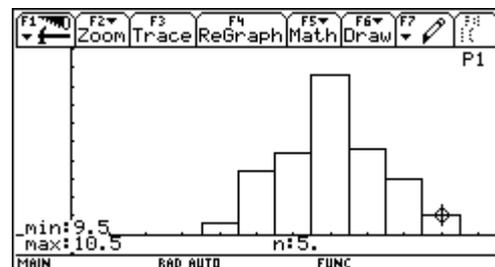


Fig. 8

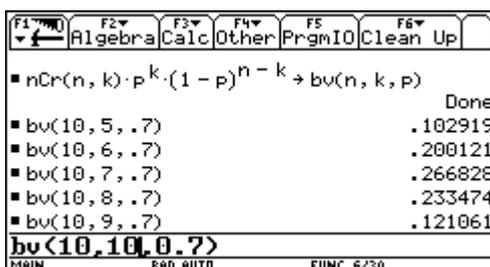
Im Histogramm können die absoluten Häufigkeiten (abs. H.) von 0, 1, 2, ..., 9, 10 Siegen im Graphik-Fenster im TRACE-Mode bestimmt werden. Durch Division der absoluten Häufigkeiten (rel. H.) durch 128 erhält man dann die relativen Häufigkeiten.

### Auswertung des Histogramms

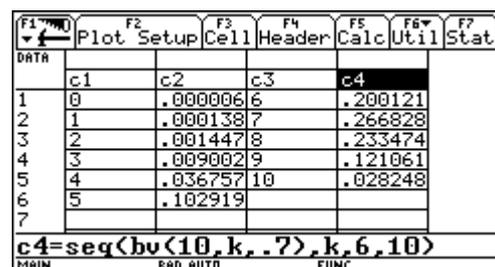
Siege	0	1	2	3	4	5	6	7	8	9	10	Σ
abs.H.	0	0	0	0	3	17	22	43	23	15	5	128
rel.H.	0	0	0	0	0.023	0.070	0.172	0.336	0.180	0.117	0.039	1

b) Es soll mit der Formel der Binomialverteilung die **Wahrscheinlichkeiten P(X=i)** (i ist die Anzahl der gewonnenen Sätze bei 10 gespielten Sätzen) berechnet und die erhaltenen Ergebnisse mit den Ergebnissen der Simulation verglichen werden.

geg.  $n = 10, p = 0,7$       ges.  $P(X = i) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{(n-k)}$ ,  $k = 0, 1, 2, \dots, 10$



Berechnung im HOME-Fenster



Berechnung im Daten-Editor

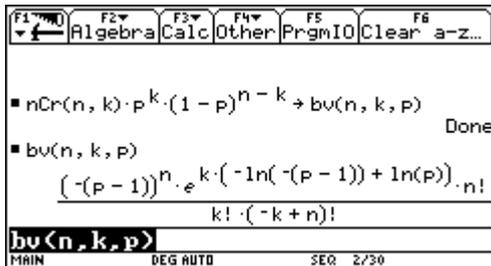
## Vergleich der Wahrscheinlichkeiten von Modell und Simulation

Siege	0	1	2	3	4	5	6	7	8	9	10	Σ
Modell	0	0	0	0.01	0.04	0.10	0.20	0.27	0.23	0.12	0.03	
Simul.	0	0	0	0	0.02	0.07	0.17	0.34	0.18	0.12	0.04	1

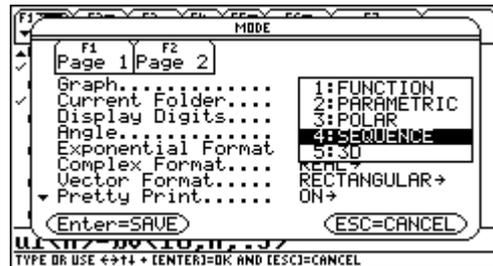
### 4. Graphische Darstellung der Binomialverteilung

Es ist vorteilhaft die Formel der Binomialverteilung im HOME-Fenster unter einem Namen wie z.B.  $bv(n,k,p)$  zu speichern. Ruft man dann  $bv(n,k,p)$  auf, so sieht man, dass der TI-92 zur Berechnung der Binomialverteilung die e-Funktion verwendet.

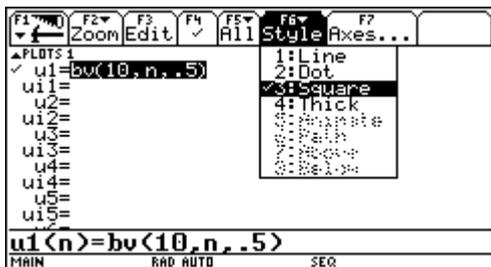
Da die Binomialverteilung eine diskrete Verteilung ist, kann die graphische Darstellung leicht im **SEQUENCE-Mode** durchgeführt werden. Die Werte von  $P(X=k)$  können sowohl im Tabellen-Fenster als auch im Graphik-Fenster (TRACE-Mode) abgelesen werden.



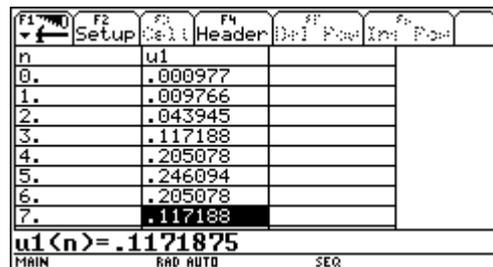
Berechnung mit Hilfe der e-Funktion



Einstellung des SEQUENCE-Mode

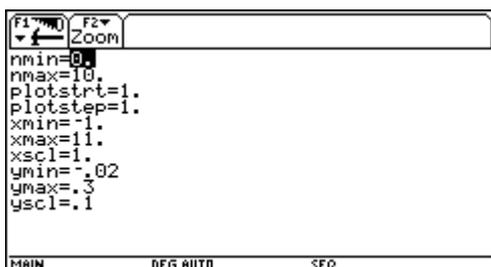


Eingabe im SEQUENCE-Fenster

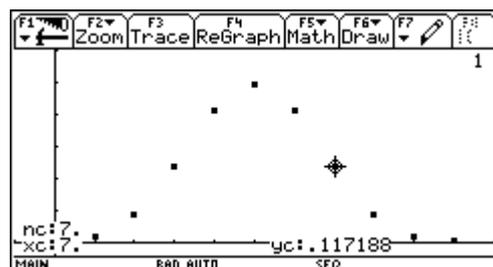


BV-Tabelle für  $p=0.5$ ,  $n=10$ ,  $k=0,1,2,\dots$

Im SEQUENCE-Mode ist zu beachten, dass für das **k** in der Formel stets **n** als Laufvariable genommen werden muss!



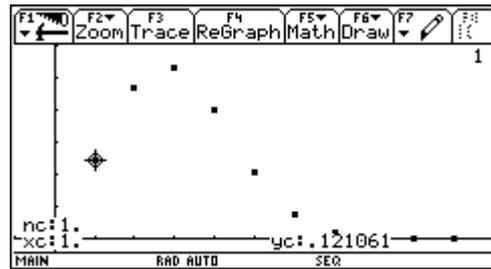
Einstellung des Graphik-Fensters  $p=0.5$



TRACE-Mode:  $bv(10,7,0.5) = 0.117$

F1	F2	F3	F4	F5	F6
Setup	Cell	Mode	Del	Pol	Int
n	u1				
0.	.028248				
1.	.121061				
2.	.233474				
3.	.266828				
4.	.200121				
5.	.102919				
6.	.036757				
7.	.009002				
n=0.					
MAIN	RAD AUTO	SEQ			

BV-Tabelle für p=0.3, n=10, k=0,1,2,...



TRACE-Mode:  $bv(10,1,0.3) = 0.121$

Die beiden Binomialverteilungen  $bv(10, k, 0.5)$  und  $bv(10, k, 0.3)$  haben ein unterschiedliches Aussehen. Sie besitzen an verschiedenen Stellen ihr Maximum (ihren Erwartungswert), die erste für  $k=5$  und die zweite für  $k=3$ , und „streu“ auch verschieden um diese.

### 5. Ableitung von $\mu$ und $\sigma$ einer Binomialverteilung bei gegebenem n

G. MALLE (1991, S.162) schlägt im Lehrplankommentar vor, von den mit Hilfe der relativen Häufigkeit formulierten Formeln für Mittelwert und Varianz einer Häufigkeitsverteilung (5. Klasse AHS) auszugehen, um damit die Formeln für **Erwartungswert** und **Varianz** einer diskreten Wahrscheinlichkeitsverteilung zu motivieren:

$$\mu = E(X) = \sum x_i \cdot P(X = x_i) \quad \sigma^2 = V(X) = \sum (x_i - \mu)^2 \cdot P(X = x_i)$$

G. MALLE (1991, S.163) meint weiter, dass es im einfachsten Falle genüge, die entsprechenden Formeln für die Binomialverteilung den Schülern einfach mitzuteilen

$$\mu = n \cdot p \quad \sigma^2 = n \cdot p \cdot (1 - p)$$

oder für  $n \leq 3$  mit beliebigem p aus der Definition von  $E(X)$  und  $V(x)$  mit der Hand herleiten zu lassen.

ASPETSBERGER / SCHLÖGELHOFER (1996, S.95f) zeigen, dass der TI-92 in der Lage ist, die Formeln für die Binomialverteilung für  $n = 20$  und höher problemlos aus der Definition herzuleiten. Allerdings ist die Rechenzeit bei der Berechnung der Varianz bedeutend länger als bei der des Erwartungswertes. Die Rechenzeit bei der Varianz kann allerdings reduziert werden, wenn man die zweite Formel für die Varianz aus der Formelsammlung nimmt:

$V(X) = \sum k^2 \cdot P(X = x_i) - \mu^2 = E(X^2) - [E(X)]^2$  (KRAFT/BÜRGER/UNFRIED/HASCHKOVITZ, Mathematische Formelsammlung, Verlag Hölder-Pichler-Tempsky, Wien, 1994).

F1	F2	F3	F4	F5	F6
Algebra	Calc	Other	PrgmIO	Clean	Up
20					
■	$\sum_{k=0}^{20} (k \cdot bv(20, k, p))$				20 · P
50					
■	$\sum_{k=0}^{50} (k \cdot bv(50, k, p))$				50 · P
$\Sigma(k * bv(50, k, p), k, 0, 50)$					
MAIN	RAD AUTO	FUNC 2/30			

Berechnung von  $E(X)$  für  $n=20$  u.  $n=50$

F1	F2	F3	F4	F5	F6
Algebra	Calc	Other	PrgmIO	Clean	Up
20					
■	$\sum_{k=0}^{20} ((k - 20 \cdot p)^2 \cdot bv(20, k, p))$				-20 · P · (P - 1)
20					
■	$\sum_{k=0}^{20} (k^2 \cdot bv(20, k, p)) - (20 \cdot p)^2$				-20 · P · (P - 1)
$\dots *bv(20, k, p), k, 0, 20) - (20p)^2$					
MAIN	RAD AUTO	FUNC 4/30			

Berechnung von  $V(X)$  auf zwei Arten



c) Anzahl der Schwarzfahrer im Mittel ? - Bereich mit 90% Wahrscheinlichkeit ?

Die Zufallsvariable X gibt die Anzahl der Schwarzfahrer unter 300 Fahrgästen an. X wird als binomialverteilt mit  $n = 300$  und  $p = 0.06$  angenommen.  $E(X) = \mu = 300 \cdot 0.06 = 18$ .

Der Kontrolleur wird im Mittel 18 Schwarzfahrer pro Tag antreffen.

**Gesucht ist ein symmetrisches Intervall  $[\mu - z, \mu + z]$ ,  $z \in \mathbb{N}$ , so dass die Summe der Einzelwahrscheinlichkeiten von  $P(\mu - z)$  bis  $P(\mu + z) \geq 0,90$  ist ! ( $\mu = 18$ )**

M. KOTH löst dieses Beispiel (S. 165) dadurch, dass sie die Binomialverteilung durch die Normalverteilung annähert. Mit dem TI-92 kann aber das Problem mit Hilfe des Befehls **cumSum(liste1)** gelöst werden.

**TI-92 Befehl:** (TI-92 Handbuch 1996,S.386)

**cumSum(Liste1)** → Liste    Eingabe: cumSum({1, 2, 3, 4})    Ausgabe: {1 3 6 10}

Der Befehl gibt eine Liste der kumulierten Summen der Elemente aus Liste1 zurück, wobei bei Element 1 begonnen wird.

Es sind also solange die folgenden Teilsummen der Wahrscheinlichkeiten zu bilden, bis 0,90 erreicht oder erstmals überschritten wird:

1. Teilsumme =  $P(X=18) = 0,09654$
  2. Teilsumme =  $P(X=18) + P(X=17) + P(X=19)$
  3. Teilsumme = 2. Teilsumme +  $P(X=16) + P(X=20)$
- .....

Lösungsidee: Erzeuge im Dateneditor in einer **Spalte (c2)**  $P(X=17), P(X=16), \dots$  und in der anderen **Spalte (c3)**  $P(X=19), P(X=20), \dots$   
**so lange bis die Teilsummen größer oder gleich  $0,90 - P(X=18) = 0,80346$  sind!**

	F2	F3	F4	F5	F6	F7
	Plot	Setup	Cell	Header	Calc	Util
DATA						
	c1	c2	c3	c4	c5	
1	17	.0962	19	.09146	.18765	
2	16	.09021	20	.08202	.35988	
3	15	.07934	21	.0698	.50903	
4	14	.0652	22	.0565	.63073	
5	13	.04982	23	.04359	.72414	
6	12	.03523	24	.03212	.79149	
7	11	.02292	25	.02263	.83704	
<b>c5=cumSum(c2+c4)</b>						

In Spalte c5 sind die Teilsummen !

	F2	F3	F4	F5	F6
	Algebra	Calc	Other	PrgmIO	Clean Up
	cumSum({1 2 3 4})		{1 3 6 10}		
24	$\sum_{k=12}^{24} \text{bv}(300, k, .06)$		.888029		
25	$\sum_{k=11}^{25} \text{bv}(300, k, .06)$		.933581		
<b><math>\Sigma(\text{bv}(300, k, 0.06), k, 11, 25)</math></b>					

richtiges Intervall: [11, 25]

Die Zahl der pro Tag angetroffenen Schwarzfahrer liegt mit (mindestens) 90%iger Wahrscheinlichkeit im Bereich [11, 25].

## 7. Die Poissonverteilung als Sonderfall der Binomialverteilung

Wenn bei einer Anwendung die Zahl  $n$  der Teilversuche sehr groß, hingegen die Wahrscheinlichkeit  $p$  der einzelnen Versuche sehr klein ist, so dass der Limes für  $n \rightarrow \infty$  von  $n \cdot p \rightarrow \lambda$  mit  $\lambda$  als endliche reelle Zahl geht, so kann die Binomialverteilung durch die Poissonverteilung angenähert werden. Diese Näherung ist für  $0 < p < 0,1$  sehr gut und kann auch für  $0,9 < p < 1$  durch Anwendung auf die Gegenwahrscheinlichkeit verwendet werden.

A. ENGEL (S. 118f) schlägt einen Weg für die Gewinnung der Formel vor, der mit dem TI-92+ sehr gut gegangen werden kann (Bild 1 u. 2). Verwendet man jedoch einen normalen TI-92, so erscheint die Meldung „Circular definition“ (Bild 3). In Bild 4 wird gezeigt, wie man das auftretende Problem durch nachträgliche Substitution mit  $a = k + 1$  lösen kann.

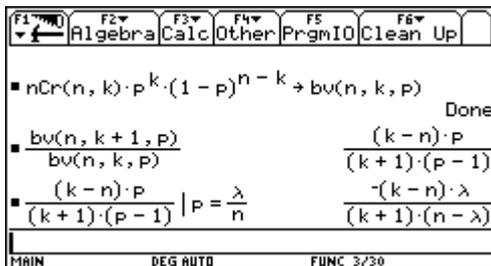


Bild 1:  $bv(k+1)/bv(k)$  mit dem TI-92+

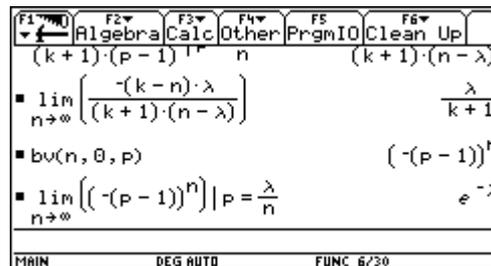


Bild 2:  $bv(k+1)/bv(k) \rightarrow \lambda/(k+1)$

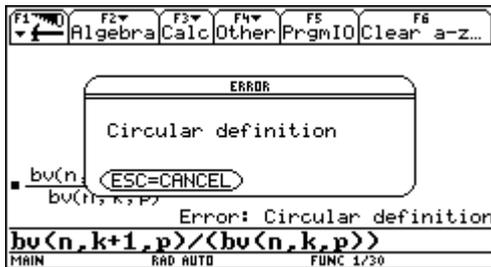


Bild 3: Probleme beim normalen TI-92

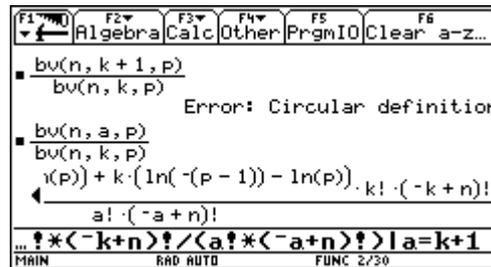
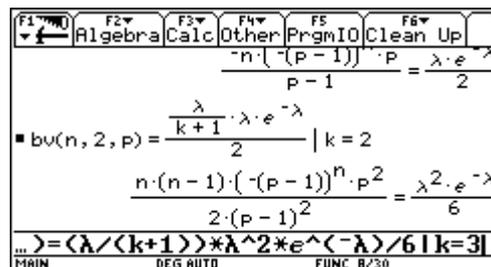
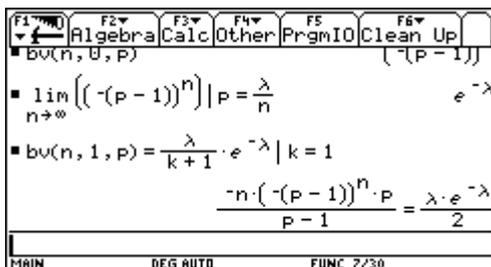


Bild 4: Substitution  $a = k+1 \rightarrow$  Bild1, 2

<b>Rekursionsformel:</b> $\frac{bv(n, k+1, p)}{bv(n, k, p)} = \frac{\lambda}{k+1} \quad \text{mit } bv(n, 0, p) = e^{-\lambda}$
---



Durch Einsetzen für  $k = 1, 2, 3, \dots$  ergibt sich als Formel für die Poissonverteilung:

$pv(\lambda, k) = \frac{\lambda^k}{k!} \quad \text{mit } \lambda = \mu = n \cdot p \text{ als Erwartungswert und Varianz } \sigma^2$
--

## Beispiel

4% aller Fluggäste, die Plätze reservieren erscheinen nicht. Die Fluggesellschaft weiß dies und verkauft 75 Flugkarten für die 73 verfügbaren Plätze.

Wie groß ist die Wahrscheinlichkeit, dass alle Fluggäste Platz bekommen ?

Aufgabe aus KIRSCHENHOFER/ARNOLD, Aufgabensammlung zur Wahrscheinlichkeitsrechnung mit didaktischen Beiträgen, Band II, Institut für Mathematik, Universität Linz, Nr. 2.16, S. 32.

Die Wahrscheinlichkeit, dass ein Fluggast erscheint, beträgt 0,96, dass er nicht erscheint 0,04. Somit sind bei der Binomialverteilung zwei Varianten möglich: Es bekommen alle Fluggäste einen Platz, wenn höchstens 73 Fluggäste erscheinen oder wenn mindestens 2 der 75 Fluggäste nicht erscheinen.

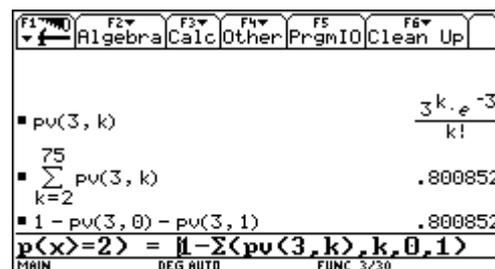
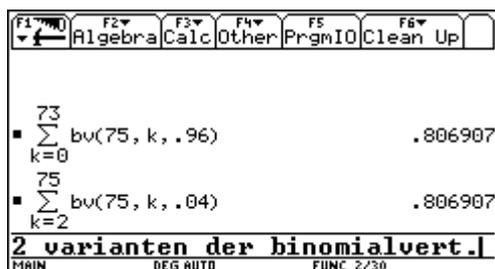
Berechnung von  $\mu$  bzw.  $\lambda$ :

$n = 75$  für  $p = 0,96$  ist  $\mu = 75 \cdot 0,96 = 72$

für  $p = 0,04$  ist  $\mu = \lambda = 75 \cdot 0,04 = 3$

Lösung als Binomialverteilung  $bv(n,k,p)$

Lösung als Poissonverteilung  $pv(\lambda, k)$



Die erscheinenden Fluggäste bekommen einen Platz mit ca. 80%iger Wahrscheinlichkeit.

Binomialverteilung  $p = 80,7\%$

Poissonverteilung  $p = 80,1\%$

## 8. Darstellung der Binomialverteilung als Histogramm und als Wahrscheinlichkeitsdichtepolygon

Da die Binomialverteilung eine diskrete Verteilung ist, ist die adäquate Darstellung das **Punkt-** bzw. das **Stabdiagramm**. Leider ist das sehr anschauliche Stabdiagramm – die Summe der Stablängen ist 1 – mit dem TI-92 nur annähernd als Staffeldbild mit schmalen Staffeln darstellbar (Bild 4).

Daher erscheint es sinnvoll, auch im Hinblick auf die Vorbereitung der Normalverteilung, das **Staffeldbild (Histogramm) mit der Staffeldbreite 1** zu verwenden (Bild 6). Als Maß für die Wahrscheinlichkeit gilt nun die Summe der Flächeninhalte der Rechtecke (Gesamtsumme 1). Verbindet man die Breitenmitten der Rechtecke (Bild 8), so erhält man ein flächengleiches Polygon, ein **Wahrscheinlichkeitsdichtepolygon** (Bild 10). Es stellt sich beim Betrachten des Bildes die Frage, ob nicht das Polygon durch den Graphen einer Funktion ersetzt werden könnte. (→ Kapitel 9)

**Darstellungsformen der Binomialverteilung mit  $n = 20$ ,  $p = \frac{1}{2}$  und  $k = 0, 1, 2, \dots, 20$ .**



Bild 1: Einstellung Scatter

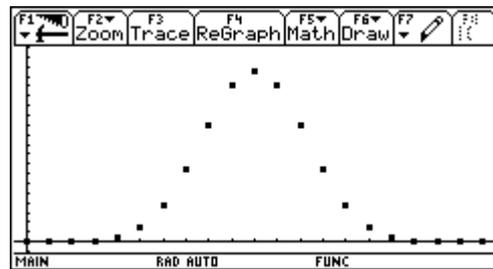


Bild 2: Punktdiagramm



Bild 3: Einstellung Histogramm

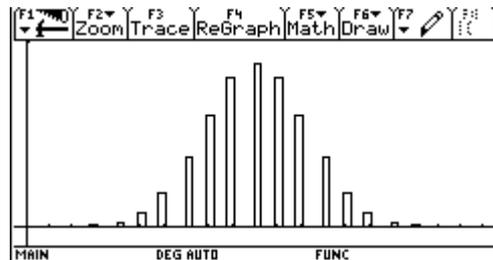


Bild 4: „Stabdiagramm“ mit Breite 0,3

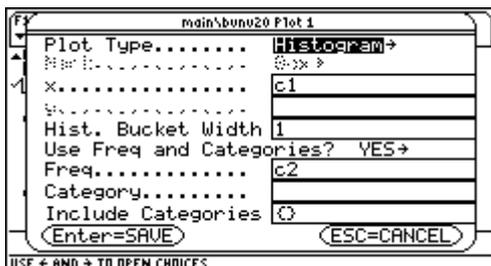


Bild 5: Einstellung Frequences YES

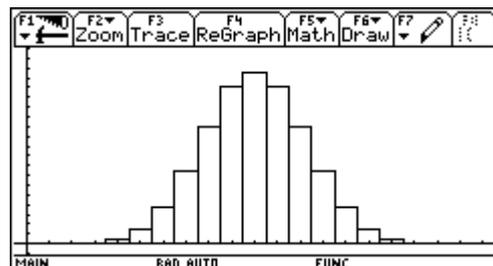


Bild 6: Histogramm mit Breite = 1

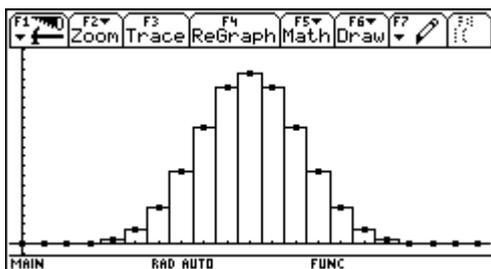


Bild 7: Punkte in der Breitenmitte

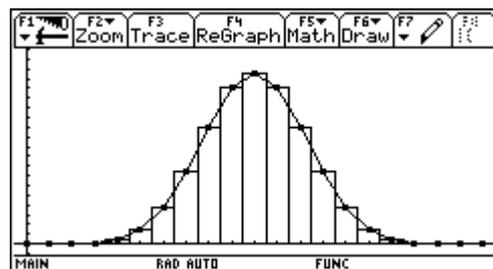


Bild 8: Breitenmitten verbunden (xy-line)



Bild 9: xy-line (Streckenzug)

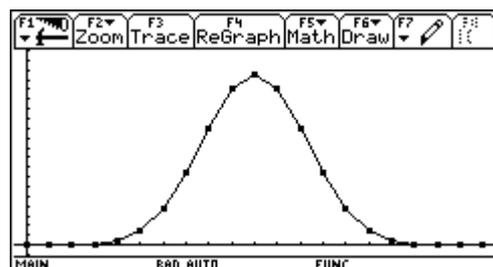
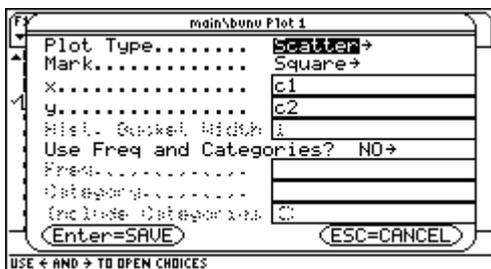


Bild 10: **Wahrscheinlichkeitsdichtepolygon**

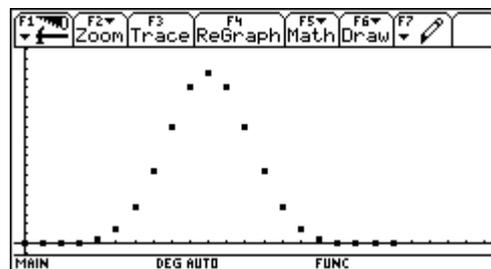
## 9. Der Zugang zur Normalverteilung über die Binomialverteilung

Da die Normalverteilung als Grenzwert der Binomialverteilung aufgefasst werden kann, kann sie für große  $n$  durch die Normalverteilung angenähert werden. Dieser Weg wird auch von REICHEL/HANISCH/MÜLLER (1989, S.196ff) gewählt. Die Verwendung des TI-92 in der 7. Klasse ermöglicht, die Schüler so an Punktdiagramme und Histogramme der Binomialverteilung zu gewöhnen, dass mit der Zeit die Frage aktuell wird, ob nicht durch die erhaltenen Punkte („Glockengestalt“) der Graph einer Ausgleichsfunktion gelegt werden kann. Der im Folgenden skizzierte Weg ist im Unterricht erprobt worden (vgl. O. WURNIG, 1993, S.ff).

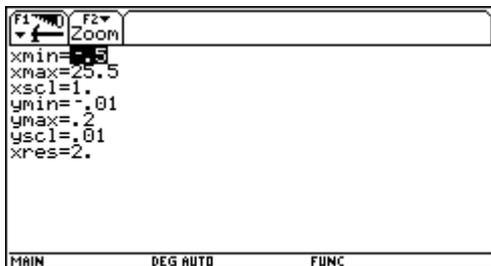
Zuerst werden im Graphikfenster im FUNCTION- Mode drei verschiedene Binomialverteilungen für  $n = 20, 25, 30$  und  $p = 0.5$  bei gleicher WINDOW – Einstellung gezeichnet.



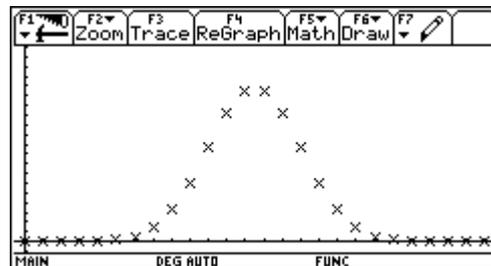
Definition des Graphiktyps SCATTER



$n = 20, p = 0.5$



Einstellung für alle  $n = 20, 25, 30$

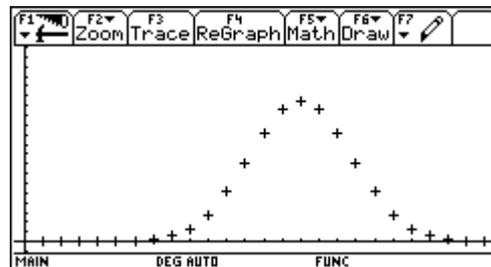


$n = 25, p = 0.5$

	c1	c2	c3	c4
1	0	9.54E-7	2.98E-8	9.3E-10
2	1	.000019	7.45E-7	2.79E-8
3	2	.000181	.000009	4.05E-7
4	3	.001087	.000069	.000004
5	4	.004621	.000377	.000026
6	5	.014786	.001583	.000133
7	6	.036964	.005278	.000553

$c4 = \text{seq}(b0(30, k, .5), k, 0, 30)$

$c2 (n = 20), c3 (n = 25), c4 (n = 30)$

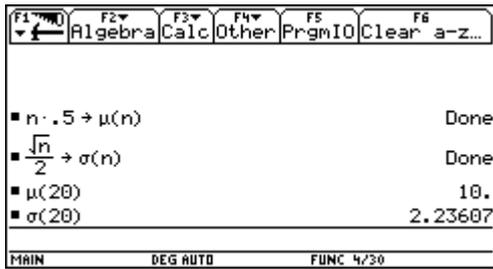


$n = 30, p = 0.5$

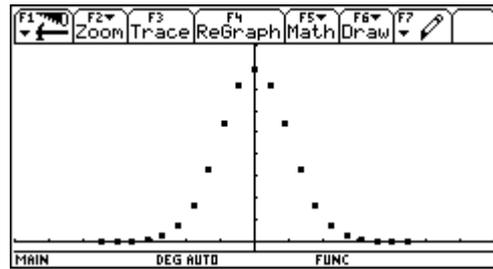
Vergleicht man die drei graphischen Darstellungen für  $n = 20, 25, 30$  so sieht man,

- dass alle drei Bilder symmetrisch sind ( $p = 0.5$ ),
- dass mit wachsendem  $n$ 
  - der Erwartungswert  $\mu = n.p$  immer größer,  $f(\mu)$  hingegen immer kleiner wird,
  - die Standardabweichung  $\sigma = \sqrt{n.p.(1-p)} = n^{0.5}/2$  ständig zunimmt.

Es wird daher für alle drei Darstellungen eine **Standardisierung** auf  $\mu = 0$  und  $\sigma = 1$  durchgeführt. Dies gelingt mit Hilfe der Transformation  $k \rightarrow (k-\mu)/\sigma$  und  $f(k) \rightarrow \sigma.f(k)$ .

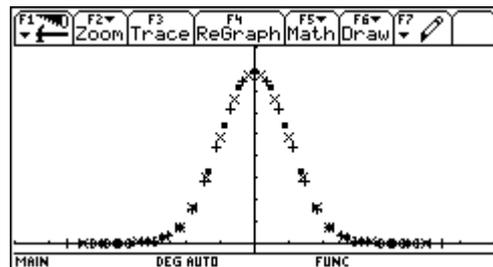
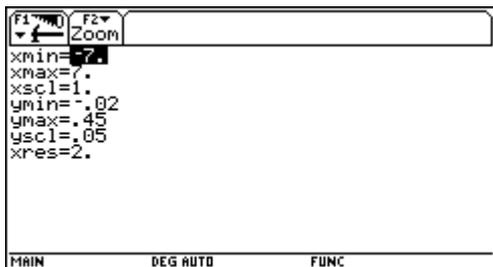


Standardisierung für n = 20



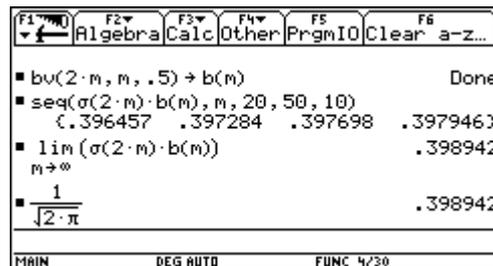
standardisierte Binomialverteilung n=20

Das folgende Bild enthält alle Punkte der standardisierten Darstellung für n = 20, 25, 30.



Betrachtet man dieses Bild, so stellt sich sofort die Frage, welcher Funktionsgraph diese Punkte möglichst gut annähern kann. Es ist somit eine Funktion f zu bestimmen, die folgende Eigenschaften besitzt:

- $f(x) \geq 0$  für alle  $x$  und  $f(0) = d$  (Abschnitt auf der  $y$ -Achse);
- der Graph von  $f$  ist glockenförmig und symmetrisch zur  $x$ -Achse;
- die von der Kurve und der  $x$ -Achse eingeschlossene Fläche hat den Inhalt 1.



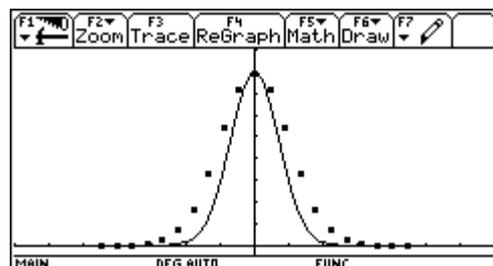
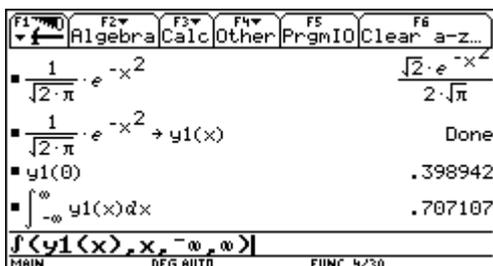
**Berechnung von  $f(0) = d = \sigma(n) \cdot b(\mu)$**

z.B.  $n=40, \mu=40 \cdot 0.5=20 \rightarrow d \approx 0.396$

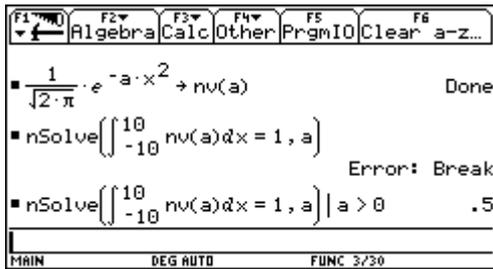
**Gibt es für  $n \rightarrow \infty$  einen Grenzwert ?**

$n \rightarrow 2n, \mu \rightarrow n$  ergibt  $d = \frac{1}{\sqrt{2\pi}}$

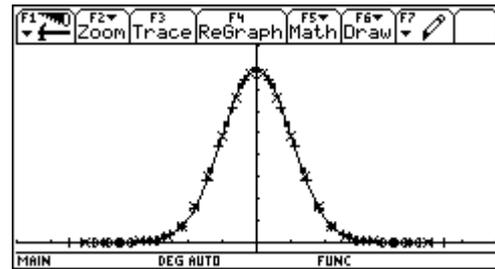
Wie man im Bild erkennen kann, ergibt die Funktion  $y_1(x) = d \cdot e^{-x^2}$  eine glockenförmige, zur  $x$ -Achse symmetrische Funktion, aber der Flächeninhalt von  $-\infty$  nach  $+\infty$  ist 0.7 und nicht 1.



Um die Gestalt der Kurve bei gleichbleibendem  $d$  so zu verändern, dass der Flächeninhalt von  $-\infty$  nach  $+\infty$  wirklich 1 wird, erhält der Exponent einen noch genauer zu bestimmenden Korrekturfaktor  $a$ . Da die angesetzte Gleichung mit dem Integral wegen der Grenzen  $-\infty$  und  $+\infty$  nicht gelöst werden kann, werden die Grenzen durch  $-10$  und  $+10$  ersetzt. Ergebnis:  **$a = 0.5$**

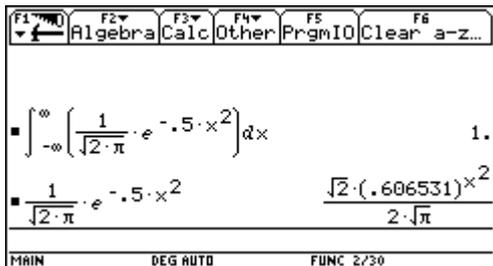


Berechnung von a

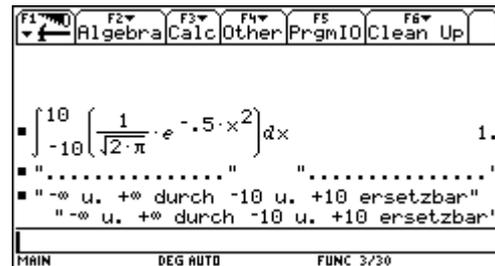


Die berechnete Kurve passt !

Es soll diese neue Funktion  $y_1(x) = d \cdot e^{(-0.5x^2)}$  mit  $d=1/\sqrt{2\pi}$  kontrolliert werden, ob sie wirklich von  $-\infty$  bis  $+\infty$  bzw. von  $-10$  bis  $+10$  den Flächeninhalt 1 besitzt.



Der Flächeninhalt unter der Kurve ist 1.



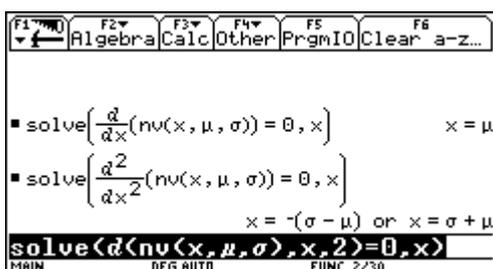
Der Flächeninhalt von  $-10$  bis  $+10$  ist 1.

## 12. Diskussion der GAUSSschen Glockenkurve

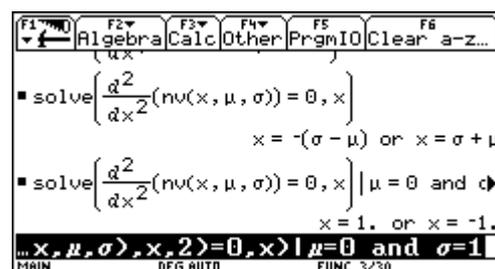
Die erhaltene standardisierte Funktion ist aber nur der Sonderfall mit  $\mu = 0$  und  $\sigma = 1$  der Wahrscheinlichkeitsdichtefunktion  $nv(x, \mu, \sigma)$ :

$$nv(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-0,5 \cdot ((x - \mu) / \sigma)^2}$$

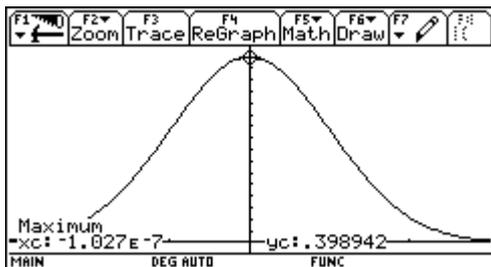
Die durch diese Dichtefunktion festgelegte stetige Verteilung heißt **Normalverteilung** mit den Parametern  $\mu$  und  $\sigma$ . **Der Graph von  $nv(x, \mu, \sigma)$  heißt GAUSSsche Glockenkurve.** Von ihr können durch Differenzieren im Home-Fenster immer das **Maximum** und die **Wendepunkte** berechnet werden. Im Graphik-Fenster können diese Punkte auch, aber nur für bekanntes  $\mu$  und  $\sigma$  bestimmt werden.



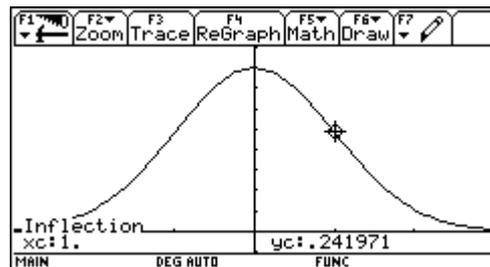
Maximum/Wendepunkte von  $nv(x, \mu, \sigma)$



Wendepunkte für  $\mu=0$  und  $\sigma=1$



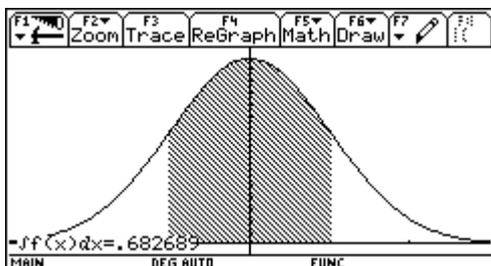
Der TI-92 gibt hier  $-1,027 \cdot 10^{-7}$  statt 0 aus!



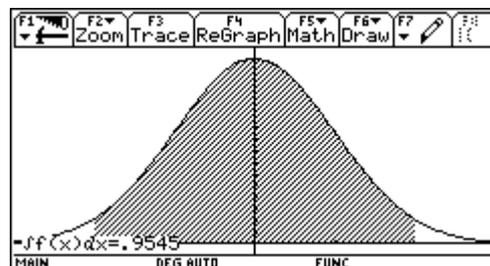
Einer der beiden Wendepunkte von  $nv(0,1)$

Die Tatsache, dass die Glockenkurve bei  $(\mu - \sigma)$  u.  $(\mu + \sigma)$ , bei der standardisierten Form bei  $+1$  und  $-1$ , Wendepunkte besitzt, sollte beim Erstellen von Skizzen stets berücksichtigt werden.

Um eine Wahrscheinlichkeit, d.h. einen Flächeninhalt, gut abschätzen zu können, ist es wichtig zu wissen, wie groß der Flächeninhalt um  $\mu$  im symmetrischen  $\sigma$ -,  $2\sigma$ - und  $3\sigma$ -Bereich ist.



Flächeninhalt im  $\sigma$ -Bereich 68,27%



Flächeninhalt im  $2\sigma$ -Bereich 95,45%

Im  $3\sigma$ -Bereich beträgt der Flächeninhalt bereits 99,9937%.

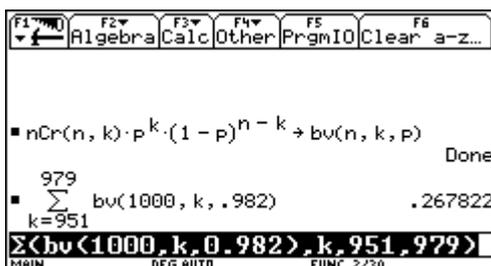
## 11. Eine Aufgabe - vier Lösungsvarianten

*Die Säuglingssterblichkeit, d.h. die Wahrscheinlichkeit, innerhalb des ersten Lebensjahres zu sterben, betrage in einem bestimmten Land 1.8%.*

*Wie groß ist die Wahrscheinlichkeit, daß von 1000 zufällig gewählten Säuglingen mehr als 950 und weniger als 980 Kinder den ersten Geburtstag erleben?*

Aufgabe aus REICHEL/MÜLLER/HANISCH, Lehrbuch der Mathematik 8, Nummer 432c, Seite 124, Verlag Holder-Pichler-Tempsky, Wien, 1992.

### 1. Variante: Lösung als Binomialverteilung $bv(n,k,p)$

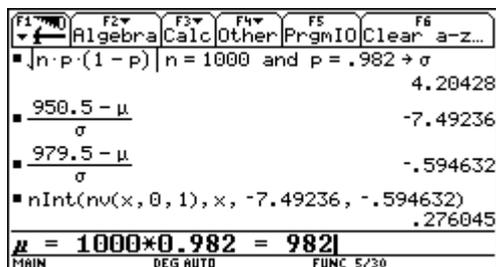


BV:  $n = 1000$   $p = 0.982$

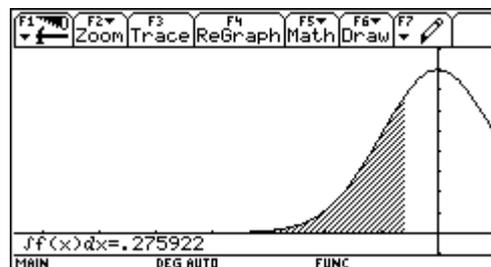
Die Wahrscheinlichkeit, dass von 1000 Säuglingen mehr als 950 und weniger als 980 den ersten Geburtstag erleben ist 26.8%.

Bei großem  $n$  kann die Binomialverteilung durch die Normalverteilung angenähert werden, jedoch ist eine Stetigkeitskorrektur vorzunehmen: d.h. statt  $[951, 979] \rightarrow [950.5, 979.5]$ .

## 2. Variante: Lösung mit der standardisierten Normalverteilung $nv(0,1)$

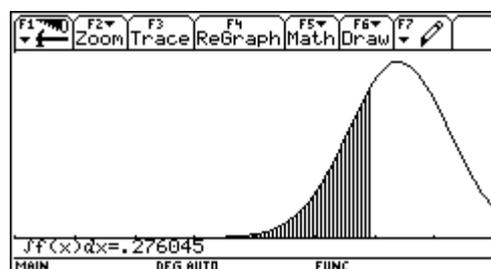
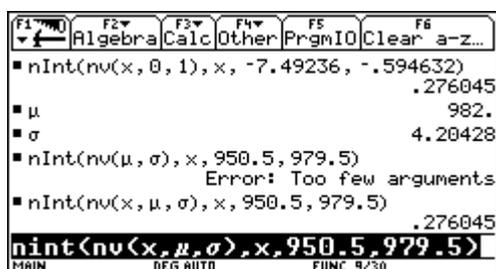


Ergebnis: 27.6%



Ergebnis: 27.6%

## 3. Variante: Lösung mit der Normalverteilung $nv(\mu, \sigma)$



Der TI-92 berechnet somit mit der BV 26.8% und mit der NV beide Male 27.6%.

Die Lösung dieses Beispiels zeigt, dass sich solche Aufgaben zur Binomialverteilung mit dem TI-92 durch bloße Summation mit  $\Sigma$  lösen lassen. Es zeigt sich aber auch deutlich, dass die **Binomialverteilung** unter Beachtung der Stetigkeitskorrektur **durch die Normalverteilung angenähert** werden kann:

$$\text{Faustregel: } \sigma = \sqrt{(n \cdot p \cdot (1-p))} > 3$$

## 4. Variante: Lösung mit die Poissonverteilung $pv(\lambda, k)$

Bei großem  $n$  und kleinem  $p$  ( $p < 0,1$ ) kann die Binomialverteilung auch durch die Poissonverteilung angenähert werden. Bei großem  $p$  ( $p > 0,9$ ) kann das Problem durch Berechnung der Wahrscheinlichkeit des Gegenereignisses gelöst werden.

Umformulierte Fragestellung:

Wie groß ist die Wahrscheinlichkeit, dass von 1000 zufällig gewählten Säuglingen weniger als 21 Kinder den ersten Geburtstag nicht erleben?

F1	F2	F3	F4	F5	F6
Algebra	Calc	Other	PrgmIO	Clean Up	
$PV(\lambda, k) = \frac{\lambda^k \cdot e^{-\lambda}}{k!}$					
$\sum_{k=0}^{20} PV(18, k)$					.73072
$1 - .73072015572374$					.26928
MAIN DEG AUTO FUNC 3/30					

$$PV: n = 1000p = 0,018$$

$$\lambda = n \cdot p = 1000 \cdot 0,018 = 18$$

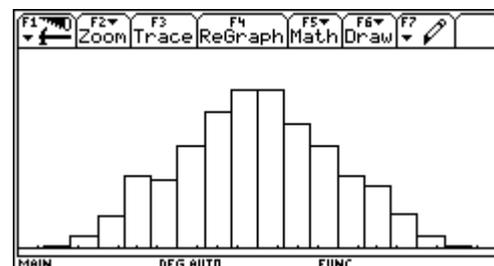
Die Wahrscheinlichkeit, dass von 1000 Säuglingen weniger als 21 den ersten Geburtstag nicht erleben, ist 73,1%, die Wahrscheinlichkeit des Gegenereignisses somit 26,9%.

## 12. Der Zugang zur Normalverteilung über die Augenzahlsumme dreier Würfel

Damit die Normalverteilung nicht nur als Grenzverteilung der Binomialverteilung betrachtet wird, wählen BRENNER/LESKY/VOGEL 1980 als **Einstiegsproblem das mehrmalige Würfeln**. Sie begründen ihre Wahl (S. 108) damit, dass sich sowohl beim mehrmaligen Würfeln als auch bei der mehrmaligen Durchführung des Bernoulli-Experimentes (zur Erzeugung von Binomialverteilungen) die Zufallsvariablen als Summe von identisch verteilten Zufallsvariablen wiedergeben lassen und das offenbar zur Folge habe, dass sich mit zunehmender Versuchszahl die Graphen der Wahrscheinlichkeitsfunktionen immer mehr einer Glockenkurve annähern. Diese **Glockenkurve ist charakteristisch für den Graphen der Dichtefunktion einer Normalverteilung**. Mit dem TI-92 lässt sich das Experiment des mehrmaligen Werfens dreier Würfel rasch simulieren. Es sollen daher drei Würfel unabhängig voneinander 864mal mit Hilfe des TI-92 Zufallsgenerators gewürfelt und anschließend bezogen auf die Augenzahlsummen das Histogramm erstellt werden.

F1	F2	F3	F4	F5	F6	F7
Plot	Setup	Cell	Header	Calc	Util	Stat
DATA						
	c1	c2	c3	c4	c5	
1	2	4	5	11		
2	2	6	2	10		
3	6	6	3	15		
4	5	4	1	10		
5	5	6	3	14		
6	2	1	6	9		
7	2	3	1	6		
<b>c4=c1+c2+c3</b>						
MAIN DEG AUTO FUNC						

c4 Augenzahlsumme dreier Würfel



Histogramm für die Augenzahlsummen

Die absoluten Häufigkeiten der Augenzahlsummen 3, 4, 5, ..., 17, 18 werden im Graphikfenster mit Hilfe des TRACE-Mode bestimmt.

3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	9	24	54	51	77	103	120	119	94	77	54	46	25	9	1

Werden die absoluten Häufigkeiten durch die Anzahl der Würfel (**n = 864**) dividiert, so erhält man die relativen Häufigkeiten.

3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1/n	9/n	24/n	54/n	51/n	77/n	103/n	120/n	119/n	94/n	77/n	54/n	46/n	25/n	9/n	1/n

Nun erfolgt die Eingabe der 16 relativen Häufigkeiten im Daten-Editor, die Definition des Graphiktyps **SCATTER** und danach die Ausgabe im Graphikfenster mit 16 Punkten.

	c1	c2	c3	c4
1	3	1/864		
2	4	1/96		
3	5	1/36		
4	6	1/16		
5	7	17/288		
6	8	77/864		
7	9	103/864		

**c1=seq(i,i,3,18)**

Eingabe im Daten-Editor

```

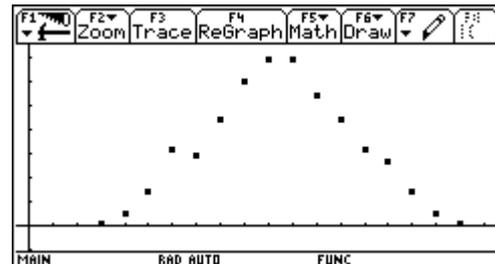
main\wurf3r3 Plot 6
Plot Type..... Scatter→
Mark..... Square→
x..... c1
y..... c2
Hist. Graph: Mark:
Use Freq and Categories? NO→
Freq.....
Category.....
Include Categories: C
Enter=SAVE ESC=CANCEL
  
```

Definition des Graphiktyps

```

MAIN RAD AUTO FUNC
Zoom
xmin=3.5
xmax=19.5
xsc1=1.
ymin=-.02
ymax=.15
ysc1=.02
xres=2.
  
```

Fenstereinstellung



Verteilung der 16 Augenzahlsummen

Da die Punkte (fast) symmetrisch in Glockenform liegen, soll mit Hilfe von  $\mu$  und  $\sigma$  mit der **Dichtefunktion**  $nv(x, \mu, \sigma)$  die dazu passende **Gaußsche Glockenkurve** gezeichnet werden.

	c1	c2	c3	c4
1	3	1/864	1/288	.064423
2	4	1/96	1/24	.434791
3	5	1/36	5/36	.828297
4	6	1/16	3/8	1.24359
5	7	17/288	119/288	.706922
6	8	77/864	77/108	.539605
7	9	103/864	103/96	.25434

**c3=c1\*c2**

	STAT VARS	
c1	$\bar{x}$	=.653791
1	$\Sigma x$	=10.460648
2	$\Sigma x^2$	=10.912517
3	$Sx$	=.521117
4	nStat	=16.
5	minX	=.003472
6	q1	=.157986
7	medStat	=.587963
8		
9		434

**c3=c1\*c2**

Im Daten-Editor können mit dem Funktionsaufruf F6 statistische Berechnungen durchgeführt werden. Der **Erwartungswert**  $\mu = \Sigma x_i \cdot p_i$  kann unter  $\Sigma x = 10.460648$  abgelesen werden.

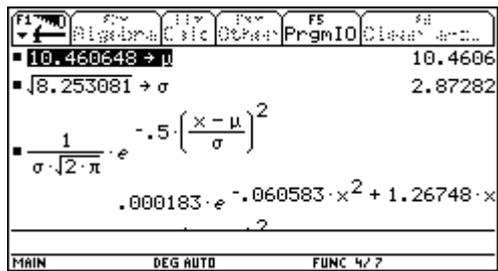
	c1	c2	c3	c4
1	3	1/864	1/288	.064423
2	4	1/96	1/24	.434791
3	5	1/36	5/36	.828297
4	6	1/16	3/8	1.24359
5	7	17/288	119/288	.706922
6	8	77/864	77/108	.539605
7	9	103/864	103/96	.25434

**c4=(c1-10.460648)^2\*c2**

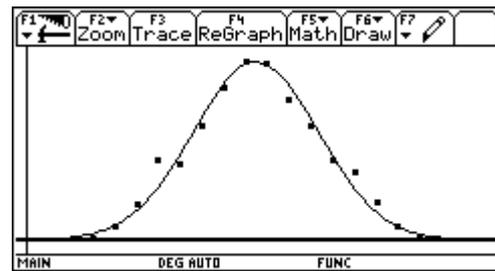
	STAT VARS	
c1	$\bar{x}$	=.515818
1	$\Sigma x$	=8.253081
2	$\Sigma x^2$	=6.488154
3	$Sx$	=.385666
4	nStat	=16.
5	minX	=.029472
6	q1	=.160065
7	medStat	=.492527
8		
9		434

**c4=(c1-10.460648)^2\*c2**

Die errechnete **Varianz**  $\sigma^2 = \Sigma (x_i - \mu)^2 \cdot p_i$  ist unter  $\Sigma x = 8.253081$  zu finden.



$nv(x, \mu, \sigma)$



Gaußsche Glockenkurve

Wie man sieht, passt sich die Gaußsche Glockenkurve recht gut an die dargestellten Punkte an.

## Darstellung im MODELL

Da die Augensummenzahlen aller drei Würfel im Modell gleichverteilt sind, ist die Anzahl der verschiedenen Möglichkeiten des Zustandekommens der Augensummen 3, 4, 5, ..., 17, 18, z.B. mit Hilfe eines Baumdiagramms zu bestimmen. Die relativen Häufigkeiten erhält man, wenn die absoluten Häufigkeiten, das sind die günstigen Möglichkeiten, durch die **Anzahl aller Möglichkeiten** ( $n = 216 = 6 \cdot 6 \cdot 6$ ) dividiert.

So kann die **Augensumme 10** auf folgende **27 Arten** erzielt werden: 1+3+6, 1+4+5, 1+5+4, 1+6+3, 2+2+6, 2+3+5, 2+4+4, 2+5+3, 2+6+2, 3+1+6, 3+2+5, 3+3+4, 3+4+3, 3+5+3, 3+6+1, 4+1+5, 4+2+4, 4+3+3, 4+4+2, 4+5+1, 5+1+4, 5+2+3, 5+3+2, 5+4+1, 6+1+3, 6+2+2, 6+3+1.

In der folgenden Tabelle wird ein **Vergleich der absoluten Häufigkeiten** der Augensummenzahlen von Simulation und Modell vorgenommen. Da bei der Simulation mit den TI-Würfeln 864mal gewürfelt wurde, müssen im Modellversuch die absoluten Häufigkeiten mit 4 multipliziert werden: aus  $864 = 216 \cdot 4$  folgt daher für die **Augensumme 10**  $27 \cdot 4 = 108$ .

**Vergleich von Simulation (2. Zeile) und Modell (3. Zeile) bei 864 Würfeln**

<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>
1	9	24	54	51	77	103	120	119	94	77	54	46	25	9	1
4	12	24	40	60	84	100	108	108	100	84	60	40	24	12	4

**Tabelle der relativen Häufigkeiten für den Modellversuch (n = 216)**

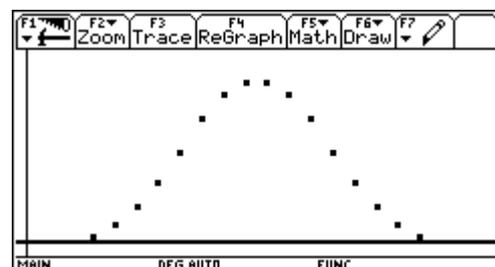
<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>
1/n	3/n	6/n	10/n	15/n	21/n	25/n	27/n	27/n	25/n	21/n	15/n	10/n	6/n	3/n	1/n

Nach der Eingabe der relativen Häufigkeiten im Daten-Editor und der Definition des Graphiktyps SCATTER, ergibt sich im Graphik-Fenster das folgende Bild. Die Punkte liegen im Modell streng symmetrisch in Glockenform. Um die passende **Dichtefunktion  $nv(x, \mu, \sigma)$**  zu erhalten, ist wiederum  $\mu$  und  $\sigma$  zu bestimmen.

	F1	F2	F3	F4	F5	F6	F7
	Plot	Setup	Cell	Header	Calc	Util	Stat
DATA	c1	c2	c3	c4			
1	3	1/216	1/72	.260417			
2	4	1/72	1/18	.586806			
3	5	1/36	5/36	.840278			
4	6	5/108	5/18	.9375			
5	7	5/72	35/72	.850694			
6	8	7/72	7/9	.607639			
7	9	25/216	25/24	.260417			

**c2=listc2**

$c3 = c1 \cdot c2$                        $c4 = (c1 - \mu)^2 \cdot c2$



c1 als x-Werte, c2 als y-Werte

Der **Erwartungswert  $\mu$**  ergibt sich als **Summe der c3-Spalte** und die **Varianz  $\sigma^2$**  als **Summe der c4-Spalte**.

F1		F2		F3		F4		F5		F6		F7	
P1		STAT VARS										Stat	
DATA	c1	$\bar{x}$	=	6.5625									
1	3	$\Sigma x$	=	10.5								0417	
2	4	$\Sigma x^2$	=	10.693287								6806	
3	5	$Sx$	=	5.03499								0278	
4	6	nStat	=	16.								75	
5	7	minX	=	.013889								0694	
6	8	q1	=	1.875								7639	
7	9	medStat	=	5.90278								0417	
c3=c1*c2													

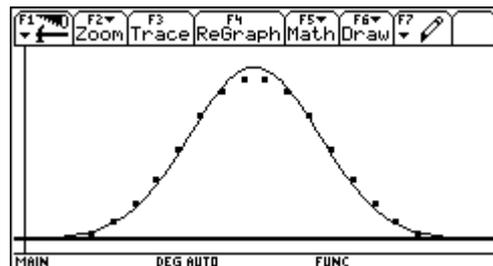
$$\mu = \Sigma x = 10.5$$

F1		F2		F3		F4		F5		F6		F7	
P1		STAT VARS										Stat	
DATA	c1	$\bar{x}$	=	5.46875									
1	3	$\Sigma x$	=	8.75								0417	
2	4	$\Sigma x^2$	=	6.31766								6806	
3	5	$Sx$	=	3.19636								0278	
4	6	nStat	=	16.								75	
5	7	minX	=	.03125								0694	
6	8	q1	=	2.60417								7639	
7	9	medStat	=	5.97222								0417	
c4=(c1-10.5)^2*c2													

$$\sigma^2 = \Sigma x = 8.75$$

F1		F2		F3		F4		F5		F6		F7	
Zoom		Edit		All		Style							
▲ PLOTS													
Plot 2: $x:c1$ b:1													
Plot 1: $x:c2$ b:1													
$y1=1/6$													
$y2 = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-.5 \cdot \left(\frac{x - \mu}{\sigma}\right)^2}$													
y3=nv   $\mu=10.5$ and $\sigma=8.75$													
y4=													
y3(x)=nv   $\mu=10.5$ and $\sigma=8.75$													

$$nv(x, \mu, \sigma)$$



Gaußsche Glockenkurve

Die **Bedeutung der Normalverteilung** begründet sich darin, dass bei vielen praktischen Anwendungen die i. a. unbekannte Wahrscheinlichkeitsverteilung der auftretenden Zufallsgrößen sehr gut durch die Normalverteilung beschrieben werden kann.

„Ein gemeinsames Merkmal solcher Zustandsgrößen besteht darin, dass sie sich durch Überlagerung einer großen Anzahl weitgehend voneinander unabhängiger zufälliger Effekte ergeben, wobei jedoch jeder dieser Effekte nur einen - im Verhältnis zur Summe der anderen - unbedeutenden Einfluß auf die betrachtete Zustandsgröße besitzt.“

G. MAIBAUM (1971, S.187)

Mit Hilfe von **DERIVE** können Aufgabenstellungen zur Normalverteilung ohne Transformation und Tabelle **direkt mit der Integralrechnung** gelöst werden. Es sollte aber immer eine Skizze mit der GAUSSschen Glockenkurve für die Dichtefunktion  $nv(x, \mu, \sigma)$  verlangt werden, wobei der Flächeninhalt als Maß für die Wahrscheinlichkeit zu schraffieren ist.

Beim **TI-92** kann ganz analog wie bei **DERIVE** vorgegangen werden, wobei die **Wahrscheinlichkeit als Flächeninhalt im Graphikfenster** möglichst oft dargestellt werden sollte. Es ist aber im Graphik-Fenster zu beachten, dass zur Berechnung nur endliche Grenzen verwendet werden dürfen. Die passenden Werte für die WINDOW-Einstellung erhält man mit Hilfe des Tabellen-Fenster.

Beim **TI-92 mit 128 kB** liefert das Integrieren mit Grenzen wie  $+\infty$  und  $-\infty$  im HOME-Fenster nicht immer ein Ergebnis, wohl aber mit endlichen Grenzen. Da bei der numerischen Berechnung des Integrals die Größe des verfügbaren Speichers eine wichtige Rolle spielt, ist vor allem beim TI-92 mit 128 kB zu empfehlen, um die Rechenzeit zu verkürzen, häufig die standardisierte Dichtefunktion  $nv(x, 0, 1)$  mit den transformierten Grenzen zu nehmen.

### 13. Eine Musteraufgabe zur Normalverteilung

Automotoren haben eine mittlere Lebensdauer von 100000km mit einer Standardabweichung von 20000km.

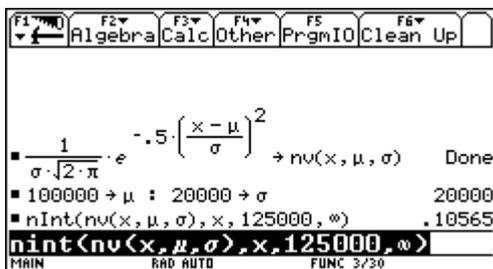
- Wieviel Prozent der erzeugten Motoren haben eine Lebensdauer von mindestens 125000km ?
- Bei wieviel Prozent der Motoren weicht die Lebensdauer um mehr als 15000km vom Mittelwert ab ?
- Auf wieviel km muß durch Verbesserung des Motors seine mittlere Lebensdauer erhöht werden, damit bei gleichbleibenden  $\sigma$  mindestens 80% der Motoren 90000km laufen ?

Mache bei allen drei Rechnungen zuerst eine Skizze mit Hilfe der Glockenkurve!

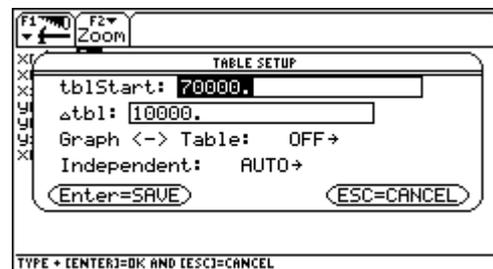
Aufgabe aus FIALA / MOSER, Mathematik Maturaufgaben, Nummer W 19, Seite 314, Verlag Hölder-Pichler-Tempsky, Wien, 3. Auflage (1. Auflage 1982).

Die Zufallsvariable  $X$  gibt die **Lebensdauer der Automotoren** in km an.  $X$  wird als normalverteilt mit  $\mu = 100000$  und  $\sigma = 20000$  angenommen.

a)  $P(X > 125000) = ?$  Lösung mit dem TI-92+ ohne Standardisierung



numerische Integration ( $\infty$  obere Grenze)



Einstellung der Tabelle

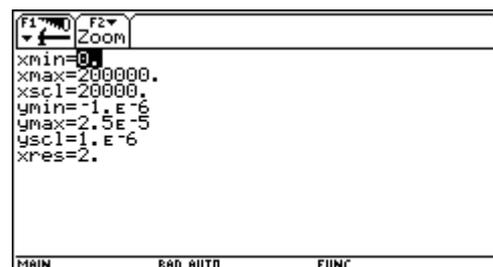
Es erreichen somit ca. 10.6% der erzeugten Motoren mindestens 125000 km.

Der TI-92 ermöglicht diese berechnete Fläche graphisch darzustellen. Um die richtige Fenstereinstellung für das Graphik-Fenster zu finden, ist es günstig,  $nv(x, \mu, \sigma)$  zu tabellieren.

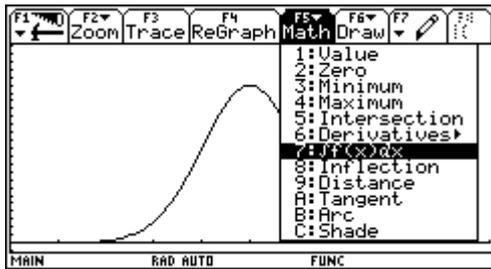
x	y1
70000.	.000006
80000.	.000012
90000.	.000018
100000.	.00002
110000.	.000018
120000.	.000012
130000.	.000006
140000.	.000003

x=70000.

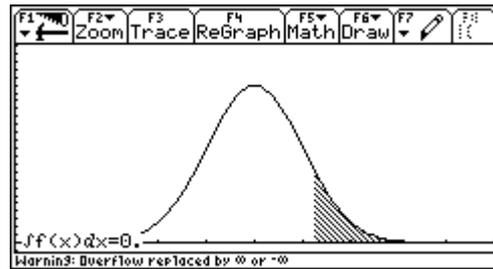
Ablesung der Tabellenwerte



Einstellung des Graphik-Fensters



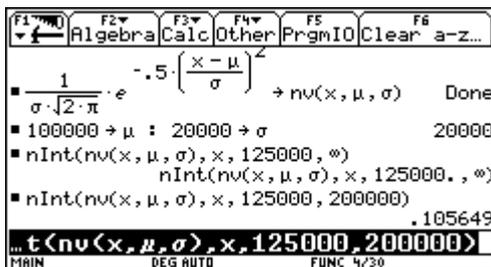
Aufruf zur Flächeninhaltsberechnung:  
Eingabe [125000, 200000]



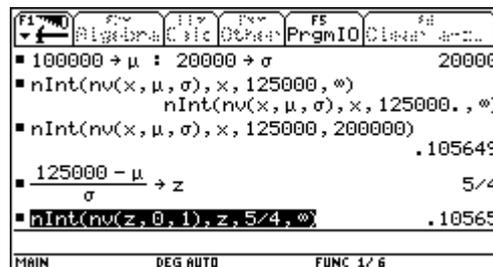
Die Fläche wird richtig schraffiert, kann  
aber zahlenmäßig nicht berechnet werden.

a)  $P(X > 125000) = ?$

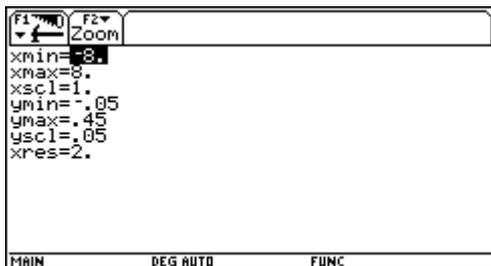
Lösung mit dem TI-92 (128 kB)



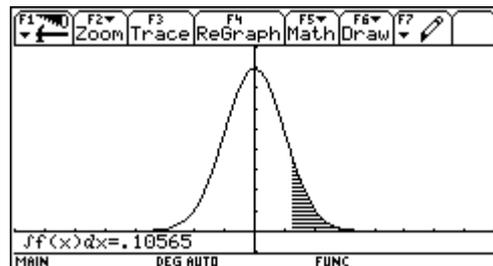
Lösung ohne Standardisierung



Lösung mit Standardisierung



symmetrische Einstellung des G-Fensters

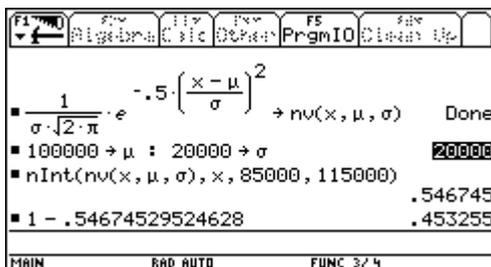


Schraffierung und richtige Berechnung

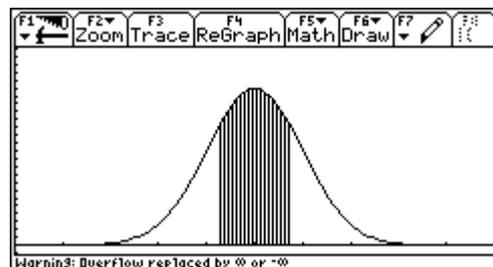
Für Schraffierung u. Berechnung im Graphik-Fenster wurden als Grenzen 1.25 und 8 gewählt.

b)  $1 - P(85000 < x < 115000) = ?$

Lösung mit dem TI-92+ ohne Standardisierung



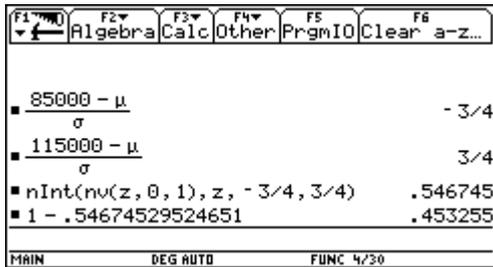
Lösung im HOME-Fenster



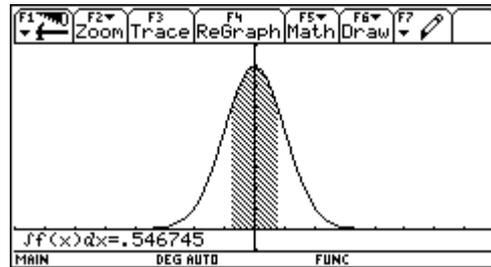
nur Schraffierung im Graphikfenster

Bei ca. 45.3% weicht die Lebensdauer der Motoren um mehr als 15000 km ab.

**b)  $1 - P(85000 < x < 115000) = ?$  Lösung mit dem TI-92 (128 kB) mit Standardisierung**



transformierte Grenzen:  $[-3/4, +3/4]$

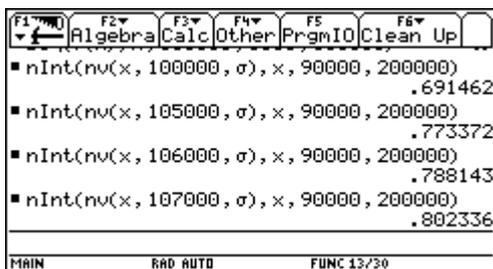


numerische Lösung im Graphik-Fenster

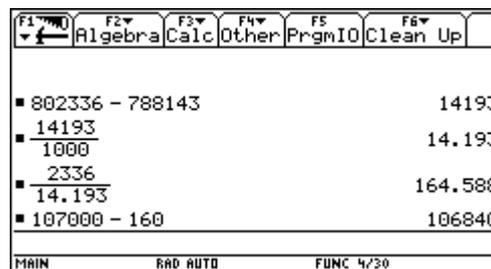
Die Berechnung konnte im HOME- und GRAPHIK-Fenster durchgeführt werden.

**c) Für welches  $\mu$  ist  $P(X > 90000) = 0.80$  ? ( $\sigma$  bleibt 20000)**

Lösungsweg mit dem TI-92+ durch gezieltes Probieren mit  $nv(x, \mu, \sigma)$  und interpolieren.  
 Es können vom TI-92 für verschieden angenommene  $\mu$  wie 100000, 105000, 106000, ... die Flächeninhalte von 90000 bis  $\infty$  (abgeändert auf 200000) berechnet werden, bis 0.80 knapp unter- bzw. überschritten wird. Anschließend ist linear zu interpolieren.

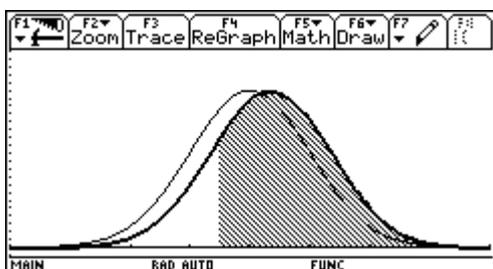


Berechnung der Flächeninhalte

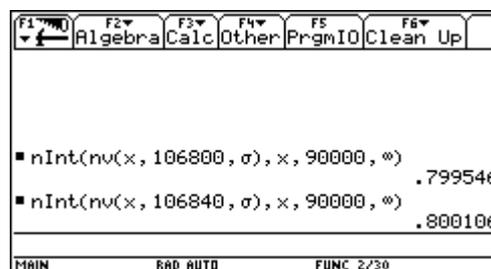


lineare Interpolation

Die mittlere Lebensdauer der Motoren müßte auf ca. 106840 km erhöht werden.



$nv(x, 100000, \sigma)$  und  $nv(x, 106800, \sigma)$  sind übereinander gelegt und die gesuchte Fläche von 0,80 schraffiert,



Durch die Berechnung des Flächeninhaltes von 90000 bis  $\infty$  wird noch einmal überprüft, ob 0,80 knapp überschritten wird.

## 14 Beispiele zur Normalverteilung

Die folgenden Beispiele sind eine exemplarische Zusammenstellung von Aufgaben zur Wahrscheinlichkeitsrechnung, die zum Teil aus Lehrbüchern der AHS-Oberstufe entnommen ist und für deren Lösung eine Näherung durch die Normalverteilung verwendet wird.

Auf die Behandlung von Standardaufgaben, die in allen Lehrbüchern zu finden sind und die auch im klassischen Mathematikunterricht immer wieder gerechnet werden, wurde hier bewusst verzichtet. Es wurden nur solche Beispiele aufgenommen, die ohne CAS System nicht oder nur schwer lösbar sind.

### Bei der Lösung der Beispiele finden die folgenden Strategien Verwendung :

- Direktes Lösen einer Gleichung mit dem Befehl solve oder nsolve
- Ansatz der Gaußfläche direkt als Integral bzw. als numerisch berechnetes Integral mit dem Befehl nInt.
- Verwendung der Tabellenfunktion des Rechners und Suchen einer Lösung.
- Grafische Darstellung einer Funktion und Ablesen des Ergebnisses.
- Verwendung der Standardisierungsformel.

Alle diese Techniken werden an einigen Beispielen vorgestellt. Oft führen mehrere Wege zum Ziel, jedoch mit unterschiedlichem Aufwand bzw. mit erheblich differierenden Rechenzeiten. Welche Methode gewählt wird bleibt dem Benutzer vorbehalten.

### Verwendete Formeln und Funktionen

Binomialverteilung mit den Parametern n, k und p	bv(n,k,p)
Summe über mehrere Werte für k	bvsum(n,p,a,e) wobei a und e den Anfangswert und den Endwert der Summation darstellen
Gauß'sche Formel der Normalverteilung	nv(x,μ,σ)
Gauß'sche Fläche	nvfl(μ,σ,a,e) wobei a und e den Anfangswert und den Endwert der Integration darstellen
Gauß'sche Fläche – numerisch berechnet	nvfln(μ,σ,a,e)

Alle diese Funktionen können entweder direkt im Home Screen oder mit dem Funktioneneditor eingegeben werden.

### 1. Beispiel : Ein Einstiegsbeispiel zur Binomialverteilung.

Wie groß ist die Wahrscheinlichkeit, daß beim 1000-maligen Würfeln mindestens 120 und höchstens 155 6er vorkommen?

Wir rechnen  $bvsum(1000,1/6,120,155)$ ; der Rechner benötigt **mehr als 2 Minuten**, das Ergebnis lautet 0,171892.

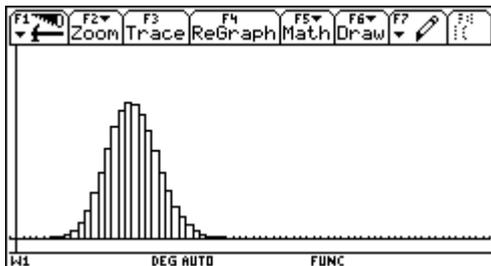
Wir untersuchen die gleiche Aufgabe mit  $n = 200$  und dem Intervall  $[30,45]$ . Das Ergebnis von 0,75 benötigt etwa 20s.

Für Rechnungen mit großen Zahlen ist die Summenbildung eine sehr zeitaufwendige Methode – der Rechner ist zu langsam. Man sucht daher eine bessere Variante.

Zur Veranschaulichung rechnen wir nochmals für  $n = 100$ ,  $I = [10,18]$

$$bvsum(n,p,10,18) \rightarrow 0.675$$

Zur besseren Veranschaulichung stellen wir die Verteilung grafisch dar.



Durch die Balkenbreite 1 lässt sich die Wahrscheinlichkeit als Fläche darstellen, man verwendet die bekannte Formel von Gauß.

$$nv(x, \mu, \sigma) = \frac{1}{\sqrt{(2\pi) \cdot \sigma}} \cdot e^{-0,5 \cdot ((x - \mu) / \sigma)^2}$$

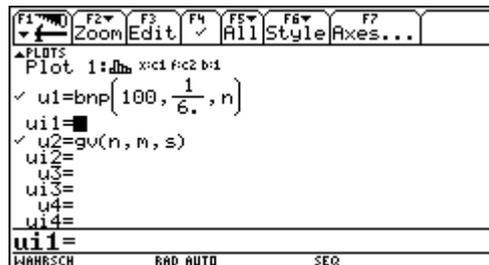
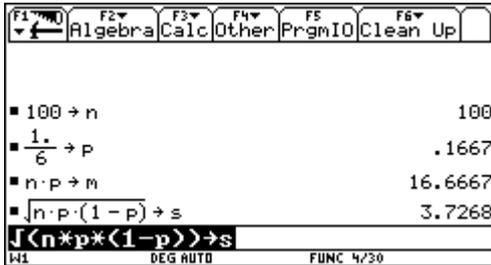
Diese Formel wird in den Rechner eingespeichert und mit  $nv(x,m,s)$  (Normal-Verteilung) bezeichnet.

Die Buchstaben  $\mu$  und  $\sigma$  sind etwas mühsam im Rechner einzugeben. Zur Vereinfachung werden im TI in weiterer Folge immer die Buchstaben **m** und **s** verwendet.

Wir wollen nun die Güte dieser Formel untersuchen. Dazu schalten wir aus dem Funktionsmodus in den Sequence Mode um.

Zuerst legen wir im Home-Screen einige Parameter fest, um die Eingabe später zu vereinfachen.

Es sei  $n = 100$  und  $p = 1/6$ . Dann geben wir die Funktionen im *Sequencemode* ein, erstellen eine Tabelle und lassen die Funktionen grafisch darstellen.



n	u1	u2			
10.000	.0214	.0216			
11.000	.0350	.0337			
12.000	.0520	.0489			
13.000	.0703	.0660			
14.000	.0874	.0829			
15.000	.1002	.0969			
16.000	.1065	.1053			
17.000	.1052	.1066			

n=10.  
DEG AUTO SEQ

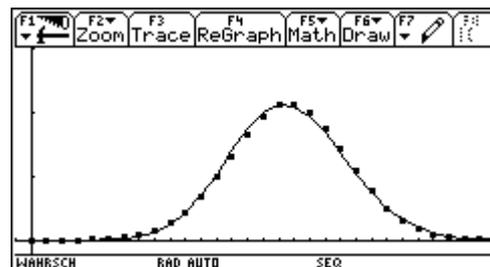
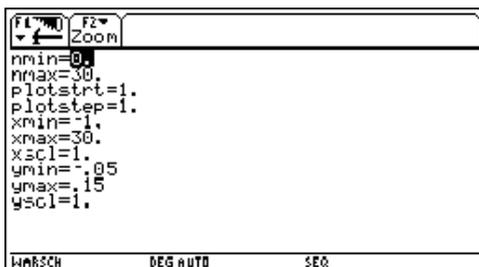
Start : 10

Δtbl : 1

u1 : Line

u2 : Sqare

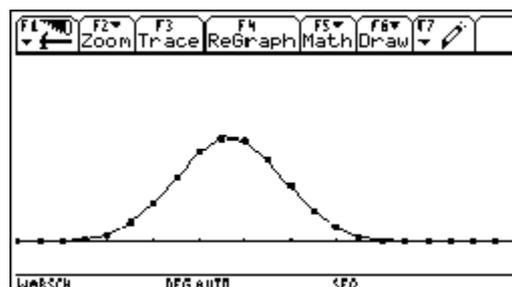
Dabei stellt die Funktion  $u1$  die Werte der Binomialverteilung und  $u2$  die Werte der Normalverteilung dar. Dabei ist zu beachten, dass im Sequencemode der Parameter stets mit  $n$  bezeichnet werden muss.



Für  $n = 1000$  ergibt sich mit dem Startwert 100 und  $\Delta tbl : 10$  (Dauert sehr lange !!!!!)

n	u1	u2			
120.	7.7E-6	.00001			
130.	.00021	.00027			
140.	.00249	.00262			
150.	.01263	.01245			
160.	.02926	.02885			
170.	.03226	.03252			
180.	.01753	.01785			
190.	.00485	.00477			

n=180.  
DEG AUTO SEQ



d.h. Die Treppenfunktion der Binomialverteilung kann durch eine stetige Funktion angenähert werden. Diese Annäherung ist umso besser, je größer  $n$  ist.

**Aus der Summe der Einzelwahrscheinlichkeiten wird die Fläche unter der Kurve.**

## 2. Beispiel :

Wie groß ist die Wahrscheinlichkeit, beim 100-maligen Würfeln mindestens 10 und höchstens 18 6er zu erhalten?

Zuerst rechnen wir als Binomialverteilung, dann als Normalverteilung.

The screenshot shows a TI-84 Plus calculator screen with the following data:

Calculation	Result
$\frac{1}{6} \rightarrow p$	.1667
$n \cdot p \rightarrow m$	16.6667
$\sqrt{n \cdot p \cdot (1 - p)} \rightarrow s$	3.7268
$bvsum(n, p, 10, 18)$	.6752
$\int_{10}^{18} nv(x, m, s) dx$	.6029

The bottom of the screen shows the function definition:  $f(nv(x, m, s), x, 10, 18)$ . The status bar at the bottom indicates "DEG AUTO" and "SEQ 6/30".

Zur Vereinfachung der Berechnung definiert man die Funktion

$$\int (nv(x, \mu, \sigma), x, a, e) \rightarrow nvfl(\mu, \sigma, a, e) \quad \text{Gauß-Fläche}$$

Damit erhält man für das obige Beispiel :  $nvfl(\mu, \sigma, 10, 18) \rightarrow 0.6029$

Der Einsatz der Stetigkeitskorrektur bringt dabei eine deutliche Verbesserung des Ergebnisses

$$\underline{nvfl(\mu, \sigma, 9.5, 18.5) \rightarrow 0.6614}$$

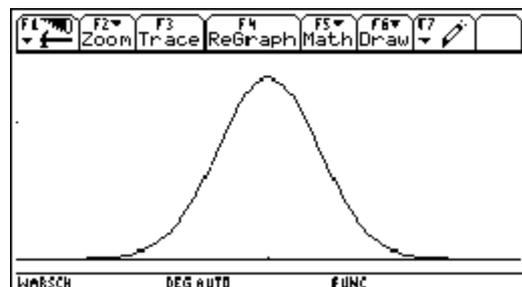
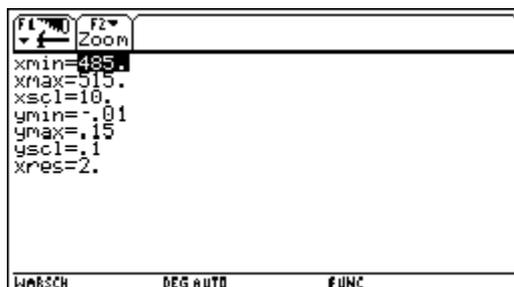
### 3. Beispiel :

Eine Abfülleinrichtung ist auf eine Abfüllmenge von 500g eingestellt. Die Standardabweichung beträgt 3g. Stelle die Verteilungsfunktion grafisch dar.

Berechne die Wahrscheinlichkeit, dass die Abfüllmenge

- höchstens 506g beträgt,
- nicht unter 497g sinkt,
- um nicht mehr als 3g vom Mittelwert abweicht. Stelle den Bereich grafisch dar !
- Wie groß ist die Toleranz  $d$  zu wählen, damit 90% der abgefüllten Pakete in diesem Bereich liegen.

Für die grafische Darstellung – zur besseren Veranschaulichung - definiert man zuerst die Variablen  $m$  und  $s$  im Homescreen und die Funktionswerte der Normalverteilung im Funktionen - Editor  $nv(x,m,s) \rightarrow y1(x)$  und lässt diese Funktion mit den angegebenen



Fensterwerten zeichnen.

a)  $\mu = 500, \sigma = 3$

$nvfl(\mu, \sigma, -\infty, 506) = 0,97725$

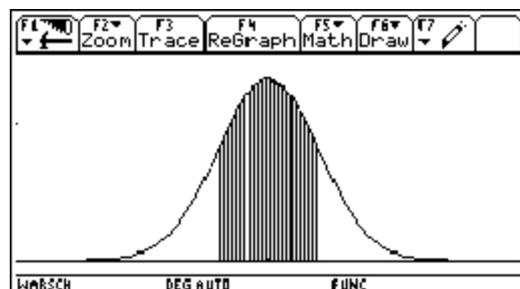
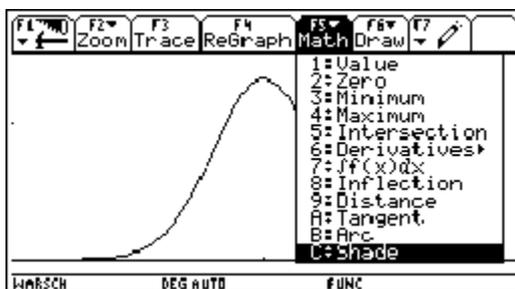
Dieser Ansatz ist zwar formal richtig, praktisch aber sinnlos und führt auch im Rechner immer wieder zu kryptischen Fehlermeldungen. Außerdem dauert die Berechnung sehr lange.

Bessere Ansätze sind  $nvfl(\mu, \sigma, 0, 506)$  oder  $nvfl(\mu, \sigma, 480, 506)$ . Letzterer Ansatz berücksichtigt eine  $\sigma$ -Grenze und liefert sehr rasch ein brauchbares Ergebnis.

b)  $nvfl(\mu, \sigma, 497, \infty) = 0,84135 = \triangleq 84,14\%$  oder besser  $nvfl(\mu, \sigma, 497, 520)$

c)  $nvfl(\mu, \sigma, 497, 503) = 0,97725$

Für die Grafik wählt man  $F5/Shade$ ,



bestätigt *Above X axis* und gibt die untere und obere Grenze ein.

d) Ansatz :  $nvfl(\mu, \sigma, 500 - d, 500 + d)$

$nsolve(\dots = 0.9, d)$  ergibt für  $d$  den Wert  $d = 4,93g$

Diese Rechnung braucht relativ lange (ca. 4 min).

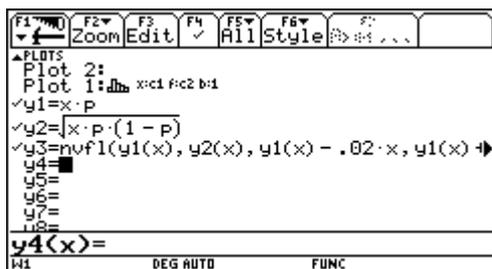
Eine andere Möglichkeit wäre die Verwendung einer Tabelle und das Herantasten an die Lösung. Diese Technik wird in den folgenden Beispielen verwendet.



Es können aber auch *andere Lösungswege* gewählt werden.

- **Versuch mit der Tabelle**

Im Funktionen-Editor muss als Variable immer x verwendet werden. Es soll die Stichprobengröße berechnet werden, somit wird diese mit x angesetzt. Definiert man die notwendigen Parameter sowie die Funktion *nvfl* wie angegeben, so erhält man diese Tabelle :



x	y1	y2	y3
1400.	700.	18.7083	.865519
1500.	750.	19.3649	.878665
1600.	800.	20.	.890401
1700.	850.	20.6155	.900902
1800.	900.	21.2132	.910314
1900.	950.	21.7945	.918764
2000.	1000.	22.3607	.926362
2100.	1050.	22.9129	.933202

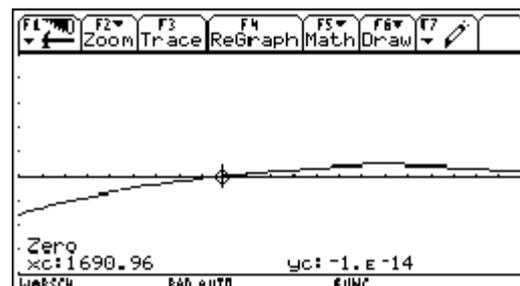
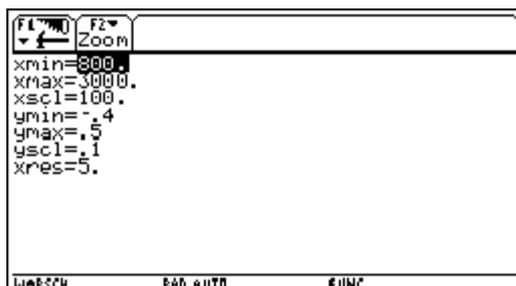
x=1400.

Daraus lässt sich eine **Stichprobengröße von etwa 1700** ablesen.

Führt man dieselbe Aufgabe für  $p = 0.2$  durch, so erhält man eine Stichprobengröße von  $n = 1100$ . (Übung)

- **Grafische Lösung**

Wir definieren die Funktion  $y4(x) = y3(x) - 0,90$ , und lassen diese Funktion mit geeigneten Fensterparametern darstellen. Von dieser Funktion suchen wir die Nullstellen.



Wir erhalten das "gleiche Ergebnis" wie oben.

*Probleme:* Der Rechner zeichnet sehr lange. Wenn man den richtigen Bereich nicht erraten kann, dann ist diese Variante eher ungünstig.

## 5. Beispiel :

Ein Meinungsforschungsinstitut befragt 2000 zufällig ausgewählte Personen und stellt fest, dass 31,5% davon für die Partei A sind. Mit 95,5%-iger Sicherheit soll eine Aussage über den Anteil  $p$  der A-Sympathisanten in der wahlberechtigten Bevölkerung gemacht werden.

$X$  sei der Anteil der A-Freunde unter den 2000 befragten.  $X$  wird als binomialverteilt mit den Parametern  $n=2000$  und  $p = 0,315$  angenommen.

Wir berechnen  $m$  und  $s$  und jenes symmetrische Intervall  $[m-d, m+d]$  für das gilt  $W(m-d < X < m+d) = 0,955$ .

Die Lösung der Integralgleichung mit dem Befehl *nSolve* ist sehr aufwendig, innerhalb einer vernünftigen Zeit erhält man kein Ergebnis.

F1	F2	F3	F4	F5	F6
Algebra	Calc	Other	PrgmIO	Clean Up	
<ul style="list-style-type: none"> <li>▪ .315 → p <span style="float: right;">.315</span></li> <li>▪ 2000 → n <span style="float: right;">2000</span></li> <li>▪ n · p → m <span style="float: right;">630.</span></li> <li>▪ <math>\sqrt{n \cdot p \cdot (1-p)} \rightarrow s</math> <span style="float: right;">20.77</span></li> <li>▪ <math>\text{nvf1}(m, s, m-d, m+d)</math></li> <li>▪ <math>.0192 \cdot \int_{630.-d}^{d+630.} ((.9988)(x-630.)^2) dx</math></li> <li>▪ <math>\text{nvf1}((x-630.)^2), x, 630., -d, d+630.)</math></li> </ul>					
<small>W1 DEG AUTO FUNC 5/30</small>					

Wir versuchen einen Ansatz über die Tabelle. Da wir nicht wissen, was herauskommt setzen wir für die Tabellenparameter vorerst einmal grobe Werte an und erhalten :

F1	F2	F3	F4	F5	F6
Algebra	Calc	Other	PrgmIO	Clear	a-z...
<p style="text-align: center;">TABLE SETUP</p> <ul style="list-style-type: none"> <li>tblStart: 0.</li> <li>tbl: 10.</li> <li>Graph &lt;-&gt; Table: OFF→</li> <li>Independent: AUTO→</li> </ul> <p style="text-align: center;">Enter=SAVE      ESC=CANCEL</p> <p>... lve(gf1(m, s, m-d, m+d) = .955, d)</p> <p style="font-size: small;">TYPE = CENTER) = OK AND ESC = CANCEL</p>					

F1	F2	F3	F4	F5	F6
Setup	Cell	Mode	Defl Pos	Ini Pos	
x	43				
0.	0.				
10.	.369751				
20.	.664328				
30.	.851297				
40.	.945834				
50.	.979106				
60.	.979106				
70.	.979106				
x=0.					
Warning: Overflow replaced by ∞ or -∞					

Der gesuchte Wert liegt somit zwischen 40 und 50. Damit lässt sich die Tabelle verfeinern.

F1	F2	F3	F4	F5	F6
Setup	Cell	Mode	Defl Pos	Ini Pos	
x	43				
40.	.945834				
41.	.951577				
42.	.956801				
43.	.96154				
44.	.96583				
45.	.969703				
46.	.973194				
47.	.976331				
x=40.					
WARSCH DEG AUTO FUNC					

Man erkennt :  $d = 42$ .

Somit gilt  $W(630-42 \leq X \leq 630+42) = 0,955$  bzw.  **$W(588 < X < 672) = 0,955$** , d.h. zu **95.5%** liegt der Anteil der A-Wähler zwischen **29,4%** und **33,6%**.

Sehr hilfreich kann auch der Umweg über die **Standardisierung** sein.

F1	F2	F3	F4	F5	F6
Algebra	Calc	Other	PrgmIO	Clean Up	
<ul style="list-style-type: none"> <li>▪ <math>\text{nvf1}(0, 1, -d, d) = .955</math></li> <li>▪ <math>\frac{\sqrt{2} \cdot \int_0^d ((.6065)^{x^2}) dx}{\sqrt{\pi}} = .955</math></li> <li>▪ <math>\text{nvf1}(0, 1, -d, d) = .955</math></li> </ul>					
<small>W1 DEG AUTO FUNC 1/30</small>					

Löst man diese Gleichung mit *nSolve* nach  $d$  und substituiert man zurück erhält man das gleiche Ergebnis wie oben.

Die benötigte Zeit ist dabei wesentlich kürzer.

## 6. Beispiel :

Vor einer Bundespräsidentenwahl, bei der nur zwei Personen kandidieren, wird unter 3000 Personen eine Meinungsbefragung durchgeführt, von denen 1512 den Kandidaten A bevorzugen.

a) Kann mit 99%-iger Sicherheit gesagt werden, dass A die absolute Mehrheit erreichen wird?

b) In welchem Bereich wird der Anteil von A zu 99% liegen?

c) Mit mindestens wie vielen Stimmen kann A mit 99%iger Sicherheit rechnen?

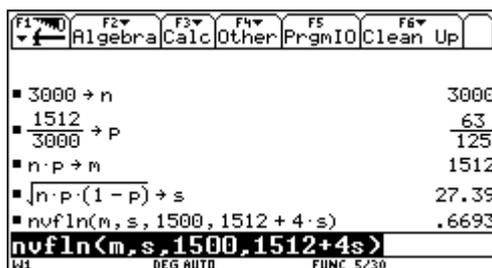
a) Es ist zu untersuchen, ob folgendes gilt  $W(X \geq 1500) \geq 0,99$

Zuerst werden die Parameter eingegeben.

Der Ansatz **nvfl(m,s,1500,1512+4s)** wäre zwar logisch, man kommt aber nicht recht weiter, weil der Rechner zu lange braucht – außerdem treten beim Lösen der Gleichung Fehlermeldungen (Overflow) auf.

Ein anderer Ansatz könnte mit Hilfe der **numerischen Integration** funktionieren.

$$\mathbf{nInt(nv(t,m,s),t, a, b) \rightarrow nvfln(m,s,a,b)}$$



3000 → n	3000
1512/3000 → p	63/125
n · p → m	1512
$\sqrt{n \cdot p \cdot (1 - p)} \rightarrow s$	27.39
nvfln(m, s, 1500, 1512 + 4 · s)	.6693
nvfln(m,s,1500,1512+4*s)	

Funktioniert innerhalb vernünftiger Zeit.

Nimmt man als obere Grenze **1512+4s**, so ist man auf jeden Fall im sicheren Bereich.

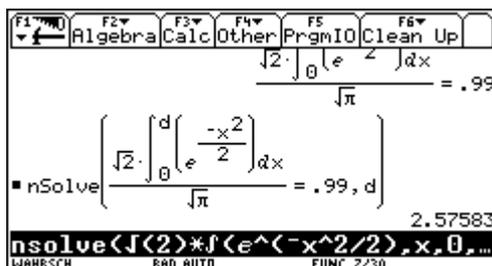
**Kandidat A wird nur mit 67%-iger Wahrscheinlichkeit die absolute Mehrheit erreichen.**

b) Gesucht ist wieder ein symmetrisches Intervall um den Erwartungswert, so dass gilt:

$$\underline{\underline{W(m-t < X < m+t) = 0.99}}$$

- Eine mögliche Vorgangsweise wäre wie im obigen Beispiel mit Hilfe der Tabelle.

- Ein andere Ansatz wäre der Versuch über die **Standardisierung**:  $W(-d < X < d) = 0.99$   
d.h. im TI  $\mathbf{nvfl(0,1,-d,d) = 0.99}$



$\frac{\sqrt{2} \cdot \int_0^d (e^{-x^2/2}) dx}{\sqrt{\pi}} = .99$
nSolve
$\frac{\sqrt{2} \cdot \int_0^d (e^{-x^2/2}) dx}{\sqrt{\pi}} = .99, d$
2.57583
nsolve(sqrt(2)*int(e^(-x^2/2),x,0,...

Löst man diese Gleichung mit **nSolve** nach d, so erhält man

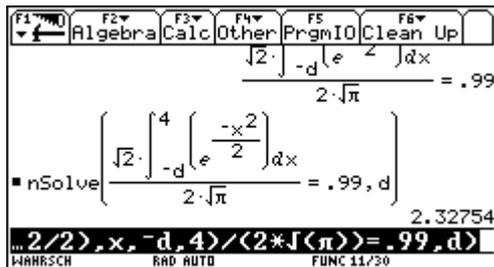
$$\mathbf{d = 2.57583}$$

Nach Rücksubstituieren ergibt sich für die untere Grenze **1441** (1441.46) und für die obere Grenze **1583** (1582.54).

Die Anzahl der A-Wähler liegt somit mit 99%-iger Sicherheit zwischen 1441 und 1583 d.h. zwischen 48,03% und 52,76%. Es kann mit 99%iger Sicherheit keine absolute Mehrheit vorhergesagt werden.

c) Zu Lösen ist die Gleichung  $W(X \geq t) = 0.99$

- Ein Ansatz mit Hilfe der Standardisierung könnte etwa so aussehen :  $\text{nvfl}(0,1,-d,4) = 0.99$ , wobei 4 wieder für die obere Grenze ( $4 \cdot \sigma$ ) steht.



Der Befehl **nSolve** liefert dieses Ergebnis, dauert allerdings relativ lange.

$$d = 2.33$$

Durch Zurücksostituieren ergibt sich als unterer Wert 1448.26, was 48.27% entspricht.

**Mit 99%-iger Sicherheit erhält Kandidat A mehr als 48,27% der Stimmen.**

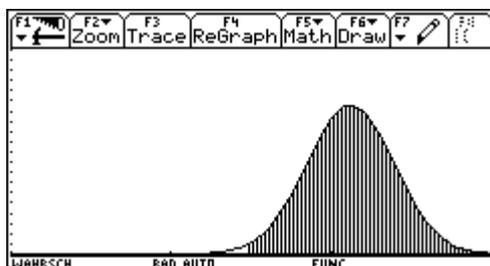
- Eine andere Lösungsmöglichkeit wäre wieder mit Hilfe einer Tabelle.

x	y				
-3.	.99862				
-2.9	.9981				
-2.8	.99741				
-2.7	.9965				
-2.6	.99531				
-2.5	.99376				
-2.4	.99177				
-2.3	.98924				

x = -2.3

Herantasten an die Lösung durch Verfeinern der Tabelle !!

- Grafische Darstellung



F5/Shade

## 7. Beispiel :

In einem Wahlkreis werden 600 Wahlberechtigte hinsichtlich ihres beabsichtigten Wahlverhaltens befragt, von denen sich 32 für die Partei Z entscheiden. Mit welcher Wahrscheinlichkeit kann prognostiziert werden, ob diese Partei die 5% Hürde überspringen wird?

Zu berechnen ist die Wahrscheinlichkeit  $W(X \geq 0.05 \cdot n)$

- Erster Versuch:  $n=600$ ,  $m$  und  $s$  berechnen und die Flächenfunktion ansetzen.

F1	F2	F3	F4	F5	F6
Algebra	Calc	Other	PrgmIO	Clean	Up
600 → n					600
$\frac{32}{600} \rightarrow p$					.0533
$n \cdot p \rightarrow m$					32.
$\sqrt{n \cdot p \cdot (1-p)} \rightarrow s$					5.504
$nvf1(m, s, .05 \cdot n, n)$					Error: Break
<b><math>nvf1(m, s, .05n, n)</math></b>					

Dauert sehr lange, ergibt auch Fehlermeldungen - *Overflow* !

- Zweiter Versuch mit der **Standardisierung**.

F1	F2	F3	F4	F5	F6
Algebra	Calc	Other	PrgmIO	Clean	Up
$\sqrt{n \cdot p \cdot (1-p)} \rightarrow s$					5.504
$nvf1(m, s, .05 \cdot n, n)$					Error: Break
$\frac{.05 \cdot n - m}{s} \rightarrow t1$					-.3634
$\frac{n - m}{s} \rightarrow t2$					103.2
$nvf1(0, 1, t1, t2)$					Error: Break
$nvf1n(0, 1, t1, t2)$					.6418
<b><math>nvf1n(0, 1, t1, t2)</math></b>					

Berechnung der neuen Grenzen.

Liefert einen *Overflow* und kein Ergebnis !

Mit der numerischen Integration erhalten wir den Wert 0,6418 .

- Dritter Versuch:  $nvf1(0, 1, t1, t2) = nvf1(0, 1, t1, 0) + nvf1(0, 1, 0, t2)$

F1	F2	F3	F4	F5	F6
Algebra	Calc	Other	PrgmIO	Clean	Up
$\frac{.05 \cdot n - m}{s} \rightarrow t1$					-.3634
$\frac{n - m}{s} \rightarrow t2$					103.2
$nvf1(0, 1, t1, t2)$					Error: Break
$nvf1n(0, 1, t1, t2)$					.6418
$nvf1(0, 1, t1, 0)$					.1418
$.14183805620933 + .5$					.6418
<b><math>.14183805620933 + .5</math></b>					

Da  $t2$  sehr groß ist, kann für den zweiten Summanden 0.5 angenommen werden. Der erste Summand wird vom Rechner blitzschnell berechnet – man erhält insgesamt den Wert 0.6418.

Geht sehr schnell !!

**Die Partei Z überspringt die 5% Hürde zu 64%.**

## Abschließende Bemerkungen

Gerade über dieses Thema ist schon sehr viel geschrieben worden. Hier lässt sich mit Hilfe des Rechners die Qualität und die Wirklichkeitsnähe der gestellten Aufgaben stark verbessern. Beispiele, die früher wegen ihres großen Rechenaufwandes nicht gegeben oder behandelt werden konnten, sind hier vielfach kein Problem mehr. Auch die Einschränkung auf kleine Zahlen fällt weg, allerdings wird durch die technische Konzeption des Rechners bald seine Kapazitätsgrenze erreicht.

Ähnliche Beispiele finden sich in vielen Ergänzungen zu Schulbüchern und persönlichen Abhandlungen, sowie in den Artikeln auf der ACDCA - Homepage <http://www.acdca.ac.at>

## Literatur

APETSBERGER, K. / SCHLÖGELHOFER, F.: Der TI-92 im Mathematikunterricht. Texas Instruments, 1996.

KIRSCHENHOFER, W. / ARNOLD, H.: Aufgabensammlung zur Wahrscheinlichkeitsrechnung mit didaktischen Beiträgen, Band II. Institut für Mathematik, Universität Linz, 1979.

KOTH, M.: 100 Maturaaufgaben mit vollständigen Lösungen. Verlag Hölder-Pichler-Tempsky, Wien, 1993.

BRENNER, J. / LESKY, P. / VOGEL, A.: Wahrscheinlichkeitsrechnung und Statistik. Verlag Ernst Klett, Stuttgart, 1980.

ENGEL, A.: Wahrscheinlichkeitsrechnung und Statistik, Band 1. Ernst Klett Verlag, Stuttgart, 1973.

FIALA, F. / MOSER, W.: Mathematik Maturaaufgaben. Verlag Hölder-Pichler-Tempsky, Wien, 1982.

FISCHER, R.: Darstellen und Analysieren von Daten und Beziehungsstrukturen. In: LEITNER, L. / BENEDIKT, E. (Hgb.): Mathematik AHS-Oberstufe Kommentar. Österreichischer Bundesverlag, Wien, 1991.

KRAFT/BÜRGER/UNFRIED/HASCHKOVITZ: Mathematische Formelsammlung. Verlag Hölder-Pichler-Tempsky, Wien, 1994.

MAIBAUM, G.: Wahrscheinlichkeitsrechnung. Verlag Volk und Wissen, Berlin, 1971.

MALLE, G.: Wahrscheinlichkeitsrechnung und Statistik in der 7. Und 8. Klasse. In: LEITNER, L. / BENEDIKT, E. (Hgb.): Mathematik AHS-Oberstufe Kommentar. Österreichischer Bundesverlag, Wien, 1991.

REICHEL, H.-C. / HANISCH, G. / MÜLLER, R.: Wahrscheinlichkeitsrechnung und Statistik. Verlag Hölder-Pichler-Tempsky, Wien, 1989.

REICHEL, H.-C. / MÜLLER, R.: Mathematik mit dem TI-92. Verlag Hölder-Pichler-Tempsky, Wien, 1997.

TI-92 Handbuch, Texas Instruments, 1995, 1996.

WURNIG, O.: Von der Binomial- zur Normalverteilung mit Hilfe des Computers. ÖMG Didaktik-Reihe, Heft 21, Wien, 1993.