

(Gymnasiale Oberstufe: Mathematik Grundkurs, 3-stündig)

Grundlagen

Wahrscheinlichkeiten, also mathematische Größen p ($0 \leq p \leq 1$, p reell) treten im Rahmen von Zufallsexperimenten als Maß für die Sicherheit bzw. Unsicherheit eines Ergebnisses bzw. Ereignisses in Erscheinung treten. Zufallsexperimente (Zufallsversuche, Zufallsvorgänge) sind mathematisch modellierte Prozesse, die auf der (endlichen) Wiederholung (Mehrstufigkeit) einer gleichen festgelegten Versuchssituation (Merkmale, Versuchsausgänge) beruhen, wobei die (abzählbar-endlichen) möglichen Ergebnisse einer solchen Versuchsdurchführung ebenso wie die Ergebnismwahrscheinlichkeiten (als relative Häufigkeiten [Gesetz der großen Zahlen]) bekannt sind. Zufallsexperimente lassen sich durch sog. Wahrscheinlichkeitsbäume (aus Knoten, Verzweigungen [Ausgänge, Merkmalsausprägungen], Kanten [Zweige] und Pfaden [Äste]) darstellen, die Ergebnisse und Wahrscheinlichkeiten anzeigen. Zufallsexperimente, die auf Ergebnisse mit immer derselben Wahrscheinlichkeit hinführen, heißen Laplace-Experimente. Ergebnisse sind Elementarereignisse, Ereignisse sind Zusammenfassungen von Ergebnissen (Mengenlehre der Ereignisse), die Wahrscheinlichkeit eines Ereignisses errechnet sich gemäß den Pfadregeln (Multiplikation von Wahrscheinlichkeiten innerhalb eines Pfades, Addition von Wahrscheinlichkeiten verschiedener Pfade); die Wahrscheinlichkeit aller Ergebnisse eines Zufallsexperiments stellt eine Wahrscheinlichkeitsverteilung dar; die Summe der Wahrscheinlichkeiten aller Ergebnisse ergibt das sichere Ereignis. Aus diesen Sachverhalten folgt die Axiomatik der Wahrscheinlichkeiten mit den daraus abgeleiteten Formeln des Additionssatzes und des Gegenereignisses (De Morgansche Regeln). Eine Zufallsvariable ist eine Funktion, die Ereignissen des Zufallsexperiments eine reelle Zahl zuordnet. Bzgl. der Zufallsvariablen lassen sich Aussagen zu Erwartungswert und Standardabweichung treffen.

Für Wahrscheinlichkeiten p und Ereignisse $A, B, \dots \subset S$ eines Zufallsexperiments gilt:

Ereignisse	Wahrscheinlichkeiten
Leeres/unmögliches Ereignis \emptyset	$p(\emptyset) = 0$
Sicheres Ereignis S	$p(S) = 1$
Ereignis A	$0 \leq p(A) \leq 1$
Ereignisse A, B disjunkt $\rightarrow A \cup B$ als Ereignis „A oder B“	$p(A \cup B) = p(A) + p(B)$
Ereignisse $A, B \rightarrow A \cup B$ als Ereignis „A oder B“, $A \cap B$ als Ereignis „A und B“	Additionssatz für Wahrscheinlichkeiten: $p(A \cup B) = p(A) + p(B) - p(A \cap B)$
Ereignisse $A, B, B \subset A \rightarrow A \setminus B$ als Ereignis „A ohne B“	$p(A \setminus B) = p(A) - p(B)$
Ereignis $A \rightarrow$ Gegenereignis $A^c = S \setminus A$ als Ereignis „nicht A“	$p(A) + p(A^c) = 1, p(A^c) = 1 - p(A)$

Ereignisse, Wahrscheinlichkeiten

Eine (diskrete) Zufallsvariable (Zufallsgröße) X ist eine reelle Abbildung zur Beschreibung eines Zufallsexperiments. Sie ordnet daher den (abzählbar-endlichen) möglichen Ergebnissen des Versuchs eine reelle Zahl zu, also: $X(a_1) = x_1, X(a_2) = x_2, \dots, X(a_n) = x_n$ mit Ergebnissen a_1, a_2, \dots, a_n . Es ergibt sich die durch die Zufallsvariable abgeleitete Wahrscheinlichkeitsverteilung:

$$x_1: p(X=x_1), x_2: p(X=x_2), \dots, x_n: p(X=x_n)$$

mit $p(X=x_1)+p(X=x_2)+\dots+p(X=x_n) = 1$ und daraus der Erwartungswert

$$E(X) = \mu = x_1p(X=x_1) + x_2p(X=x_2) + \dots + x_np(X=x_n)$$

(als durchschnittlicher [mittlerer] Wert der Zufallsvariablen) sowie die Standardabweichung

$$\sigma = \sqrt{\text{Var}(X)} = \sqrt{(x_1 - \mu)^2 p(X = x_1) + (x_2 - \mu)^2 p(X = x_2) + \dots + (x_n - \mu)^2 p(X = x_n)}$$

(als Maß für die Abweichung vom Erwartungswert). Zufallsvariablen X sind durch ihren Erwartungswert μ und ihre Standardabweichung σ charakterisierbar. Aufbauend auf der Wahrscheinlichkeitsverteilung und der daraus abgeleiteten Wahrscheinlichkeitsfunktion $f(x_i) = P(X=x_i)$ ($i=1, \dots, n$) ergibt sich die Verteilungsfunktion $F(x) = p(X \leq x)$ als Summe der Wahrscheinlichkeiten der Ausprägungen der Zufallsvariablen X , die kleiner gleich einer vorgegebenen reellen Zahl x sind.

Ein Wahrscheinlichkeitsbaum dient der grafischen Aufbereitung und Beschreibung eines n -stufigen Zufallsexperiments mit k_1, k_2, \dots Verzweigungen von Merkmalen (Ausgänge, Merkmalsausprägungen, je Stufe) (n, k_1, k_2, \dots als natürliche Zahlen) und besteht aus einer dem Zufallsversuch und dessen (kausaler) Entscheidungshierarchie entsprechenden (gerichteten) Anordnung von Knoten und Kanten.

Wahrscheinlichkeitsbaum (Merkmale: Merkmal A (Ausgänge a, c), Merkmal B (Ausgänge b, d); zweistufig):

	Merkmal A	Merkmal B	◀ Merkmale				
		p(b) b	>	p(a; b) = p(a)p(b)	1	◀ Pfad (a;b) ◀	
	p(a) a					Ereignis E: p(A) = p(a;b)+p(c;d) = p(a)p(b)+ p(c)p(d)	
		1-p(b) d	>	p(a; d) = p(a)(1-p(b))	2		◀ Pfad (a;d)
		1-p(d) b	>	p(c; b) = p(c)(1-p(d))	3		◀ Pfad (c;b)
	p(c) c						
		p(d) d	>	p(c; d) = p(c)p(d)	4	◀ Pfad (c;d) ◀	
Wurzel ▲ Knoten, Kanten ▲ Blätter							
	1. Stufe	2. Stufe	Summe:	1			
	▲ Wahrscheinlichkeiten, Merkmalsausprägungen			▲ Wahrscheinlichkeitsverteilung		▲ Ergebnisse	

Die Knoten stellen die (Zwischen-) Ergebnisse des Zufallsversuchs dar gemäß der jeweiligen Durchführung des Experiments, die Kanten sind mit den dem jeweiligen Ergebnis entsprechenden Wahrscheinlichkeiten versehen. Die Wahrscheinlichkeiten im Wahrscheinlichkeitsbaum ergeben sich aus kombinatorischen Überlegungen (Urnenmodell mit/ohne Zurücklegen u.a.). Die Summe der Wahrscheinlichkeiten an den Kanten, die von einem Knoten ausgehen, ist 1.

Die Aufeinanderfolge von Kanten zwischen dem Anfangsknoten des Wahrscheinlichkeitsbaums (Wurzel) und den Endknoten (Blättern) heißt Pfad. Für ein Ereignis A gelten hinsichtlich der das Ereignis charakterisierenden Pfade und deren Wahrscheinlichkeiten die Pfadregeln:

- Entlang eines zum Ereignis E gehörenden Pfades werden die Wahrscheinlichkeiten der Pfadkanten multipliziert. Es ergibt die Wahrscheinlichkeit des Pfades (Pfad: O -> p(a) -> a -> p(b) -> b ...: $p(a;b;...) = p(a) \cdot p(b) \cdot \dots$).
- Die multiplizierten Wahrscheinlichkeiten von allen zum Ereignis E gehörenden Pfaden werden addiert. Es ergibt sich die Wahrscheinlichkeit des Ereignisses E (Pfade: O -> p(a) -> a -> p(b) -> b ...; O -> p(c) -> c -> p(d) -> d ...; ...: $p(E) = p(a;b;...) + p(c;d;...) + \dots$).

Ein Pfad in einem Wahrscheinlichkeitsbaum stellt ein Elementarereignis dar. Die Summe der Wahrscheinlichkeiten aller Elementarereignisse ist folglich 1. Wird ein Ereignis E durch mehr Pfade charakterisiert als sein Gegenereignis E^c , so empfiehlt sich, die Wahrscheinlichkeit des Gegenereignisses zu berechnen, so dass die Formel $p(E) = 1 - p(E^c)$ zum Tragen kommt.

Je nach Fragestellung können vorgegebene Wahrscheinlichkeitsbäume umorganisiert, reduziert und „eingeklappert“ werden, indem bestimmte Merkmalsausprägungen zusammengefasst und somit die Zahl der Kanten vermindert werden. Allgemein ist zur Anzahl der Pfade in einem Wahrscheinlichkeitsbaum zu sagen, dass bei n -maligem Durchführen eines Zufallsversuchs mit gleichen k Ausgängen je Stufe die Pfadanzahl k^n beträgt; bei unterschiedlicher Anzahl von Ausgängen k_1, k_2, \dots, k_n ist die Pfadanzahl $k_1 \cdot k_2 \cdot \dots \cdot k_n$.

Bedingte Wahrscheinlichkeiten, Vierfeldertafeln

Gegeben sei der folgende zweistufige Wahrscheinlichkeitsbaum mit je zwei Ausgängen für die zwei Merkmale:

Wahrscheinlichkeitsbaum (Merkmale: Merkmal 1 (Ausgänge A, A⁻), Merkmal 2 (Ausgänge B, B⁻); zweistufig):

	1. Merkmal	2. Merkmal		
		$p_A(B) \quad \mathbf{B}$	>	$p(A; B) = p(A \cap B)$ 1
	$p(A) \quad \mathbf{A}$			
		$p_A(B^-) \quad \mathbf{B^-}$	>	$p(A; B^-) = p(A \cap B^-)$ 2
		$p_{A^-}(B) \quad \mathbf{B}$	>	$p(A^-; B) = p(A^- \cap B)$ 3
	$p(A^-) \quad \mathbf{A^-}$			
		$p_{A^-}(B^-) \quad \mathbf{B^-}$	>	$p(A^-; B^-) = p(A^- \cap B^-)$ 4
		Summe:		1

Gemäß dem Wahrscheinlichkeitsbaum ermitteln sich die bedingten Wahrscheinlichkeiten, d.h. die Wahrscheinlichkeiten, bei denen das bedingende Ereignis „herausgerechnet“ wurde, bei Bedingung A und Ereignis B als:

$$p_A(B) = \frac{p(A \cap B)}{p(A)}.$$

Es gilt noch: $p(A \cap B) = p(A) \cdot p_A(B)$. Hat das bedingende Ereignis A keinen Einfluss auf das Ereignis B, ist also: $p_A(B) = p(B)$, so gilt: $p(A \cap B) = p(A) \cdot p(B)$ und damit die stochastische Unabhängigkeit der Ereignisse sowie $p(A \cap B) \neq p(A) \cdot p(B)$ bei stochastischer Abhängigkeit.

Aus obigem Wahrscheinlichkeitsbaum ergibt sich zudem die Vierfeldertafel der Wahrscheinlichkeiten:

	B	B⁻	
A	$p(A \cap B)$	$p(A \cap B^-)$	$p(A)$
A⁻	$p(A^- \cap B)$	$p(A^- \cap B^-)$	$p(A^-)$
	$p(B)$	$p(B^-)$	1

die auch auf absolute Häufigkeiten zu beziehen ist:

	B	B⁻	
A	n_{11}	n_{12}	n_A
A⁻	n_{21}	n_{22}	$n - n_A$
	n_B	$n - n_B$	n

oder um die bedingten Wahrscheinlichkeiten ergänzt werden kann:

	B	B⁻			
A	$p(A \cap B)$	$p(A \cap B^-)$	$p(A)$	$p_A(B)$	$p_A(B^-)$ 1
A⁻	$p(A^- \cap B)$	$p(A^- \cap B^-)$	$p(A^-)$	$p_{A^-}(B)$	$p_{A^-}(B^-)$ 1
	$p(B)$	$p(B^-)$	1		
	$p_B(A)$	$p_B(A^-)$			
	$p_B(A)$	$p_B(A^-)$			
	1	1			

Bernoulli-Experiment

Ein Bernoulli-Experiment ist ein Zufallsexperiment mit zwei Ausgängen (T = Treffer, N = Nichttreffer), der Grundwahrscheinlichkeit p als Trefferwahrscheinlichkeit, der Anzahl n der Experimentwiederholung „mit Zurücklegen“. Die Zufallsvariable X gibt die Anzahl der Treffer bei n -maliger Wiederholung des Experiments an. Es gelten auf Grund der Pfadregeln für Wahrscheinlichkeitsbäume (Multiplikation der Wahrscheinlichkeiten entlang eines Pfades, Addition der [multiplizierten] Wahrscheinlichkeiten verschiedener Pfade) die Trefferwahrscheinlichkeiten der Bernoulli-Formel:

$$p(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

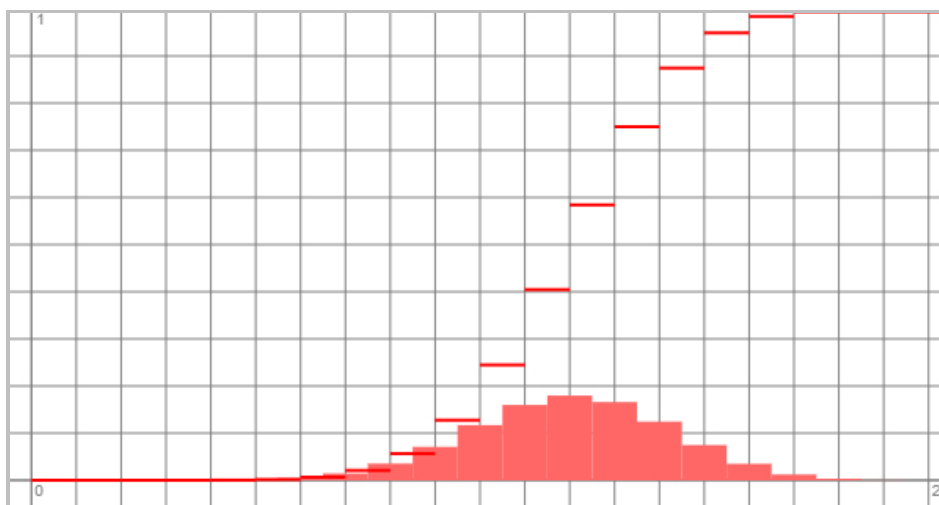
Bernoulli-Formel

mit den Binomialkoeffizienten $\binom{n}{k} = \frac{n(n-1)(n-2) \cdot \dots \cdot (n-k+1)}{1 \cdot 2 \cdot \dots \cdot k} = \frac{n!}{k!(n-k)!}$ (als Anzahl der Pfade mit gleicher Wahrscheinlichkeit $p^k(1-p)^{n-k}$) und weiter:

$$\begin{aligned} p(X=0) &= (1-p)^n \\ p(X=n) &= p^n \\ p(X \leq k) &= p(X=0) + p(X=1) + \dots + p(X=k) = 1 - p(X > k) \\ p(X < k) &= p(X \leq k-1) = 1 - p(X \geq k) \\ p(X \geq k) &= 1 - p(X \leq k-1) \\ p(X > k) &= p(X \geq k+1) = 1 - p(X \leq k) \\ p(k_1 \leq X \leq k_2) &= p(X=k_1) + \dots + p(X=k_2) = p(X \leq k_2) - p(X \leq k_1 - 1) \\ p(k_1 < X \leq k_2) &= p(X=k_1+1) + \dots + p(X=k_2) = p(X \leq k_2) - p(X \leq k_1) \\ p(k_1 \leq X < k_2) &= p(X=k_1) + \dots + p(X=k_2-1) = p(X \leq k_2-1) - p(X \leq k_1-1) \\ p(k_1 < X < k_2) &= p(X=k_1+1) + \dots + p(X=k_2-1) = p(X \leq k_2-1) - p(X \leq k_1) \end{aligned}$$

Rechenregeln für Wahrscheinlichkeiten der Binomialverteilung

Es gelten die „Übersetzungen“: \leq : höchstens, \geq : mindestens, wenigstens, $<$: weniger als, $>$: mehr als. Die Binomialverteilung $p(X=k)$, $k=0, 1, \dots, n$, kann durch ein Histogramm dargestellt werden:



Binomialverteilung B(20; 0,6): $p(X=k)$, $p(X \leq k)$, $k=0, 1, \dots, 20$

Es gilt weiter hinsichtlich des Erwartungswerts $E(X)$ beim Bernoulli-Experiment:

$$E(X) = \mu = np.$$

Hinsichtlich der Standardabweichung σ beim Bernoulli-Experiment folgt:

$$\sigma = \sqrt{np(1-p)}.$$

Es folgt noch hinsichtlich Trefferwahrscheinlichkeit und Versuchsanzahl:

$$p = 1 - \frac{\sigma^2}{\mu}, \quad n = \frac{\mu^2}{\mu - \sigma^2} = \frac{\mu}{p}.$$

Es gelten dann die sog. σ -Regeln für die 1σ -, 2σ - und 3σ -Intervalle der Binomialverteilung:

$$\begin{aligned} p(\mu - \sigma \leq X \leq \mu + \sigma) &\approx 68,27\% \text{ (1}\sigma\text{-Intervall)} \\ p(\mu - 2\sigma \leq X \leq \mu + 2\sigma) &\approx 95,45\% \text{ (2}\sigma\text{-Intervall)} \\ p(\mu - 3\sigma \leq X \leq \mu + 3\sigma) &\approx 99,73\% \text{ (3}\sigma\text{-Intervall)}, \end{aligned}$$

d.h.: ungefähr mit 68,27%, 95,45% und 99,73% Wahrscheinlichkeit liegt die Anzahl der Treffer in den entsprechenden σ -Intervallen.

Beim Rechnen mit Wahrscheinlichkeiten der $B(n, p)$ -verteilten Zufallsgröße X eines Bernoulli-Experiments sind Ausdrücke der Form $p(X=k)$ und $p(X \leq k)$ zu verwenden, wenn:

1. n, p vorgegeben $\Rightarrow p(X=k), p(X \leq k)$ berechnen (\rightarrow u.a. Binomialverteilung als Histogramm)
2. $p, p(X \leq k) \leq p^*$ o.ä. vorgegeben \Rightarrow Versuchsanzahl n ermitteln
3. $n, p(X=k) \leq p^*, p(X \leq k) \leq p^*$ o.ä. vorgegeben \Rightarrow Trefferwahrscheinlichkeit p ermitteln (p kann auch diskrete Werte [Brüche, Dezimalzahlen] annehmen)

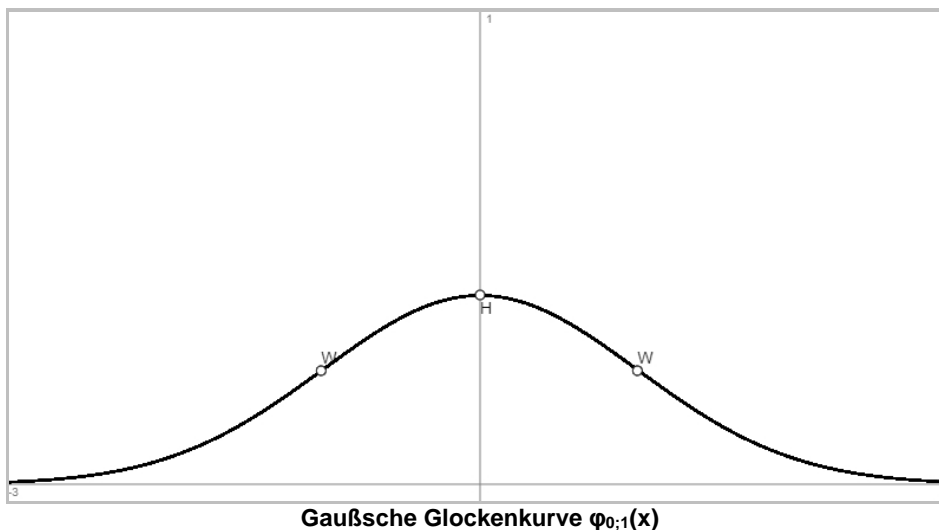
Normalverteilung

Die Gaußsche Glockenkurve ist eine von reellen Parametern μ und σ (>0) abhängige Schar von reellwertigen Funktionen $\varphi_{\mu,\sigma}(x)$:

$$\varphi_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Die Glockenkurve hat folgende Eigenschaften: $\varphi_{\mu,\sigma}(x)$ ist definiert auf allen reellen Zahlen und dort positiv: $\varphi_{\mu,\sigma}(x) > 0$; $\varphi_{\mu,\sigma}(x)$ ist achsensymmetrisch zur Senkrechten $x = \mu$; $\varphi_{\mu,\sigma}(x)$ besitzt einen Hochpunkt an der Stelle $x_H = \mu$ mit $H(\mu | \frac{1}{\sigma\sqrt{2\pi}})$; $\varphi_{\mu,\sigma}(x)$ besitzt zwei Wendepunkte an den Stellen

$x_W = \mu \pm \sigma$ mit $W_1(\mu - \sigma | \frac{1}{\sigma\sqrt{2\pi}})$, $W_2(\mu + \sigma | \frac{1}{\sigma\sqrt{2\pi}})$; mit $x \rightarrow \pm\infty$ gilt: $\varphi_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \rightarrow 0 = y$ (x -Achse) als waagerechter Asymptote.



Die Eigenschaften der Funktionen $\varphi_{\mu,\sigma}(x)$ führen zu deren Verwendung in der Stochastik, stellt doch jede Funktion $\varphi_{\mu,\sigma}(x)$ die Wahrscheinlichkeitsdichte (Gauß-Funktion) einer sog. normalverteilten (gauß-verteilten) stetigen Zufallsvariablen X dar: $X \sim N(\mu, \sigma^2)$ als Normalverteilung mit Erwartungswert μ und Standardabweichung σ . Der Parameter μ verschiebt die Dichtefunktion (nach rechts: $\mu > 0$; nach links: $\mu < 0$), der Parameter σ (>0) steht für die Breite der Dichtefunktion (schmäler: $\sigma < 1$; breiter: $\sigma > 1$). Ist $\mu=0$ und $\sigma=1$, so liegt die Standardnormalverteilung $N(0; 1)$ vor. Die Verteilungsfunktion $\Phi_{\mu,\sigma}(x)$ der normalverteilten Zufallsvariablen wird als Integralfunktion über die Dichte $\varphi_{\mu,\sigma}(x)$ definiert:

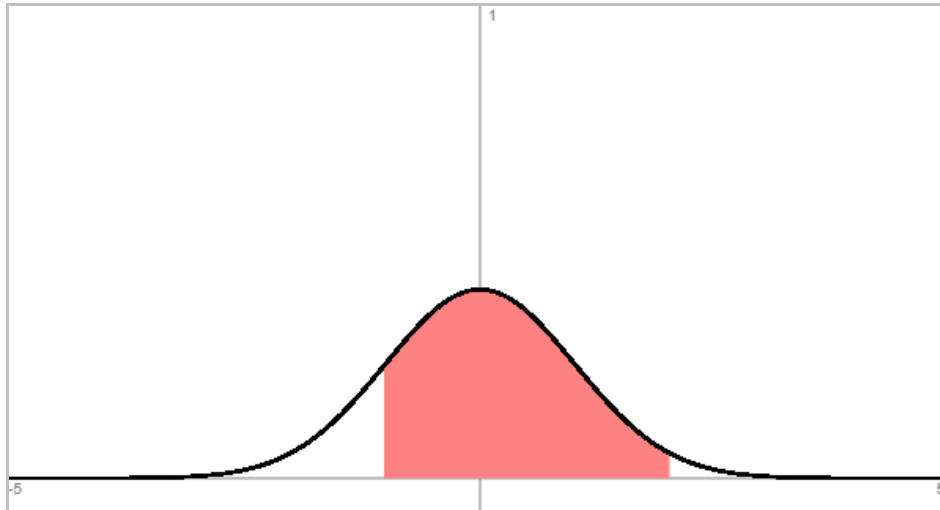
$$\Phi_{\mu,\sigma}(x) = \int_{-\infty}^x \varphi_{\mu,\sigma}(z) dz = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(z-\mu)^2}{2\sigma^2}} dz.$$

Für die $N(\mu, \sigma^2)$ -verteilte stetige Zufallsvariable X gelten dann die folgenden Rechenregeln:

$$\begin{aligned}
p(X=a) &= 0 \\
p(a \leq X) &= p(a < X) = 1 - \Phi_{\mu, \sigma}(a) \\
p(a \leq X \leq b) &= p(a < X \leq b) = p(a \leq X < b) = p(a < X < b) = \Phi_{\mu, \sigma}(b) - \Phi_{\mu, \sigma}(a) \\
p(X \leq b) &= p(X < b) = \Phi_{\mu, \sigma}(b) \\
p(X \leq \mu) &= p(\mu \leq X) = 0,5 \\
p(X \leq \mu - a) &= p(X \geq \mu + a) = \Phi_{\mu, \sigma}(\mu - a) \\
p(\mu - a \leq X \leq \mu + a) &= 2 \cdot p(\mu \leq X \leq \mu + a) = 2 \cdot \Phi_{\mu, \sigma}(\mu + a) - 1 \\
p(\mu - \sigma \leq X \leq \mu + \sigma) &= \Phi_{\mu, \sigma}(\mu + \sigma) - \Phi_{\mu, \sigma}(\mu - \sigma) \approx 0,683 \\
p(\mu - 2\sigma \leq X \leq \mu + 2\sigma) &= \Phi_{\mu, \sigma}(\mu + 2\sigma) - \Phi_{\mu, \sigma}(\mu - 2\sigma) \approx 0,954 \\
p(\mu - 3\sigma \leq X \leq \mu + 3\sigma) &= \Phi_{\mu, \sigma}(\mu + 3\sigma) - \Phi_{\mu, \sigma}(\mu - 3\sigma) \approx 0,997 \quad (\sigma\text{-Intervalle der Normalverteilung})
\end{aligned}$$

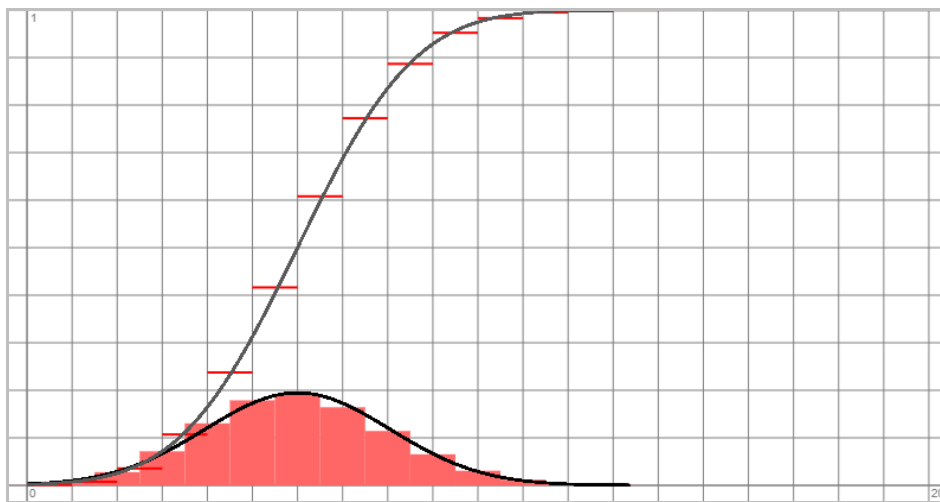
Rechenregeln für Wahrscheinlichkeiten der Normalverteilung

$p(a \leq X \leq b)$ z.B. gibt damit die Wahrscheinlichkeit an, dass die Zufallsvariable X einen Wert zwischen a und b annimmt, a, b reell.



Normalverteilung $N(0; 1)$ und Wahrscheinlichkeit $p(-1 \leq X \leq 2)$

Die Normalverteilung ist die stetige Zufallsgröße zu den diskreten Größen der Binomialverteilung. Insofern können Binomialverteilungen $B(n, p)$ mit Erwartungswert $\mu = np$ und Standardabweichung $\sigma = (np(1-p))^{1/2} > 3$ durch eine Normalverteilung $N(\mu, \sigma^2)$ mit denselben Parametern angenähert werden; die σ -Regeln für Binomial- und Normalverteilung sind hierfür ein Beispiel.



Binomialverteilung $B(20; 0,3)$, Normalverteilung $N(6; 4,2)$: Dichte und Verteilung

Als Näherung zwischen den Wahrscheinlichkeiten der $B(n, p)$ -verteilten Zufallsgröße X und den Wahrscheinlichkeiten der Normalverteilung $N(\mu, \sigma^2)$ gilt:

$$p(a \leq X \leq b) \approx \Phi_{\mu, \sigma}(b+0,5) - \Phi_{\mu, \sigma}(a-0,5), \quad 0 \leq a \leq b.$$

Weiter kann die Trefferwahrscheinlichkeit p in einem Bernoulli-Experiment mit $B(n, p)$ -verteilter Zufallsgröße X die Wahrscheinlichkeit $p = p(a \leq Y)$, $p(Y \leq b)$, $p(a \leq Y \leq b)$ o.ä. einer $N(\mu, \sigma^2)$ -verteilten Zufallsgröße Y sein.