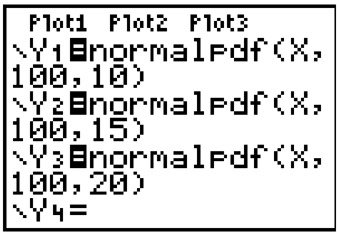
<u>Übersicht Normalverteilung</u> <u>Ac</u>

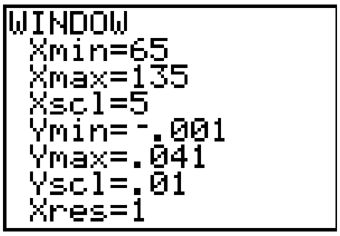
Die Normalverteilung wird unter anderem verwendet, um

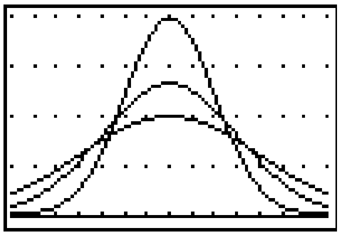
a) die Binomialverteilung zu approximieren und
b) die Verteilung von Zufallsgrößen in Naturwissenschaften, Medizin etc. zu beschreiben .

Die Normalverteilung ist definiert durch $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0{,}5\left(\frac{x-\mu}{\sigma}\right)^2}$

Diese Funktion f hat die Eigenschaften $f(x) \geq 0$ und $\int_{-\infty}^{\infty} f(x)dx = 1$, und zwar <u>unabhängig</u> vom Erwartungswert μ und der Standardabweichung σ (siehe Grafik).

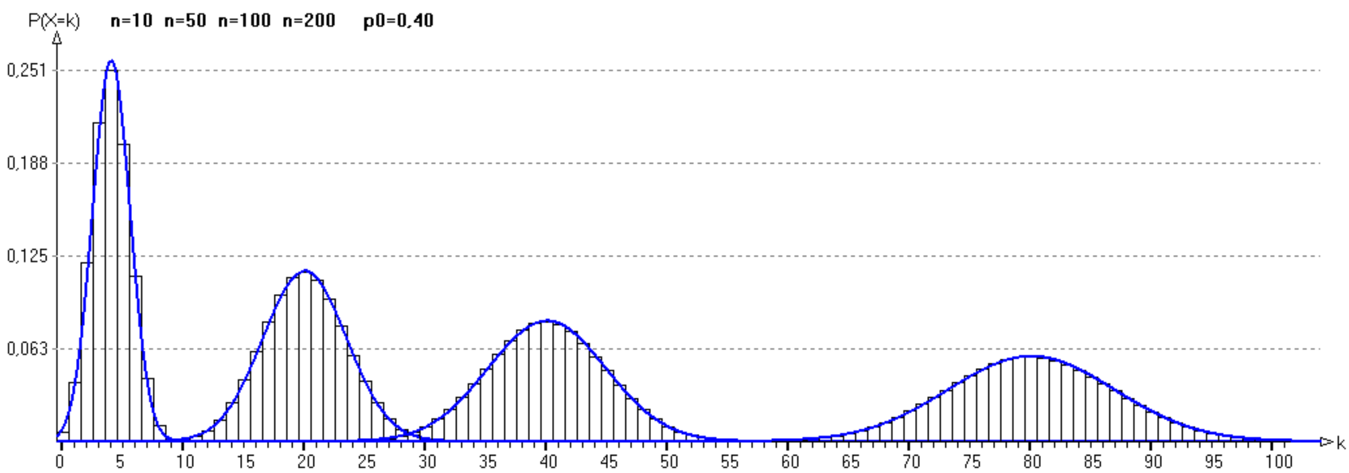






In der Grafik gilt stets μ=100, aber σ nimmt die Werte 10, 15, 20 an .
Man erkennt, dass σ ein Maß für die Breite der Normalverteilung ist.

Zu a): Für eine Binomialverteilung gilt: $\mu = n \cdot p$ und $\sigma = \sqrt{n \cdot p \cdot (1-p)}$.
Hier hat man also für verschiedene n und p auch verschiedene Normalverteilungen (siehe Grafik unten) .
Mit zunehmendem n werden auch μ und σ größer.
Sind 2 der Größen n, μ, σ bekannt, so kann man die dritte Größe errechnen (siehe Formeln oben) !



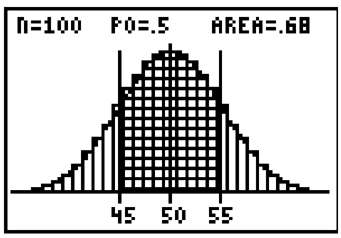
Für die Verteilung ganz links gilt μ=4 und σ=1,55.
Es folgen: μ=20; σ=3,46 μ=40; σ=4,9 μ=80; σ=6,93

In der σ-Umgebung [μ-σ ; μ+σ] liegen ca. 68,3% der gesamten Verteilung .

Dies zeigt auch die TI-Grafik rechts .
Die σ-Umgebung ist dort
[50-5 ; 50+5] = [45 ; 55] .

Der Flächeninhalt ist
$P(45 \leq X \leq 55) = 0{,}729$
(mit binomcdf berechnet) und

$P(45 \leq X \leq 55) = 0{,}6827$
(mit normalcdf berechnet,
ohne Stetigkeitskorrektur).



area=.68 bezieht sich hier auf die Normalverteilung !

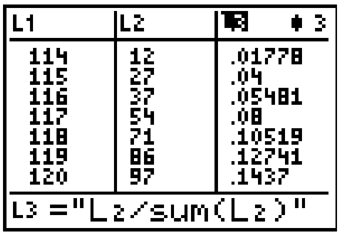
Zu b): Hier ist die Sache ganz anders als bei der Approximation der Binomialverteilung, denn man ermittelt die Normalverteilung anhand von Messwerten (Häufigkeiten), z.B. den Körpergrößen von Menschen. Eine Anzahl n liegt in der Regel nicht vor ! Für die Häufigkeiten berechnet man zunächst den Mittelwert $\bar{x}$ sowie die Standardabweichung σ . Diese setzt man in die Formel $\frac{1}{\sigma\sqrt{2\pi}} e^{-0,5(\frac{x-\bar{x}}{\sigma})^2}$ bzw. **normalpdf(X, μ, σ)** ein, man verwendet also $\bar{x}$ als Schätzwert für μ .

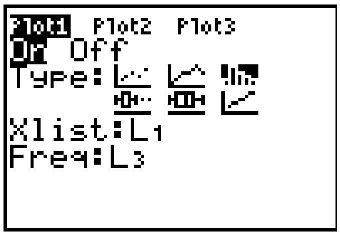
Schließlich betrachtet man die Güte der Approximation .
μ und σ hängen also hier <u>nicht</u> von n und p ab, sondern von den gegebenen Messwerten !

<u>Beispiel Körpergröße von Kindern:</u>

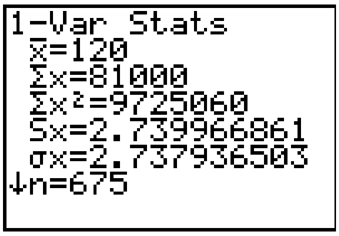
Größe in cm	abs. Häufigkeit
114	12
115	27
116	37
117	54
118	71
119	86
120	97
121	86
122	76
123	55
124	39
125	24
126	11

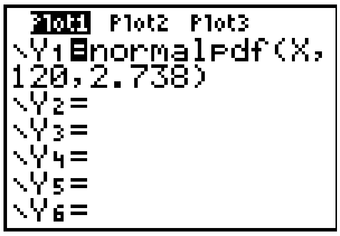
Gesucht ist eine „passende“ Normalverteilung.

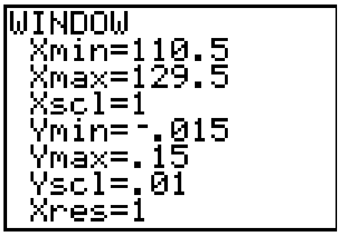




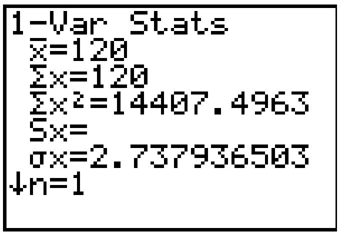
1-Var Stats L1,L2
liefert:



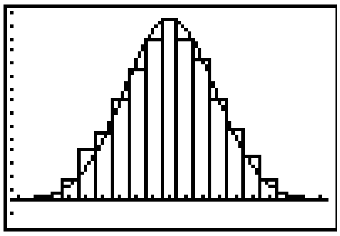




Alternativ:
1-Var Stats L1,L3



(hier fehlt SX, da **<u>rel.</u>** H. verarbeitet wurden)



<u>Berechnung einiger Wahrscheinlichkeiten:</u>

(Die Vorgehensweise ist laut Schroedel-Buch ähnlich wie bei der Approximation der Binomialverteilung; d.h.: es wird mit einer „Stetigkeitskorrektur“ gerechnet.
Allerdings gibt es hier im Gegensatz zur Binomialverteilung auch „gebrochene“ k- bzw. x-Werte !)

<u>Vereinbarung:</u> x=119cm bedeutet $x \in [118,5cm ; 119,5cm[$!!

P(kleiner als 120cm) = P(X≤119,5) = normalcdf(0,119.5,120,2.738) = 42,8%
P(kleiner als 120,5cm) = P(X≤119,5) = normalcdf(0,119.5,120,2.738) = 42,8%
P(größer als 117cm) = P(X≥117,5) = 1- P(X≤117,5) = 1-normalcdf(0,117.5,120,2.738) = 81,9%
P(größer als 117,5cm) = P(X≥117,5) = 1- P(X≤117,5) = 1-normalcdf(0,117.5,120,2.738) = 81,9%
P(größer als 116,5cm) = P(X≥116,5) = 1- P(X≤116,5) = 1-normalcdf(0,116.5,120,2.738) = 89,9%
P(zwischen 114cm und 123cm) = P(114,5≤X≤122,5) = normalcdf(114.5,122.5,120,2.738) = 79,7%