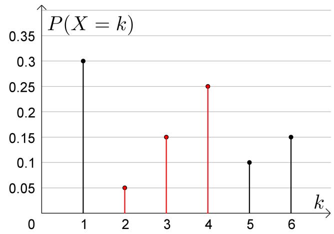
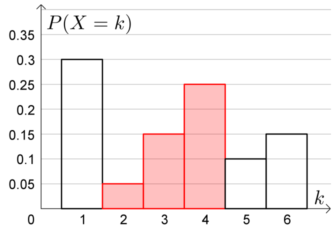
Wahrscheinlichkeiten als Flächeninhalte

Eine Zufallsvariable X kann die Werte 1, 2, 3, 4, 5 oder 6 annehmen.
Die zugehörigen Wahrscheinlichkeiten sind rechts in einem Stabdiagramm dargestellt. Trage die Wahrscheinlichkeiten (in %) unten in die Tabelle ein.



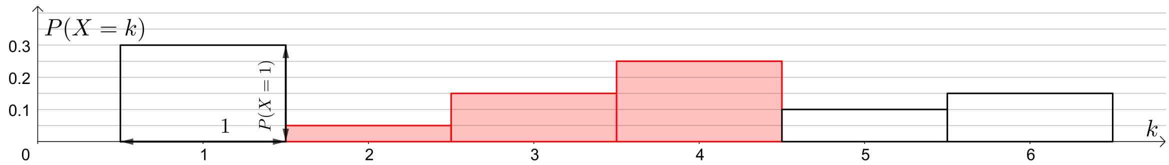
k	1	2	3	4	5	6
$P(X=k)$	30 %	5 %	15 %	25 %	10 %	15 %

Wir wandeln das Stabdiagramm in ein bestimmtes Säulendiagramm um: Dazu ersetzen wir jeden Stab jeweils durch ein Rechteck mit derselben Höhe und der *Breite* 1. Das Ergebnis ist rechts dargestellt.



Der **Flächeninhalt** jeder Säule ist damit gleich groß wie die entsprechende **Wahrscheinlichkeit**.

Wir können die beiden Achsen unabhängig voneinander skalieren.
Die *tatsächlichen* Breiten und Höhen der Säulen – und damit der Flächeninhalt – bleiben unverändert:

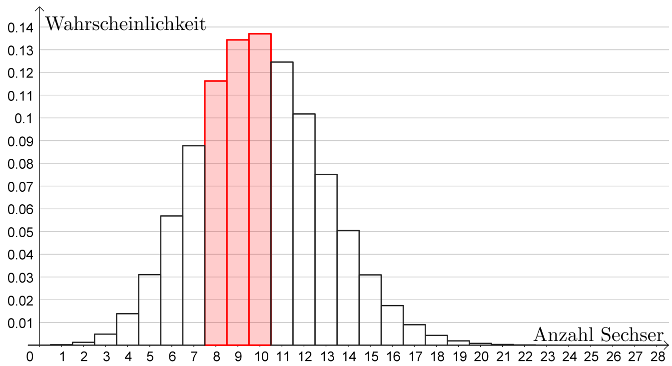


Die Wahrscheinlichkeit $P(2 \leq X \leq 4)$ ist in jedem der drei Bilder rot hervorgehoben. Es gilt:

$$P(2 \leq X \leq 4) = 5\,\% + 15\,\% + 25\,\% = 45\,\%$$

Binomialverteilung

Wir würfeln n Mal mit einem fairen 6-seitigen Würfel.
Die Zufallsvariable S_n gibt die Anzahl der geworfenen Sechser an.
Rechts unten sind die Wahrscheinlichkeiten bei $n = 60$ Würfen in einem Säulendiagramm dargestellt.
Gesucht ist die Wahrscheinlichkeit, dabei mindestens 8 Sechser, aber höchstens 10 Sechser zu würfeln.



1) Markiere rechts die zugehörige Fläche.

2) Ermittle die Wahrscheinlichkeit mithilfe des Säulendiagramms näherungsweise:

$$P(8 \leq S_{60} \leq 10) \approx 39\,\%$$

3) Ermittle die Wahrscheinlichkeit mit Technologieeinsatz.
Binomialverteilung: $n = 60$, $p = \frac{1}{6}$

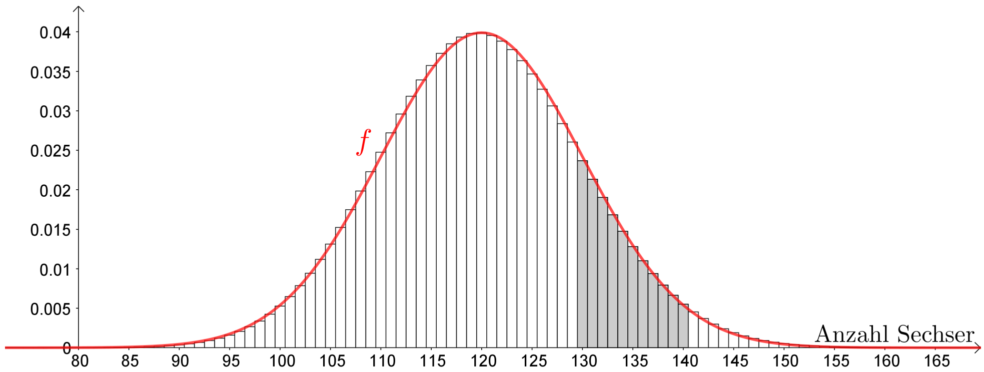
$$P(8 \leq S_{60} \leq 10) = 38{,}75...\,\%$$

4) Trage Zahlen richtig in die Kästchen ein:

$$P(8 \leq S_{60} \leq 10) = \underbrace{\binom{60}{8} \cdot \left(\frac{1}{6}\right)^{8} \cdot \left(\frac{5}{6}\right)^{52}}_{P(S_{60}=8)} + \underbrace{\binom{60}{9} \cdot \left(\frac{1}{6}\right)^{9} \cdot \left(\frac{5}{6}\right)^{51}}_{P(S_{60}=9)} + \underbrace{\binom{60}{10} \cdot \left(\frac{1}{6}\right)^{10} \cdot \left(\frac{5}{6}\right)^{50}}_{P(S_{60}=10)}$$

Von der Binomialverteilung zur Normalverteilung

Bei $n = 720$ Würfen nimmt das Säulendiagramm von S_{720} die folgende Form an:



Animationen:
y-Achse variabel skaliert
y-Achse fix skaliert

Der Flächeninhalt der grau markierten Säulen ist die Wahrscheinlichkeit $P(130 \leq S_{720} \leq 140)$.

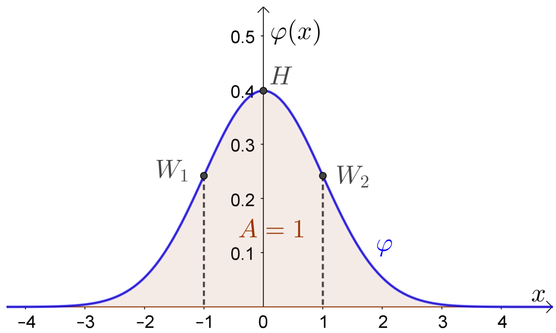
Wir haben auch den Graphen einer ganz bestimmten Funktion f eingezeichnet.
Dieser Graph schmiegt sich an das glockenförmige Profil des Säulendiagramms an.
Wie würdest du mithilfe von f diese Wahrscheinlichkeit *näherungsweise* berechnen?

$$P(130 \leq S_{720} \leq 140) \approx \int_{130}^{140} f(x)\,\mathrm{d}x \qquad \text{oder} \qquad \int_{129{,}5}^{140{,}5} f(x)\,\mathrm{d}x$$

Dichtefunktion der Standardnormalverteilung

Die Funktion φ mit $\varphi(x) = \dfrac{1}{\sqrt{2\cdot\pi}} \cdot e^{-\frac{1}{2}\cdot x^2}$ heißt **Dichtefunktion** der **Standardnormalverteilung**.

Die Dichtefunktion φ hat die folgenden Eigenschaften:



1) Alle Funktionswerte sind positiv. Warum?

$\dfrac{1}{\sqrt{2\cdot\pi}} > 0$ und $e^z = (2{,}718...)^z > 0$ für alle $z \in \mathbb{R}$

Also ist $\varphi(x)$ als Produkt positiver Zahlen positiv.

2) Der größte Funktionswert ist an der Stelle $x = 0$.

$$\varphi'(x) = \underbrace{\frac{1}{\sqrt{2\cdot\pi}} \cdot e^{-\frac{1}{2}\cdot x^2} \cdot (-x)}_{\text{Kettenregel}} = \varphi(x)\cdot(-x)$$

3) Die beiden Wendestellen sind $x = -1$ und $x = 1$.

$$\varphi''(x) = \underbrace{\varphi'(x)\cdot(-x) + \varphi(x)\cdot(-1)}_{\text{Produktregel}} = \varphi(x)\cdot x^2 - \varphi(x) = \varphi(x)\cdot\left(x^2-1\right) = \varphi(x)\cdot(x+1)\cdot(x-1)$$

4) Aus $x^2 = (-x)^2$ folgt $\varphi(x) = \varphi(-x)$ für alle $x \in \mathbb{R}$.

Der Graph ist deshalb symmetrisch zur senkrechten Achse. φ ist eine gerade Funktion.

5) Aus $\lim\limits_{n\to\infty} e^{-n} = \lim\limits_{n\to\infty} \frac{1}{e^n} = 0$ folgt: $\lim\limits_{x\to\infty} \varphi(x) = 0$ bzw. $\lim\limits_{x\to-\infty} \varphi(x) = 0$

6) Auch wenn φ keine elementare Stammfunktion hat, kann man folgende Eigenschaft zeigen:

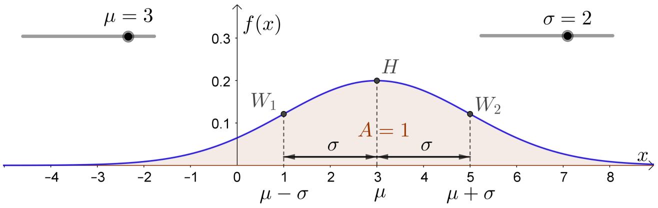
$$A = \int_{-\infty}^{\infty} \varphi(x)\,\mathrm{d}x = 1 = 100\,\%$$

Dichtefunktion der Normalverteilung

Die Funktion $\boldsymbol{f}$ mit

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2}$$

heißt **Dichtefunktion** der **Normalverteilung** mit **Erwartungswert** $\boldsymbol{\mu}$ und **Standardabweichung** $\boldsymbol{\sigma > 0}$.



Die beiden Funktionsgleichungen

$$\varphi(x) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot x^2} \quad \text{und} \quad f(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2}$$

sind eng miteinander verknüpft. Tatsächlich entsteht der Graph von f aus dem Graphen von φ in 3 Schritten:

i) Horizontale Skalierung um den Faktor σ:

$$f_1(x) = \varphi\left(\frac{x}{\sigma}\right) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{x}{\sigma}\right)^2}$$

ii) Horizontale Verschiebung um μ Einheiten nach rechts:

$$f_2(x) = f_1(x - \mu) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2}$$

iii) Vertikale Skalierung um den Faktor $\frac{1}{\sigma}$:

$$f(x) = \frac{f_2(x)}{\sigma} = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2}$$

Die Dichtefunktion f hat die folgenden Eigenschaften:

1) Alle Funktionswerte sind positiv.

2) Der größte Funktionswert ist an der Stelle $x = \mu$.

3) Die beiden Wendestellen sind $x = \mu - \sigma$ und $x = \mu + \sigma$.

4) Es gilt $f(\mu + x) = f(\mu - x)$ für alle $x \in \mathbb{R}$.
Also ist der Graph symmetrisch zur senkrechten Gerade $x = \mu$.

5) Es gilt: $\lim\limits_{x \to \infty} f(x) = 0$ bzw. $\lim\limits_{x \to -\infty} f(x) = 0$

6) Der gesamte Flächeninhalt zwischen der x-Achse und dem Funktionsgraphen ist 1.

7) Eine Veränderung von μ bewirkt eine Verschiebung des Graphen in x-Richtung.

8) Je größer σ ist, desto kleiner ist der größte Funktionswert und desto größer ist die Entfernung der beiden Wendestellen.

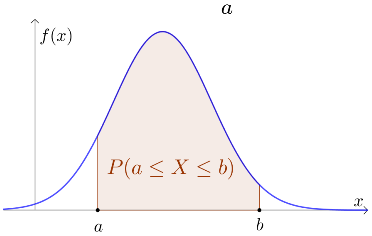
9) Je kleiner σ ist, desto größer ist der größte Funktionswert und desto kleiner ist die Entfernung der beiden Wendestellen.

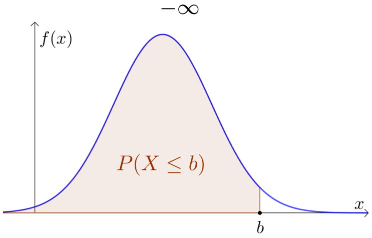
10) f ist die Dichtefunktion der Standardnormalverteilung, wenn $\mu = 0$ und $\sigma = 1$ ist.

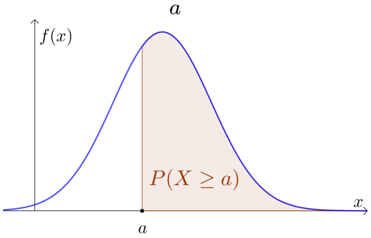
Normalverteilte Zufallsvariable

f ist die Dichtefunktion der Normalverteilung mit Erwartungswert μ und Standardabweichung $\sigma > 0$.
Eine **Zufallsvariable** $\boldsymbol{X}$ heißt **normalverteilt** mit **Erwartungswert** $\boldsymbol{\mu}$ und **Standardabweichung** $\boldsymbol{\sigma}$, wenn die folgenden Gleichungen für alle reellen Zahlen $a \le b$ gelten:

$$P(a \le X \le b) = \int_a^b f(x)\,\mathrm{d}x \qquad P(X \le b) = \int_{-\infty}^b f(x)\,\mathrm{d}x \qquad P(X \ge a) = \int_a^{\infty} f(x)\,\mathrm{d}x$$







Eine normalverteilte Zufallsvariable kann also jede reelle Zahl als Wert annehmen.

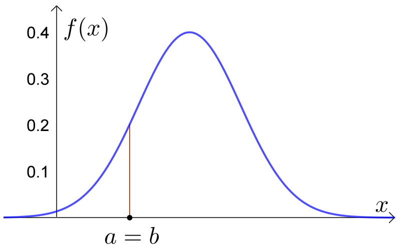
$f(a) \neq P(X = a)$ MmF

X ist eine normalverteilte Zufallsvariable mit Dichtefunktion f.

Wie wahrscheinlich ist es, dass X *genau* den Wert $a \in \mathbb{R}$ annimmt?

Für diese Wahrscheinlichkeit gilt:

$$P(X = a) = P(a \leq X \leq a) = \int_a^a f(x)\,\mathrm{d}x = 0$$



Wenn X eine normalverteilte Zufallsvariable ist, dann gilt also:

$$P(X \leq a) = P(X < a) + P(X = a) = P(X < a)$$

Das ist anders als bei der Binomialverteilung.

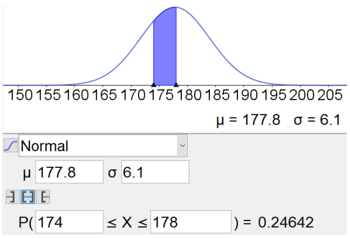
Körpergröße MmF

Die Körpergröße X von 42-jährigen Männern ist annähernd normalverteilt mit Erwartungswert $\mu = 177{,}8\,\text{cm}$ und Standardabweichung $\sigma = 6{,}1\,\text{cm}$. Ein 42-jähriger Mann wird zufällig ausgewählt.

a) Berechne mit Technologieeinsatz die Wahrscheinlichkeit, dass seine Körpergröße …

… im Intervall $[174\,\text{cm}; 178\,\text{cm}]$ liegt.

$$P(174 \leq X \leq 178) = 0{,}2464... = 24{,}64...\,\%$$

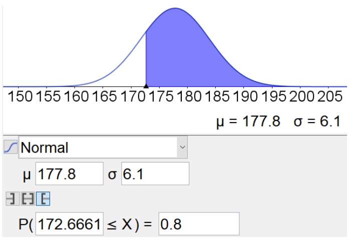


… größer als 190 cm ist.

$$P(X > 190) = P(X \geq 190) = 0{,}0227... = 2{,}27...\,\%$$

… kleiner als 178 cm ist.

$$P(X < 178) = P(X \leq 178) = 0{,}5130... = 51{,}30...\,\%$$



b) Welche Körpergröße übertrifft er mit der Wahrscheinlichkeit 80 %?

$$P(X > a) = 80\,\% = 0{,}8 \quad \Longrightarrow \quad a = 172{,}66...\,\text{cm}$$

Welche Körpergröße übertrifft er mit der Wahrscheinlichkeit 45 % *nicht*?

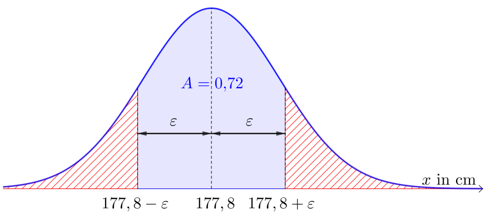
$$P(X \leq b) = 45\,\% = 0{,}45 \quad \Longrightarrow \quad b = 177{,}03...\,\text{cm}$$

c) In welchem um μ symmetrischen Intervall liegt die Körpergröße mit der Wahrscheinlichkeit 72 %?

Dieses Intervall heißt auch zweiseitiger 72 %-Zufallsstreubereich für einen Einzelwert von X.

Rechts ist die zugehörige Fläche blau markiert. Welchen Inhalt haben die beiden rot schraffierten Flächen links und rechts vom um μ symmetrischen Intervall jeweils?

$$\frac{1 - 0{,}72}{2} = 0{,}14$$



Berechne die beiden Intervallgrenzen mit Technologieeinsatz:

$$P(X \leq 171{,}2...) = 14\,\% \qquad P(X \geq 184{,}3...) = 14\,\%$$

Die Körpergröße befindet sich also mit der Wahrscheinlichkeit 72 %

im Intervall $[171{,}2...\,\text{cm}; 184{,}3...\,\text{cm}]$.

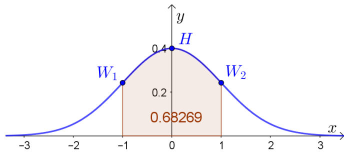
σ-Umgebungen

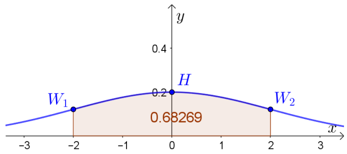
Die Zufallsvariable X ist normalverteilt mit Erwartungswert μ und Standardabweichung $\sigma > 0$.
Man kann zeigen, dass die Wahrscheinlichkeit

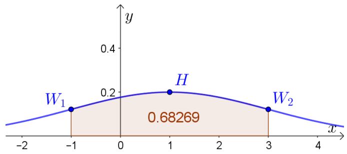
$$P(\mu - k \cdot \sigma \leq X \leq \mu + k \cdot \sigma)$$

nicht von μ bzw. σ abhängt, sondern nur von $k \geq 0$.

Zum Beispiel ist in jedem der drei Bilder rechts $k = 1$.







Ermittle die folgenden Wahrscheinlichkeiten mit Technologieeinsatz.
Runde jeweils auf eine Nachkommastelle.

1) $P(\mu - 1 \cdot \sigma \leq X \leq \mu + 1 \cdot \sigma) \approx 68{,}3\,\%$

2) $P(\mu - 2 \cdot \sigma \leq X \leq \mu + 2 \cdot \sigma) \approx 95{,}4\,\%$

3) $P(\mu - 3 \cdot \sigma \leq X \leq \mu + 3 \cdot \sigma) \approx 99{,}7\,\%$

σ-Umgebungen

Die Zufallsvariable X ist normalverteilt mit $\mu = 40\,\text{g}$ und $\sigma = 2\,\text{g}$. Du erzeugst eine große Stichprobe mit Werten aus dieser Normalverteilung. Beantworte die folgenden Fragen ohne Technologieeinsatz.

a) In welchem um μ symmetrischen Intervall sollten rund 95 % der Werte liegen?

b) Ungefähr wieviel Prozent der Werte sollten kleiner als 38 g sein?

a) $[\mu - 2 \cdot \sigma; \mu + 2 \cdot \sigma] = [36\,\text{g}; 44\,\text{g}]$

b) $\approx 16\,\%$ (Außerhalb der $1 \cdot \sigma$-Umgebung liegen rund 32 % der Werte.)

Standardisierung

Die Zufallsvariable X ist normalverteilt mit Erwartungswert μ und Standardabweichung $\sigma > 0$.

Dann ist die Zufallsvariable $\boldsymbol{Z = \dfrac{X - \mu}{\sigma}}$ standardnormalverteilt – also mit $\mu_Z = 0$ und $\sigma_Z = 1$.

Das heißt: Erfüllen die Intervallgrenzen von $[x_1; x_2]$ und $[z_1; z_2]$ die Gleichungen

$$z_1 = \frac{x_1 - \mu}{\sigma} \qquad \text{und} \qquad z_2 = \frac{x_2 - \mu}{\sigma},$$

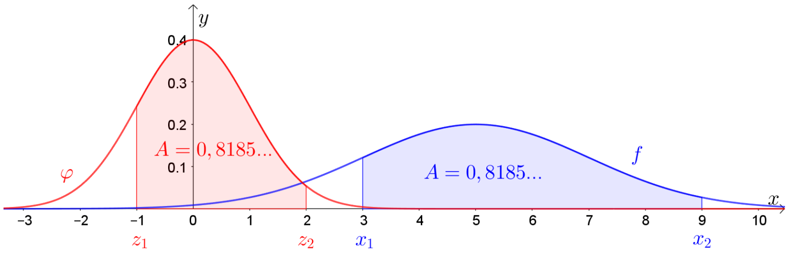
dann gilt: $P(x_1 \leq X \leq x_2) = P(z_1 \leq Z \leq z_2)$ Insbesondere gilt: $P(\mu - k \cdot \sigma \leq X \leq \mu + k \cdot \sigma) = P(-k \leq Z \leq k)$

Für die normalverteilte Zufallsvariable X im Bild rechts unten gilt $\mu = 5$ und $\sigma = 2$.
Trage Zahlen richtig in die Kästchen ein.

$x_1 = 3 \implies z_1 = \dfrac{3 - 5}{2} = -1$

$x_2 = 9 \implies z_2 = \dfrac{9 - 5}{2} = 2$

$P(3 \leq X \leq 9) = P(-1 \leq Z \leq 2)$



Allgemein gilt nämlich: $\underbrace{\displaystyle\int_{x_1}^{x_2} \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2} \mathrm{d}x}_{=P(x_1 \leq X \leq x_2)} = \underbrace{\displaystyle\int_{z_1}^{z_2} \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot z^2} \mathrm{d}z}_{P(z_1 \leq Z \leq z_2)}$ (Substitution $z = \frac{x-\mu}{\sigma}$ mit $\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{1}{\sigma}$)

Erwartungswert gesucht

Die normalverteilte Zufallsvariable X hat die Standardabweichung $\sigma = 2$. Weiters gilt: $P(X \leq 8) = 0{,}62$

Wie groß ist der Erwartungswert μ von X?

Dazu berechnen wir die entsprechende Intervallgrenze z der standardnormalverteilte Zufallsvariable Z:

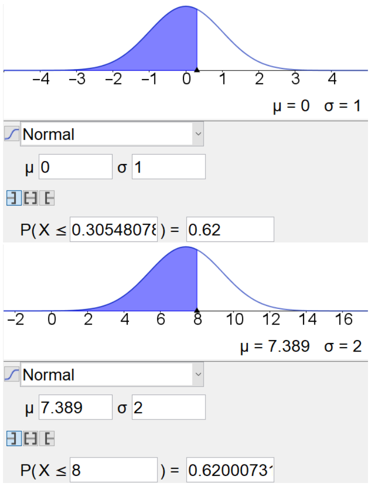
$$P(Z \leq z) = 0{,}62 \implies z = 0{,}305\,48\ldots$$

Die Intervallgrenze $x = 8$ von X entspricht also der Intervallgrenze $z = 0{,}305\,48\ldots$ von Z.

Setze in $z = \dfrac{x-\mu}{\sigma}$ ein, und berechne den Erwartungswert μ von X.

$$z \cdot \sigma = x - \mu \implies \mu = x - z \cdot \sigma = 7{,}389\ldots$$

Rechts kontrollieren wir das Ergebnis mit Technologieeinsatz.



Annäherung: Binomialverteilung – Normalverteilung

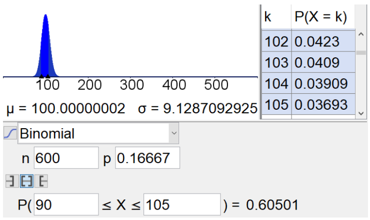
Du würfelst 600 Mal mit einem fairen 6-seitigen Würfel.
Wie wahrscheinlich ist es, dass du dabei mindestens 90, aber höchstens 105 Sechser würfelst?
Die Zufallsvariable X gibt die Anzahl der geworfenen Sechser an.

1) X ist binomialverteilt mit $n = 600$ und $p = \frac{1}{6}$.

$P(90 \leq X \leq 105) = 60{,}50\ldots\%$

Erwartungswert: $E(X) = n \cdot p = 100$

Standardabweichung: $\sigma(X) = \sqrt{n \cdot p \cdot (1-p)} = 9{,}128\ldots$

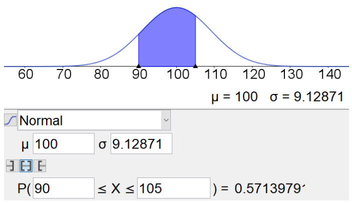


2) Wir können diese Wahrscheinlichkeit auch durch eine normalverteilte Zufallsvariable X' mit gleichem Erwartungswert und gleicher Standardabweichung annähern. Satz von Moivre-Laplace

Erwartungswert: $\mu = E(X) = 100$

Standardabweichung: $\sigma = \sigma(X) = 9{,}128\ldots$

$P(90 \leq X' \leq 105) = 57{,}13\ldots\%$

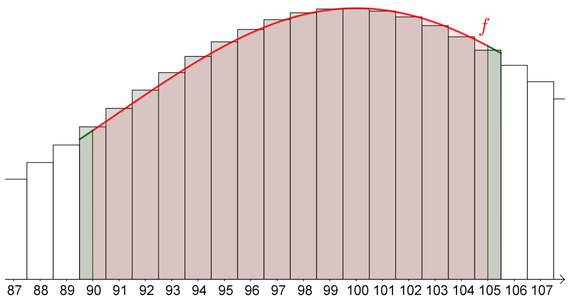


Rechts unten ist das Säulendiagramm der binomialverteilten Zufallsvariable X dargestellt.

Wodurch entsteht der (relativ große) Unterschied zwischen $P(90 \leq X \leq 105)$ und $\displaystyle\int_{90}^{105} f(x)\,\mathrm{d}x$?

Eine bessere Annäherung erhalten wir durch Anwendung der sogenannten Stetigkeitskorrektur:

$$P(89{,}5 \leq X' \leq 105{,}5) = 60{,}15\ldots\%$$



Warum nähert man die Binomialverteilung an? GeoGebra scheitert beim Versuch der exakten Berechnung von $\binom{100\,000}{42\,000}$ zurecht.
Bei $n = 100\,000$ wird die Binomialverteilung aber relativ gut durch die Normalverteilung angenähert.