# Stochastik

### Grundbegriffe der Wahrscheinlichkeitsrechnung

### Zufallsexperiment

z.B.: "Das Glücksrad wird zweimal Gedreht."

### **Ergebnis**

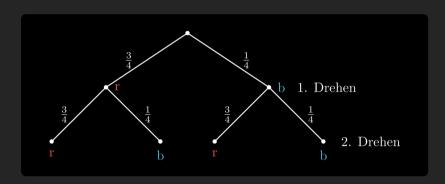
z.B.: "zuerst Rot, dann Blau"



### Ergebnismenge

$$S = \{rr, rb, br, bb\}$$

### **Baumdiagramm**



### **Ereignis**

z.B.: "mindestens einmal blau drehen"

### **Ereignismenge**

 $E = \{rb, br, bb\}$ 

### Gegenereignis

 $\overline{E}$ : "nie blau drehen"  $\overline{E}=\{rr\}$ 

 $P(E) = 1 - P(\overline{E}) = 1 - P(rr) = 1 - \frac{9}{16} = \frac{7}{16}$ 

### **Produktregel**

 $P(rb) = P(r) \cdot P(b) = \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16}$ 

### Summenregel

 $P(E)=P(rb)+P(br)+P(bb)=rac{7}{16}$ 

### Zufallsvariable

X: "Gewinn in €"

Xkann die Werte -1,0,4 annehmen

### Wahrs chein lich keits verteilung

Gewinn $k$	-1		4
P(X=k)	$\frac{9}{16}$	$\frac{7}{16}$	$\frac{1}{16}$

### Erwartungswert von X

$$E(X) = -1 \cdot rac{9}{16} + 0 \cdot rac{7}{16} + 4 \cdot rac{1}{16} = -rac{5}{16} \quad \Rightarrow \ pprox -0.31(\epsilon)$$

### Varianz

$$\mathrm{Var}(X) = (-1 - (-\frac{5}{16}))^2 \cdot \frac{9}{16} + (0 - (-\frac{5}{16}))^2 \cdot \frac{7}{16} + (4 - (-\frac{5}{16}))^2 \cdot \frac{1}{16} \approx 1{,}465(\epsilon^\circ)$$

### Standardabweichung

$$\sigma = \sqrt{\mathrm{Var}(X)} pprox \sqrt{1,465} = 1,21(\epsilon)$$

### Ziehen mit Zurücklegen



Werden aus einer Urne mit n verschiedene Kugeln, k Stück herausgezogen, nach jedem Zug wieder zurückgelegt und in ihrer Reihenfolge notiert, so kann die Anzahl an verschiedenen **Reihenfolgen** berechnet werden mit:

n'

### **Beispiel**

 $\hbox{Ein W\"urfel mit $6$ Seiten wird $2$ mal gew\"urfelt. Wie viele m\"ogliche Zahlenpaare gibt es? }$ 

Lösung:

$$6^2 = 36$$

### Ziehen ohne Zurücklegen



Werden aus einer Urne mit n verschiedenen Kugeln, k Stück nacheinander herausgezogen ohne diese zurückzulegen und in ihrer Reihenfolge notiert, so kann die Anzahl an verschiedenen **Reihenfolgen** berechnet werden mit:

$$\frac{n!}{(n-k)!}$$

#### Beispiel

 $Ein \ Wettrennen \ hat \ 8 \ Teilnehmer. \ Für \ die \ schnellsten \ 3 \ gibt \ es \ Preise. \ Wie \ viele \ M\"{o}glichkeiten \ gibt \ es, \ die \ Preise \ zu \ vergeben?$ 

Lösung:

$$\frac{8!}{(8-3)!} = \frac{8!}{5!} = \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 8 \cdot 7 \cdot 6 = 336$$

 $\Rightarrow$  Es gibt 336 verschieden Möglichkeiten die Preise zu vergeben

### Ziehen mit einem Griff



Werden aus einer Urne mit n verschiedenen Kugeln, k Stück mit einem Griff (gleichzeitig) herausgezogen, so kann man die Anzahl verschiedener **Möglichkeiten** berechnet werden mit:

$$\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}$$

Taschenrechner:  $n \operatorname{nCr} k$ 

#### Spezialfälle

$$egin{pmatrix} n \ 0 \end{pmatrix} = egin{pmatrix} n \ n \end{pmatrix} = 1$$

$$egin{pmatrix} n \ 1 \end{pmatrix} = egin{pmatrix} n \ n-1 \end{pmatrix} = n$$

#### **Beispiel**

Ein Verein hat 10 Mitglieder. Eine Mannschaft besteht aus 4 Spielern. Wie viele mögliche Mannschaften hat der Verein?

Lösung:

$$\binom{10}{4} = \frac{10!}{(10-4)! \cdot 4!} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot \underline{6 \cdot 5 \cdot 4 \cdot 8 \cdot 2 \cdot 1}}{\underline{6 \cdot 5 \cdot 4 \cdot 8 \cdot 2 \cdot 1} \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1} = 210$$

 $\Rightarrow$  Der Verein hat 210 mögliche Mannschaften



Ein Trick zum berechnen per Hand ist es, die k größten Zahlen aus n! miteinander zu multiplizieren und durch k! zu teilen. In dem Beispiel oben kommt man dadurch schnell auf den letzten Term mit je 4 Faktoren im Zähler und Nenner. Diese kann man dann oft noch weiter kürzen.

### Erwartungswert, Varianz und Standardabweichung

9

Für eine Zufallsvariable X, die die Werte  $x_1,x_2,\ldots,x_n$  annehmen kann, definiert man den **Erwartungswert**  $\mu$  von X als:

$$E(X) = \sum_{i=1}^n x_i \cdot P(X=x_1)$$

 ${\cal E}(X)$  gibt an, welcher Wert für X im Durchschnitt auf lange Sicht zu erwarten ist.

Ein Spiel heißt **fair**, wenn der Gewinn = 0 ist.

9

Die **Varianz** 

$$\operatorname{Var}(X) = \sum_{i=1}^n (x_i - E(X))^2 \cdot P(X = x_i)$$

und die Standardabweichung

$$\sigma = \sqrt{\mathrm{Var}(X)}$$

sind Maße für die durchschnittliche Abweichung von X zum Erwartungswert.

#### **Beispiel**

### **Bedingte Wahrscheinlichkeit**



Für zwei Ereignisse A und B mit  $P(A) \neq 0$  ist die **bedingte Wahrscheinlichkeit**, die Wahrscheinlichkeit für das Eintreten von B unter der Bedingung, dass A schon eingetreten ist.

$$P_A(B) = rac{P(A \cap B)}{P(A)}$$

### **Beispiel**

U: "Einkauf ist unter 10€." K: "Einkauf wird mit Karte bezahlt."

$$P(U)=0.4 \quad P(U\cap K)=0.3$$

$$P_U(K) = rac{P(U \cap K)}{P(U)} = rac{0.3}{0.4} = 0.75$$

⇒ Es werden 75% der Einkäufe unter 10€ mit Karte bezahlt.

### Stochastische Unabhängigkeit

#### **Definition**



Zwei Ereignisse A und B heißen **stochastisch unabhängig**, wenn:

$$P(A \cap B) = P(A) \cdot P(B)$$

#### Satz



Zwei Ereignisse A und B sind genau dann stochastisch unabhängig, wenn:

$$P_A(B) = P(B)$$

Denn:

$$P_A(B) = \frac{P(A \cap B)}{P(A)} = \underbrace{P(A) \cdot P(B)}_{P(A)} = P(B)$$

### **Beispiel**

In einem Betrieb hat 570 Mitarbeiter. Davon fahren  $\frac{2}{3}$  mit einem Pkw zur Arbeit. Von den 190 auswärtigen Mitarbeitern kommen 177 mit einem Pkw. Untersuche F: "Fährt mit einem Pkw" und A: "Kommt von Auswärts" auf stochastische Unabhängigkeit.

$$P(F) = rac{2}{3}$$
 
$$P(A) = rac{190}{570} = rac{1}{3}$$
 
$$P(F \cap A) = rac{177}{570} \approx 0,311$$

$$P(F) \cdot P(A) = \frac{2}{9} \neq P(F \cap A)$$
  
 $\Rightarrow F$  und  $A$  sind stochastisch abhänigig

### Bernoulli-Experiment und Bernoulli-Kette



Ein Zufallsversuch mit genau zwei möglichen Ergebnissen heißt Bernoulli-Experiment.

Wenn man ein Bernoulli-Experiment n-mal wiederholt, sodass die Durchführungen voneinander unabhängig sind, so spricht man von einer **Bernoulli-Kette** der Länge n.

#### **Beispiel**

5-mal eine Münze werfen ist eine Bernoulli-Kette der Länge 5, da es genau zwei Ergebnisse gibt (Kopf oder Zahl) und ein einzelner Münzwurf keinen Einfluss auf die anderen Münzwürfe hat.

### Die Formel von Bernoulli

•

Für eine Bernoulli-Kette der Länge n und Trefferwahrscheinlichkeit p gilt für die Anzahl an Treffern k:

$$P_p^k(X=k) = inom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Man schreibt auch  $B_{n:p}(k)$  und sagt X ist **binomialverteilt**.

9

Für den Erwartungswert E(X) gilt:

$$E(X) = \mu = n \cdot p$$

### **Beispiel**

10-facher Münzwurf

$$n=10$$
  $p=rac{1}{2}$   $X$ : "Anzahl Zahl (binomialverteilt)"

genau 5 mal Zahl

$$\underbrace{P_{\frac{1}{2}}^{10}(X=5)}_{X=0} = \binom{10}{5} \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^{10-5} \approx 0,246$$

höchstens 5 mal Zahl

$$\begin{split} \underbrace{P_{\frac{1}{2}}^{10}(X \leq 5)} &= P_{\frac{1}{2}}^{10}(X = 0) + P_{\frac{1}{2}}^{10}(X = 1) + P_{\frac{1}{2}}^{10}(X = 2) \\ &+ P_{\frac{1}{2}}^{10}(X = 3) + P_{\frac{1}{2}}^{10}(X = 4) + P_{\frac{1}{2}}^{10}(X = 5) \\ &= \sum_{k=0}^{5} P_{\frac{1}{2}}^{10}(X = k) \approx 0,\!623 \end{split}$$

### **Umformungen**

$$P_p^n(X\geq a)=1-P_p^n(X\leq a-1)$$
  $P_p^n(b\leq X\leq c)=P_p^n(X\leq c)-P_p^n(X\leq b-1)$ 

### Standardabweichung einer Binomialverteilung



X ist eine binomialverteilte Zufallsgröße.

$$\sigma = \sqrt{n \cdot p \cdot (1-p)}$$
 $\mathrm{Var}(X) = n \cdot p \cdot (1-p)$ 

### **Beispiel**

$$n=10$$
  $p=0,4$   $\sigma=\sqrt{10\cdot0,4\cdot0,6}pprox1,55$ 

### Zweiseitiger Hypothesentest

### **Beispiel**



Es soll eine Münze überprüft werden:

**Nullhypothese**  $H_0: p=\frac{1}{2}$  (Münze ist in Ordnung)

**Gegenhypothese** (Alternative)  $H_1: p 
eq rac{1}{2}$  (Münze ist verbogen)

**Signifikanzniveau** (maximale Irrtumswahrscheinlichkeit)  $\hat{\alpha}=5\%$ 

Gesuchte Grenzen  $g_1$  und  $g_2$ :

$$P(x \leq g_1) \leq rac{\hat{lpha}}{2} = 2.5\% P(x \geq g_2) \geq rac{\hat{lpha}}{2} = 2.5\%$$

 $q_1$  soll dabei möglichst groß und  $q_2$  möglichst klein sein.

n = 50; X: Anzahl Kopf (binomial verteilt)

#### **Tabellenkalkulation**

#### **Linksseitiger Test:**

 $g_1$  ist die größtmögliche Zahl mit  $P(X \leq g_1) \leq 2{,}5\%$ .

$g_1$	$P^{50}_{\frac12}(X\leq g_1)$	
17	0,016	kleiner als $2,\!5\%$
18	0,032	größer als $2{,}5\%$

$$\Rightarrow g_1 = 17$$

#### **Rechtsseitiger Test:**

 $g_2$  ist die kleinstmöglichste Zahl mit  $P(X \ge g_2) \le 2.5\%$  bzw. die kleinstmöglichste Zahl mit  $P(X \le g_2 - 1) \ge 97.5\%$ 

$g_2 - 1$	$P_{\frac{1}{2}}^{50}(X \le g_2 - 1)$	
31	0,968	kleiner als $97,\!5\%$
32	0,984	größer als $97,\!5\%$

$$\Rightarrow g_2 = 32 + 1 = 33$$



**Ablehnungsbereich der Nullhypothese**: {0; 1; ...; 17} und {33; ...; 50} Die Nullhypothese wird verworfen, wenn maximal 17 oder mindestens 33 mal Kopf geworfen wird (Münze ist verbogen).

**Annahmebereich der Nullhypothese**: {18; ...; 32} Die Nullhypothese wird angenommen, wenn mindestens 18 und höchstens 32 mal Kopf geworfen wird (Münze ist in Ordnung).

Irrtumswahrscheinlichkeit: Die Wahrscheinlichkeit, dass die man denkt, dass die Münze verbogen ist, obwohl sie in Ordnung ist. Sie beträgt:

$$lpha=P_{rac{1}{2}}^{50}(X\leq \underbrace{17}_{q_1})+P_{rac{1}{2}}^{50}(X\geq \underbrace{33}_{q_2})pprox 3,3\%\leq \hat{lpha}$$

### **Einseitiger Hypothesentest**

Stichprobenumfang n und Signifikanzniveau  $\hat{\alpha}$  werden festgelegt.

#### **Linksseitiger Test**

Nullhypothese:  $H_0: p > p_0$ 

Gegenhypothese:  $H_1: p < p_0$ 

**Ablehnungsbereich:**  $A = \{0, 1, \dots, g\}$ , wobei g die größte Natürliche Zahl mit  $P_{p_0}^n(X \leq g) \leq \hat{\alpha}$ 

#### **Rechtsseitiger Test**

Nullhypothese:  $H_0: p \leq p_0$ 

Gegenhypothese:  $H_1: p > p_0$ 

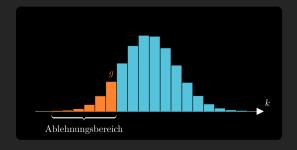
**Ablehnungsbereich:**  $A = \{g, g+1, \dots, n\}$ , wobei g die kleinste natürliche Zahl mit  $P^n_{p_0}(X\geq g)=1-P^n_{p_0}(X\leq g-1)\leq \hat{lpha}$ 

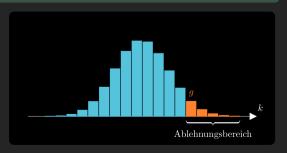


Man führt eine Stichprobe vom Umfang n durch.

#### **Entscheidungsregel:**

Wenn das Stichprobenergebnis im Ablehnungsbereich liegt, wird  $H_0$  verworfen. Ansonsten wird  $H_0$  angenommen.





### Fehler 1. und 2. Art

7	
N	~

	$H_0$ ist wahr	$H_0$ ist falsch
$H_0$ wird verworfen	Fehler 1. Art	
$H_0$ wird nicht verworfen	Fehler 1. Art	
		Fehler 2. Art

- Die Wahrscheinlichkeit für den Fehler 1. Art ist die Irrtumswahrscheinlichkeit  $\alpha$ .
- Die Wahrscheinlichkeit für den Fehler 2. Art ist  $\beta$  und lässt sich nur berechnen, wenn die echte Trefferwahrscheinlichkeit gegeben ist.
- Wenn man n vergrößert und  $\hat{\alpha}$  beibehält, dann wird  $\beta$  kleiner
- Wenn man n beibehält und  $\hat{\alpha}$  verkleinert, dann wird  $\beta$  größer
- $\alpha$  kann verkleinert werden, indem man  $\hat{\alpha}$  verkleinert
- $oldsymbol{\circ} \quad eta$  kann verkleinert werden, indem man n vergrößert



In anderen Worten: Der Fehler 1. Art ist, wenn man im Ablehnungsbereich landet, obwohl die Nullhypothese wahr ist. Der Fehler 2 Art ist, wenn man nicht im Ablehnungsbereich landet, obwohl die Nullhypothese falsch ist.

#### **Beispiel**

Angenommen, eine Firma testet, ob eine neue Batterie eine durchschnittliche Laufzeit von mindestens 10 Stunden hat.

 $H_0$ : Die Batterie hält **weniger** als 10 Stunden.

 $H_1$ : Die Batterie hält **mindestens** 10 Stunden.

#### Fehler 1. Art

Der Test ergibt, dass die Laufzeit **mindestens** 10 Stunden beträgt ( $H_0$  wird verworfen), obwohl sie in Wirklichkeit **weniger als** 10 Stunden beträgt ( $H_0$  ist wahr).

→ Die Firma akzeptiert schlechte Batterien.

#### Fehler 2.Art

Der Test ergibt, dass die Laufzeit **weniger als** 10 Stunden beträgt ( $H_0$  wird angenommen), obwohl sie in Wirklichkeit **mindestens** 10 Stunden beträgt ( $H_0$  ist falsch).

→ Die Firma wirft gute Batterien weg.

#### Berechnen von lpha und eta

$$n = 50 \quad \hat{lpha} = 5\% \ H_0 = p \geq 0, 6 \quad H_1 = p < 0, 6$$

$$P_{0,6}^{50}(X \leq g) \leq 5\%$$
  $\Rightarrow q = 23$ 

Um  $\beta$  zu berechnen brauchen wir die tatsächliche Wahrscheinlichkeit. Ich gebe sie hier als p=0.55 vor.

$$lpha = P_{0.50}^{50}(X \le 23) pprox 0,0314 \ eta = P_{0.55}^{50}(X > 23) = 1 - P_{0.55}^{50}(X \le 23) pprox 0,872$$

### Wahl der Nullhypothese



Bei einem Hypothesentest wird der Fehler, dass die Nullhypothese  $H_0$  aufgrund des Stichprobenergebnisses fälschlicherweise verworfen wird. kontrolliert

Seine Wahrscheinlichkeit ist stets **höchstens** so groß wie das Signifikanzniveau  $\hat{\alpha}$ .

#### 1. Faustregel

Man wählt die Nullhypothese so, dass der Fehler 1. Art derjenige ist, den man vermeiden möchte.

#### 2. Faustregel

Man wählt die Behauptung, die statistisch gestützt werden soll, als Alternative.

### **Beispiel**

Eine Zeitung behauptet, dass sich maximal 85% der Autofahrer angurten. Ein Autoclub bezweifelt dies.

#### Aus der 2. Faustregel folgt:

Der Autoclub möchte zeigen, dass sich mehr als 85% angurten. Daher wählt man  $H_1: p>0.85$ . Daraus folgt  $H_0: p\leq 0.85$ .

#### Dadurch ist die 1. Faustregel auch erfüllt:

Der schlimmere Fehler ist der 1. Art, da der Autoclub der Zeitung vorwirft falsche Informationen zu verbreiten. Beim Fehler 2. Art würde im schlimmsten Fall unnötigerweise eine Kampagne zur Förderung des Anschnallens gestartet werden.



Die 1. Faustregel folgt meist aus der 2. Faustregel. Daher ist es oft am einfachsten, nach der 2. Faustregel zu handeln, da diese auch einfacher ist und man nicht argumentieren muss, welcher Fehler schlimmer ist.

### Normalverteilung und Gaußsche Glockenfunktion

Eine Zufallsgröße X heißt **normalverteilt** mit den Parametern  $\mu$  (Erwartungswert) und  $\sigma$  (Standardabweichung), wenn für zwei reellen Zahlen a und b mit a < b gilt: Die Wahrscheinlichkeit, dass X Werte zwischen a und b annimmt, ist

$$P(a \leq X \leq b) = \int_a^b arphi_{\mu;\sigma}(x) \, dx$$

Man sagt kurz: X ist  $N_{u:\sigma}$ -verteilt. Die Normalverteilung ist eine **stetige** Verteilung.

 $\P$  Man nennt die Funktion  $arphi_{u:\sigma}$  mit

$$arphi_{\mu;\sigma}(x) = rac{1}{\sigma\sqrt{2\pi}} \cdot e^{-rac{(x-\mu)^2}{2\sigma^2}}$$

Gaußsche Glockenfunktion und ihren Graphen Gaußsche Glockenkurve.

- $igg arphi_{\mu;\sigma}$  hat die Extremstelle  $\mu$  und die Wendestellen  $\mu\pm\sigma$  und ist die **Dichtefunktion** einer  $N_{\mu;\sigma}$ -verteilten Zufallsgröße.
- Fs gilt

$$P(X=a) = \int_a^a arphi_{\mu;\sigma}(x)\,dx = 0$$
  
somit gilt auch  
 $P(X \le a) = P(X \le a)$ 

Zusätzlich gilt:

$$arphi_{\mu;\sigma}(x)>0$$
 für alle  $x\in\mathbb{R}$ 

$$\int_{-\infty}^{+\infty} arphi_{\mu;\sigma}(x)\,dx = 1$$

### Beispiele

$$egin{aligned} \mu &= 100 \quad \sigma = 15 \ P(70 \leq X < 110) = \int_{70}^{110} arphi_{100;15} \, dx pprox 0,725 \ P(X < 85) = \int_{-\infty}^{85} arphi_{100;15} \, dx pprox 0,159 \end{aligned}$$

### μ und σ aus gegebenen Daten ermitteln



Gemessene Daten können häufig durch eine Normalverteilung mit Mittelwert  $\mu$  und Standardabweichung  $\sigma$  modelliert werden. Für die Messwerte  $x_1, x_2, \ldots, x_n$  gilt:

$$egin{array}{cccc} \mu &= rac{1}{n} \sum_{i=1}^n x_i & & \sigma = & \hline rac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \ & & igvee \end{array}$$

### Beispiel

Punkte in der nächsten Mathe Klausur:

 Messwerte
 11
 13
 15
 12
 14
 12

$$\mu = \frac{1}{6} \cdot (11 + 13 + \dots + 12)$$

$$= \frac{1}{6} \cdot 77$$

$$\approx 12.83$$

$$\sigma = \sqrt{\frac{1}{6} \cdot ((11 - 12,83)^2 + (13 - 12,83)^2 \dots + (12 - 12,83)^2)}$$

$$= \sqrt{\frac{1}{6} \cdot 10,83}$$

$$\approx 1,34$$

 $\Rightarrow$  Die Messwerte können durch eine  $N_{12,83;1,34}$ -verteilte Zufallsgröße modelliert werden.

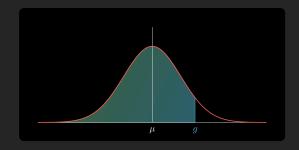
### Umkehrung der Normalverteilung



Für eine  $N_{\mu;\sigma}$ -verteilte Zufallsgröße X lässt sich für eine vorgegebene Wahrscheinlichkeit p durch den WTR-Befehl **invNormal** die Zahl g mit  $P(X \le g) = p$  bestimmen.

### **Beispiel**

$$\mu = 250$$
  $\sigma = 20$   $p = 0.7$   
invNormal  $\Rightarrow q \approx 260,49$ 



## Normalverteilung zur Annäherung einer diskreten Zufallsgröße

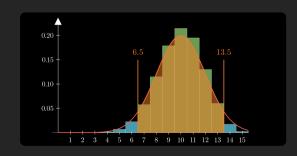


Nutzt man die Normalverteilung um die Wahrscheinlichkeit für eine diskrete Zufallsgröße näherungsweise zu bestimmen, muss man die Integralgrenzen entsprechend anpassen.

### **Beispiel**

Notenpunkte 0..15 sind  $N_{10.2}$ -verteilt.

$$\underbrace{P(7 \leq X \leq 13)}_{\substack{7,8,9,10,11,12,13\\7 \text{ districts Worte}}} = \underbrace{\int_{6,5}^{13,5}}_{\substack{\text{Interval}\\\text{Interval}\\\text{Interval}}} \varphi_{10;2}(x) \ dx \approx 0.92$$



### Sigma-Regeln



Für eine  $N_{\mu;\sigma}$ -verteilte Zufallsgröße gelten folgende Sigma-Regeln

Intervall $\emph{I}$	$P(X \in I)$
$[\mu-\sigma;\mu+\sigma]$	pprox 0,683
$[\mu-2\sigma;\mu+2\sigma]$	pprox 0,954
$[\mu-3\sigma;\mu+3\sigma]$	pprox 0,997

Diese Regeln gelten auch näherungsweise für eine Binomialverteilung.

