VORKURS

15.10.2021

BINOMIALVERTEILUNG I

Eine der wichtigsten Wahrscheinlichkeiten in der Mathematik resultieren aufgrund der **Binomialverteilung**. Dabei handelt es sich zum einen um unabhängige Ereignisse und zum anderen existieren nur zwei Möglichkeiten („Treffer“ oder „Niete“).

Da es sich um ein Experiment **mit Zurücklegen** handelt, bleiben die Wahrscheinlichkeiten gleich.

Eine solche Versuchsreihe wird auch **Bernoulli-Prozess** genannt.

Die **Zufallsvariable** X beschreibt hier die Anzahl des günstigen Ereignis bei n Wiederholungen, so dass sich die Wahrscheinlichkeit als Produkt aus den möglichen Kombinationen sowie den günstigen $P(A)$ und ungünstigen $P(\bar{A})$ Wahrscheinlichkeiten ergibt.

Es gilt:

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

Anzahl der gewünschten Treffer

Anzahl der Kombinationen

Wahrscheinlichkeit eines Treffer

Wahrscheinlichkeit einer Niete

Handelt es sich um eine solche diskrete Zufallsvariable, dann nennt man die dazugehörige Verteilung auch **binomialverteilt** und schreibt $X \sim B(n;p)$.

Die gesuchten Wahrscheinlichkeiten (gerade auch kumuliert) können direkt mit Hilfe der Tabelle der **Binomialverteilung** abgelesen werden.

BINOMIALVERTEILUNG II

Neben dem Vorteil, dass sich ein Großteil von Fragen aus der Wahrscheinlichkeit auf die Situation Treffer oder Niete reduzieren lässt, können aufgrund der binomialverteilten Zufallsvariable X auch verschiedene wichtige Werte der Statistik sehr einfach bestimmt werden.

Erwartungswert: $E(X) = n \cdot p$

Varianz: $Var(X) = n \cdot p \cdot (1 - p)$

Beispiel:

Wir haben eine gezinkte Münze, die mit einer Wahrscheinlichkeit von 60% Wappen anzeigt. Nun stellt sich die Frage, wie hoch die Wahrscheinlichkeit für siebenmal Wappen bei zehnmaligen Werfen ist.

$$n = 10;\ \ k = 7;\ P(W) = 0{,}6;\quad P(Z) = 1 - P(W) = 0{,}4$$

Es handelt sich somit um eine binomialverteilte Zufallsvariable, so dass $X \sim B(10; 0{,}6)$ gilt.

$$P(X = 7) = \binom{10}{7} \cdot 0{,}6^7 \cdot 0{,}4^3 = 0{,}215$$

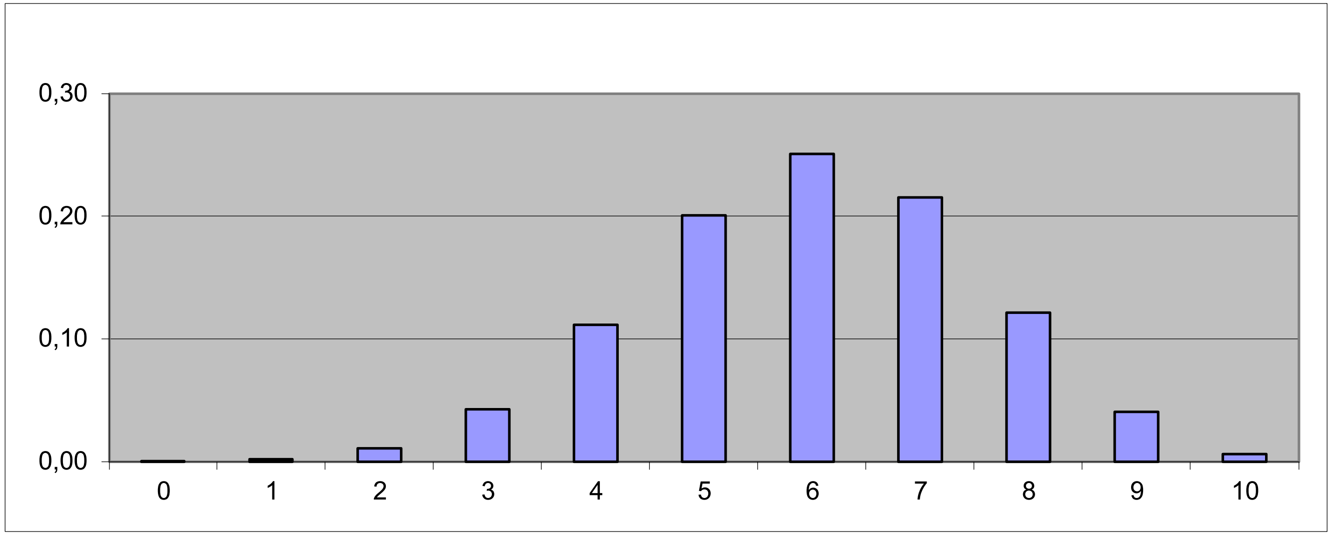
BINOMIALVERTEILUNG III

Für das letzte Beispiel mit $X \sim B(10; 0{,}6)$ und $k = 7$ ergibt sich weiter:

Erwartungswert: $E(X) = n \cdot p = 10 \cdot 0{,}6 = 6$

Varianz: $Var(X) = n \cdot p \cdot (1 - p) = 10 \cdot 0{,}6 \cdot 0{,}4 = 2{,}4$

Graphisch bildet die Binomialverteilung immer die sogenannte **Gaußsche Glockenkurve** und sieht aufgrund unseres letzten Beispiels so aus:



BINOMIALVERTEILUNG IV

Bei einem Glücksspiel gewinnt man mit einer Wahrscheinlichkeit von 5%, sofern man aus 10 Kugeln zwei markierte Exemplare zieht.

$$n = 10, k = 2, p = 0{,}05, q = 1 - p = 0{,}95$$

Via Formel: $P(X = 2) = \binom{10}{2} \cdot 0{,}05^2 \cdot 0{,}95^8 = 0{,}0746$

Wertetafel zur Binomialverteilung $P(X=k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$

n	k	p 0,01	0,05	0,1	$\frac{1}{6}$	0,2	0,25	0,3	$\frac{1}{3}$	0,4	0,5	*k*	*n*
	0	0,9044	0,5987	0,3487	0,1615	0,1074	0,0563	0,0282	0,0173	0,0060	0,0010	*10*	
	1	0,0914	0,3151	0,3874	0,3230	0,2684	0,1877	0,1211	0,0867	0,0403	0,0098	*9*	
	2	0,0042	0,0746	0,1937	0,2907	0,3020	0,2816	0,2335	0,1951	0,1209	0,0439	*8*	
	3	0,0001	0,0105	0,0574	0,1550	0,2013	0,2503	0,2668	0,2601	0,2150	0,1172	*7*	
	4	0,0000	0,0010	0,0112	0,0543	0,0881	0,1460	0,2001	0,2276	0,2508	0,2051	*6*	
10	5		0,0001	0,0015	0,0130	0,0264	0,0584	0,1029	0,1366	0,2007	0,2461	*5*	*10*
	6		0,0000	0,0001	0,0022	0,0055	0,0162	0,0368	0,0569	0,1115	0,2051	*4*	
	7			0,0000	0,0002	0,0008	0,0031	0,0090	0,0163	0,0425	0,1172	*3*	
	8				0,0000	0,0001	0,0004	0,0014	0,0030	0,0106	0,0439	*2*	
	9					0,0000	0,0000	0,0001	0,0003	0,0016	0,0098	*1*	
	10							0,0000	0,0000	0,0001	0,0010	*0*	

Sollten Sie direkt über das Komplementärereignis sprich Gegenwahrscheinlichkeit gehen, nutzen Sie die rechte Spalte, um die Zufallsvariable abzulesen (siehe Beispiel S.85).

BINOMIALVERTEILUNG V

Um Fragen nach Wahrscheinlichkeiten einer Zufallsvariablen zu berechnen, die mittels einer Ungleichung beschrieben wird, nutzt man die Tabelle der **kumulierten Binomialverteilungen**.

Man sollte darauf achten, dass stets eine $\leq$ - Beziehung entsteht bzw. diese erzeugen.

$$P(X > k) = 1 - P(X \leq k)$$

Beispiel:

Bei einer Produktion von Glühbirnen können mit einer Wahrscheinlichkeit von 10% farbliche Unterschiede entstehen. Sollten bei einer Stichprobe aus 100 Birnen mehr als 10 von ihnen defekt sein, erhält der Käufer einen Preisnachlass von 25%.

Sie erhalten somit: $n = 100, k > 10, p = 0{,}1, q = 1 - p = 0{,}9$

Umwandlung : $P(X > 10) = 1 - P(X \leq 10)$

Ohne die folgende Tabelle der kumulierten Binomialverteilung müsste man alle Einzelwahrscheinlichkeiten für $k = 0; 1; 2 \ldots; 9; 10$ berechnen.

BINOMIALVERTEILUNG VI

Sie erhalten über die Tabelle für $n = 100$ und $p = 0{,}1$ den Wert für $P(X \leq 10) = 0{,}5832$:

$$P(X > 10) = 1 - 0{,}5832 = 0{,}4168$$

Der Käufer erhält somit mit einer Wahrscheinlichkeit von 41,68% den Preisnachlass.

Summierte Binomialverteilung (n = 100)

$$P(X \leq k) = \sum_{i=0}^{k} \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i}$$

n	k	p 0,01	0,05	0,1	$\frac{1}{6}$	0,2	0,25	0,3	$\frac{1}{3}$	0,4	0,5	*k*	*n*
	0	0,3660	0,0059	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*99*	
	1	0,7358	0,0371	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*98*	
	2	0,9206	0,1183	0,0019	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*97*	
	3	0,9816	0,2578	0,0078	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*96*	
	4	0,9966	0,4360	0,0237	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*95*	
	5	0,9995	0,6160	0,0576	0,0004	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*94*	
	6	0,9999	0,7660	0,1172	0,0013	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	*93*	
	7	1,0000	0,8720	0,2061	0,0038	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	*92*	
	8		0,9369	0,3209	0,0095	0,0009	0,0000	0,0000	0,0000	0,0000	0,0000	*91*	
	9		0,9718	0,4513	0,0213	0,0023	0,0000	0,0000	0,0000	0,0000	0,0000	*90*	
	10		0,9885	$\boxed{0{,}5832}$	0,0427	0,0057	0,0001	0,0000	0,0000	0,0000	0,0000	*89*	
	11		0,9957	0,7030	0,0777	0,0126	0,0004	0,0000	0,0000	0,0000	0,0000	*88*	
	12		0,9985	0,8018	0,1297	0,0253	0,0010	0,0000	0,0000	0,0000	0,0000	*87*	
	13		0,9995	0,8761	0,2000	0,0469	0,0025	0,0001	0,0000	0,0000	0,0000	*86*	
	14		0,9999	0,9274	0,2874	0,0804	0,0054	0,0002	0,0000	0,0000	0,0000	*85*	
	15		1,0000	0,9601	0,3877	0,1285	0,0111	0,0004	0,0000	0,0000	0,0000	*84*	

AUFGABE

1. Ein bekannter Hersteller von Schokolade verspricht in einer Aktion, dass in jeder fünften Tafel eine besondere Überraschung versteckt sei.
   Voller Freude kauft Egon gleich 20 Tafeln der Schokolade.

   a) Wie hoch ist die Wahrscheinlichkeit, dass sich unter den 20 Tafeln genau 4 Überraschungen zu befinden?

   b) Wie hoch ist Wahrscheinlichkeit keine der Überraschungen zu finden?

   c) Tatsächlich befinden sich in den 20 Tafeln exakt drei Überraschungen.
   Wie hoch ist die Wahrscheinlichkeit, dass sich in den ersten 5 Tafeln zwei der drei Überraschungen befinden?

   d) Bei einer Stichprobe werden 100 Tafeln geprüft.
   Bestimmen Sie die Wahrscheinlichkeit mit $p = \frac{1}{5}$ dafür weniger als 10 Tafeln bzw. weniger als 20 Tafeln zu finden.

HYPERGEOMETRISCHE VERTEILUNG I

Im Prinzip handelt es sich um ein ähnliches Zufallsexperiment wie bei der Binomialverteilung mit dem Unterschied, dass es sich um ein Urnenmodell **ohne Zurücklegen** handelt.

Also ändert sich die Wahrscheinlichkeit je Ziehungsstufe (siehe Baumdiagramm).

Im Unterschied zur Binomialverteilung, wo die Grundgesamtheit aufgrund der Stichprobe unbekannt ist, kennen wir bei der Hypergeometrischen Verteilung diese.

$$P(X = k) = \frac{\binom{M}{k} \cdot \binom{N - M}{n - k}}{\binom{N}{n}}$$

N: Anzahl der Grundgesamtheit
n: Anzahl der Elemente der Stichprobe
M: Anzahl der Elemente mit der Erfolgseigenschaft

Erwartungswert: $E(X) = n \cdot \frac{M}{N}$

Varianz: $Var(X) = n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right) \cdot \frac{N-n}{N-1}$

HYPERGEOMETRISCHE VERTEILUNG II

Beispiel:

In einer Kiste befinden sich 50 Schrauben, von denen 10% defekt sind. Es werden 8 Stück mit einem Griff gezogen. Mit welcher Wahrscheinlichkeiten sind 3 defekte Schrauben darunter?

X = Anzahl an defekten Schrauben

$N = 50, n = 8, M = 5, k = 3$

Kombination Treffer　　　　Kombination Nieten

$$P(X = 3) = \frac{\binom{5}{3} \cdot \binom{50-5}{8-3}}{\binom{50}{8}}$$

alle Möglichkeiten

Berechnung von $\binom{50}{8}$:

$$\binom{50}{8} = \frac{50!}{8! \cdot (50-8)!} = 50\ nCR\ 8$$

Erwartungswert: $E(X) = n \cdot \frac{M}{N} = 8 \cdot \frac{5}{50} = 0{,}8$

Varianz: $Var(X) = n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right) \cdot \frac{N-n}{N-1} = 8 \cdot \frac{5}{50} \cdot \left(1 - \frac{5}{50}\right) \cdot \frac{50-8}{50-1} = 0{,}617$

Anmerkung: Würde sich die Aufgabenstellung ändern, so dass wir aus einer Gesamtheit eine Stichprobe mit Umfang 50 tätigen, wäre es binomialverteilt.

$$P(X = 3) = \binom{50}{3} \cdot 0{,}1^3 \cdot 0{,}9^{47}$$

AUFGABE

1. Bei einer Lieferung eines namhaften Handtaschenherstellers Gocci befinden sich in einer Kiste 100 Stück, von denen aus Erfahrung heraus 12% einen Farbfehler haben.
   Es werden gleichzeitig 6 Handtaschen selektiert und genauer untersucht.

   a) Wie hoch ist die Wahrscheinlichkeit, dass nur eine Tasche den Farbfehler hat?

   b) Mit welcher Wahrscheinlichkeit sind mindestens eine Tasche defekt?

   c) Wie hoch ist Wahrscheinlichkeit, dass höchstens 2 Taschen defekt sind?

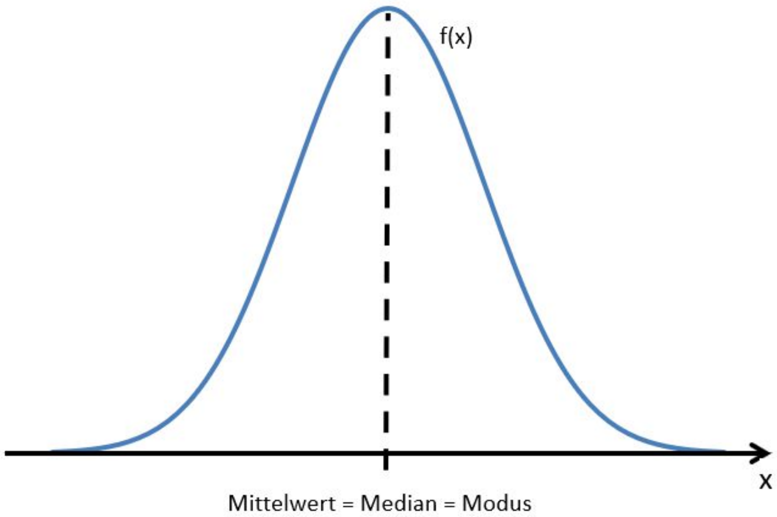
   d) Mit wie vielen Handtaschen mit Fehlern hätte man rechnen müssen?

   e) Wie groß ist die Standardabweichung innerhalb dieser Untersuchung.

NORMALVERTEILUNG I

Eine normalverteilte Funktion entspricht der Gauß'schen Glockenkurve.
Reale Verteilungen sind z.B. Längen Gewichte oder auch Testergebnisse bei Examen.

Die Normalverteilung ist von den beiden Parametern μ (Erwartungswert) und σ^2 (Varianz) abhängig und wird als $N(\mu, \sigma^2)$ oder auch $N(\mu, \sigma)$ bezeichnet.



$$f_{normal;\mu;\sigma}(X) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(X-\mu)^2}{2\sigma^2}}$$

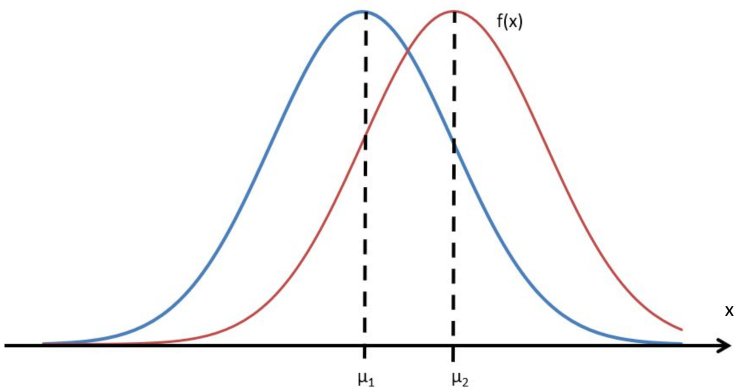
Am höchsten Punkt befindet sich der Mittelwert μ, Median $\overline{X_Z}$, Modus $\overline{X_D}$
Die Fläche unterhalb der Kurve entspricht 1.

NORMALVERTEILUNG II

Da die Funktion der Normalverteilung von den beiden Parametern Varianz als auch Mittelwert abhängig ist, ergeben sich für den Funktionsgraphen folgende Zusammenhänge:

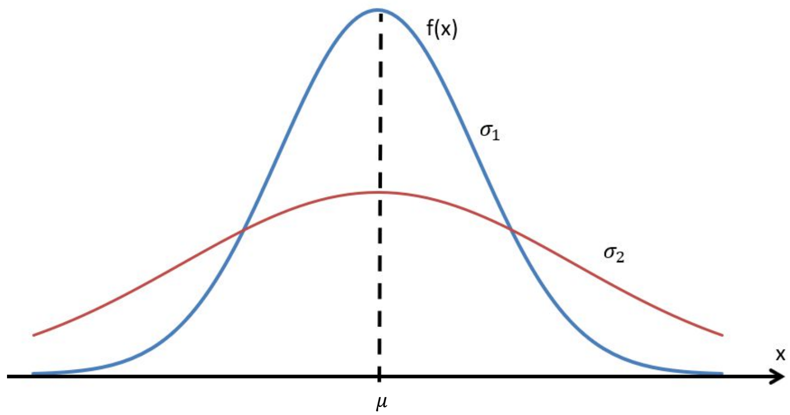
Der Mittelwert bestimmt die Lage der Kurve,
d.h. verschiebt diese nach rechts bzw. links.

$$\mu_1 < \mu_2$$

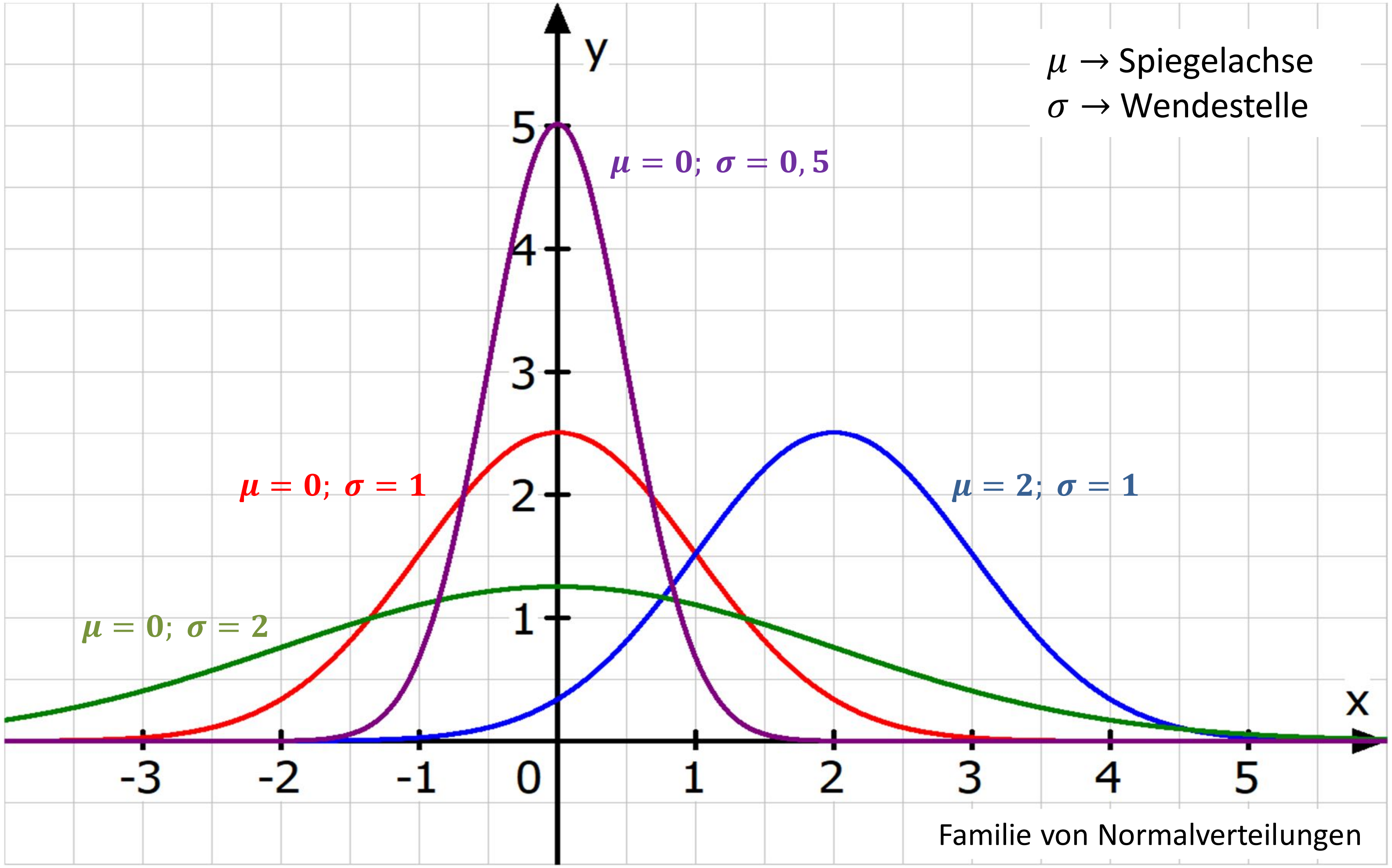


Durch die Varianz bzw. Standardabweichung wird die Breite der Kurve beschrieben,
d.h. je größer die Varianz ist,
desto breiter / flacher ist der Verlauf.

$$\sigma_1 < \sigma_2$$

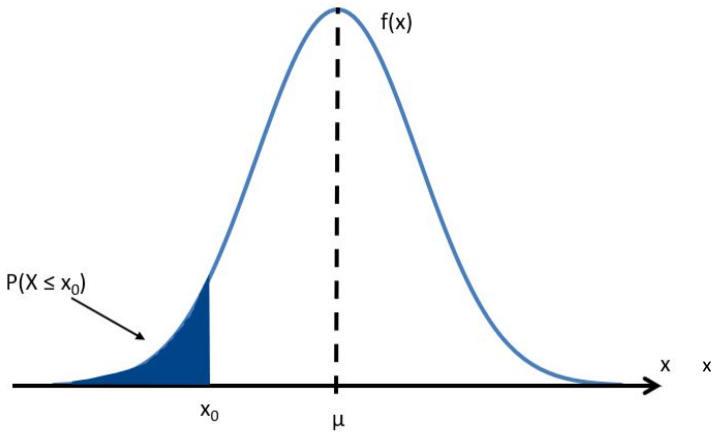


NORMALVERTEILUNG III



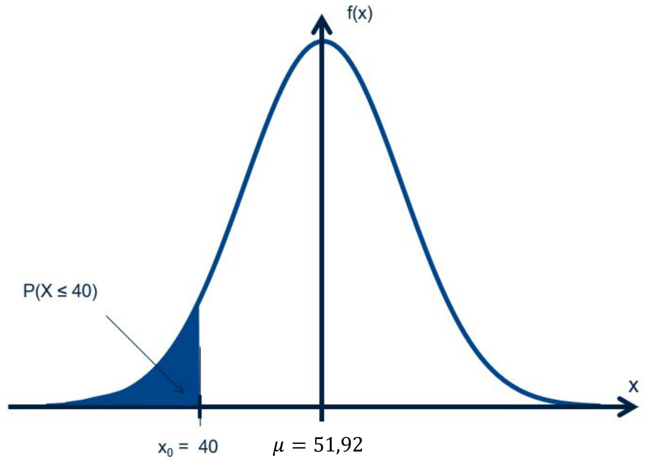
NORMALVERTEILUNG IV

Wird die Wahrscheinlichkeit einer normalverteilten Zufallsvariablen gesucht, die kleiner gleich einem definierten x_0 sein soll, so entspricht diese der Fläche unterhalb der Wahrscheinlichkeitsdichtefunktion bis zum x_0.



<u>Beispiel</u>:

Bei einer Befragung im Jahr 2015 nach dem Alter der Bürger in der Europäischen Union ($n = 30.101$), ergab sich als Mittelwert $\mu = 51{,}92$ und Standardabweichung $\sigma = 17{,}39$. Im Graphen wird die Wahrscheinlichkeit dargestellt, dass ein Befragter höchstens 40 Jahre alt ist.



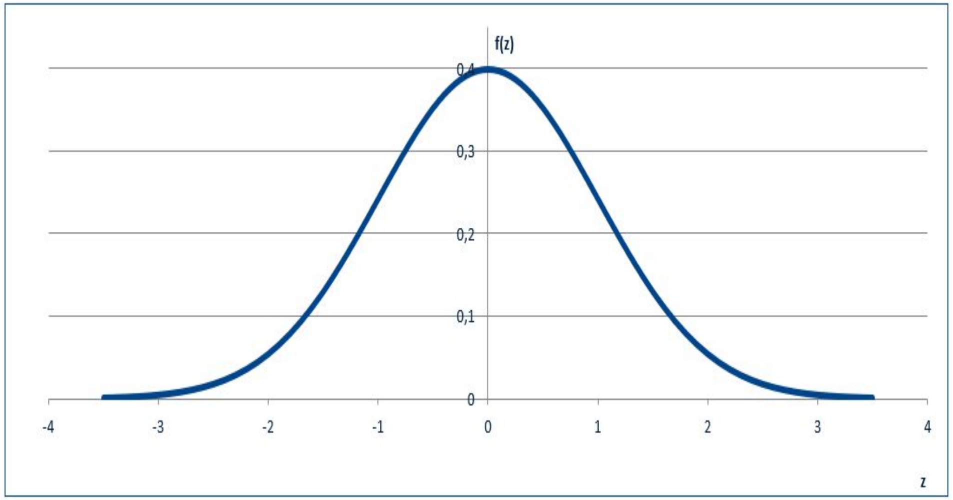
STANDARD-NORMALVERTEILUNG I

Das Problem der Normalverteilung ist, dass es aufgrund der beiden Parameter μ und σ eine Vielzahl von ähnlichen Funktionen zu einem Sachverhalt entstehen.

Um dies zu kompensieren beschränkt man sich auf einen Vertreter dieser Familie und standardisiert diesen, in dem man als Mittelwert $\mu = 0$ und als Standardabweichung $\sigma = 1$ festlegt.

$$f_{normal;0;1}(Z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}$$

$$Z = \frac{X - \mu}{\sigma} \Leftrightarrow X = \mu + \sigma \cdot Z$$



Da die Standardnormalverteilung symmetrisch zur y-Achse ist, gilt der folgende Zusammenhang:

$$P(Z \leq -z_0) = P(Z \geq z_0) = 1 - P(Z < z_0)$$

STANDARD-NORMALVERTEILUNG II

Bei einer Befragung im Jahr 2015 nach dem Alter der Bürger in der Europäischen Union ($n = 30.101$), ergab sich als Mittelwert $\mu = 51{,}92$ und Standardabweichung $\sigma = 17{,}39$.
Berechnen Sie die Wahrscheinlichkeit, dass ein Befragter höchstens 70 Jahre alt ist ($X \leq 70$).

Transformation: $Z = \frac{X-\mu}{\sigma} = \frac{70-51{,}92}{17{,}39} = 1{,}04$

Wahrscheinlichkeit: $P(X \leq 70) = P(Z \leq 1{,}04) = 0{,}8508$

z	0	1	2	3	4	5	6	7	8	9
...	...	...	...	...	...	...	...	...	...	...
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9392	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
...	...	...	...	...	...	...	...	...	...	...

STANDARD-NORMALVERTEILUNG III

Ein Pharmagroßhändler versucht seine Lagerhaltung zu optimieren.

Sollte der Vorrat eines bestimmten Impfstoffs weniger oder gleich 40 Packungen betragen, muss nachbestellt werden (Lieferzeit 3 Tage). Er will aber niemals ausverkauft sein.

Durch Untersuchungen ergibt sich ein Mittelwert von $\mu = 30$ und eine Standardabweichung von $\sigma = 10$.

Wie groß ist die Wahrscheinlichkeit, dass der Lagerbestand währen der Nachlieferung ausreicht?

X = Nachfrage in der Nachlieferzeit $\quad P(X \leq 40)$

Z-Transformation: $Z = \frac{X-\mu}{\sigma} = \frac{40-30}{10} = 1$ $\quad P(Z \leq 1) = 0{,}8413$

Die Wahrscheinlichkeit beträgt 84,13%.

Demzufolge beträgt die Wahrscheinlichkeit dafür, dass der Großhändler die Nachfrage nach Impfstoff während der Nachlieferungszeit nicht bedienen kann 15,87%

$$P(X > 40) = 1 - P(X \leq 40) = 1 - 0{,}8413 = 0{,}1587$$

STANDARD-NORMALVERTEILUNG IV

Bestimmen Sie nun die Wahrscheinlichkeit, dass der Lagerbestand während der Nachlieferung zwischen 10 und 40 Packungen liegt?

Es wird also $P(10 < X < 40) = P(X \leq 39) - P(X \leq 10)$ gesucht.

Z-Transformation: $Z_{40} = \frac{X-\mu}{\sigma} = \frac{39-30}{10} = 0{,}9$

$P(Z \leq 0{,}9) = 0{,}8159$

$Z_{10} = \frac{X-\mu}{\sigma} = \frac{10-30}{10} = -2$

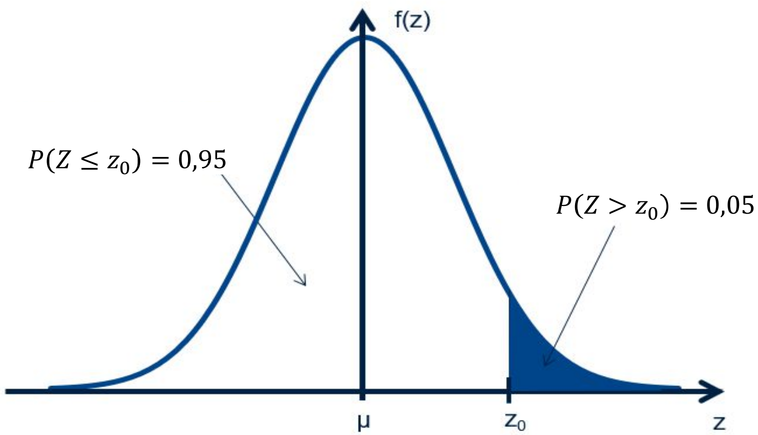
$P(Z \leq -2) = 1 - P(Z \leq 2) = 1 - 0{,}9772 = 0{,}0228$

Wahrscheinlichkeiten: $P(-2 \leq Z \leq 0{,}9) = P(X \leq 0{,}9) - P(X \leq -2) = 0{,}8159 + -0{,}0228 = 0{,}7913$

Die Wahrscheinlichkeit beträgt 79,13%.

STANDARD-NORMALVERTEILUNG V

Wenn der Großhändler nur mit einer Wahrscheinlichkeit von höchstens 5% die Nachfrage nicht bedienen können will, stellt sich die Frage ab welcher Menge bestellt werden muss?



z	0	1	2	3	4	5	6	7	8	9
...	...	...	...	...	...	...	...	...	...	...
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9392	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
...	...	...	...	...	...	...	...	...	...	...

Somit ergibt sie $z_o = 1{,}645$.

Z-Transformation (Rückwärts): $x_0 = \mu + \sigma \cdot z_0 = 30 + 10 \cdot 1{,}645 = 46{,}45$

Der Großhändler muss ab einer Menge von 47 Packungen nachbestellen.

AUFGABEN I

1. In einer Fabrik werden Tüten mit Kartoffelchips hergestellt.
   Das durchschnittliche Gewicht soll nach Angaben des Werks 200 Gramm betragen.

   Da die Tüten maschinell befüllt werden, wird dieser Wert nur mit einer Standardabweichung von 5 Gramm eingehalten.

   Wie groß ist die Wahrscheinlichkeit, dass eine Tüte mindestens 210 Gramm wiegt?

2. Das Schlachtgewicht von Schafen sei normalverteilt mit einem Mittelwert von $\mu = 60\ kg$ und einer Standardabweichung von $\sigma = 10\ kg$.

   a) Wie groß ist wie Wahrscheinlichkeit, dass das Schlachtgewicht zwischen 50kg und 65kg liegt?

   b) Die Wahrscheinlichkeit, dass das Mindestschlachtgewicht überschritten wird beträgt 5%.
      Bestimmen Sie das Mindestschlachtgewicht.

AUFGABEN II

3. Die durchschnittliche Niederschlagsmenge pro Jahr sei normalverteilt mit dem Erwartungswert $\mu = 1.300\ mm$ und der Varianz $\sigma^2 = 62.500\ mm^2$.

a) Wie groß ist die Wahrscheinlichkeit, dass es in einem Jahr mehr als $1.500\ mm$ regnet?

b) Bestimmen Sie das neunte Dezentil der Verteilung der Niederschlagsmenge und interpretieren Sie dieses.

c) Aufgrund der Klimaveränderung teigt die mittlere Niederschlagsmenge auf $1.400\ mm$. Welchen Wert muss die Varianz annehmen, damit sich das berechnete 90%-Quantil nicht ändert?