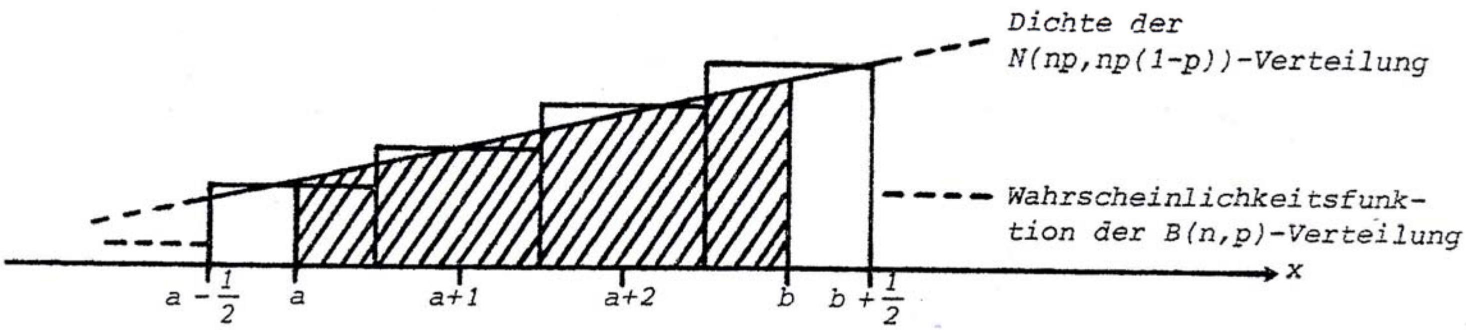
Eine Zufallsvariable X, die einer B(n,p)-Verteilung genügt, ist näherungsweise N(np,np(1-p))-verteilt:

$$(*)\quad P(a \le X \le b) \approx \Phi\left(\frac{b-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right) \qquad \text{für } a \le b$$

Diese Formel ist außerordentlich wichtig für die praktische Berechnung von Wahrscheinlichkeiten, die mit der Binomialverteilung zusammenhängen. Dabei ist die angegebene Näherung umso besser, je größer der Index n und je näher der Erfolgsparameter p dem Wert $\frac{1}{2}$ ist. Der dabei auftretende Approximationsfehler läßt sich häufig wesentlich verringern, indem man eine sogenannte *Stetigkeitskorrektur* vornimmt und (*) ersetzt durch

$$(**)\quad P(a \le X \le b) \approx \Phi\left(\frac{b+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a-\frac{1}{2}-np}{\sqrt{np(1-p)}}\right) \qquad \text{für } a \le b$$

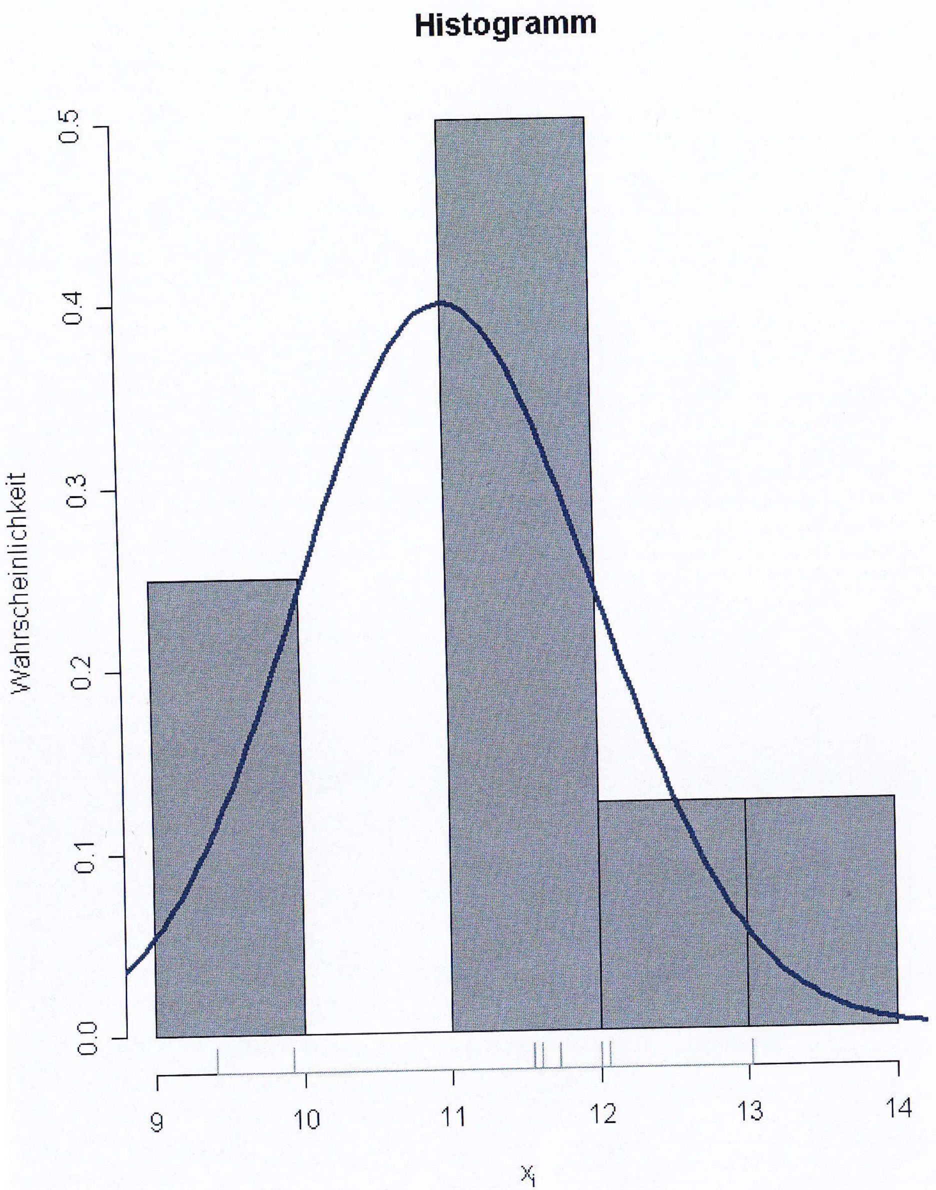
Zur Begründung dieser Stetigkeitskorrektur diene die folgende Skizze:

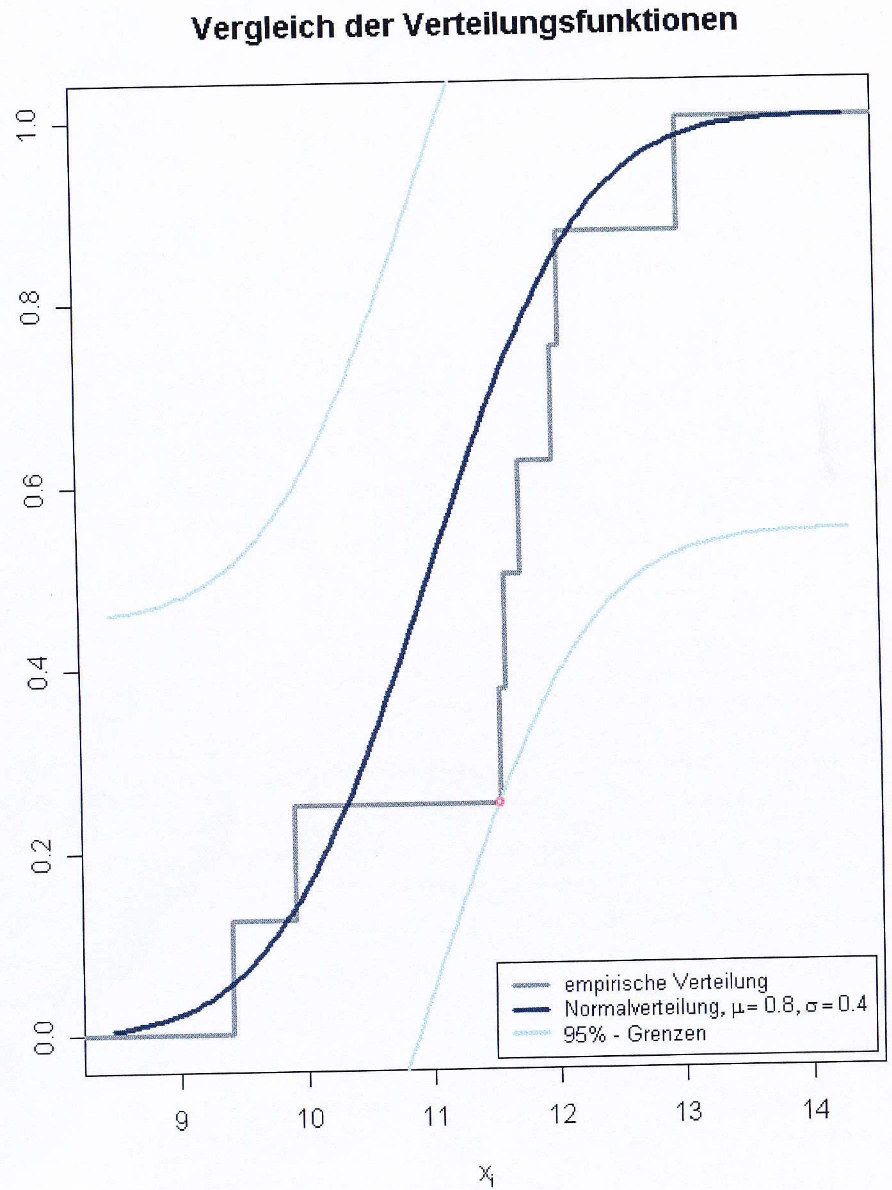


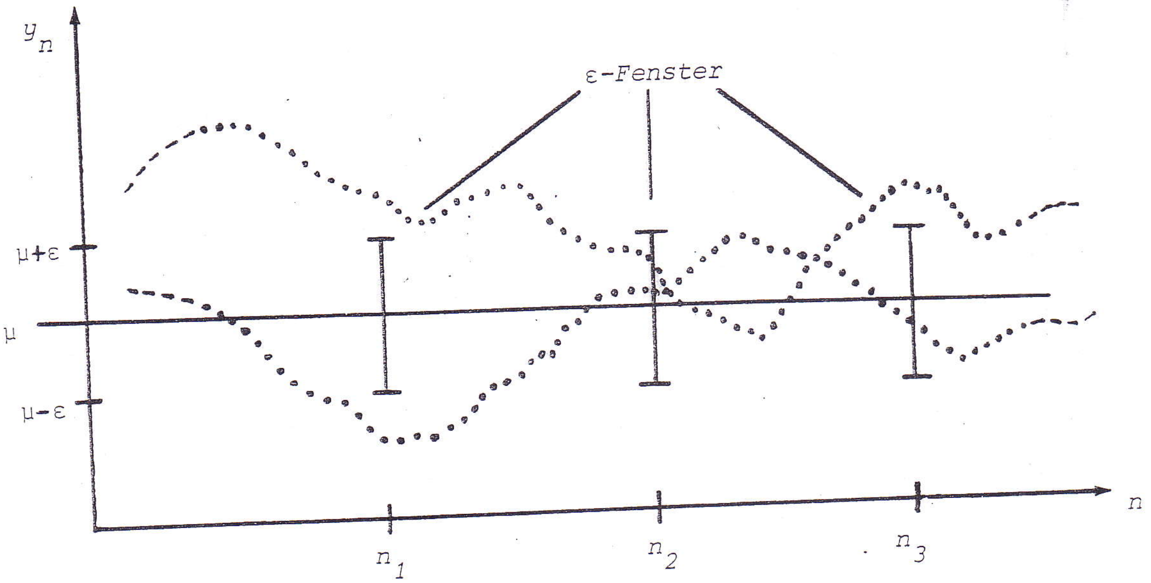
Seien a und b (a<b) zwei natürliche Zahlen, die von X mit positiver Wahrscheinlichkeit angenommen werden können. Dann wird die Wahrscheinlichkeit $P(a \le X \le b)$ durch den Gesamt-Flächeninhalt der vier gezeichneten Rechtecke dargestellt. Bei der Normalapproximation (*) wird als Näherung für diese Wahrscheinlichkeit der Flächeninhalt unter der N(np,np(1-p))-Dichte zwischen den senkrechten Begrenzungslinien in a und b verwendet (schraffierte Fläche). Die Approximation (**) nimmt zum schraffierten Flächeninhalt noch die Flächen unter der Dichte hinzu, die durch die Senkrechten in $a-\frac{1}{2}$ und a bzw. b und $b+\frac{1}{2}$ begrenzt werden und ist damit im allgemeinen eine Verbesserung der Normalapproximation (*).

Ein Gefühl für die numerische Qualität der Normalapproximationen (*) und (**) vermittelt das nachfolgende Beispiel.

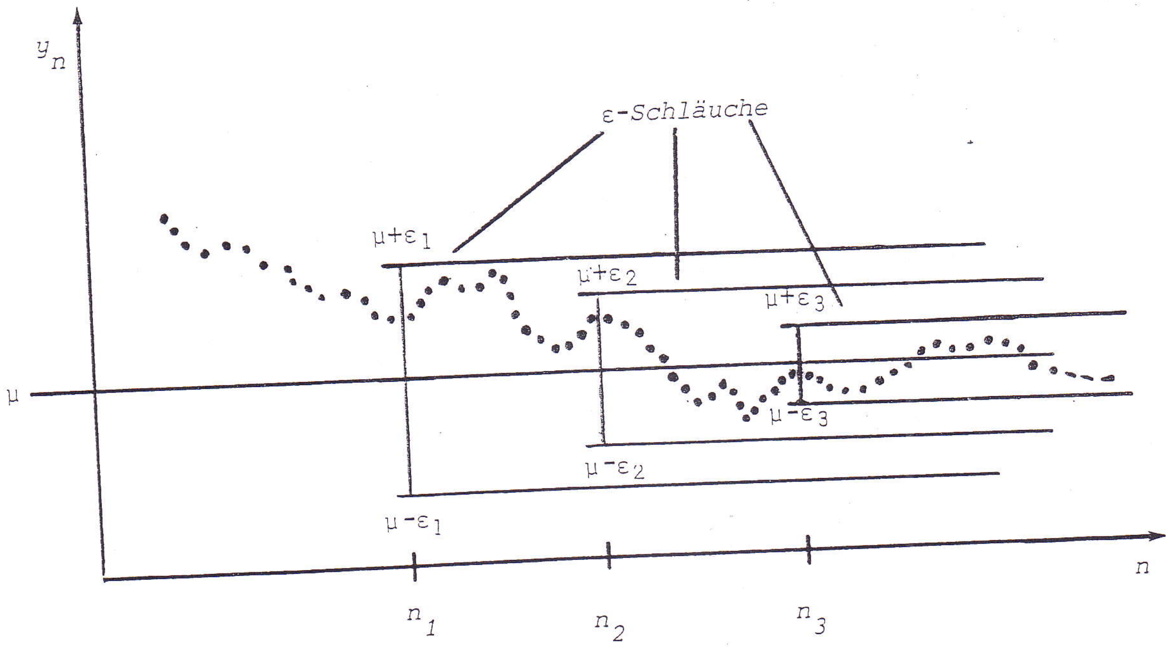
Binomialverteilung	Wahrscheinlichkeit	exakter Wert	(*)-Normalapproximation	(**)-Normalapproximation
$B(5,0.2)$	$P(1\leq X\leq 4)$	0.67200	0.49960	0.71187
$B(7,0.4)$	$P(2\leq X\leq 4)$	0.74511...	0.55418	0.74723
$B(20,0.5)$	$P(8\leq X\leq 12)$	0.73681...	0.62891	0.73645
$B(50,0.1)$	$P(4\leq X\leq 9)$	0.72522...	0.65165	0.74330
$B(100,0.01)$	$P(2\leq X\leq 5)$	0.26370...	0.15741	0.30765
$B(100,0.3)$	$P(12\leq X\leq 14)$	0.00015...	0.00020	0.00033
$B(100,0.3)$	$P(31\leq X\leq 33)$	0.23013...	0.15729	0.23405
$B(200,0.5)$	$P(95\leq X\leq 105)$	0.56325...	0.52050	0.56332
$B(500,0.1)$	$P(50\leq X\leq 55)$	0.31762...	0.27197	0.32357

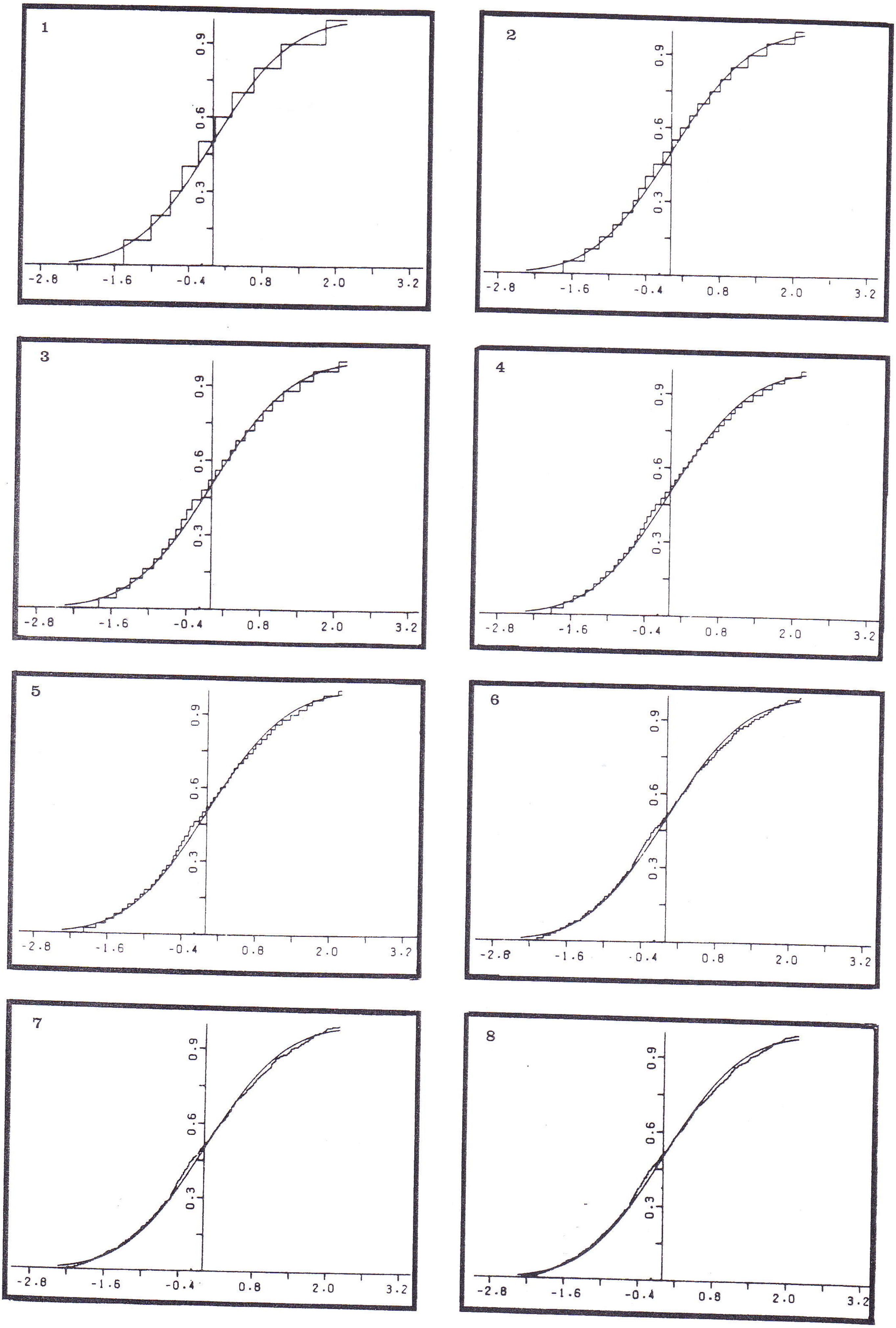

Histogramm
Wahrscheinlichkeit
0.0
0.1
0.2
0.3
0.4
0.5
9
10
11
12
13
14
x_i


Vergleich der Verteilungsfunktionen
0.0
0.2
0.4
0.6
0.8
1.0
9
10
11
12
13
14
x_i
empirische Verteilung
Normalverteilung, $\mu = 0.8$, $\sigma = 0.4$
95% - Grenzen



Diese Möglichkeit des Pfades, (mit abnehmender aber stets positiver Wahrscheinlichkeit) sich immer wieder mal von μ um mehr als ε zu entfernen, wird durch die betrachteten schwachen Gesetze der großen Zahlen nicht ausgeschlossen. Eine solche Situation würde nun doch nicht ganz unserer Vorstellung von einem sich mit wachsendem n stabilisierenden arithmetischen Mittel (bzw. einer sich stabilisierenden relativen Häufigkeit) entsprechen. Wir stellen uns unter einer solchen Stabilisierung eher die folgende Situation vor:




1
2
3
4
5
6
7
8
0.3
0.6
0.9
-2.8
-1.6
-0.4
0.8
2.0
3.2