

§23 Die Normalverteilung



Unter aller Verteilung tritt die Normalverteilung in den Anwendungen am häufigsten auf. Der Grund dafür ist der **zentrale Grenzwertungssatz**. Wir werden diesen zentralen Grenzwertungssatz und seine Auswirkungen im Detail diskutieren und uns anschließend ausführlich mit der Normalverteilung und ihren zahlreichen Anwendungen befassen.

Mit der Normalverteilung sind eine Reihe von Verteilungen eng verbunden, welche vor allem in der Statistik eine große Rolle spielen. Es handelt sich dabei um die Chi-Quadrat Verteilung, die Student T Verteilung und die Fisher F Verteilung. Wir werden diese Verteilungen sowie die zu diesen Verteilungen führenden Transformationen ausführlich besprechen und damit bereits viel Vorarbeit für im Rahmen der Statistik durchzuführende Berechnungen leisten.

Darüber hinaus befassen wir uns in diesem Kapitel auch mit der Multinormalverteilung, welche ein natürliches mehrdimensionales Analogon zur Normalverteilung darstellt.

23.1 Der zentrale Grenzwertungssatz

Der zentrale Grenzwertungssatz besagt, dass die Summe einer großen Anzahl von vollständig unabhängigen und identisch verteilten, quadratisch integrierbaren Zufallsvariablen annähernd normalverteilt ist. Bevor wir diesen Satz exakt formulieren, wollen wir ihn an einigen Spezialfällen veranschaulichen:

23.1.1 Bemerkung: In [Beispiel 20.4.8](#) haben wir gezeigt, dass sich die $\mathcal{B}[n, p]$ -Verteilung für große n durch die Normalverteilung mit den Parametern $\mu = n p$ und $\sigma = \sqrt{n p (1 - p)}$ approximieren lässt. Dieses Ergebnis lässt sich auch folgendermaßen interpretieren: Sind die Zufallsvariablen X_1, X_2, \dots, X_n vollständig unabhängig und identisch $\mathcal{B}[1, p]$ -verteilt, so ist ihre **bekanntlich** $\mathcal{B}[n, p]$ -verteilte Summe $Z_n = X_1 + X_2 + \dots + X_n$ annähernd $\mathcal{N}[n p, \sqrt{n p (1 - p)}]$ -verteilt.

Wir erläutern die Auswirkungen dieser Erkenntnis an einigen Beispielen:

23.1.2 Beispiel: Eine homogene Münze wird n mal geworfen. Man bestimme die asymptotische Verteilung der Anzahl der **Runs** und berechne damit näherungsweise die Wahrscheinlichkeit dafür, dass bei $n = 100$ Würfeln mehr als $r = 55$ Runs auftreten.



Lösung: Die Zufallsvariable X_i sei gleich 1 bzw 0 je nachdem, ob beim i -ten Wurf ein Adler bzw eine Zahl erscheint. Ist $g[x, y] = 1 - \delta_{x,y}$ (wobei $\delta_{x,y}$ das Kroneckersymbol bezeichnet), so gilt für die Anzahl Z_n der Runs

$$Z_n = 1 + g[X_1, X_2] + g[X_2, X_3] + \dots + g[X_{n-1}, X_n]$$

Nun sind aber die Zufallsvariablen $g[X_1, X_2], g[X_2, X_3], \dots, g[X_{n-1}, X_n]$ offensichtlich vollständig unabhängig und wegen

$$\mathbb{P}\{g[X_{k-1}, X_k] = 1\} = \mathbb{P}\{X_{k-1} \neq X_k\} = 1/2$$

identisch $\mathcal{B}[1, 1/2]$ -verteilt. Ihre Summe ist daher $\mathcal{B}[n - 1, 1/2]$ -verteilt. Wegen [Bemerkung 23.1.1](#) ist Z_n für große n somit näherungsweise $\mathcal{N}[\mu, \sigma]$ -verteilt mit

$$\mu = \mathbb{E}[Z_n] = 1 + (n - 1)/2 = (n + 1)/2 \quad \text{und} \quad \sigma = \sqrt{\mathbb{V}[Z_n]} = \sqrt{(n - 1)/4}$$

Für die gesuchte Wahrscheinlichkeit $\mathbb{P}\{Z_n > r\}$ gilt damit näherungsweise

```

n = 100; r = 55;
1 - CDF[NormalDistribution[(n + 1)/2, Sqrt[(n - 1)/4]], r] // N
Clear[n, r]

0.182856

```

23.1.3 Beispiel: An einer Wahl zwischen zwei Kandidaten A und B nehmen $n = 1\,000\,000$ Wähler teil. Davon kennen $a = 2000$ Wähler den Kandidaten A aus Wahlkampfveranstaltungen und stimmen geschlossen für ihn. Die übrigen Wähler sind mehr oder weniger unentschlossen und treffen ihre Entscheidung unabhängig voneinander durch Werfen einer Münze. Wie groß ist die Wahrscheinlichkeit p für einen Sieg von Kandidat A ?

▼

Lösung: Die Zufallsvariable X_i sei gleich 1 bzw 0 je nachdem, ob der i -te unentschlossene Wähler Kandidat A bzw Kandidat B wählt. Die Anzahl $Z_{n-a} = X_1 + X_2 + \dots + X_{n-a}$ der unentschlossenen Wähler, welche für Kandidat A stimmen ist **daher** $\mathcal{B}[(n-a), 1/2]$ -verteilt und **damit** annähernd $\mathcal{N}[(n-a)/2, \sqrt{(n-a)/4}]$ -verteilt. Für die von uns gesuchte Wahrscheinlichkeit $p = \mathbb{P}\{Z_{n-a} > 498\,000\}$ gilt somit näherungsweise

```

n = 1000000; a = 2000; s = 498000;
1 - CDF[NormalDistribution[(n - a)/2, Sqrt[(n - a)/4]], s] // N
Clear[n, a, s]

0.977358

```

23.1.4 Beispiel: Die Wahrscheinlichkeit $\mathbb{P}[A]$ eines Ereignisses A lässt sich bekanntlich durch die relative Häufigkeit $H[A]$ approximieren, wobei die folgende **Faustregel** gilt: "Soll der Fehler kleiner als 10^{-k} sein, so sind dazu etwa $n = 10^{2k}$ Wiederholungen erforderlich." Wir wollen diese Faustregel nun wahrscheinlichkeitstheoretisch begründen und fragen dazu: "Wie oft muss ein Zufallsexperiment unabhängig wiederholt werden, um mit einer vorgegebenen Wahrscheinlichkeit p sicher zu stellen, dass sich die relative Häufigkeit $H[A]$ eines Ereignisses von der Wahrscheinlichkeit $\mathbb{P}[A]$ dieses Ereignisses um weniger als ε unterscheidet?"

▼

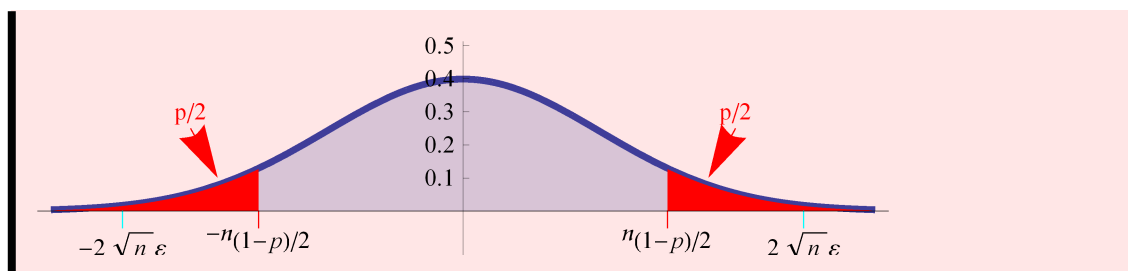
Lösung: Die Zufallsvariable X_i sei gleich 1 bzw 0 je nachdem, ob beim i -ten Versuch das Ereignis A eintritt bzw nicht eintritt. Ihre Summe $X_1 + X_2 + \dots + X_n$ ist **daher** $\mathcal{B}[n, \mathbb{P}[A]]$ -verteilt und **somit** für große n näherungsweise $\mathcal{N}[n\mathbb{P}[A], \sqrt{n\mathbb{P}[A](1-\mathbb{P}[A])}]$ -verteilt. Aus dem **Satz über die affine Transformation** von Normalverteilungen folgt damit

$$\begin{aligned}
\mathbb{P}\{ |H[A] - \mathbb{P}[A]| < \varepsilon \} &= \mathbb{P}\left\{ \left| \frac{X_1 + X_2 + \dots + X_n - n\mathbb{P}[A]}{\sqrt{n\mathbb{P}[A](1-\mathbb{P}[A])}} \right| < \frac{\sqrt{n}}{\sqrt{\mathbb{P}[A](1-\mathbb{P}[A])}} \varepsilon \right\} \geq \\
&\geq \mathbb{P}\left\{ \left| \frac{X_1 + X_2 + \dots + X_n - n\mathbb{P}[A]}{\sqrt{n\mathbb{P}[A](1-\mathbb{P}[A])}} \right| < 2\sqrt{n}\varepsilon \right\} \approx \mathcal{N}[0, 1][[-2\sqrt{n}\varepsilon, 2\sqrt{n}\varepsilon]]
\end{aligned}$$

wobei bei der Ungleichung " \geq " die Tatsache verwendet wurde, dass stets $\mathbb{P}[A](1-\mathbb{P}[A]) \leq 1/4$ gilt. Wählt man nun n so, dass $2\sqrt{n}\varepsilon \geq n_{(1+p)/2}$ ist, wobei wir mit n_q das q -Quantil der $\mathcal{N}[0, 1]$ -Verteilung bezeichnen, so gilt offenbar

$$\mathbb{P}\{ |H[A] - \mathbb{P}[A]| < \varepsilon \} \geq \mathcal{N}[0, 1][[-2\sqrt{n}\varepsilon, 2\sqrt{n}\varepsilon]] = \mathcal{N}[0, 1][[-n_{(1+p)/2}, n_{(1+p)/2}]] \geq p$$

An Hand der folgenden Zeichnung wird dieser Sachverhalt sofort klar:



Um mit einer vorgegebenen Wahrscheinlichkeit p sicher zu stellen, dass sich die relative Häufigkeit $\mathbb{H}[A]$ eines Ereignisses von der Wahrscheinlichkeit $\mathbb{P}[A]$ dieses Ereignisses um weniger als ϵ unterscheidet, muss somit

$$n \geq \frac{1}{4} (n_{(1+p)/2})^2 \epsilon^{-2}$$

gelten. Wir tabellieren den Faktor $\alpha = (n_{(1+p)/2})^2/4$ für verschiedene Werte von p mit *Mathematica* und erkennen dabei, wie unsere [Faustregel](#) zu verstehen ist: Soll etwa mit einer Wahrscheinlichkeit von $p = 0.95$ sicher gestellt werden, dass sich die relative Häufigkeit $\mathbb{H}[A]$ eines Ereignisses von der Wahrscheinlichkeit $\mathbb{P}[A]$ dieses Ereignisses um weniger als 10^{-k} unterscheidet, so sind dazu $n = 0.960365 \times 10^{2k}$ Wiederholungen erforderlich.

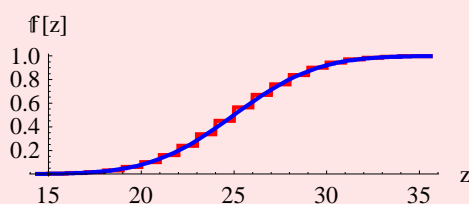
```
TableForm[Table[{p, Quantile[NormalDistribution[0, 1], (1 + p)/2]^2/4}, {p, 0.9, 0.99, 0.01}],
  TableSpacing -> {1, 5}, TableHeadings -> {None, {"p", "alpha"}}]
```

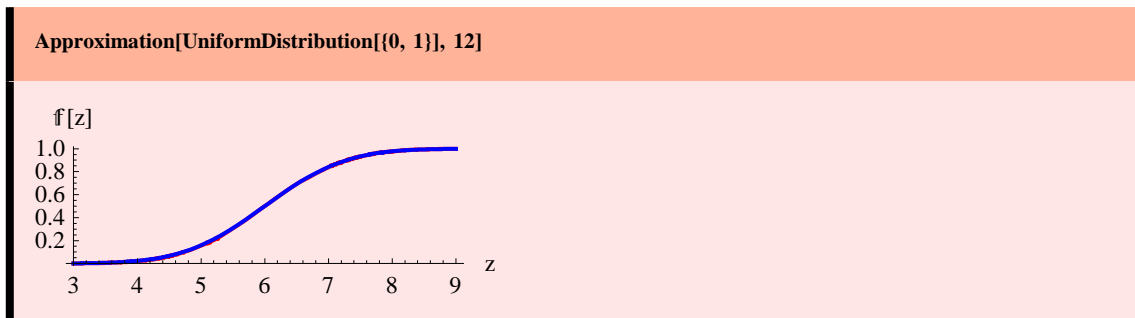
p	α
0.9	0.676386
0.91	0.718593
0.92	0.766225
0.93	0.820755
0.94	0.884346
0.95	0.960365
0.96	1.05447
0.97	1.17732
0.98	1.35297
0.99	1.65872

Nach diesem Spezialfall des zentralen Grenzwertungssatzes, mit dem gezeigt wurde, dass die Summe von vielen vollständig unabhängigen, identisch $\mathcal{B}[1, p]$ -verteilten Zufallsvariablen annähernd normalverteilt ist, wollen wir nun zeigen, dass auch die Summe von vielen vollständig unabhängigen, identisch verteilten, quadratisch integrierbaren Zufallsvariablen annähernd normalverteilt ist, wobei über die Verteilung dieser Zufallsvariablen keine weiteren Aussagen gemacht werden:

Mit dem Befehl `Approximation[distr, n]` wird graphisch gezeigt, wie gut sich die (durch Simulation auf der Basis von 10000 Wiederholungen gewonnene empirische) **Verteilungsfunktion** der Summe von n vollständig unabhängigen, identisch gemäß der (in *Mathematica* implementierten) Verteilung *distr* verteilten Zufallsvariablen durch die **Verteilungsfunktion** einer Normalverteilung, deren Erwartungswert und Streuung mit dem Erwartungswert und der Streuung der Summe dieser n Zufallsvariablen überein stimmt, approximieren lässt:

```
Approximation[DiscreteUniformDistribution[{0, 1}], 50]
```





Durch Experimentieren mit verschiedenen (diskreten und stetigen) Verteilungen gewinnt man also die Erkenntnis:

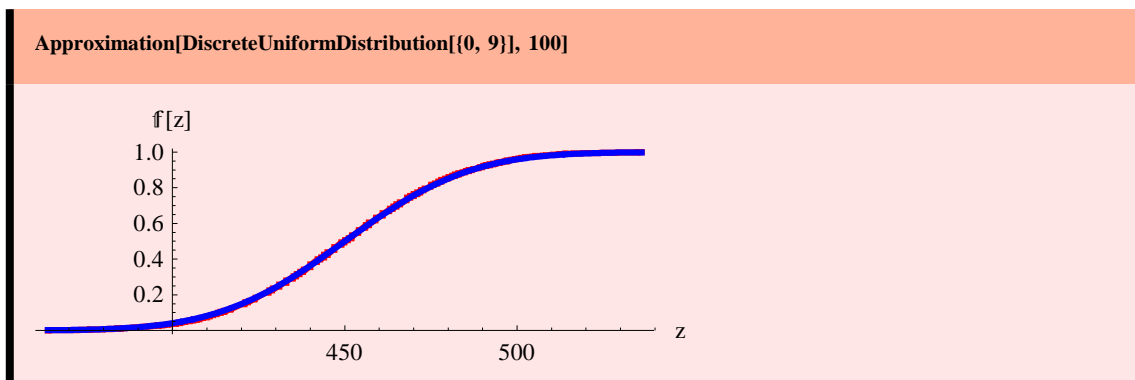
23.1.5 Bemerkung: Für großes n lässt sich die Verteilung der Summe $Z_n = X_1 + X_2 + \dots + X_n$ von n vollständig unabhängigen, identisch \mathbb{P}_X -verteilten, quadratisch integrierbaren Zufallsvariablen durch eine Normalverteilung mit den Parametern $\mu = n \mathbb{E}[X]$ und $\sigma = \sqrt{n \mathbb{V}[X]}$ approximieren. Wie groß dabei n sein soll, ist jeweils im Einzelfall zu überprüfen.

Wir erläutern diese Erkenntnis wieder an Hand von Beispielen:

23.1.6 Beispiel: Aus dem Intervall $[0, 1]$ wird zufällig eine Zahl ausgewählt. Man bestimme die Verteilung der Summe der ersten n (n groß) Ziffern ihrer Dezimalbruchentwicklung und berechne damit näherungsweise die Wahrscheinlichkeit dafür, dass die Summe der ersten $n = 100$ Ziffern der Dezimalbruchentwicklung einer zufällig aus dem Intervall $[0, 1]$ ausgewählten Zahl größer als $s = 500$ ist.

▼

Lösung: Wir bezeichnen mit X_i die i -te Ziffer der Dezimalbruchentwicklung einer zufällig aus dem Intervall $[0, 1]$ ausgewählten Zahl. Die Zufallsvariablen X_1, X_2, \dots, X_n sind offenbar vollständig unabhängig und auf der Menge $\{0, 1, 2, \dots, 9\}$ gleichverteilt. Wegen [Bemerkung 23.1.5](#) (man vergleiche auch die folgende Graphik)



ist ihre Summe $Z_n = X_1 + X_2 + \dots + X_n$ damit annähernd $\mathcal{N}[\mu, \sigma]$ -verteilt mit $\mu = 9n/2$ und $\sigma = \sqrt{33n}/2$.

```
{μ = n Mean[DiscreteUniformDistribution[{0, 9}],
σ = Sqrt[n] StandardDeviation[DiscreteUniformDistribution[{0, 9}]]}
```

$$\left\{ \frac{9n}{2}, \frac{\sqrt{33} \sqrt{n}}{2} \right\}$$

Für die gesuchte Wahrscheinlichkeit $\mathbb{P}\{Z_n > s\}$ gilt somit näherungsweise

```
n = 100; s = 500;
1 - CDF[NormalDistribution[μ, σ], s] // N
Clear[μ, σ, n, s]

0.0408614
```

23.1.7 Beispiel: Sei $f : [0, 1] \rightarrow \mathbb{R}$ eine quadratisch integrierbare Funktion und X_1, X_2, \dots, X_n vollständig

unabhängige, auf dem Intervall $[0, 1]$ gleichverteilte Zufallsvariable. Wie groß muß n mindestens sein, um mit der Wahrscheinlichkeit p sicher zu stellen, dass

$$J_n = \frac{1}{n} \sum_{i=1}^n f[X_i] \quad \text{von} \quad J = \int_0^1 f[z] dz = \mathbb{E}[f[X_1]]$$

um weniger als ε abweicht? (Diese Überlegung ist wesentlich für die **Monte-Carlo-Methode** zur näherungsweise Berechnung von Integralen, da damit eine Aussage über den Fehler gemacht werden kann.)

▼

Lösung: Wegen **Bemerkung 23.1.5** ist die Summe $f[X_1] + f[X_2] + \dots + f[X_n]$ für große n annähernd normalverteilt mit den Parametern $\mu = n \mathbb{E}[f[X_1]]$ und $\sigma = \sqrt{n \mathbb{V}[f[X_1]]}$. Aus dem **Satz über die affine Transformation** von Normalverteilungen folgt damit

$$\begin{aligned} \mathbb{P}\{|J_n - J| < \varepsilon\} &= \mathbb{P}\left\{\left| \frac{f[X_1] + f[X_2] + \dots + f[X_n] - n\mathbb{E}[f[X_1]]}{\sqrt{n \mathbb{V}[f[X_1]]}} \right| < \frac{\sqrt{n}}{\sqrt{\mathbb{V}[f[X_1]]}} \varepsilon\right\} \simeq \\ &\simeq \mathcal{N}[0, 1] \left[\left[-\frac{\sqrt{n}}{\sqrt{\mathbb{V}[f[X_1]]}} \varepsilon, \frac{\sqrt{n}}{\sqrt{\mathbb{V}[f[X_1]]}} \varepsilon \right] \right] \end{aligned}$$

Wählt man also n so, dass

$$\frac{\sqrt{n}}{\sqrt{\mathbb{V}[f[X_1]]}} \varepsilon \geq n_{(1+p)/2}$$

ist, wobei wir mit n_q wieder das q -Quantil der $\mathcal{N}[0, 1]$ -Verteilung bezeichnen, so gilt (an Hand einer Zeichnung wird dieser Sachverhalt unmittelbar klar)

$$\mathbb{P}\{|J_n - J| < \varepsilon\} \simeq \mathcal{N}[0, 1] \left[\left[-\frac{\sqrt{n}}{\sqrt{\mathbb{V}[f[X_1]]}} \varepsilon, \frac{\sqrt{n}}{\sqrt{\mathbb{V}[f[X_1]]}} \varepsilon \right] \right] \geq p$$

Um mit der vorgegebenen Wahrscheinlichkeit p sicher zu stellen, dass sich die Summe J_n vom Integral J um weniger als ε unterscheidet, muss daher

$$n \geq (n_{(1+p)/2})^2 \mathbb{V}[f[X_1]] \varepsilon^{-2} = (n_{(1+p)/2})^2 \left(\int_0^1 f[z]^2 dz - \left(\int_0^1 f[z] dz \right)^2 \right) \varepsilon^{-2}$$

sein. Ist man in der Lage, den Ausdruck $\int_0^1 f[z]^2 dz - \left(\int_0^1 f[z] dz \right)^2$ zu berechnen (oder abzuschätzen), so lässt sich aus dieser Beziehung die Größe n leicht ermitteln.

23.1.8 Beispiel: In **Beispiel 16.6.3** haben wir das Integral $\int_0^1 e^{-z^2} dz$ mit Hilfe der Monte-Carlo-Methode näherungsweise berechnet und ohne zu überlegen $n = 10^4$ gesetzt. Wie groß muss n aber tatsächlich sein, um mit einer Wahrscheinlichkeit von $p = 0.99$ sicher zu stellen, dass der mit der Monte-Carlo-Methode ermittelte Wert vom tatsächlichen Wert des Integrals um weniger als $\varepsilon = 10^{-2}$ abweicht?

▼

Lösung: In **Beispiel 23.1.7** haben wir gezeigt, dass

$$n \geq (n_{(1+p)/2})^2 \left(\int_0^1 e^{-2z^2} dz - \left(\int_0^1 e^{-z^2} dz \right)^2 \right) \varepsilon^{-2}$$

sein muss, um mit einer Wahrscheinlichkeit von $p = 0.99$ sicher zu stellen, dass der mit der Monte-Carlo-Methode ermittelte Wert vom tatsächlichen Wert des Integrals um weniger als $\varepsilon = 10^{-2}$ abweicht. Wegen

```
p = 0.99; ε = 10-2; quantil = Quantile[NormalDistribution[0, 1], (1 + p)/2];
var = Integrate[Exp[-2 z2], {z, 0, 1}] - (Integrate[Exp[-z2], {z, 0, 1}])2;
quantil2 var ε-2
Clear[p, ε, quantil, var]
```

```
2680.35
```

muss daher $n \geq 2681$ sein. In diesem Fall ist sicher gestellt, dass der mit der Monte-Carlo-Methode ermittelte Wert vom tatsächlichen Wert des Integrals in höchstens einem von 100 Fällen um mehr als 10^{-2} abweicht.

23.1.9 Beispiel: Aus [Bemerkung 23.1.5](#) folgt: Sind die Zufallsvariablen X_1, X_2, \dots, X_{12} vollständig unabhängig und im Intervall $[0, 1]$ gleichverteilt, so ist die Zufallsvariable $Z = X_1 + X_2 + \dots + X_{12}$ annähernd $\mathcal{N}[6, 1]$ -verteilt (diese Tatsache wird dazu verwendet, um aus im Intervall $[0, 1]$ gleichverteilten Zufallszahlen $\mathcal{N}[0, 1]$ -verteilte Zufallszahlen zu erzeugen). Man berechne den maximalen Unterschied zwischen der Verteilungsfunktion \mathbb{F}_Z der Zufallsvariablen Z und der Verteilungsfunktion ϕ der $\mathcal{N}[0, 1]$ -Verteilung.

▼

Lösung: Für die Berechnung der Verteilungsfunktion \mathbb{F}_Z der Zufallsvariablen $Z = X_1 + X_2 + \dots + X_{12}$ verwenden wir die Laplace-Transformation: Da die Zufallsvariablen X_1, X_2, \dots, X_{12} im Intervall $[0, 1]$ gleichverteilt und vollständig unabhängig sind, gilt für die Laplace-Transformierte \mathcal{L}_Z wegen [Satz 20.2.12](#) und [Beispiel 20.2.3](#)

$$\mathcal{L}_Z[s] = \mathcal{L}_{\mathcal{U}([0,1])}[s]^{12} = \left(\frac{1 - e^{-s}}{s}\right)^{12}$$

Die Verteilungsfunktion \mathbb{F}_Z von Z erhält man, indem man auf diese Laplace-Transformierte $\mathcal{L}_Z[s]$ den Befehl [InverseLaplaceTransform](#) anwendet und von diesem Ergebnis die Stammfunktion bildet.

```
fZ[x_] = InverseLaplaceTransform[(1 - e-s)12/s12, s, x];
FZ[z_] = Integrate[fZ[x], {x, 0, z}, Assumptions -> {z ∈ Reals}];
```

Wir vergleichen nun die auf diese Weise ermittelte Verteilungsfunktion \mathbb{F}_Z von Z mit der Verteilungsfunktion der $\mathcal{N}[6, 1]$ -Verteilung und erhalten als maximalen Unterschied dieser beiden Verteilungsfunktionen den Wert

```
Max[Table[Abs[FZ[z] - CDF[NormalDistribution[6, 1], z]], {z, 0, 12, 1/1000}]] // N]
```

```
0.00233593
```

Nach diesen Vorbereitungen wenden wir uns nun direkt dem zentralen Grenzwertungssatz zu, wobei wir auf den Beweis dieses tiefliegenden Satzes nicht näher eingehen:

23.1.10 Satz (Der zentrale Grenzwertungssatz): Sind die Zufallsvariablen X_1, X_2, \dots vollständig unabhängig, identisch verteilt und quadratisch integrierbar, so gilt für alle $z \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} \mathbb{P}\left[\left\{\frac{(X_1 + X_2 + \dots + X_n) - \mathbb{E}[X_1 + X_2 + \dots + X_n]}{\sqrt{\mathbb{V}[X_1 + X_2 + \dots + X_n]}} \leq z\right\}\right] = \phi[z]$$

wobei wir mit ϕ die Verteilungsfunktion der $\mathcal{N}[0, 1]$ -Verteilung bezeichnen. Dieser Satz drückt die bereits mehrfach angesprochene Tatsache aus, dass die Summe einer großen Anzahl von vollständig unabhängigen, identisch verteilten, quadratisch integrierbaren Zufallsvariablen annähernd normalverteilt ist.

Zu diesem Satz sind einige Ergänzungen angebracht:

- In der Sprechweise von [Definition 20.4.5](#) besagt der zentrale Grenzwertungssatz: Sind die Zufallsvariablen X_1, X_2, \dots vollständig unabhängig, identisch verteilt und quadratisch integrierbar und bezeichnet für jedes $n \in \mathbb{N}$

$$Z_n = \frac{(X_1 + X_2 + \dots + X_n) - \mathbb{E}[X_1 + X_2 + \dots + X_n]}{\sqrt{\mathbb{V}[X_1 + X_2 + \dots + X_n]}}$$

so konvergiert die Folge $(Z_n)_{n \in \mathbb{N}}$ in Verteilung gegen eine $\mathcal{N}[0, 1]$ -verteilte Zufallsvariable Z .

- Wie sich an Hand von Gegenbeispielen zeigen lässt, reichen die *paarweise* Unabhängigkeit der Zufallsvariablen X_1, X_2, \dots bzw die *einfache* Integrierbarkeit für die Gültigkeit des zentralen Grenzwertungssatzes **nicht!** aus.
- Besitzen die vollständig unabhängigen und identisch verteilten Zufallsvariablen X_1, X_2, \dots sogar ein drittes Moment, so lässt sich die Konvergenzgeschwindigkeit durch die Ungleichung von **BERRY-ESSÉEN** abschätzen:

$$\sup_{z \in \mathbb{R}} \left| \mathbb{P}\left\{ \frac{(X_1 + X_2 + \dots + X_n) - \mathbb{E}[X_1 + X_2 + \dots + X_n]}{\sqrt{\mathbb{V}[X_1 + X_2 + \dots + X_n]}} \leq z \right\} - \phi[z] \right| \leq 0.8 \frac{1}{\sqrt{n}} \frac{\mathbb{E}[|X_1 - \mathbb{E}[X_1]|^3]}{\mathbb{V}[X_1]^{3/2}}$$

- Bei der Ungleichung von BERRY-ESSÉEN handelt es sich um eine relativ schwache Abschätzung; die tatsächlichen Unterschiede sind meist wesentlich kleiner, wie eine mit Hilfe von **Approximation[distr, n]** ermittelte Graphik zeigt.
- Der zentrale Grenzwertungssatz erfuhr im Laufe der Zeit eine Reihe von Verallgemeinerungen. Vor allem die Voraussetzung, dass die quadratisch integrierbaren Zufallsvariablen X_1, X_2, \dots *identisch verteilt* sind, konnte durch wesentlich schwächere Voraussetzungen ersetzt werden. Die schwächste dieser Voraussetzungen stammt von **LINDEBERG-FELLER**. Etwas schlampig ausgedrückt besagt diese Voraussetzung von LINDEBERG-FELLER, dass die Zufallsvariablen X_1, X_2, \dots etwa "gleich groß" sein müssen, also keine dieser Zufallsvariablen die anderen Zufallsvariablen dominieren darf. Auch die Voraussetzung der *vollständigen Unabhängigkeit* konnte durch etwas schwächere Voraussetzungen ersetzt werden, worauf wir aber nicht näher eingehen wollen.

- Natürlich existiert auch ein **mehrdimensionales** Analogon des zentralen Grenzwertsatzes. Dieses besagt, dass die Summe einer großen Anzahl von vollständig unabhängigen, identisch verteilten, quadratisch integrierbaren Zufallsvektoren annähernd multinormalverteilt ist.

23.1.11 Beispiel: Man entwickle einen Befehl, mit dem sich für beliebige $n \in \mathbb{N}$ und eine beliebige in *Mathematica* implementierte Verteilung *distr* die rechte Seite der Ungleichung von BERRY-ESSÉEN berechnen lässt. Für die Gleichverteilung auf dem Intervall $[0, 1]$ vergleiche man die damit erzielte Abschätzung mit dem in [Beispiel 23.1.9](#) ermittelten maximalen Unterschied.

▼

Lösung: Der Befehl

```
BerryEsseen[n_, distr_] := 0.8 ExpectedValue[Abs[x - Mean[distr]]3, distr, x]/Sqrt[n Variance[distr]3] // N
```

liefert offensichtlich den gewünschten Wert. Verglichen mit dem in [Beispiel 23.1.9](#) ermittelten maximalen Unterschied von 0.00233593 ist die damit erzielte Abschätzung

```
BerryEsseen[12, UniformDistribution[{0, 1}]]
```

```
0.3
```

aber ziemlich schlecht.

Messgrößen setzen sich üblicherweise aus vielen vollständig unabhängigen Einflussgrößen additiv zusammen und sind wegen des zentralen Grenzwertsatzes somit in der Regel $\mathcal{N}[\mu, \sigma]$ -verteilt. Diese Tatsache war bereits lange vor der Entdeckung des zentralen Grenzwertsatzes bekannt. Man glaubte damals, dass es eben normal sei, dass Messgrößen normal-verteilt sind. Im Zusammenhang mit einer $\mathcal{N}[\mu, \sigma]$ -verteilten Messgrößen Z nennt man $\mathbb{E}[Z] = \mu$ den **systematischen Fehler** und $\mathbb{E}[|Z - \mu|] = \sqrt{2/\pi} \sigma$ den **mittleren Fehler**.

23.1.12 Beispiel: Ein Messgerät ohne systematischen Fehler besitze einen mittleren Fehler von $m = 40$. Wieviele Messungen müssen mit diesem Gerät durchgeführt werden, um mit der Wahrscheinlichkeit von $p = 0.9$ sicher zu stellen, dass der absolute Fehler in mindestens einer dieser Messungen den Wert $f = 7.5$ nicht überschreitet.

▼

Lösung: Die Zufallsvariable X_i beschreibe den Fehler der i -ten Messung. Aus der Angabe entnimmt man, dass die Zufallsvariablen X_1, X_2, \dots vollständig unabhängig und $\mathcal{N}[0, \sqrt{\pi/2} m]$ -verteilt sind. Gesucht ist nun die kleinste Zahl $n \in \mathbb{N}$ mit der Eigenschaft

$$\begin{aligned} p &\leq \mathbb{P}\{\{\text{Min}[|X_1|, |X_2|, \dots, |X_n|] \leq f\}\} = 1 - \mathbb{P}\{\{\text{Min}[|X_1|, |X_2|, \dots, |X_n|] > f\}\} = \\ &= 1 - \mathbb{P}\{\{|X_1| > f\}\} \mathbb{P}\{\{|X_2| > f\}\} \dots \mathbb{P}\{\{|X_n| > f\}\} = 1 - (2 \mathbb{P}\{\{|X_1| > f\}\})^n \end{aligned}$$

Mit Hilfe von *Mathematica* lässt sich leicht zeigen, dass dazu mindestens $n = 19$ Messungen durchgeführt werden müssen:

```
m = 40; f = 7.5; sigma = Sqrt[pi/2] m;
Table[{n, 1 - (2 CDF[NormalDistribution[0, sigma], -f])n}, {n, 10, 20}]
Clear[m, f, sigma]
```

```
{{10, 0.71807}, {11, 0.751598}, {12, 0.781138},
 {13, 0.807166}, {14, 0.830098}, {15, 0.850303}, {16, 0.868106},
 {17, 0.883791}, {18, 0.897611}, {19, 0.909787}, {20, 0.920515}}
```


23.2 Die Normalverteilung $\mathcal{N}[\mu, \sigma]$

Wir fassen die bereits bekannten Eigenschaften der Normalverteilung $\mathcal{N}[\mu, \sigma]$ zusammen:

23.2.1 Bemerkung: Die Normalverteilung $\mathcal{N}[\mu, \sigma]$ besitzt den Träger

$$\mathbb{T} = \mathbb{R}$$

die Verteilungsdichte

$$f[z] = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\mu)^2/(2\sigma^2)}$$

und die Verteilungsfunktion

$$F[z] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^z e^{-(t-\mu)^2/(2\sigma^2)} dt = \frac{1}{2} \left(1 + \operatorname{Erf}\left[\frac{z-\mu}{\sqrt{2}\sigma}\right] \right)$$

Eine $\mathcal{N}[\mu, \sigma]$ -verteilte Zufallsvariable Z besitzt den Erwartungswert

$$E[Z] = \mu$$

und die Varianz

$$V[Z] = \sigma^2$$

Für Normalverteilungen gilt die Faltungsformel

$$\mathcal{N}[\mu, \sigma] * \mathcal{N}[\nu, \tau] = \mathcal{N}[\mu + \nu, \operatorname{Sqrt}[\sigma^2 + \tau^2]]$$

Weiters besagt der Satz über die affine Transformation von Normalverteilungen: Ist die Zufallsvariable Z normalverteilt mit den Parametern μ und σ , so ist die Zufallsvariable $Y = \xi Z + \eta$ normalverteilt mit den Parametern $\xi\mu + \eta$ und $\sigma|\xi|$. Daraus folgt speziell: Ist die Zufallsvariable Z normalverteilt mit den Parametern μ und σ , so ist die Zufallsvariable $Y = (Z - \mu)/\sigma$ normalverteilt mit den Parametern 0 und 1 (man spricht in diesem Zusammenhang von der Standardisierung der Normalverteilung).

Es folgen wieder einige Beispiele, mit denen gezeigt werden soll, wie mit der Normalverteilung gearbeitet wird:

23.2.2 Beispiel: Eine Walze wird zugelassen, wenn ihr Durchmesser X der Bedingung $0.96 \leq X \leq 1.04$ genügt. Erfahrungsgemäß sind 3.5 % der Walzen zu dick und 1.5 % der Walzen zu dünn. Unter der Annahme, dass X normalverteilt ist, berechne man den Erwartungswert μ und die Streuung σ von X .

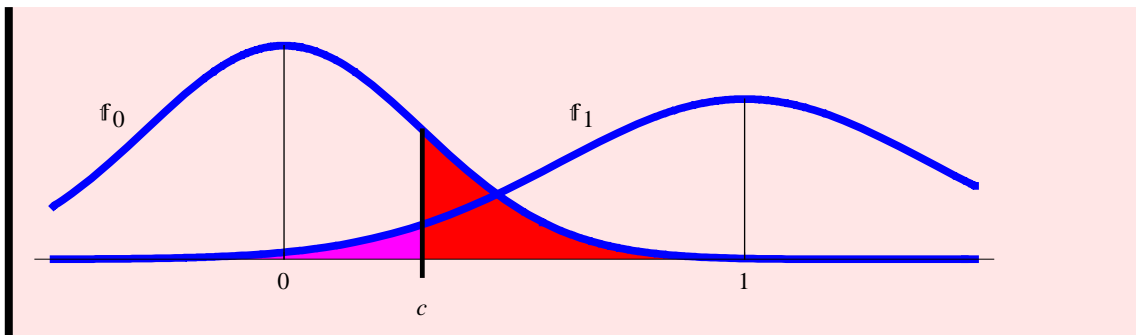
▼

Lösung: Wir haben die beiden Parameter μ und σ der Verteilung $\mathbb{P}_X = \mathcal{N}[\mu, \sigma]$ von X so zu bestimmen, dass die beiden Bedingungen $\mathbb{P}\{X > 1.04\} = 0.035$ und $\mathbb{P}\{X < 0.96\} = 0.015$ erfüllt sind. Diese Aufgabe lässt sich mühelos mit Hilfe des Befehls `FindRoot` mit den naheliegenden Startwerten $\mu_0 = 1$ und $\sigma_0 = 0.01$ lösen:

```
FindRoot[{1 - CDF[NormalDistribution[μ, σ], 1.04] == 0.035, CDF[NormalDistribution[μ, σ], 0.96] == 0.015}, {{μ, 1}, {σ, 0.01}}
```

```
{μ → 1.0036, σ → 0.0200904}
```

23.2.3 Beispiel: Ein Sender sendet die Signale 0 und 1 aus, die bei der Übermittlung verrauscht werden und beim Empfänger als $\mathcal{N}[0, \sigma_0]$ - bzw. $\mathcal{N}[1, \sigma_1]$ -verteilte Signale eintreffen. Der Empfänger identifiziert ein Signal als 0, wenn das eintreffende Signal kleiner als eine gewisse Schranke c ist und als 1, wenn das eintreffende Signal größer als diese Schranke c ist. Wie groß soll c gewählt werden, damit die Fehlerwahrscheinlichkeit minimal wird.



23.2.4 Beispiel: Ist die Zufallsvariable Z normalverteilt mit den Parametern μ und σ , so nennt man das Intervall $[\mu - k\sigma, \mu + k\sigma]$ den k -Sigma-Bereich. Man berechne für beliebige $k \in \mathbb{N}$ die Wahrscheinlichkeit $w[k]$ des k -Sigma-Bereichs.

▼

Lösung: Wegen des [Satzes über die affine Transformation](#) von Normalverteilungen gilt für alle $k \in \mathbb{N}$

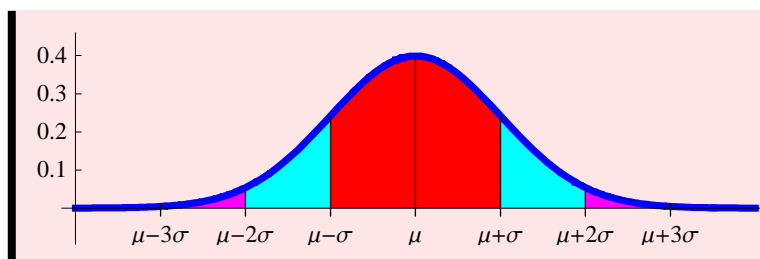
$$w[k] = \mathbb{P}\{Z \in [\mu - k\sigma, \mu + k\sigma]\} = \mathbb{P}\{(Z - \mu)/\sigma \in [-k, k]\} = \phi[k] - \phi[-k]$$

wobei wir mit ϕ die Verteilungsfunktion der $\mathcal{N}[0, 1]$ -Verteilung bezeichnen. (Wir erkennen dabei, dass die Wahrscheinlichkeit $w[k]$ des k -Sigma-Bereiches **nicht!** von den Parametern μ und σ abhängt). Diese Wahrscheinlichkeit lässt sich leicht tabellieren (für die Praxis ist in erster Linie die Wahrscheinlichkeit des Ein-, Zwei- und Drei-Sigma-Bereichs von Bedeutung):

```
w[k_] := N[CDF[NormalDistribution[0, 1], k] - CDF[NormalDistribution[0, 1], -k], 10]
TableForm[Table[{k, w[k]}, {k, 1, 6}], TableHeadings -> {None, {"k", "w[k]"}}, TableSpacing -> {1, 5}]
```

k	$w[k]$
1	0.6826894921
2	0.9544997361
3	0.9973002039
4	0.9999366575
5	0.999994267
6	0.999999980

Die folgende Skizze veranschaulicht den Ein-, Zwei- und Drei-Sigma-Bereich zusammen mit den zugehörigen Wahrscheinlichkeiten $w[k]$ (dabei handelt es sich um die entsprechenden Flächen unter der Verteilungsdichte):



23.2.5 Beispiel: Ein Schiff passiert eine 10 km breite Meerenge. An den beiden Ufern dieser Meerenge befinden sich Leuchttürme, deren Leuchtfeuer man X bzw Y km weit sehen kann. Unter der Annahme, dass die beiden Zufallsvariablen X und Y unabhängig und $\mathcal{N}[7, 1]$ -verteilt sind, berechne man die Wahrscheinlichkeit p dafür, dass von einem Schiff, das die Meerenge Z km vom linken Ufer entfernt passiert, beide Leuchtfeuer gesehen werden können, wenn man annimmt, dass Z im Intervall $[0, 10]$ gleichverteilt ist?

▼

Lösung: Unter Verwendung des [Satzes von der totalen Wahrscheinlichkeit](#) in differenzieller Form und der Tatsache, dass die drei Zufallsvariablen X , Y und Z vollständig unabhängig sind, ergibt sich für die von uns gesuchte Wahrscheinlichkeit

$$\begin{aligned} p &= \mathbb{P}[\{X > Z\} \cap \{Y > 10 - Z\}] = \int_{-\infty}^{\infty} \mathbb{P}[\{X > Z\} \cap \{Y > 10 - Z\} \mid \{Z = z\}] f_Z[z] dz = \\ &= \int_{-\infty}^{\infty} \mathbb{P}[\{X > z\}] \mathbb{P}[\{Y > 10 - z\}] f_Z[z] dz = \int_0^{10} (1 - F_X[z]) (1 - F_Y[10 - z]) 10^{-1} dz \end{aligned}$$

Wir werten dieses Integral mit Hilfe von [NIntegrate](#) aus und erhalten

```
NIntegrate[(1 - CDF[NormalDistribution[7, 1], z]) (1 - CDF[NormalDistribution[7, 1], 10 - z]), {z, 0, 10}]/
10
0.400021
```

23.2.6 Beispiel: Bei einer Werbeaktion eines Versandhauses sollen die ersten 1000 Einsender einer Bestellung eine Damen- bzw Herrenarmbanduhr als Geschenk erhalten. Wir nehmen an, dass sich beide Geschlechter gleichermaßen von dem Angebot angesprochen fühlen. Wieviele Damen- bzw Herrenarmbanduhren sollte das Versandhaus vorrätig haben, damit mit einer Wahrscheinlichkeit von mindestens 98 % alle 1000 Einsender eine passende Uhr erhalten?

▼

Lösung: Wir bezeichnen mit X die Anzahl der Damen unter den ersten 1000 Einsendern. Die Zufallsvariable X ist offenbar $\mathcal{B}[1000, 1/2]$ -verteilt und damit wegen [Bemerkung 23.1.1](#) annähernd $\mathcal{N}[500, \sqrt{250}]$ -verteilt. Falls das Versandhaus a Damen- bzw Herrenarmbanduhren vorrätig hat, so werden mit einer Wahrscheinlichkeit von

$$\mathbb{P}[\{X \leq a\} \cap \{1000 - X \leq a\}] = \mathbb{P}[\{1000 - a \leq X \leq a\}] = F_X[a] - F_X[1000 - a]$$

alle 1000 Einsender eine passende Uhr erhalten. Wir berechnen mit Hilfe von *Mathematica* diese Wahrscheinlichkeit für einige Werte von a und erkennen, dass $a \geq 537$ sein muss, damit alle 1000 Einsender mit einer Wahrscheinlichkeit von mindestens 98 % eine passende Uhr bekommen.

```
Table[{a, CDF[NormalDistribution[500, Sqrt[250]], a] - CDF[NormalDistribution[500, Sqrt[250]], 1000 - a] // N}, {
{{530, 0.94222}, {531, 0.950076}, {532, 0.957015},
{533, 0.963121}, {534, 0.968472}, {535, 0.973143}, {536, 0.977204},
{537, 0.980721}, {538, 0.983754}, {539, 0.986359}, {540, 0.988588}}
```

23.2.7 Beispiel: Ein Unternehmen hat insgesamt $n = 1000$ Aktien ausgegeben. Ihre Besitzer entscheiden sich bei jeder Aktie mit einer Wahrscheinlichkeit von $0 < p < 1$ zum Verkauf. Diese Entscheidungen finden unabhängig voneinander statt. Der Markt kann $s = 50$ Aktien aufnehmen, ohne dass der Kurs fällt. Wie groß darf p höchstens sein, damit der Kurs mit einer Wahrscheinlichkeit von 90 % nicht fällt.

▼

Lösung: Wir bezeichnen mit X die Anzahl der Aktien, die zum Verkauf angeboten werden. Die Zufallsvariable X ist offenbar $\mathcal{B}[n, p]$ -verteilt und damit wegen [Bemerkung 32.1.1](#) annähernd $\mathcal{N}[n p, \sqrt{n p (1 - p)}]$ -verteilt. Um mit einer Wahrscheinlichkeit von 90 % sicher zu stellen, dass der Kurs der Aktie nicht fällt, darf p höchstens so groß sein, dass $\mathbb{P}\{X \leq s\} = 0.90$ ist. Wir werten diese Beziehung unter Verwendung von [Solve](#) aus:

```
n = 1000; s = 50;
Solve[CDF[NormalDistribution[n p, Sqrt[n p (1 - p)]], s] == 0.90, p]
Clear[n, s]

{{p -> 0.0418818}}
```

23.2.8 Beispiel: Die Zufallsvariable Z sei $\mathcal{N}[0, \sigma]$ -verteilt. Man bestimme die Streuung σ so, dass die Wahrscheinlichkeit des Ereignisses $\{a < Z < b\}$ mit $0 < a < b$ maximal ist.

▼

Lösung: Wir haben die Streuung σ so zu bestimmen, dass

$$p[\sigma] = \mathbb{P}\{a < Z < b\} = \mathbb{F}_Z[b] - \mathbb{F}_Z[a]$$

maximal ist. Wir differenzieren dazu $p[\sigma]$ nach σ , ermitteln mit Hilfe von [Solve](#) jenen Wert von σ , für den diese Ableitung gleich Null ist und berücksichtigen außerdem mit Hilfe von [FullSimplify](#) die Tatsache, dass $0 < a < b$ vorausgesetzt wurde:

```
FullSimplify[Solve[D[CDF[NormalDistribution[0, sigma], b] - CDF[NormalDistribution[0, sigma], a], sigma] == 0, sigma],
0 < a < b]

{{sigma -> -1/2 * sqrt((2 a^2 - 2 b^2) / (Log[a/b]) )}, {sigma -> 1/2 * sqrt((2 a^2 - 2 b^2) / (Log[a/b]) )}}
```

Für den gesuchten Wert von σ (σ muss positiv sein) gilt damit

$$\sigma = \sqrt{2 b^2 - 2 a^2} / (\text{Log}[b] - \text{Log}[a])$$

23.2.9 Beispiel: Wie stark darf eine $\mathcal{N}[0, \sigma]$ -verteilte Zufallsvariable Z maximal streuen, wenn sie mit einer Wahrscheinlichkeit von mindestens 0.999 innerhalb des Toleranzbereiches $[-a, a]$ liegen soll?

▼

Lösung: Wir haben die Streuung σ so zu bestimmen, dass

$$p[\sigma] = \mathbb{P}\{-a \leq Z \leq a\} = \mathbb{F}_Z[a] - \mathbb{F}_Z[-a] = 0.999$$

gilt. Mit Hilfe von [Solve](#)

```
Solve[CDF[NormalDistribution[0,  $\sigma$ ], a] - CDF[NormalDistribution[0,  $\sigma$ ], -a] == 0.999,  $\sigma$ ]
```

```
{{ $\sigma \rightarrow 0.303903 a$ }}
```

ergibt sich für σ der Wert $\sigma = 0.303903 a$.

23.2.10 Beispiel: Ein zufälliger Punkt der Ebene besitze unabhängige, $\mathcal{N}[0, 1]$ -verteilte Koordinaten X und Y . Man ermittle die Verteilung seiner Polarkoordinaten R und Φ und zeige, dass R und Φ unabhängig sind.

▼

Lösung: Da X und Y unabhängig und $\mathcal{N}[0, 1]$ -verteilt sind, gilt für alle $x, y \in \mathbb{R}$ offenbar

$$f_{X,Y}[x, y] = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

Bezeichnet nun für alle $r \geq 0$ $K_r = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq r\}$ einen Kreis mit Radius r und für alle $\varphi \in [0, 2\pi[$ $S_\varphi = \{(x, y) \in \mathbb{R}^2 \mid \text{Arg}[x + yi] \leq \varphi\}$ einen Sektor mit Winkel φ , so erhält man (Übergang zu Polarkoordinaten)

$$F_R[r] = \mathbb{P}[\{(X, Y) \in K_r\}] = \iint_{K_r} f_{X,Y}[x, y] dx dy = \int_0^r \left(\int_0^{2\pi} \frac{1}{2\pi} e^{-\rho^2/2} \rho d\psi \right) d\rho = 1 - e^{-r^2/2}$$

sowie

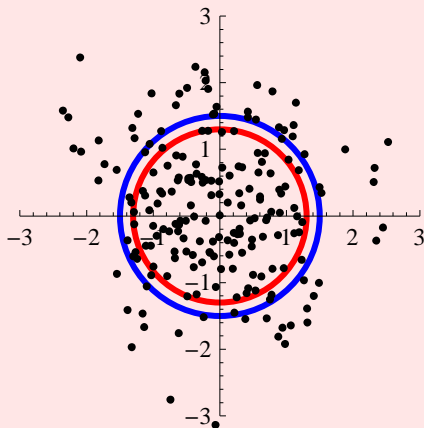
$$F_\Phi[\varphi] = \mathbb{P}[\{(X, Y) \in S_\varphi\}] = \iint_{S_\varphi} f_{X,Y}[x, y] dx dy = \int_0^\varphi \left(\int_0^\infty \frac{1}{2\pi} e^{-\rho^2/2} \rho d\rho \right) d\psi = \frac{\varphi}{2\pi}$$

und damit

$$\begin{aligned} F_{R,\Phi}[r, \varphi] &= \mathbb{P}[\{(X, Y) \in K_r \cap S_\varphi\}] = \iint_{K_r \cap S_\varphi} f_{X,Y}[x, y] dx dy = \\ &= \int_0^r \left(\int_0^\varphi \frac{1}{2\pi} e^{-\rho^2/2} \rho d\psi \right) d\rho = (1 - e^{-r^2/2}) \frac{\varphi}{2\pi} = F_R[r] F_\Phi[\varphi] \end{aligned}$$

Die Tatsache, dass die beiden Zufallsvariablen R und Φ unabhängig sind und Φ im Intervall $[0, 2\pi]$ gleichverteilt ist, lässt sich auch graphisch veranschaulichen: Simuliert man n derartige Punkte, so ergibt sich ein annähernd kreisförmiger Punktschwarm. Die dabei in den (schmalen) Kreisring mit den Radien $r_1 < r_2$ fallenden Punkte sind unabhängig von diesen Radien innerhalb dieses Kreisrings gleichverteilt:

```
n = 200; r1 = 1.3; r2 = 1.5;
p1 = Graphics[Table[{PointSize[0.02], Point[RandomReal[NormalDistribution[0, 1], 2]]}, {n}]];
c1 = Graphics[{Red, Thickness[0.015], Circle[{0, 0}, r1]}];
c2 = Graphics[{Blue, Thickness[0.015], Circle[{0, 0}, r2]}];
Show[{c1, c2, p1}, Axes -> True, PlotRange -> {{-3, 3}, {-3, 3}}, AspectRatio -> Automatic]
Clear[n, r1, r2, p1, c1, c2]
```



23.2.11 Beispiel: Wir betrachten ein in einem Gefäß eingeschlossenes Gas der Temperatur T , dessen Moleküle die Masse m besitzen. Unter der Annahme, dass die Komponenten V_1, V_2, V_3 des

Geschwindigkeitsvektors \vec{V} eines zufällig ausgewählten Moleküls vollständig unabhängig und $\mathcal{N}[0, \sigma]$ -verteilt sind, wobei $\sigma = \sqrt{\kappa T/m}$ ist, zeige man die Gültigkeit des Gesetzes $P V = R T$. Dabei bezeichnet P den Druck, V das Volumen eines Mols dieses Gases, T die absolute Temperatur, κ die Boltzmann Konstante und $R = N \kappa$ die universelle Gaskonstante, wobei N die Loschmidt'sche Zahl (also die Anzahl der Moleküle in einem Mol) bezeichnet.

▼

Lösung: Der Druck P des Gases ist gleich dem Erwartungswert der Impulse, die einem Flächenstück der Wand des Gefäßes vom Flächeninhalt Eins während eines Zeitintervalls der Länge Eins übermittelt werden. Wir nehmen an, dass die Zusammenstöße ideal elastisch sind. Stößt ein Molekül mit der Masse m und der Geschwindigkeit $|\vec{V}| = v$ unter dem Winkel $\Theta = \vartheta$ (mit der Normalen der Wand) an die Wand des Gefäßes, so übermittelt dieses Molekül einen Impuls von der Größe $2 m v \cos[\vartheta]$. Ein Molekül mit der Geschwindigkeit $|\vec{V}| = v$ und einer Richtung, die mit der Normalen der Wand den Winkel $\Theta = \vartheta$ einschließt, stößt während der Zeit $[t, t + 1]$ genau dann auf einen Bereich K der Wand, welcher den Flächeninhalt Eins besitzt, wenn sich dieses Teilchen im Zeitpunkt t in einem schiefen Zylinder befindet, dessen Grundfläche der Bereich K und dessen Höhe gleich $v \cos[\vartheta]$ ist. Nimmt man an, dass die Moleküle gleichmäßig im Gefäß verteilt sind, so ist die Wahrscheinlichkeit für dieses Ereignis gleich $v \cos[\vartheta]/W$, wobei W das Volumen des Gefäßes bezeichnet. Damit ergibt sich für den bedingten Erwartungswert $\mathbb{E}[I | \{|\vec{V}| = v\} \cap \{\Theta = \vartheta\}]$ jenes Impulses I , den ein Molekül mit der Geschwindigkeit $|\vec{V}| = v$ und dem Winkel $\Theta = \vartheta$ (mit $0 \leq \vartheta \leq \pi/2$), der Wand des Gefäßes übermittelt

$$\mathbb{E}[I | \{|\vec{V}| = v\} \cap \{\Theta = \vartheta\}] = \frac{2 m v^2 \cos[\vartheta]^2}{W} = \frac{4}{W} \frac{m v^2}{2} \cos[\vartheta]^2$$

Analog zu [Beispiel 23.2.11](#) lässt sich zeigen, dass die Geschwindigkeit $|\vec{V}|$ und die Flugrichtung eines zufällig herausgegriffenen Moleküls unabhängig sind und dass diese Flugrichtung im Raum gleichverteilt ist. In [Beispiel 15.3.11](#) haben wir die Verteilungsfunktion des Winkels Θ , den diese Flugrichtung mit einer vorgegebenen Richtung (in unserem Fall mit der Normalen auf die Wand) einschließt, ermittelt und dabei für alle $\vartheta \in \mathbb{T}_\Theta = [0, \pi]$ die Formel $F_\Theta[\vartheta] = (1 - \cos[\vartheta])/2$ erhalten. Für den Erwartungswert der Zufallsvariablen $Z = 1_{[0, \pi/2]} \cos[\Theta]^2$ ergibt sich damit

```
Integrate[Cos[ϑ]^2 D[(1 - Cos[ϑ])/2, ϑ], {ϑ, 0, π/2}]
```

```
1  
--  
6
```

Wegen [Beispiel 19.3.9](#) und [Bemerkung 18.2.2](#) erhalten wir damit für den mittleren Impuls, den ein zufällig herausgegriffenes Molekül der Wand übermittelt

$$\begin{aligned} \mathbb{E}[I] &= \mathbb{E}[\mathbb{E}[I | \{|\vec{V}| = \bullet_1\} \cap \{\Theta = \bullet_2\}] \circ \{|\vec{V}|, \Theta\}] = \mathbb{E}\left[\frac{4}{W} \frac{m |\vec{V}|^2}{2} \cos[\Theta]^2\right] = \\ &= \frac{4}{W} \mathbb{E}\left[\frac{m |\vec{V}|^2}{2}\right] \mathbb{E}[\cos[\Theta]^2] = \frac{4}{W} \frac{3 \kappa T}{2} \frac{1}{6} = \frac{\kappa T}{W} \end{aligned}$$

Da in einem Mol Gas N Moleküle vorhanden sind, übt eine Gasmenge von n Molen auf die Wand des Gefäßes damit den Druck

$$P = n N \mathbb{E}[I] = \frac{n N \kappa T}{W} = \frac{N \kappa T}{V} = \frac{R T}{V}$$

aus, wobei $V = W/n$ das Molvolumen bezeichnet.

23.2.12 Beispiel: Die beiden Zufallsvariablen X und Y seien unabhängig und $\mathcal{N}[0, 1]$ -verteilt. Man ermittle die Verteilungsdichte ihres Produkts $U = X Y$ sowie ihres Quotienten $V = X/Y$.

▼

Lösung: Wegen des [Satzes von der totalen Wahrscheinlichkeit](#) in differenzieller Form sowie der Regel über das [Einsetzen einer Bedingung](#) gilt für alle $u, v \in \mathbb{R}$

$$\begin{aligned} f_U[u] &= \frac{1}{du} \mathbb{P}[\{X/Y \in [u, u + du]\}] = \frac{1}{du} \int_{-\infty}^{\infty} \mathbb{P}[\{X \in [\frac{u}{y}, \frac{u+du}{y}]\}] \mathbb{P}[\{Y \in [y, y + dy]\}] = \\ &= \int_{-\infty}^{\infty} f_X[\frac{u}{y}] \frac{1}{|y|} f_Y[y] dy \end{aligned}$$

und

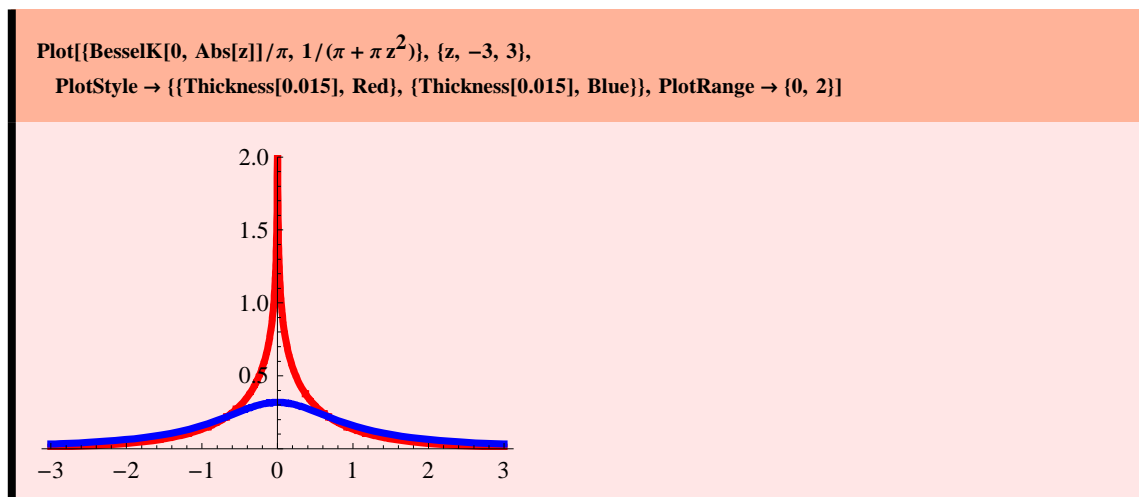
$$\begin{aligned} f_V[v] &= \frac{1}{dv} \mathbb{P}[\{X/Y \in [v, v + dv]\}] = \frac{1}{dv} \int_{-\infty}^{\infty} \mathbb{P}[\{X \in [v y, (v + dv) y]\}] \mathbb{P}[\{Y \in [y, y + dy]\}] = \\ &= \int_{-\infty}^{\infty} f_X[v y] |y| f_Y[y] dy \end{aligned}$$

Wertet man diese Integrale mit Hilfe von *Mathematica* aus, so erhält man die beiden gesuchten Verteilungsdichten ([BesselK](#) bezeichnet dabei die modifizierte Besselfunktion zweiter Ordnung):

```
FullSimplify[Integrate[PDF[NormalDistribution[0, 1], u/y]/Abs[y] PDF[NormalDistribution[0, 1], y], {y, -∞, ∞}],
BesselK[0, Abs[u]]
π

FullSimplify[Integrate[PDF[NormalDistribution[0, 1], v y] Abs[y] PDF[NormalDistribution[0, 1], y], {y, -∞, ∞}], v
1
π + π v2
```

Wir zeichnen abschließend diese beiden (offenbar symmetrischen) Verteilungsdichten (bei der roten Kurve handelt es sich um die Verteilungsdichte von $U = X/Y$, bei der blauen Kurve handelt es sich um die Verteilungsdichte von $V = X/Y$):



23.3 Die Chi-Quadrat Verteilung $\chi^2[n]$

Wir beginnen mit einem für die Chi-Quadrat Verteilung zentralen Satz:

23.3.1 Satz: Sind die Zufallsvariablen X_1, X_2, \dots, X_n vollständig unabhängig und identisch $\mathcal{N}[0, 1]$ -verteilt, so ist die Zufallsvariable

$$Z^{(n)} = X_1^2 + X_2^2 + \dots + X_n^2$$

Chi[n]-verteilt. Den Parameter n nennt man den **Freiheitsgrad** der Chi[n]-Verteilung.

▼

Beweis: Aus dem **Satz über die Quadrateigenschaft** zusammen mit der **Faltungsformel** für Chi-Quadrat Verteilungen folgt unmittelbar, dass die Zufallsvariable $Z^{(n)} = X_1^2 + X_2^2 + \dots + X_n^2$ Chi[n]-verteilt ist.

Wir fassen die bereits bekannten Eigenschaften der Chi-Quadrat Verteilung Chi[n] zusammen, wobei wir die Beziehung $\text{Chi}[n] = \mathcal{Gamma}[n/2, 2]$ verwenden (man vergleiche dazu **Satz 22.3.8**):

23.3.2 Bemerkung: Die Chi-Quadrat Verteilung $\text{Chi}[n]$ besitzt den **Träger**

$$\mathbb{T} = [0, \infty[$$

die **Verteilungsdichte**

$$f[z] = \begin{cases} \frac{1}{\Gamma[n/2]} 2^{-n/2} e^{-z/2} z^{n/2-1} & \text{für } z \geq 0 \\ 0 & \text{sonst} \end{cases}$$

und die **Verteilungsfunktion**

$$F[z] = \begin{cases} 0 & \text{für } z < 0 \\ \frac{1}{\Gamma[n/2]} \int_0^z 2^{-n/2} e^{-t/2} t^{n/2-1} dt = \Gamma_r[n/2, 0, z/2] & \text{für } z \geq 0 \end{cases}$$

Eine $\text{Chi}[n]$ -verteilte Zufallsvariable Z besitzt den **Erwartungswert**

$$\mathbb{E}[Z] = n$$

und die **Varianz**

$$\mathbb{V}[Z] = 2n$$

Für Chi-Quadrat Verteilungen gilt die **Faltungsformel**

$$\text{Chi}[m] * \text{Chi}[n] = \text{Chi}[m+n]$$

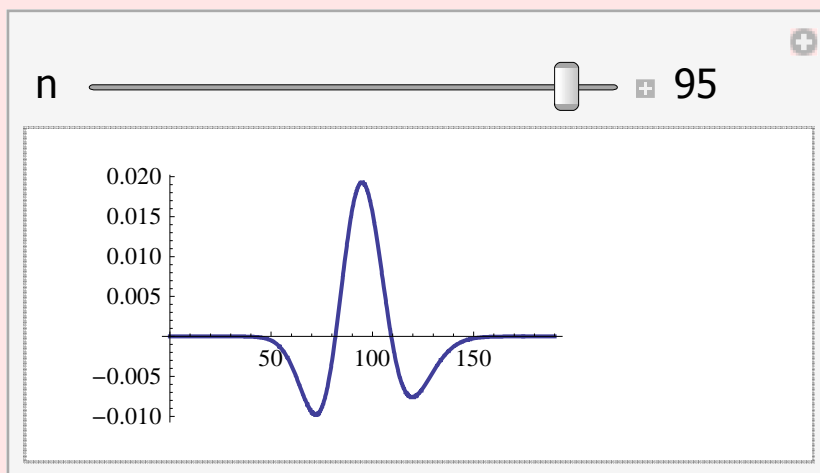
[Beispiel 19.3.6](#) und [Beispiel 19.3.9](#) sind typische Anwendungen der Chi-Quadrat Verteilung. In den beiden folgenden Beispielen werden Beziehungen zwischen der Chi-Quadrat Verteilung und anderen Verteilungen aufgezeigt:

23.3.3 Beispiel: Aus dem [zentralen Grenzwertungssatz](#) folgt, dass sich für großes $n \in \mathbb{N}$ die $\text{Chi}[n]$ -Verteilung durch die $\mathcal{N}[n, \sqrt{2n}]$ -Verteilung approximieren lässt. Man ermittle, wie groß n dabei sein muss, damit der maximale Unterschied zwischen diesen beiden Verteilungsfunktionen kleiner als 0.01 ist.

▼

Lösung: Wir plotten zuerst die Differenz dieser beiden Verteilungsfunktionen im Bereich $0 \leq z \leq 2n$

```
Manipulate[Plot[CDF[ChiSquareDistribution[n], z] - CDF[NormalDistribution[n, Sqrt[2 n]], z],
  {z, 0, 2 n}, PlotStyle -> Thickness[0.01], ImageSize -> {200, 100}, PlotRange -> All,
  {n, 10, 100, 1, Appearance -> "Labeled"}]
```



und erkennen, dass der maximale Unterschied dieser beiden Verteilungsfunktionen in der Nähe von $z = n$ liegt. Mit Hilfe von [FindMaximum](#) und $z = n$ als Startwert lässt sich diese maximale Differenz leicht ermitteln. Durch Experimentieren mit verschiedenen Werten von n zeigt sich, dass für $n \geq 354$ der maximale Unterschied dieser beiden Verteilungsfunktionen kleiner als 0.01 ist.

```
n = 354;
FindMaximum[CDF[ChiSquareDistribution[n], z] - CDF[NormalDistribution[n, Sqrt[2 n]], z], {z, n}]
Clear[n]
```

```
{0.00999634, {z -> 353.833}}
```

23.3.4 Beispiel: Man zeige: Bezeichnet G_m die Verteilungsfunktion der Chi-Quadrat Verteilung mit m Freiheitsgraden und F_λ die Verteilungsfunktion der Poissonverteilung mit Parameter λ , so gilt für alle $z > 0$ und alle $n \in \mathbb{N}$ die Beziehung $1 - G_{2n}[z] = F_{z/2}[n - 1]$.

▼

Lösung: Wir beweisen für alle $n \in \mathbb{N}$ die Gültigkeit dieser Beziehung mit Hilfe von *Mathematica*

```
FullSimplify[1 - CDF[ChiSquareDistribution[2 n], z] == CDF[PoissonDistribution[z/2], n - 1],
  {n ∈ Integers, n > 0}]
```

```
True
```

Die Chi-Quadrat Verteilung findet in der Statistik bei den sogenannten Chi-Quadrat-Tests Verwendung. Man kann nämlich zeigen, dass gewisse Testgrößen näherungsweise einer Chi-Quadrat Verteilung genügen. Ohne auf den

Beweis näher einzugehen, erwähnen wir in diesem Zusammenhang den folgenden Satz, den wir beim Chi-Quadrat Anpassungstest verwenden werden.

23.3.5 Satz: Sind die Zufallsvariablen X_1, X_2, \dots, X_n vollständig unabhängig und identisch \mathbb{P}_X verteilt und ist $\{A_1, A_2, \dots, A_r\}$ eine Partition des Trägers \mathbb{T}_X , so ist die Zufallsvariable

$$Z = \sum_{k=1}^r \frac{n(H_k - p_k)^2}{p_k}$$

welche ein Maß für die Abweichung der relativen Häufigkeiten $H_k = |\{i \mid 1 \leq i \leq n, X_i \in A_k\}|/n$ von den theoretischen Wahrscheinlichkeiten $p_k = \mathbb{P}_X[A_k]$ der Ereignisse A_k darstellt, asymptotisch (also für $n \rightarrow \infty$) Chi-Quadrat verteilt mit $r - 1$ Freiheitsgraden.

Wir veranschaulichen die Aussage dieses Satzes an einem Beispiel:

23.3.6 Beispiel: Wird mit einem homogenen Würfel n mal gewürfelt und für jedes $k \in \{1, 2, \dots, 6\}$ die relative Häufigkeit H_k für das Auftreten der Augenzahl k mit der theoretischen Wahrscheinlichkeit $p_k = 1/6$ verglichen, so ist die Zufallsvariable

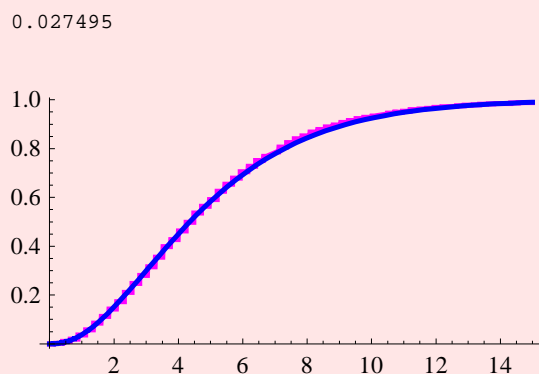
$$Z = \sum_{k=1}^6 \frac{n(H_k - p_k)^2}{p_k}$$

wegen [Satz 23.3.5](#) für große Werte von n annähernd Chi[5]-verteilt. Man ermittle mittels Simulation für einen Stichprobenumfang von $n = 50$ den maximalen Unterschied zwischen der [empirischen Verteilungsfunktion](#) \hat{F}_Z von Z und der Verteilungsfunktion F der Chi-Quadrat Verteilung mit 5 Freiheitsgraden.

▼

Lösung: Wir ermitteln durch Simulation zusammen mit dem Befehl [EmpiricalCDF](#) die empirische Verteilungsfunktion \hat{F}_Z der Zufallsvariablen Z und berechnen für alle [Stellen](#) z , in denen diese empirische Verteilungsfunktion eine [Sprungstelle](#) aufweist, die Differenzen $|\hat{F}_Z[z] - F[z]|$ und $|\hat{F}_Z[z-] - F[z]|$ zwischen dieser empirischen Verteilungsfunktion und der Verteilungsfunktion F der Chi[5]-Verteilung. Von allen so ermittelten Differenzen bestimmen wir das Maximum. Zur Veranschaulichung zeichnen wir außerdem diese [empirische Verteilungsfunktion](#) zusammen mit der [Verteilungsfunktion der Chi\[5\]-Verteilung](#) in eine gemeinsame Zeichnung:

```
n = 50; m = 10 000;
daten = Table[Apply[Plus, 6 n (Rest[BinCounts[RandomInteger[{1, 6}, n]]/n - 1/6)^2], {m}];
z = Union[daten];
Max[Join[Table[Abs[EmpiricalCDF[daten, z[[i]]] - CDF[ChiSquareDistribution[5], z[[i]]], {i, 1, Length[z]}],
Table[Abs[EmpiricalCDF[daten, z[[i - 1]]] - CDF[ChiSquareDistribution[5], z[[i]]], {i, 1, Length[z]}]]];
Plot[{EmpiricalCDF[daten, z], CDF[ChiSquareDistribution[5], z]}, {z, 0, 15},
PlotStyle -> {{Thickness[0.01], Magenta}, {Thickness[0.01], Blue}}, AspectRatio -> 0.5]
Clear[n, m, daten, z]
```



Der folgende tiefliegende Satz ist Grundlage für viele Anwendungen der Statistik. Zu seiner Formulierung ist die Schreibweise der Matrizenrechnung erforderlich: Matrizen werden mit großen, fetten Buchstaben bezeichnet; für

die Einheitsmatrix verwenden wir den Buchstaben \mathbf{E} ; Zeilenvektoren bezeichnen wir mit $\vec{x} = \{x_1, x_2, \dots, x_n\}$, den zugehörigen Spaltenvektor bezeichnen wir mit $\vec{x}^t = \{x_1, x_2, \dots, x_n\}^t$; das skalare Produkt von Vektoren bzw. Matrizen bezeichnen wir mit " \cdot "; unter dem Rang $\text{Rg}[A]$ einer Matrix A versteht man die Anzahl der linear unabhängigen Zeilenvektoren dieser Matrix.

23.3.7 Satz von COCHRAN: Sind die Zufallsvariablen X_1, X_2, \dots, X_n vollständig unabhängig und identisch $\mathcal{N}[0, 1]$ -verteilt und sind $\Gamma_1, \Gamma_2, \dots, \Gamma_k$ symmetrische $n \times n$ Matrizen mit den beiden Eigenschaften

$$\Gamma_1 + \Gamma_2 + \dots + \Gamma_k = \mathbf{E} \quad \text{und} \quad \text{Rg}[\Gamma_1] + \text{Rg}[\Gamma_2] + \dots + \text{Rg}[\Gamma_k] = n$$

so sind die Zufallsvariablen Z_1, Z_2, \dots, Z_k mit

$$Z_i = \{X_1, X_2, \dots, X_n\} \cdot \Gamma_i \cdot \{X_1, X_2, \dots, X_n\}^t$$

vollständig unabhängig. Außerdem genügt für alle $i \in \{1, 2, \dots, k\}$ die Zufallsvariable Z_i einer Chi-Quadrat-Verteilung mit $\text{Rg}[\Gamma_i]$ Freiheitsgraden.

▼

Beweis: Wir beschränken uns der Einfachheit halber auf einen Beweis dieses Satzes für den Fall $k = 2$:

a) Ist Γ_1 eine symmetrische $n \times n$ Matrix mit Rang $\text{Rg}[\Gamma_1] = r$, so existiert bekanntlich eine orthogonale $n \times n$ Matrix \mathbf{P} (eine Matrix \mathbf{P} heißt orthogonal, wenn $\mathbf{P}^t \cdot \mathbf{P} = \mathbf{E}$ gilt) mit der Eigenschaft

$$\mathbf{P}^t \cdot \Gamma_1 \cdot \mathbf{P} = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_r, 0, \dots, 0]$$

wobei $\lambda_1, \lambda_2, \dots, \lambda_r$ die von 0 verschiedenen Eigenwerte der Matrix Γ_1 bezeichnen. Damit gilt aber

$$\mathbf{P}^t \cdot \Gamma_2 \cdot \mathbf{P} = \mathbf{P}^t \cdot (\mathbf{E} - \Gamma_1) \cdot \mathbf{P} = \mathbf{E} - \mathbf{P}^t \cdot \Gamma_1 \cdot \mathbf{P} = \text{diag}[1 - \lambda_1, 1 - \lambda_2, \dots, 1 - \lambda_r, 1, \dots, 1]$$

Nun hat aber die Matrix Γ_2 und damit auch die Matrix $\mathbf{P}^t \cdot \Gamma_2 \cdot \mathbf{P}$ voraussetzungsgemäß den Rang $n - r$, also muss $\lambda_1 = \lambda_2 = \dots = \lambda_r = 1$ sein.

b) Die Zufallsvariablen Y_1, Y_2, \dots, Y_n mit $\{Y_1, Y_2, \dots, Y_n\} = \{X_1, X_2, \dots, X_n\} \cdot \mathbf{P}$ sind wegen [Bemerkung 23.6.6](#) vollständig unabhängig und identisch $\mathcal{N}[0, 1]$ -verteilt. Auf Grund der [Familieneigenschaft](#) sind damit die beiden Zufallsvariablen

$$\begin{aligned} Z_1 &= \{X_1, X_2, \dots, X_n\} \cdot \Gamma_1 \cdot \{X_1, X_2, \dots, X_n\}^t = \\ &= \{Y_1, Y_2, \dots, Y_n\} \cdot \mathbf{P}^t \cdot \Gamma_1 \cdot \mathbf{P} \cdot \{Y_1, Y_2, \dots, Y_n\}^t = Y_1^2 + Y_2^2 + \dots + Y_r^2 \end{aligned}$$

und

$$\begin{aligned} Z_2 &= \{X_1, X_2, \dots, X_n\} \cdot \Gamma_2 \cdot \{X_1, X_2, \dots, X_n\}^t = \\ &= \{Y_1, Y_2, \dots, Y_n\} \cdot \mathbf{P}^t \cdot \Gamma_2 \cdot \mathbf{P} \cdot \{Y_1, Y_2, \dots, Y_n\}^t = Y_{r+1}^2 + Y_{r+2}^2 + \dots + Y_n^2 \end{aligned}$$

unabhängig und genügen wegen [Satz 23.3.1](#) einer Chi-Quadrat-Verteilung mit r bzw. $n - r$ Freiheitsgraden.

Mit dem folgenden Satz wird exemplarisch gezeigt, wie sich der Satz von Cochran anwenden lässt. Im Rahmen der Statistik werden wir das hier vorgestellte Ergebnis oft verwenden:

23.3.8 Satz: Sind die Zufallsvariablen X_1, X_2, \dots, X_n vollständig unabhängig und identisch $\mathcal{N}[0, 1]$ -verteilt und bezeichnet $\bar{X}^{(n)} = (X_1 + X_2 + \dots + X_n)/n$ ihren **empirischen Mittelwert**, so ist die Zufallsvariable

$$Q_X^{(n)} = \sum_{i=1}^n (X_i - \bar{X}^{(n)})^2$$

$\text{Chi}[n - 1]$ -verteilt. Außerdem sind die beiden Zufallsvariablen $\bar{X}^{(n)}$ und $Q_X^{(n)}$ unabhängig.

▼

Beweis: Wir verwenden zum Nachweis dieser Aussage den [Satz von Cochran](#) und setzen

$$\Gamma_1 = \begin{pmatrix} 1/n & 1/n & \dots & 1/n \\ 1/n & 1/n & \dots & 1/n \\ \vdots & \vdots & \ddots & \vdots \\ 1/n & 1/n & \dots & 1/n \end{pmatrix} \quad \text{und} \quad \Gamma_2 = \mathbf{E} - \Gamma_1$$

Die beiden Matrizen Γ_1 und Γ_2 sind offensichtlich symmetrisch. Außerdem gilt $\text{Rg}[\Gamma_1] = 1$ (alle Zeilen von Γ_1 sind gleich) und $\text{Rg}[\Gamma_2] = n - 1$ (subtrahiert man von den ersten $n - 1$ Spalten der Matrix Γ_2 die letzte Spalte (diese Operation ändert den Rang einer Matrix bekanntlich nicht) so erhält man eine Matrix, bei der links oben die $(n - 1) \times (n - 1)$ Einheitsmatrix steht und deren Zeilen sich zur Nullzeile addieren). Nun ist aber

$$\{X_1, X_2, \dots, X_n\} \cdot \Gamma_1 \cdot \{X_1, X_2, \dots, X_n\}^t = n (\bar{X}^{(n)})^2$$

und

$$\{X_1, X_2, \dots, X_n\} \cdot \Gamma_2 \cdot \{X_1, X_2, \dots, X_n\}^t = \sum_{i=1}^n X_i^2 - n (\bar{X}^{(n)})^2 = \sum_{i=1}^n (X_i - \bar{X}^{(n)})^2 = Q_X^{(n)}$$

womit gezeigt ist, dass $Q_X^{(n)}$ einer $\text{Chi}[n - 1]$ -Verteilung genügt und die beiden Zufallsvariablen $n (\bar{X}^{(n)})^2$ und $Q_X^{(n)}$ und damit auch die beiden Zufallsvariablen $\bar{X}^{(n)}$ und $Q_X^{(n)}$ unabhängig sind.

23.4 Die Student T Verteilung $\mathcal{T}[n]$

Wir beginnen wieder mit einem für die Student T Verteilung zentralen Satz:

23.4.1 Satz: Sind die beiden Zufallsvariablen X und Y unabhängig und gilt $X \approx \mathcal{N}[0, 1]$ und $Y \approx \text{Chi}[n]$, so ist die Zufallsvariable

$$Z = \frac{X}{\sqrt{Y}} \sqrt{n}$$

$\mathcal{T}[n]$ -verteilt. Den Parameter n nennt man dabei den **Freiheitsgrad** der $\mathcal{T}[n]$ -Verteilung.

▼

Beweis: Aus dem [Satz von der totalen Wahrscheinlichkeit](#) in differenzieller Form zusammen mit der Eigenschaft über das [Einsetzen einer Bedingung](#) ergibt sich für alle $z \in \mathbb{R}$

$$\begin{aligned} f_Z[z] &= \frac{1}{dz} \mathbb{P}[\{Z \in [z, z + dz]\}] = \\ &= \frac{1}{dz} \int_0^\infty \mathbb{P}[\{\frac{X}{\sqrt{Y}} \sqrt{n} \in [z, z + dz] \mid \{Y = y\}\}] \mathbb{P}[\{Y \in [y, y + dy]\}] = \\ &= \frac{1}{dz} \int_0^\infty \mathbb{P}[\{X \in [\sqrt{y/n} z, \sqrt{y/n} (z + dz)]\}] \mathbb{P}[\{Y \in [y, y + dy]\}] = \\ &= \int_0^\infty f_X[\sqrt{y/n} z] \sqrt{y/n} f_Y[y] dy \end{aligned}$$

Wir werten dieses Integral mit Hilfe von *Mathematica* aus und zeigen gleichzeitig, dass es sich bei $f_Z[z]$ um die Verteilungsdichte der Student T Verteilung mit n Freiheitsgraden handelt:

```
FullSimplify[Integrate[PDF[NormalDistribution[0, 1], Sqrt[y/n] z] Sqrt[y/n] PDF[ChiSquareDistribution[n], y], {y,
PDF[StudentTDistribution[n], z], {n > 0, z ∈ Reals}]
```

True

Zusammen mit [Satz 23.3.8](#) folgt daraus der für die Statistik wichtige

23.4.2 Satz: Sind die Zufallsvariablen X_1, X_2, \dots, X_n vollständig unabhängig und identisch $\mathcal{N}[0, 1]$ -verteilt und bezeichnet

$$\bar{X}^{(n)} = \frac{1}{n} (X_1 + X_2 + \dots + X_n) \quad \text{bzw.} \quad V_X^{(n)} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}^{(n)})^2$$

ihren **empirischen Mittelwert** bzw ihre **empirische Varianz**, so ist die Zufallsvariable

$$T_X^{(n)} = \frac{\bar{X}^{(n)}}{\sqrt{V_X^{(n)}}} \sqrt{n}$$

$\mathcal{T}[n-1]$ -verteilt.

▼

Beweis: Wegen der [Faltungsformel](#) zusammen mit dem [Satz über die affine Transformation](#) von Normalverteilungen ist die Zufallsvariable $\sqrt{n} \bar{X}^{(n)}$ offenbar $\mathcal{N}[0, 1]$ -verteilt. Wegen [Satz 23.3.8](#) genügt die Zufallsvariable

$$Q_X^{(n)} = (n-1) V_X^{(n)}$$

einer $\text{Chi}[n-1]$ -Verteilung. Außerdem sind die beiden Zufallsvariablen $\sqrt{n} \bar{X}^{(n)}$ und $Q_X^{(n)}$ unabhängig. Damit ist die Zufallsvariable

$$T_X^{(n)} = \frac{\bar{X}^{(n)}}{\sqrt{V_X^{(n)}}} \sqrt{n} = \frac{\sqrt{n} \bar{X}^{(n)}}{\sqrt{Q_X^{(n)}}} \sqrt{n-1}$$

wegen [Satz 23.4.1](#) $\mathcal{T}[n-1]$ -verteilt.

Wir fassen wieder die bereits bekannten Eigenschaften der Student T Verteilung $\mathcal{T}[n]$ zusammen:

23.4.3 Bemerkung: Die Student T Verteilung $\mathcal{T}[n]$ besitzt den Träger

$$\mathbb{T} = \mathbb{R}$$

die Verteilungsdichte

$$f[z] = \frac{1}{\sqrt{n} B[n/2, 1/2]} \left(\frac{n}{n+z^2} \right)^{(n+1)/2}$$

und die Verteilungsfunktion

$$F[z] = \frac{1}{\sqrt{n} B[n/2, 1/2]} \int_{-\infty}^z \left(\frac{n}{n+t^2} \right)^{(n+1)/2} dt = \begin{cases} (1 - B_r[\frac{n}{n+z^2}, 1, \frac{n}{2}, \frac{1}{2}])/2 & \text{für } z < 0 \\ (1 + B_r[\frac{n}{n+z^2}, 1, \frac{n}{2}, \frac{1}{2}])/2 & \text{für } z \geq 0 \end{cases}$$

Eine $\mathcal{T}[n]$ -verteilte Zufallsvariable Z besitzt im Fall $n \geq 2$ den Erwartungswert

$$E[Z] = 0$$

und im Fall $n \geq 3$ die Varianz

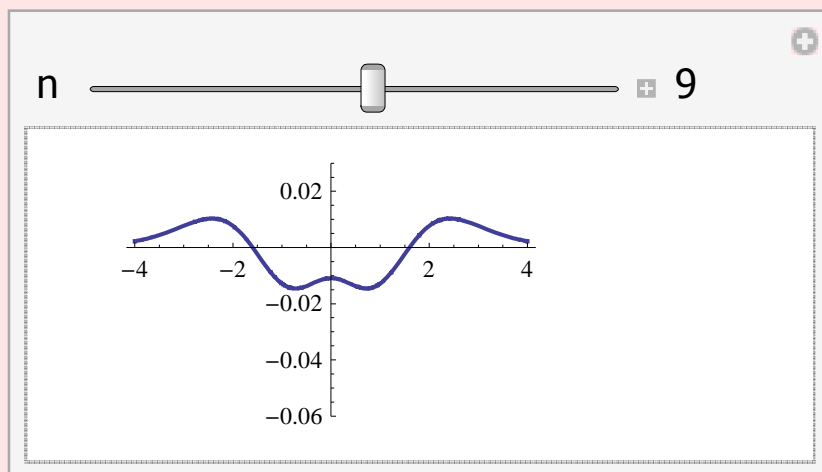
$$V[Z] = n/(n-2)$$

23.4.4 Beispiel: Wie man an Hand einer Zeichnung leicht erkennt, unterscheidet sich für große n die Verteilungsdichte der $\mathcal{T}[n]$ -Verteilung nur wenig von der Verteilungsdichte der $\mathcal{N}[0, 1]$ -Verteilung. Man ermittle, wie groß n sein muss, damit der maximale Unterschied zwischen den Verteilungsdichten dieser beiden Verteilungen kleiner als 0.01 ist.



Lösung: Wir plotten zuerst die Differenz dieser beiden Verteilungsdichten im Bereich $-4 \leq z \leq 4$

```
Manipulate[Plot[PDF[StudentTDistribution[n], z] - PDF[NormalDistribution[0, 1], z], {z, -4, 4},
  PlotStyle -> Thickness[0.01], ImageSize -> {200, 100}, PlotRange -> {-0.06, 0.03}],
  {n, 2, 15, 1, Appearance -> "Labeled"}]
```



und erkennen, dass der größte Unterschied dieser beiden Verteilungsdichten in der Nähe von $z = 1$ liegt. Mittels [FindMinimum](#) und $z = 1$ als Startwert lässt sich der maximale Unterschied leicht ermitteln. Durch Experimentieren mit verschiedenen Werten von n zeigt sich, dass bereits ab $n = 14$ der maximale Unterschied dieser beiden Verteilungsdichten kleiner als 0.01 ist.

```

n = 14;
FindMinimum[PDF[StudentTDistribution[n], z] - PDF[NormalDistribution[0, 1], z], {z, -1}]
Clear[n]

{-0.0094994, {z -> -0.730486}}

```

23.5 Die Fisher F Verteilung $\mathcal{F}[m, n]$

Wir beginnen ebenfalls mit einem für die Fisher F Verteilung zentralen Satz:

23.5.1 Satz: Sind die beiden Zufallsvariablen X und Y unabhängig und gilt $X \approx \text{Chi}[m]$ und $Y \approx \text{Chi}[n]$, so ist die Zufallsvariable

$$Z = \frac{X/m}{Y/n}$$

$\mathcal{F}[m, n]$ -verteilt. Die Parameter m und n heißen **Freiheitsgrade** der $\mathcal{F}[m, n]$ -Verteilung.

▼

Beweis: Aus dem [Satz von der totalen Wahrscheinlichkeit](#) in differenzieller Form zusammen mit der Eigenschaft über das [Einsetzen einer Bedingung](#) ergibt sich für alle $z > 0$

$$\begin{aligned}
 f_Z[z] &= \frac{1}{dz} \mathbb{P}\{Z \in [z, z + dz]\} = \\
 &= \frac{1}{dz} \int_0^\infty \mathbb{P}\left\{\frac{X/m}{Y/n} \in [z, z + dz] \mid \{Y = y\}\right\} \mathbb{P}\{Y \in [y, y + dy]\} = \\
 &= \frac{1}{dz} \int_0^\infty \mathbb{P}\{X \in [(y m/n) z, (y m/n)(z + dz)]\} \mathbb{P}\{Y \in [y, y + dy]\} = \\
 &= \int_0^\infty f_X[(y m/n) z] (y m/n) f_Y[y] dy
 \end{aligned}$$

Wir werten dieses Integral mit Hilfe von *Mathematica* aus und zeigen gleichzeitig, dass es sich bei $f_Z[z]$ um die Verteilungsdichte der Fisher F Verteilung mit m und n Freiheitsgraden handelt:

```

FullSimplify[Integrate[PDF[ChiSquareDistribution[m], (y m/n) z] (y m/n) PDF[ChiSquareDistribution[n], y], {y, 0,
PDF[FRatioDistribution[m, n], z], {m > 0, n > 0, z > 0}]

```

```
True
```

Dieser Satz hat die folgende für die Statistik wichtige Konsequenz:

23.5.2 Satz: Sind die Zufallsvariablen X_1, X_2, \dots, X_m und Y_1, Y_2, \dots, Y_n vollständig unabhängig und identisch $\mathcal{N}[0, 1]$ -verteilt und bezeichnen

$$V_X^{(m)} = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X}^{(m)})^2 \quad \text{bzw.} \quad V_Y^{(n)} = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y}^{(n)})^2$$

die **empirischen Varianzen** der Zufallsvariablen X_1, X_2, \dots, X_m bzw. Y_1, Y_2, \dots, Y_n , so ist ihr Quotient

$$F_{X,Y}^{(m,n)} = V_X^{(m)} / V_Y^{(n)}$$

$\mathcal{F}[m-1, n-1]$ -verteilt.

▼

Beweis: Wegen [Satz 23.3.8](#) sind die beiden Zufallsvariablen $Q_X^{(m)} = (m-1) V_X^{(m)}$ bzw. $Q_Y^{(n)} = (n-1) V_Y^{(n)}$ Chi-Quadrat verteilt mit $m-1$ bzw. $n-1$ Freiheitsgraden. Außerdem sind diese beiden Zufallsvariablen auf Grund der [Familieneigenschaft](#) unabhängig. Wegen [Satz 23.5.1](#) ist der Quotient

$$F_{X,Y}^{(m,n)} = \frac{V_X^{(m)}}{V_Y^{(n)}} = \frac{Q_X^{(m)}/(m-1)}{Q_Y^{(n)}/(n-1)}$$

der beiden empirischen Varianzen somit $\mathcal{F}[m-1, n-1]$ -verteilt.

Wir fassen die bereits bekannten Eigenschaften der Fisher F Verteilung $\mathcal{F}[m, n]$ wieder zusammen:

23.5.3 Bemerkung: Die Fisher F Verteilung $\mathcal{F}[m, n]$ besitzt den Träger

$$\mathbb{T} = [0, \infty[$$

die Verteilungsdichte

$$f[z] = \begin{cases} \frac{1}{B[m/2, n/2]} m^{m/2} n^{n/2} z^{m/2-1} (n+mz)^{-(m+n)/2} & \text{für } z \geq 0 \\ 0 & \text{sonst} \end{cases}$$

und die Verteilungsfunktion

$$F[z] = \begin{cases} 0 & \text{für } z < 0 \\ \frac{m^{m/2} n^{n/2}}{B[m/2, n/2]} \int_0^z t^{m/2-1} (n+mt)^{-(m+n)/2} dt = B_r\left[\frac{mz}{n+mz}, \frac{m}{2}, \frac{n}{2}\right] & \text{für } z \geq 0 \end{cases}$$

Eine $\mathcal{F}[m, n]$ -verteilte Zufallsvariable Z besitzt im Fall $n \geq 3$ den Erwartungswert

$$E[Z] = n/(n-2)$$

und im Fall $n \geq 5$ die Varianz

$$V[Z] = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$$

Weiters folgt unmittelbar aus Satz 23.5.1: Ist die Zufallsvariable $Z \mathcal{F}[m, n]$ -verteilt, so genügt ihr reziproker Wert $Y = 1/Z$ einer $\mathcal{F}[n, m]$ -Verteilung.

Abschließend erwähnen wir noch einige Beziehungen zwischen der Fisher F Verteilung und anderen Verteilungen:

23.5.4 Satz: Ist die Zufallsvariable $Z \mathcal{T}[n]$ -verteilt, so ist die Zufallsvariable $Y = Z^2 \mathcal{F}[1, n]$ -verteilt.

▼

Beweis: Genügt die Zufallsvariable Z einer $\mathcal{T}[n]$ -Verteilung, so gilt für alle $y > 0$

$$\begin{aligned} f_Y[y] &= \frac{1}{dy} \mathbb{P}\{[Z^2 \in [y, y+dy]]\} = \\ &= \frac{1}{dy} \mathbb{P}\{[Z \in [\sqrt{y}, \sqrt{y} + \frac{dy}{2\sqrt{y}}]] \cup [Z \in [-\sqrt{y} - \frac{dy}{2\sqrt{y}}, -\sqrt{y}]]\} = f_Z[\sqrt{y}]/\sqrt{y} \end{aligned}$$

wobei wir die Tatsache $\sqrt{y+dy} \approx \sqrt{y} + dy/(2\sqrt{y})$ verwendet haben (Taylorentwicklung) und außerdem berücksichtigt haben, dass die Verteilungsdichte der Student T Verteilung eine gerade Funktion ist. Mit Hilfe von *Mathematica* zeigen wir nun, dass $Y = Z^2$ einer $\mathcal{F}[1, n]$ -Verteilung genügt:

```
FullSimplify[PDF[StudentTDistribution[n], Sqrt[y]]/Sqrt[y] == PDF[FRatioDistribution[1, n], y], {y > 0, n > 0}]
```

```
True
```

23.5.5 Satz: Ist die Zufallsvariable $Z \text{Beta}[m/2, n/2]$ -verteilt, so ist die Zufallsvariable $Y = nZ/(m-mZ)$ $\mathcal{F}[m, n]$ -verteilt.

▼

Beweis: Genügt die Zufallsvariable Z einer $\text{Beta}[m/2, n/2]$ -Verteilung, so gilt für alle $y > 0$

$$\begin{aligned} f_Y[y] &= \frac{1}{dy} \mathbb{P}\left\{\left\{\frac{nZ}{m-mZ} \in [y, y+dy]\right\}\right\} = \\ &= \frac{1}{dy} \mathbb{P}\left\{Z \in \left[\frac{my}{n+my}, \frac{my}{n+my} + \frac{mndy}{(n+my)^2}\right]\right\} = f_Z\left[\frac{my}{n+my}\right] \frac{mn}{(n+my)^2} \end{aligned}$$

wobei wir verwendet haben (Taylorentwicklung), dass offenbar

$$\frac{m(y+dy)}{n+m(y+dy)} \approx \frac{my}{n+my} + \frac{mndy}{(n+my)^2}$$

gilt. Mit Hilfe von *Mathematica* zeigen wir nun, dass $Y = nZ/(m - mZ)$ einer $\mathcal{F}[1, n]$ -Verteilung genügt:

```
FullSimplify[PDF[BetaDistribution[m/2, n/2], my/(n+my)] mn/(n+my)^2 == PDF[FRatioDistribution[m, n], y]
{y > 0, n > 0, m > 0}]
```

```
True
```

23.6 Die Multinormalverteilung $\mathcal{MN}[\vec{\mu}, \Sigma]$

Das mehrdimensionale Analogon des zentralen Grenzwertungssatzes besagt, dass die Summe einer großen Anzahl von vollständig unabhängigen, identisch verteilten, quadratisch integrierbaren Zufallsvektoren annähernd multinormalverteilt ist. Wir fassen die bereits bekannten Eigenschaften der Multinormalverteilung $\mathcal{MN}[\vec{\mu}, \Sigma]$ wieder zusammen:

:23.6.1 Bemerkung: Sei $\vec{\mu} = \{\mu_1, \mu_2, \dots, \mu_r\} \in \mathbb{R}^r$ und sei $\Sigma = \{\{\sigma_{ik}\}\} \in \mathbb{R}_r^r$ eine symmetrische, positiv definite Matrix. Die Multinormalverteilung $\mathcal{MN}[\vec{\mu}, \Sigma]$ besitzt den Träger

$$\mathbb{T} = \mathbb{R}^r$$

die Verteilungsdichte

$$f[\vec{z}] = \frac{1}{(2\pi)^{r/2} \sqrt{\text{Det}[\Sigma]}} \text{Exp}\left[-\frac{1}{2}(\vec{z} - \vec{\mu}) \cdot \Sigma^{-1} \cdot (\vec{z} - \vec{\mu})^t\right]$$

und die Verteilungsfunktion

$$F[\vec{z}] = \frac{1}{(2\pi)^{r/2} \sqrt{\text{Det}[\Sigma]}} \int_{-\infty}^{z_1} \int_{-\infty}^{z_2} \dots \int_{-\infty}^{z_r} \text{Exp}\left[-\frac{1}{2}(\vec{x} - \vec{\mu}) \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu})^t\right] dx_1 dx_2 \dots dx_r$$

Ein $\mathcal{N}[\vec{\mu}, \Sigma]$ -verteilter Zufallsvektor $\vec{Z} = \{Z_1, Z_2, \dots, Z_r\}$ besitzt den Erwartungswertvektor

$$\mathbb{E}[\vec{Z}] = \{\mathbb{E}[Z_i]\}_{i \in \{1, 2, \dots, r\}} = \vec{\mu}$$

und die Kovarianzmatrix

$$\mathbb{K}[\vec{Z}] = \{\mathbb{K}[Z_i, Z_k]\}_{i, k \in \{1, 2, \dots, r\}} = \Sigma$$



Achtung! Ist $r = 1$, so stimmt die Multinormalverteilung $\mathcal{MN}[\vec{\mu}, \Sigma]$ mit $\vec{\mu} = \{\mu\}$ und $\Sigma = \{\{\sigma^2\}\}$ mit der Normalverteilung $\mathcal{N}[\mu, \sigma]$ überein. Man beachte, dass bei der Multinormalverteilung der Parameter Σ der Kovarianzmatrix entspricht, während bei der Normalverteilung der Parameter σ der Streuung (also der Wurzel der Varianz) entspricht.

```
FullSimplify[CDF[MultinormalDistribution[{\mu}, {\sigma^2}], {z}] == CDF[NormalDistribution[\mu, \sigma], z],
{\sigma > 0}]
True
```

Analog zum [Satz über die affine Transformation](#) von eindimensionalen Normalverteilungen gilt auch für Multinormalverteilungen ein entsprechender Satz:

23.6.2 Satz (Affine Transformation von Multinormalverteilungen): Ist der Zufallsvektor \vec{Z} multinormalverteilt mit dem Mittelwertsvektor $\vec{\mu}$ und der Kovarianzmatrix Σ , so ist für alle $k \in \{1, 2, \dots, r\}$, alle $\vec{v} \in \mathbb{R}^k$ und alle Matrizen $\Gamma \in \mathbb{R}_r^k$ mit vollem Rang der Zufallsvektor $\vec{Y} = \vec{v} + \vec{Z} \cdot \Gamma$ multinormalverteilt mit dem Mittelwertsvektor $\vec{v} + \vec{\mu} \cdot \Gamma$ und der Kovarianzmatrix $\Gamma^t \cdot \Sigma \cdot \Gamma$.

▼

Beweis: Wir beschränken uns auf den Fall $k = r$ (da Γ vollen Rang besitzen soll, ist die Matrix Γ invertierbar). Für alle $\vec{y} \in \mathbb{R}^r$ gilt dann unter Verwendung der Transformation $\vec{x} = (\vec{u} - \vec{v}) \cdot \Gamma^{-1}$ (für die Funktionalmatrix dieser Transformation gilt offenbar $\partial \vec{x} / \partial \vec{u} = \Gamma^{-1}$)

$$\begin{aligned} \mathbb{F}_{\vec{Y}}[\vec{y}] &= \mathbb{P}[\{\vec{Y} \leq \vec{y}\}] = \mathbb{P}[\{\vec{v} + \vec{Z} \cdot \Gamma \leq \vec{y}\}] = \mathbb{P}[\{\vec{Z} \leq (\vec{y} - \vec{v}) \cdot \Gamma^{-1}\}] = \\ &= \frac{1}{(2\pi)^{r/2} \sqrt{\text{Det}[\Sigma]}} \int_{-\infty}^{(\vec{y} - \vec{v}) \cdot \Gamma^{-1}} \text{Exp}\left[-\frac{1}{2}(\vec{x} - \vec{\mu}) \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu})^t\right] |d\vec{x}| = \\ &= \frac{1}{(2\pi)^{r/2} \sqrt{\text{Det}[\Sigma]} \text{Det}[\Gamma]} \int_{-\infty}^{\vec{y}} \text{Exp}\left[-\frac{1}{2}((\vec{u} - \vec{v}) \cdot \Gamma^{-1} - \vec{\mu}) \cdot \Sigma^{-1} \cdot ((\vec{u} - \vec{v}) \cdot \Gamma^{-1} - \vec{\mu})^t\right] |d\vec{u}| = \\ &= \frac{1}{(2\pi)^{r/2} \sqrt{\text{Det}[\Gamma^t \cdot \Sigma \cdot \Gamma]}} \int_{-\infty}^{\vec{y}} \text{Exp}\left[-\frac{1}{2}(\vec{u} - (\vec{v} + \vec{\mu} \cdot \Gamma)) \cdot (\Gamma^t \cdot \Sigma \cdot \Gamma)^{-1} \cdot (\vec{u} - (\vec{v} + \vec{\mu} \cdot \Gamma))^t\right] |d\vec{u}| \end{aligned}$$

Damit ist gezeigt, dass der Zufallsvektor $\vec{Y} = \vec{v} + \vec{Z} \cdot \Gamma$ mit dem Mittelwertsvektor $\vec{v} + \vec{\mu} \cdot \Gamma$ und der Kovarianzmatrix $\Gamma^t \cdot \Sigma \cdot \Gamma$ multinormalverteilt ist.

Aus diesem Satz folgt unmittelbar

23.6.3 Bemerkung: Ist der Zufallsvektor $\{Z_1, Z_2, \dots, Z_r\}$ multinormalverteilt mit dem Mittelwertsvektor $\{\mu_1, \mu_2, \dots, \mu_r\}$ und der Kovarianzmatrix $\{\{\sigma_{ik}\}\}_{i,k \in \{1,2,\dots,r\}}$, so ist für alle $1 \leq i_1 < i_2 < \dots < i_s \leq r$ der Zufallsvektor $\{Z_{i_1}, Z_{i_2}, \dots, Z_{i_s}\}$ multinormalverteilt mit dem Mittelwertsvektor $\{\mu_{i_1}, \mu_{i_2}, \dots, \mu_{i_s}\}$ und der Kovarianzmatrix $\{\{\sigma_{ik}\}\}_{i,k \in \{i_1, i_2, \dots, i_s\}}$.

▼

Beweis: Diese Bemerkung folgt unmittelbar aus dem [Satz über die affine Transformation](#) von Multinormalverteilungen mit der $r \times s$ -Matrix

$$\Gamma = \begin{pmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \end{pmatrix}^t$$

$\begin{matrix} \uparrow & & \uparrow & & \uparrow \\ i_1\text{-te} & & i_2\text{-te} & & i_s\text{-te} \\ \text{Spalte} & & \text{Spalte} & & \text{Spalte} \end{matrix}$

Ein wichtiger Spezialfall dieser Bemerkung besagt

23.6.4 Bemerkung: Ist der Zufallsvektor $\{Z_1, Z_2, \dots, Z_r\}$ multinormalverteilt mit dem Mittelwertsvektor $\{\mu_1, \mu_2, \dots, \mu_r\}$ und der Kovarianzmatrix $\{\{\sigma_{ik}\}\}_{i,k \in \{1,2,\dots,n\}}$, so ist dessen i -te Komponente Z_i normalverteilt mit den Parametern μ_i und $\sqrt{\sigma_{ii}}$.

Wir [wissen](#) bereits, dass unabhängige Zufallsvariable stets unkorreliert sind, umgekehrt aber unkorrelierte Zufallsvariable nicht notwendig unabhängig sein müssen. Weiß man aber, dass die gemeinsame Verteilung dieser Zufallsvariablen eine Multinormalverteilung ist, so gilt diese Umkehrung sehr wohl:

23.6.5 Bemerkung: Genügt der Zufallsvektor $\vec{Z} = \{Z_1, Z_2, \dots, Z_r\}$ einer $\mathcal{MN}[\vec{\mu}, \Sigma]$ -Verteilung, so gilt: Die Zufallsvariablen Z_1, Z_2, \dots, Z_r sind genau dann vollständig unabhängig, wenn die Kovarianzmatrix Σ eine Diagonalmatrix ist, die Zufallsvariablen Z_1, Z_2, \dots, Z_r also paarweise unkorreliert sind.

▼

Beweis: Ist $\vec{\mu} = \{\mu_1, \mu_2, \dots, \mu_r\} \in \mathbb{R}^r$ und ist $\Sigma = \text{Diag}[\sigma_{11}, \sigma_{22}, \dots, \sigma_{rr}]$ eine Diagonalmatrix mit lauter positiven Einträgen in der Diagonale, so gilt für alle Vektoren $\vec{z} = \{z_1, z_2, \dots, z_r\} \in \mathbb{R}^r$

$$\begin{aligned} f_{\vec{Z}}[\vec{z}] &= \frac{1}{(2\pi)^{r/2} \sqrt{\text{Det}[\Sigma]}} \text{Exp}\left[-\frac{1}{2}(\vec{z} - \vec{\mu}) \cdot \Sigma^{-1} \cdot (\vec{z} - \vec{\mu})^t\right] = \\ &= \frac{1}{\sqrt{2\pi\sigma_{11}} \sqrt{2\pi\sigma_{22}} \dots \sqrt{2\pi\sigma_{rr}}} e^{-(z_1 - \mu_1)^2 / (2\sigma_{11})} e^{-(z_2 - \mu_2)^2 / (2\sigma_{22})} \dots e^{-(z_r - \mu_r)^2 / (2\sigma_{rr})} \end{aligned}$$

Die gemeinsame Verteilungsdichte der Zufallsvariablen Z_1, Z_2, \dots, Z_r ist damit wegen [Bemerkung 23.6.4](#) gleich dem Produkt ihrer Marginalverteilungsdichten, womit wegen [Bemerkung 19.1.2](#) gezeigt ist, dass die Zufallsvariablen Z_1, Z_2, \dots, Z_r vollständig unabhängig sind.

Von theoretischem Interesse ist die folgende Bemerkung

23.6.6 Bemerkung: Sind die Zufallsvariablen Z_1, Z_2, \dots, Z_r vollständig unabhängig und $\mathcal{N}[0, 1]$ -verteilt und ist $\mathbf{P} \in \mathbb{R}_r^r$ eine orthogonale Matrix (eine Matrix $\mathbf{P} \in \mathbb{R}_r^r$ heißt bekanntlich orthogonal, wenn $\mathbf{P}^t \cdot \mathbf{P} = \mathbf{E}$ gilt), so sind auch die Zufallsvariablen Y_1, Y_2, \dots, Y_r mit $\{Y_1, Y_2, \dots, Y_r\} = \{Z_1, Z_2, \dots, Z_r\} \cdot \mathbf{P}$ vollständig unabhängig und $\mathcal{N}[0, 1]$ -verteilt.

▼

Beweis: Der Zufallsvektor $\vec{Z} = \{Z_1, Z_2, \dots, Z_r\}$ ist wegen [Bemerkung 23.6.5](#) $\mathcal{MN}[\vec{0}, \mathbf{E}]$ -verteilt. Auf Grund des [Satz über die affine Transformation](#) von Multinormalverteilungen sowie der Tatsache, dass die Matrix \mathbf{P} orthogonal ist, genügt auch der Zufallsvektor $\vec{Y} = \{Y_1, Y_2, \dots, Y_r\} = \vec{Z} \cdot \mathbf{P}$ einer $\mathcal{MN}[\vec{0}, \mathbf{E}]$ -Verteilung, also sind wegen [Bemerkung 23.6.5](#) auch die Zufallsvariablen Y_1, Y_2, \dots, Y_r vollständig unabhängig und $\mathcal{N}[0, 1]$ -verteilt.

Von besonderem Interesse ist der Fall $r = 2$, also der zweidimensionalen Normalverteilung:

23.6.7 Bemerkung: Ist der Zufallsvektor $\vec{Z} = \{Z_1, Z_2\}$ zweidimensional normalverteilt mit dem Mittelwertsvektor $\vec{\mu} = \{\mu_1, \mu_2\}$ und der Kovarianzmatrix Σ , so gilt

a) Die Kovarianzmatrix Σ hat die Gestalt

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}$$

wobei σ_1^2 bzw. σ_2^2 den Varianzen und ρ dem Korrelationskoeffizient der beiden Zufallsvariablen Z_1 und Z_2 entsprechen.

b) Für die gemeinsame Verteilungsdichte $f_{\vec{Z}}$ der beiden Zufallsvariablen Z_1 und Z_2 gilt offenbar

$$f_{\vec{Z}}[\vec{z}] = \frac{1}{2\pi \sigma_1 \sigma_2 \sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)} \left(\frac{(z_1 - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{z_1 - \mu_1}{\sigma_1} \frac{z_2 - \mu_2}{\sigma_2} + \frac{(z_2 - \mu_2)^2}{\sigma_2^2} \right)\right]$$

Wie bereits in [Beispiel 17.6.3](#) festgestellt wurde, liefern Schnitte mit vertikalen Ebenen nicht normierte Gauß'sche Glockenkurven und Schnitte mit horizontalen Ebenen gedrehte Ellipsen mit dem Mittelpunkt $\vec{\mu}$.

c) Für die bedingte Verteilungsdichte $f_{Z_2 | \{Z_1=z_1\}}$ gilt wegen [Bemerkung 18.2.3](#)

$$f_{Z_2 | \{Z_1=z_1\}}[z_2] = \frac{1}{\sqrt{2\pi} \sigma_2 \sqrt{1-\rho^2}} \exp\left[-\frac{1}{2\sigma_2^2(1-\rho^2)} \left(z_2 - \left(\mu_2 + \frac{\rho \sigma_2}{\sigma_1} (z_1 - \mu_1) \right) \right)^2\right]$$

Damit ist gezeigt, dass die Zufallsvariable Z_2 unter der Bedingung $\{Z_1 = z_1\}$ normalverteilt ist mit dem Erwartungswert $\mu_2 + \rho \sigma_2 (z_1 - \mu_1) / \sigma_1$ und der Streuung $\sigma_2 \sqrt{1-\rho^2}$. Geometrisch erhält man diese bedingte Verteilungsdichte von Z_2 unter der Bedingung $\{Z_1 = z_1\}$, indem man die gemeinsame Verteilungsdichte von Z_1 und Z_2 mit der vertikalen Ebene $x = z_1$ schneidet und die dabei entstehende Schnittkurve normiert.

d) Für den bedingten Erwartungswert $\mathbb{E}[Z_2 | \{Z_1 = z_1\}]$ gilt damit

$$\mathbb{E}[Z_2 | \{Z_1 = z_1\}] = \mu_2 + \frac{\rho \sigma_2}{\sigma_1} (z_1 - \mu_1)$$

▼

Beweis: a) Diese Bedeutung der Parameter der Kovarianzmatrix Σ folgt unmittelbar aus [Bemerkung 23.6.4](#) und der Definition des [Korrelationskoeffizienten](#) ρ .

c) Die bedingte Verteilungsdichte $f_{Z_2 | \{Z_1=z_1\}}[z_2] = f_{Z_1, Z_2}[z_1, z_2] / f_{Z_1}[z_1]$ ermitteln wir mit Hilfe von *Mathematica*:

```
 $\mu = \{\mu_1, \mu_2\}; \Sigma = \{\{\sigma_1^2, \rho \sigma_1 \sigma_2\}, \{\rho \sigma_1 \sigma_2, \sigma_2^2\}\};$ 
FullSimplify[PDF[MultinormalDistribution[ $\mu, \Sigma$ ], {z1, z2}]/PDF[NormalDistribution[ $\mu_1, \sigma_1$ ], z1], {-1 <  $\rho$  < 1,  $\sigma_1$  > 0}
Clear[ $\mu, \Sigma$ ]
```

$$e^{-\frac{(z_2 \sigma_1 - \mu_2 \sigma_1 + (-z_1 + \mu_1) \rho \sigma_2)^2}{2(-1 + \rho^2) \sigma_1^2 \sigma_2^2}} \frac{1}{\sqrt{2\pi} \sqrt{1-\rho^2} \sigma_2}$$

Die Eigenschaft c) kann zur Simulation von $MN[\{\mu_1, \mu_2\}, \{\{\sigma_1^2, \rho \sigma_1 \sigma_2\}, \{\rho \sigma_1 \sigma_2, \sigma_2^2\}\}]$ -verteilten Zufallspunkten $\{z_1, z_2\} \in \mathbb{R}^2$ herangezogen werden: Man erzeugt dazu zuerst die $\mathcal{N}[\mu_1, \sigma_1]$ -verteilte Zufallszahl z_1 und anschließend die $\mathcal{N}[\mu_2 + \rho \sigma_2 (z_1 - \mu_1) / \sigma_1, \sigma_2 \sqrt{1-\rho^2}]$ -verteilte Zufallszahl z_2 . Eine allgemeine Möglichkeit zur Simulation von multinormalverteilten Zufallspunkten $\vec{z} \in \mathbb{R}^r$ wird in [Beispiel 23.6.9](#) besprochen.

Es folgen wieder einige Beispiele, mit denen gezeigt wird, wie mit der Multinormalverteilung gearbeitet wird:

23.6.8 Beispiel (Der Log-Trig-Algorithmus): Man zeige: Sind die beiden Zufallsvariablen X_1 und X_2 unabhängig und im Intervall $[0, 1]$ gleichverteilt, so sind die beiden Zufallsvariablen

$$Z_1 = \sqrt{-2 \operatorname{Log}[X_1]} \operatorname{Cos}[2 \pi X_2] \quad \text{und} \quad Z_2 = \sqrt{-2 \operatorname{Log}[X_1]} \operatorname{Sin}[2 \pi X_2]$$

unabhängig und $\mathcal{N}[0, 1]$ -verteilt. Dieses Ergebnis liefert einen Algorithmus, mit dem im Intervall $[0, 1]$ gleichverteilte Zufallszahlen in normalverteilte Zufallszahlen transformiert werden können. (Dieser Algorithmus ist aber für praktische Zwecke nicht geeignet, da die Berechnung der dabei auftretenden Funktionen sehr zeitaufwändig ist.)

▼

Lösung: Die Abbildung $\vec{g} : [0, 1]^2 \rightarrow \mathbb{R}^2$ mit

$$\vec{z} = \vec{g}[\vec{x}] = \vec{g}[\{x_1, x_2\}] = \{\sqrt{-2 \operatorname{Log}[x_1]} \operatorname{Cos}[2 \pi x_2], \sqrt{-2 \operatorname{Log}[x_1]} \operatorname{Sin}[2 \pi x_2]\}$$

ist bijektiv und besitzt die Umkehrabbildung $\vec{g}^{-1} : \mathbb{R}^2 \rightarrow [0, 1]^2$ mit

$$\vec{x} = \vec{g}^{-1}[\vec{z}] = \vec{g}^{-1}[\{z_1, z_2\}] = \{\operatorname{Exp}[-(z_1^2 + z_2^2)/2], \operatorname{ArcTan}[z_2/z_1]/2 \pi\}$$

Berücksichtigt man die bekannte Tatsache, dass der infinitesimale Bereich $\vec{g}^{-1}[[\vec{z}, \vec{z} + d\vec{z}]]$ den Flächeninhalt $|\operatorname{Det}[\partial \vec{x}/\partial \vec{z}]| |d\vec{z}|$ besitzt, wobei wir mit $\partial \vec{x}/\partial \vec{z}$ die Funktionalmatrix der Funktion $\vec{x} = \vec{g}^{-1}[\vec{z}]$ an der Stelle \vec{z} bezeichnen, sowie die Tatsache, dass \vec{X} im Bereich $[0, 1]^2$ gleichverteilt ist, so gilt für alle $\vec{z} \in \mathbb{R}^2$

$$\begin{aligned} f_{\vec{Z}}[\vec{z}] &= \frac{1}{|d\vec{z}|} \mathbb{P}[\{\vec{Z} \in [\vec{z}, \vec{z} + d\vec{z}]\}] = \frac{1}{|d\vec{z}|} \mathbb{P}[\{\vec{g}^{-1}[\vec{Z}] \in \vec{g}^{-1}[[\vec{z}, \vec{z} + d\vec{z}]]\}] = \\ &= \frac{1}{|d\vec{z}|} f_{\vec{X}}[\vec{g}^{-1}[\vec{z}]] |\operatorname{Det}[\partial \vec{x}/\partial \vec{z}]| |d\vec{z}| = \frac{1}{2 \pi} \operatorname{Exp}[-(z_1^2 + z_2^2)/2] \end{aligned}$$

23.6.9 Beispiel: Man entwickle einen auf dem [Satz über die affine Transformation](#) von Multinormalverteilungen beruhenden Algorithmus, mit dem sich $\mathcal{MN}[\vec{0}, \mathbf{E}]$ -verteilte Zufallspunkte $\vec{x} \in \mathbb{R}^r$ in $\mathcal{MN}[\vec{\mu}, \mathbf{\Sigma}]$ -verteilte Zufallspunkte $\vec{z} \in \mathbb{R}^r$ transformieren lassen. Da sich $\mathcal{MN}[\vec{0}, \mathbf{E}]$ -verteilte Zufallspunkte leicht erzeugen lassen (man muss dazu nur entsprechend viele $\mathcal{N}[0, 1]$ -verteilte Zufallszahlen erzeugen), verwendet man diesen Algorithmus häufig zu Erzeugung von $\mathcal{MN}[\vec{\mu}, \mathbf{\Sigma}]$ -verteilten Zufallspunkten.

▼

Lösung: Die Matrix $\mathbf{\Sigma}$ ist symmetrisch und positiv definit. Sie ist damit diagonalisierbar, wobei alle Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_r$ positiv sind und die zugehörigen Eigenvektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r$ paarweise orthonormiert gewählt werden können. Bezeichnet man mit \mathbf{P}^t jene orthogonale Matrix, deren Spalten aus diesen orthonormierten Eigenvektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r$ der Matrix $\mathbf{\Sigma}$ bestehen, so gilt bekanntlich $\mathbf{P} \cdot \mathbf{\Sigma} \cdot \mathbf{P}^t = \operatorname{Diag}[\lambda_1, \lambda_2, \dots, \lambda_r]$.

Ist $\mathbf{\Gamma} = \operatorname{Diag}[\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_r}] \cdot \mathbf{P}$ und \vec{X} ein $\mathcal{MN}[\vec{0}, \mathbf{E}]$ -verteilter Zufallsvektor, so ist der Zufallsvektor $\vec{Z} = \vec{\mu} + \vec{X} \cdot \mathbf{\Gamma}$ wegen des [Satzes über die affine Transformation](#) multinormalverteilt mit dem Mittelwertsvektor $\vec{\mu} + \vec{0} \cdot \mathbf{\Gamma} = \vec{\mu}$ und der Kovarianzmatrix $\mathbf{\Gamma}^t \cdot \mathbf{E} \cdot \mathbf{\Gamma} = \mathbf{P}^t \cdot \operatorname{Diag}[\lambda_1, \lambda_2, \dots, \lambda_r] \cdot \mathbf{P} = \mathbf{\Sigma}$.

Dieser Vorgang lässt sich in *Mathematica* leicht implementieren: Man berechnet dazu zuerst mit Hilfe von [Eigensystem](#) die Eigenwerte und zugehörigen Eigenvektoren der symmetrischen und positiv definiten Matrix $\mathbf{\Sigma}$, orthonormiert diese Eigenvektoren mit dem Befehl [Orthogonalize](#) und berechnet anschließend unter Verwendung des Befehls [DiagonalMatrix](#) die Matrix $\mathbf{\Gamma}$. Schließlich wendet man die Transformation $\vec{\mu} + \vec{x} \cdot \mathbf{\Gamma}$ auf eine Liste \vec{x} von r $\mathcal{N}[0, 1]$ -verteilten Zufallszahlen an (bei dem folgenden Algorithmus wird zusätzlich noch geprüft, ob die Matrix $\mathbf{\Sigma}$ tatsächlich symmetrisch und positiv definit ist, wobei gegebenenfalls eine Fehlermeldung ausgegeben wird):

```

NormalverteilterZufallspunkt[μ_, Σ_] :=
Module[{r = Length[Σ], e = N[Eigensystem[Σ]], Γ},
  Γ = DiagonalMatrix[Sqrt[e[[1]]]].Orthogonalize[e[[2]]];
  If[Σ != Transpose[Σ], Print["Σ nicht symmetrisch"], If[Positive[e[[1]]] != Table[True, {r}], Print["Σ nicht posi
  μ + Table[RandomReal[NormalDistribution[0, 1]], {r}].Γ]]

```

Beispielsweise gilt

```

μ = {2, 3, 1, 3}; Σ = {{1, 1, 1, 2}, {1, 3, 1, 2}, {1, 1, 4, 1}, {2, 2, 1, 5}};
NormalverteilterZufallspunkt[μ, Σ]
Clear[μ, Σ]

{1.13126, 3.88876, -2.23091, 2.93927}

```

Es sei erwähnt, dass *Mathematica* beim Aufruf von `RandomReal[MultinormalDistribution[μ, Σ]]` genau diesen Algorithmus verwendet.

23.6.10 Beispiel (Der k -Sigma-Bereich): Für den mit dem Mittelwertsvektor $\vec{\mu}$ und der Kovarianzmatrix Σ multinormalverteilten Zufallsvektor $\vec{Z} = \{Z_1, Z_2, \dots, Z_r\}$ berechne man die Wahrscheinlichkeit dafür, dass \vec{Z} im Ellipsoid $B_k = \{\vec{z} \in \mathbb{R}^r \mid (\vec{z} - \vec{\mu}) \cdot \Sigma^{-1} \cdot (\vec{z} - \vec{\mu}) \leq k^2\}$ liegt. Man nennt dieses Ellipsoid B_k den r -dimensionalen k -Sigma-Bereich der $\mathcal{MN}[\vec{\mu}, \Sigma]$ -Verteilung.

▼

Lösung: Wir verwenden die Bezeichnungen von [Beispiel 23.6.9](#) sowie die Transformation $\vec{x} = (\vec{z} - \vec{\mu}) \cdot \Gamma^{-1}$. Für die Funktionaldeterminante $\text{Det}[\partial \vec{z} / \partial \vec{x}]$ der dazu inversen Transformation gilt wegen $\Gamma^t \cdot \Gamma = \Sigma$

$$\text{Det}[\partial \vec{z} / \partial \vec{x}] = \text{Det}[\Gamma] = \sqrt{\text{Det}[\Sigma]}$$

und damit

$$\begin{aligned} \mathbb{P}\{\vec{Z} \in B_k\} &= \int_{B_k} \dots \int \frac{1}{(2\pi)^{r/2} \sqrt{\text{Det}[\Sigma]}} \text{Exp}\left[-\frac{1}{2}(\vec{z} - \vec{\mu}) \cdot \Sigma^{-1} \cdot (\vec{z} - \vec{\mu})\right] |d\vec{z}| = \\ &= \int_{\{\vec{x} \mid \vec{x} \cdot \vec{x} \leq k^2\}} \dots \int \frac{1}{(2\pi)^{r/2} \sqrt{\text{Det}[\Sigma]}} \text{Exp}\left[-\frac{1}{2}\vec{x} \cdot \vec{x}\right] \sqrt{\text{Det}[\Sigma]} |d\vec{x}| = \\ &= \frac{1}{(2\pi)^{r/2}} \int_0^k (\text{Exp}[-\rho^2/2] \int_{\partial K_r[\rho]} dO) d\rho = \frac{2}{2^{r/2} \Gamma[r/2]} \int_0^k \rho^{r-1} \text{Exp}[-\rho^2/2] d\rho \end{aligned}$$

wobei wir mit

$$\int_{\partial K_r[\rho]} dO = (-1)^{[r/2]} 2^r \int_0^{\pi/2} \int_0^{\pi/2} \dots \int_0^{\pi/2} \frac{\partial \{y_1, y_2, \dots, y_r\}}{\partial \{\rho, \varphi_1, \varphi_2, \dots, \varphi_{r-1}\}} d\varphi_1 d\varphi_2 \dots d\varphi_{r-1} = \frac{2\pi^{r/2}}{\Gamma[r/2]} \rho^{r-1}$$

die Oberfläche einer r -dimensionalen Kugel $K_r[\rho]$ mit Radius ρ bezeichnet haben. Die Berechnung dieses r -dimensionalen Oberflächenintegrals kann dabei unter Verwendung der r -dimensionalen Kugelkoordinaten

$$\begin{aligned} y_r &= \rho \cos[\varphi_1] \cos[\varphi_2] \cos[\varphi_3] \cos[\varphi_4] \dots \cos[\varphi_{r-3}] \cos[\varphi_{r-2}] \cos[\varphi_{r-1}] \\ y_{r-1} &= \rho \sin[\varphi_1] \cos[\varphi_2] \cos[\varphi_3] \cos[\varphi_4] \dots \cos[\varphi_{r-3}] \cos[\varphi_{r-2}] \cos[\varphi_{r-1}] \\ y_{r-2} &= \rho \sin[\varphi_1] \sin[\varphi_2] \cos[\varphi_3] \cos[\varphi_4] \dots \cos[\varphi_{r-3}] \cos[\varphi_{r-2}] \cos[\varphi_{r-1}] \\ &\dots \\ y_3 &= \rho \sin[\varphi_1] \sin[\varphi_2] \sin[\varphi_3] \cos[\varphi_{r-2}] \cos[\varphi_{r-1}] \\ y_2 &= \rho \sin[\varphi_1] \sin[\varphi_2] \sin[\varphi_3] \sin[\varphi_{r-2}] \cos[\varphi_{r-1}] \\ y_1 &= \rho \sin[\varphi_1] \sin[\varphi_2] \sin[\varphi_3] \sin[\varphi_{r-2}] \sin[\varphi_{r-1}] \end{aligned}$$

mit Hilfe der *Mathematica*-Prozedur

```
y[2] = {ρ Cos[φ[1]], ρ Sin[φ[1]]};
x[2] = {ρ, φ[1]};
y[r_] := Append[y[r - 1] Cos[φ[r - 1]], ρ Sin[φ[r - 1]]];
x[r_] := Append[x[r - 1], φ[r - 1]];
Oberfläche[r_] := 2r Integrate[Det[D[y[r], {x[r]}]], Apply[Sequence, Table[{φ[i], 0, π/2}, {i, 1, r - 1}]]]
```

erfolgen. Damit gilt beispielsweise

```
r = 5;
Oberfläche[r]
FullSimplify[Oberfläche[r] == 2 πr/2 ρr-1 / Gamma[r/2]]
Clear[r]
```

$$\frac{8 \pi^2 \rho^4}{3}$$

True

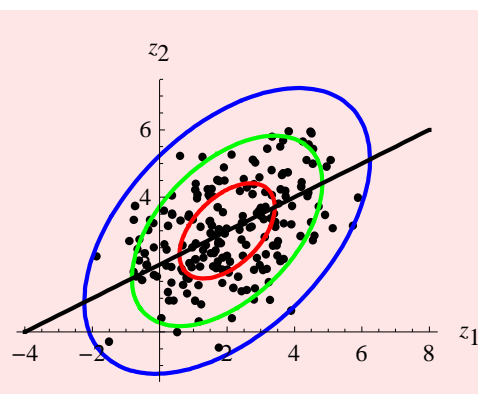
Die Berechnung der Wahrscheinlichkeit des r -dimensionalen k -Sigma-Bereichs ist damit sehr einfach:

```
Wahrscheinlichkeit[r_, k_] := 2 Integrate[ρr-1 Exp[-ρ2/2], {ρ, 0, k}]/(2r/2 Gamma[r/2]) // N
```

```
r = 2; k = 2;
Wahrscheinlichkeit[r, k]
Clear[r, k]
```

0.864665

Die folgende Zeichnung zeigt einen zweidimensionalen $MN[\vec{\mu}, \Sigma]$ -verteilten Punktschwarm zusammen mit den **Ein-**, **Zwei-** und **Drei-** Σ -Bereichen. Bei den Rändern dieser Bereiche handelt es sich um Schnitte der Verteilungsdichte der $MN[\vec{\mu}, \Sigma]$ -Verteilung mit den horizontalen Ebenen $z_3 = e^{-k^2/2}/(2\pi\sqrt{\text{Det}[\Sigma]})$ mit $k = 1, 2, 3$. Weiters ist in diese Zeichnung auch der bedingte Erwartungswert $z_2 = E[Z_2 | \{Z_1 = z_1\}]$ als Funktion von z_1 eingezeichnet. Es handelt sich dabei aber **nicht!** um die Achse dieser Schnittellipsen, sondern um eine Gerade, bei der die jeweiligen vertikalen Abstände zu den oberen und unteren Rändern dieser Schnittellipsen gleich sind.



23.6.11 Beispiel: Der $MN[\vec{\mu}, \Sigma]$ -verteilte zufällige Punkt $\vec{Z} = \{Z_1, Z_2\}$ wird orthogonal auf die durch den Ursprung gehende Gerade $y = kx$ projiziert. Man ermittle die Verteilungsdichte $f_A[a]$ des Abstandes A des projizierten Punktes vom Ursprung.

▼

Lösung: Es bezeichne $\vec{g} = \{1/\sqrt{1+k^2}, k/\sqrt{1+k^2}\}$ den Einheitsvektor in Richtung der Geraden $y = kx$. Auf Grund des [Satzes über die affine Transformation](#) von Multinormalverteilungen ist die Projektion $\vec{Z} \cdot \vec{g}^t$ von \vec{Z} auf \vec{g} normalverteilt mit dem Mittelwert

$$\nu = \vec{\mu} \cdot \vec{g}^t = (\mu_1 + k \mu_2) / \sqrt{1+k^2}$$

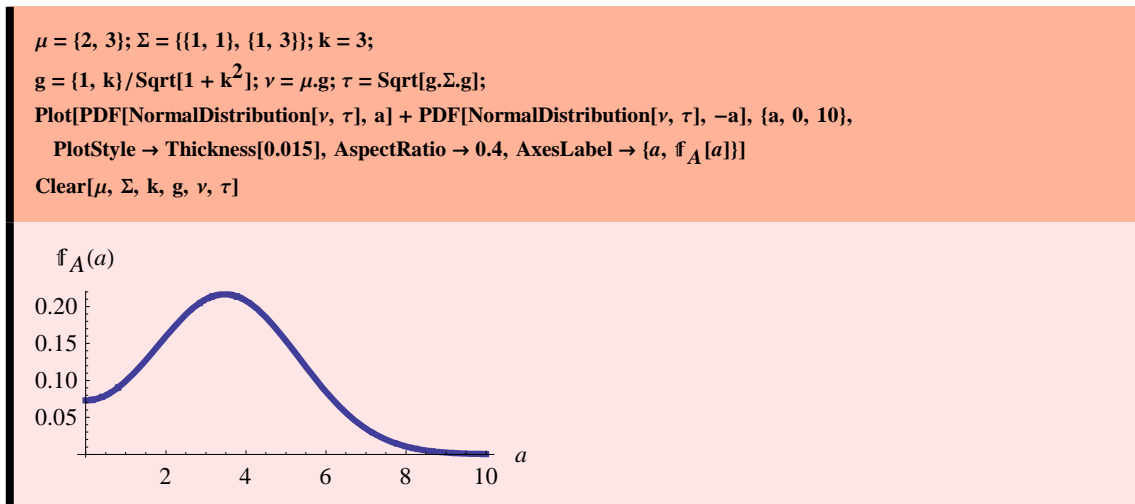
und der Streuung

$$\tau = \sqrt{\vec{g} \cdot \Sigma \cdot \vec{g}^t} = \sqrt{(\sigma_1^2 + 2k\rho\sigma_1\sigma_2 + k^2\sigma_2^2) / (1+k^2)}$$

Der Abstand $A = |\vec{Z} \cdot \vec{g}^t|$ des projizierten Punktes vom Ursprung besitzt damit die Verteilungsdichte

$$f_A[a] = \begin{cases} \varphi_{\nu,\tau}[a] + \varphi_{\nu,\tau}[-a] & \text{für } a \geq 0 \\ 0 & \text{für } a < 0 \end{cases}$$

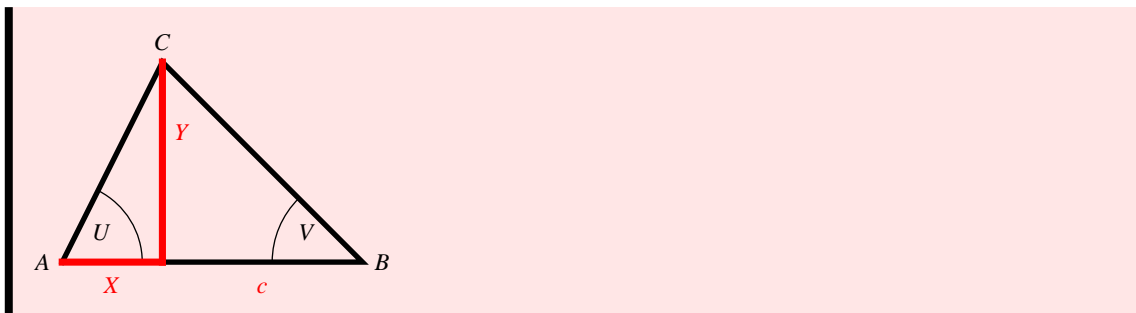
wobei wir mit $\varphi_{\nu,\tau}$ die Verteilungsdichte der Normalverteilung mit dem Mittelwert ν und die Streuung τ bezeichnen. Mit der folgenden Prozedur lässt sich diese Verteilungsdichte f_A für beliebige Werte von μ , Σ und k graphisch darstellen (die Matrix Σ muss dabei natürlich symmetrisch und positiv definit sein):



23.6.12 Beispiel: Zur Bestimmung des Standortes C eines Objekts wird dieses von den beiden c km voneinander entfernten Beobachtungspunkten A und B angepeilt und dabei die beiden Winkel U und V bestimmt. Da Winkelmessungen in der Regel mit Fehlern behaftet sind, können wir annehmen, dass es sich bei diesen beiden Winkeln U und V um Zufallsvariable handelt. Unter der Annahme, dass diese beiden Zufallsvariablen U und V unabhängig und mit den Mittelwerten α und β sowie der Streuung σ normalverteilt sind, bestimme man die gemeinsame Verteilung der Koordinaten X und Y des Punktes C .

▼

Lösung: Wir veranschaulichen die Fragestellung an Hand einer Zeichnung:



Offenbar ist $Y = X \tan[U]$ und $Y = (c - X) \tan[V]$, was

$$X = \frac{c \operatorname{Cos}[U] \operatorname{Sin}[V]}{\operatorname{Sin}[U+V]} \quad \text{und} \quad Y = \frac{c \operatorname{Sin}[U] \operatorname{Sin}[V]}{\operatorname{Sin}[U+V]}$$

zur Folge hat. Da σ (als Maß für die Genauigkeit der Messung) üblicherweise klein ist und damit U und V nur wenig von ihren Mittelwerten α und β abweichen, gilt in erster Näherung (Entwicklung von X und Y in eine Taylorreihe, welche wir nach dem ersten Glied abbrechen - diesen Vorgang nennt man **Linearisierung**)

$$X = \frac{c \operatorname{Cos}[\alpha] \operatorname{Sin}[\beta]}{\operatorname{Sin}[\alpha+\beta]} - \frac{c \operatorname{Sin}[\beta] \operatorname{Cos}[\beta]}{\operatorname{Sin}^2[\alpha+\beta]} (U - \alpha) + \frac{c \operatorname{Sin}[\alpha] \operatorname{Cos}[\alpha]}{\operatorname{Sin}^2[\alpha+\beta]} (V - \beta)$$

$$Y = \frac{c \operatorname{Sin}[\alpha] \operatorname{Sin}[\beta]}{\operatorname{Sin}[\alpha+\beta]} + \frac{c \operatorname{Sin}^2[\beta]}{\operatorname{Sin}^2[\alpha+\beta]} (U - \alpha) + \frac{c \operatorname{Sin}^2[\alpha]}{\operatorname{Sin}^2[\alpha+\beta]} (V - \beta)$$

oder in Matrixschreibweise $\{X, Y\} = \vec{v} + \{U, V\} \cdot \Gamma$ mit

$$\vec{v} = \frac{c \operatorname{Sin}[\beta]}{\operatorname{Sin}[\alpha+\beta]} \{\operatorname{Cos}[\alpha], \operatorname{Sin}[\alpha]\} - \{\alpha, \beta\} \cdot \Gamma$$

$$\Gamma = \frac{c}{\operatorname{Sin}^2[\alpha+\beta]} \begin{pmatrix} -\operatorname{Sin}[\beta] \operatorname{Cos}[\beta] & \operatorname{Sin}^2[\beta] \\ \operatorname{Sin}[\alpha] \operatorname{Cos}[\alpha] & \operatorname{Sin}^2[\alpha] \end{pmatrix}$$

Bei der gemeinsamen Verteilung $\mathbb{P}_{U,V}$ der beiden Zufallsvariablen U und V handelt es sich um eine zweidimensionale Normalverteilung mit dem Mittelwertsvektor $\{\alpha, \beta\}$ und der Kovarianzmatrix $\sigma^2 \mathbf{E}$. Wegen des **Satzes über die affine Transformation** von Multinormalverteilungen handelt es sich daher bei der gemeinsamen Verteilung $\mathbb{P}_{X,Y}$ der beiden Zufallsvariablen X und Y in erster Näherung um eine zweidimensionale Normalverteilung mit dem Mittelwertsvektor $\vec{\mu} = \vec{v} + \{\alpha, \beta\} \cdot \Gamma$ und der Kovarianzmatrix $\Sigma = \sigma^2 \Gamma^t \cdot \Gamma$.

Bisher haben wir stets vorausgesetzt, dass die Kovarianzmatrix Σ einer Multinormalverteilung $\mathcal{MN}[\vec{\mu}, \Sigma]$ symmetrisch und positiv definit ist. Ist Σ zwar symmetrisch aber nur positiv **semidefinit** (und nicht positiv definit), so spricht man von einer **ausgearteten Multinormalverteilung**.

Experimentiert man mit den dynamischen Graphiken aus **Beispiel 17.6.3**, so erkennt man, dass es sich bei einer ausgearteten zweidimensionalen Normalverteilung um eine "eindimensionale Normalverteilung auf einer Geraden" handelt. Ohne auf den Beweis im Detail einzugehen erwähnen wir in diesem Zusammenhang:

23.6.13 Satz: Ist $\Sigma \in \mathbb{R}_r^r$ eine symmetrische, positiv semidefinite Matrix mit Rang $s < r$, sind $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_s$ die normierten und paarweise orthogonalen, zu den s positiven Eigenwerten der Matrix Σ gehörenden Eigenvektoren und bezeichnet $\Gamma = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_s\} \in \mathbb{R}_r^s$ jene Matrix, deren Spalten von diesen s Eigenvektoren gebildet werden, so gilt:

a) Die ausgeartete Multinormalverteilung $\mathcal{MN}[\vec{\mu}, \Sigma]$ besitzt als Träger $\mathbb{T}_{\vec{\mu}, \Sigma}$ die lineare Mannigfaltigkeit

$$\mathbb{T}_{\vec{\mu}, \Sigma} = \{\vec{\mu} + \lambda_1 \vec{v}_1^t + \lambda_2 \vec{v}_2^t + \dots + \lambda_s \vec{v}_s^t \mid \lambda_1, \lambda_2, \dots, \lambda_s \in \mathbb{R}\}$$

b) Genügt der Zufallsvektor \vec{Z} der r -dimensionalen, ausgearteten Multinormalverteilung $\mathcal{MN}[\vec{\mu}, \Sigma]$, so genügt der Zufallsvektor $\vec{Y} = \vec{Z} \cdot \Gamma$ der s -dimensionalen, nicht ausgearteten Multinormalverteilung $\mathcal{MN}[\vec{\mu} \cdot \Gamma, \Gamma^t \cdot \Sigma \cdot \Gamma]$ (man beachte, dass es sich bei der Transformation $\vec{y} = \vec{z} \cdot \Gamma$ um die orthogonale Projektion des \mathbb{R}^r auf den von den Vektoren $\vec{v}_1^t, \vec{v}_2^t, \dots, \vec{v}_s^t$ aufgespannten Unterraum handelt).

Wir veranschaulichen diesen Satz an einem Beispiel:

23.6.14 Beispiel: Der zweidimensionale Zufallsvektor \vec{Z} genüge der ausgearteten Multinormalverteilung

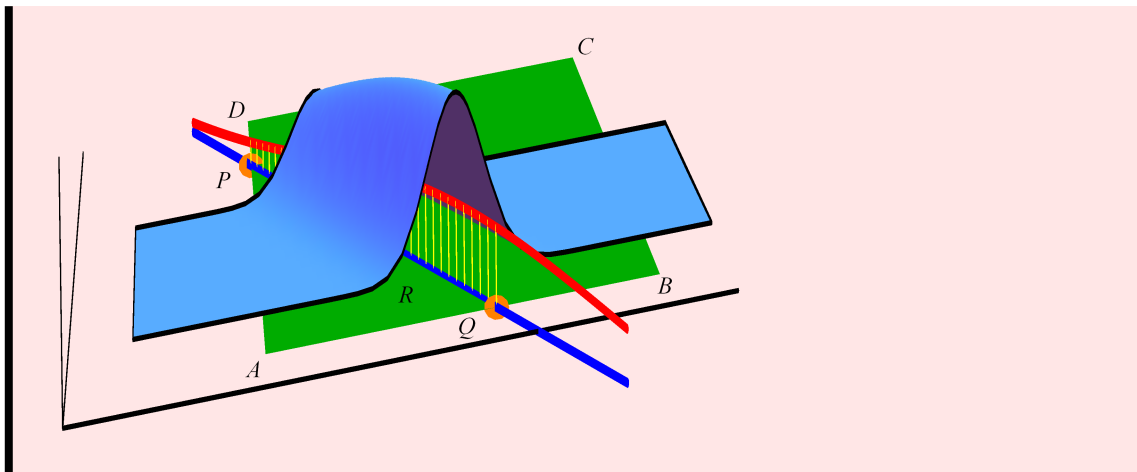
$$\mathcal{MN}[\vec{\mu}, \Sigma] \text{ mit } \vec{\mu} = \{3, 1\} \text{ und } \Sigma = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}$$

Gesucht ist die Wahrscheinlichkeit $\mathbb{P}[\{\vec{Z} \in R\}]$, wobei es sich bei der Menge R um ein Rechteck mit den Eckpunkten $A = \{1.5, 0.5\}$, $B = \{4.5, 0.5\}$, $C = \{4.5, 5.0\}$ und $D = \{1.5, 5.0\}$ handelt.



Lösung: Wir veranschaulichen die Fragestellung zuerst graphisch. In der folgenden Zeichnung ist das Rechteck R mit den Eckpunkten ABCD grün, der Träger der ausgearteten Multinormalverteilung als blaue Gerade und die Verteilungsdichte der Multinormalverteilung mit den Parametern $\vec{\mu} = \{3, 1\}$ und $\Sigma' = \{\{1, -1.95\}, \{-1.95, 4\}\}$ (es handelt sich dabei um eine nichtausgeartete Multinormalverteilung, deren Kovarianzmatrix Σ' mit der Kovarianzmatrix Σ der zur Diskussion stehenden ausgearteten Multinormalverteilung fast übereinstimmt) auszugsweise als blaue Fläche dargestellt. Bei der ausgearteten zweidimensionalen Multinormalverteilung handelt es sich um eine eindimensionale Normalverteilung auf ihrem Träger - ihre Verteilungsdichte ist rot eingezeichnet.

Die gesuchte Wahrscheinlichkeit $\mathbb{P}[\{\vec{Z} \in R\}]$ entspricht damit der gelb schraffierten Fläche unter dieser Verteilungsdichte, welche durch die beiden orangen Punkte P und Q begrenzt ist. Beim Intervall $[P, Q]$ handelt es sich dabei um den Durchschnitt des Rechtecks R mit dem Träger der ausgearteten Multinormalverteilung.



a) Man könnte die gesuchte Wahrscheinlichkeit $\mathbb{P}[\{\vec{Z} \in R\}]$ näherungsweise berechnen, indem man annimmt, dass der Zufallsvektor \vec{Z} einer nichtausgearteten Multinormalverteilung mit dem Mittelwertsvektor $\vec{\mu} = \{3, 1\}$ und einer symmetrischen und positiv definiten Kovarianzmatrix Σ' genügt, welche sich von der Kovarianzmatrix Σ nur geringfügig unterscheidet. Für zwei derartige Kovarianzmatrizen Σ_1 und Σ_2 erhält man beispielsweise

```

μ = {3, 1}; Σ1 = {{1, -1.95}, {-1.95, 4}}; Σ2 = {{1, -2}, {-2, 4.1}};
NIntegrate[PDF[MultinormalDistribution[μ, Σ1], {x, y}], {x, 1.5, 4.5}, {y, 0.5, 5}]
NIntegrate[PDF[MultinormalDistribution[μ, Σ2], {x, y}], {x, 1.5, 4.5}, {y, 0.5, 5}]
Clear[μ, Σ1, Σ2]

```

0.531815

0.530708

b) Wir wollen nun die gesuchte Wahrscheinlichkeit $\mathbb{P}[\{\vec{Z} \in R\}]$ mit Hilfe von [Satz 23.6.13](#) und den dort verwendeten Bezeichnungen berechnen: Dazu ermitteln wir zuerst mit Hilfe von [Eigensystem](#) die Eigenwerte und zugehörigen Eigenvektoren der Matrix Σ sowie die Matrix Γ (es handelt sich dabei um den Eigenvektor zum einzigen positiven Eigenwert der Matrix Σ) und berechnen den Mittelwert $\vec{\mu} \cdot \Gamma$ und die Varianz $\Gamma^t \cdot \Sigma \cdot \Gamma$ der Zufallsvariablen $Y = \vec{Z} \cdot \Gamma$ sowie die Projektionen $\vec{p} \cdot \Gamma$ und $\vec{q} \cdot \Gamma$ der Schnittpunkte P und Q des Rechtecks R mit dem Träger unserer ausgearteten Multinormalverteilung. Die gesuchte Wahrscheinlichkeit $\mathbb{P}[\{\vec{Z} \in R\}]$ entspricht damit der Wahrscheinlichkeit $\mathbb{P}[\{Y \in [\vec{q} \cdot \Gamma, \vec{p} \cdot \Gamma]\}]$ und lässt sich leicht berechnen:

```
 $\mu = \{3, 1\}; \Sigma = \{\{1, -2\}, \{-2, 4\}\}; p = \{1.5, 4\}; q = \{3.25, 0.5\};$   
 $\Gamma = \text{Eigensystem}[\Sigma][[2, 1]];$   
 $\mu.\Gamma; \Gamma.\Sigma.\Gamma;$   
 $p.\Gamma; q.\Gamma;$   
 $\text{CDF}[\text{NormalDistribution}[\mu.\Gamma, \text{Sqrt}[\Gamma.\Sigma.\Gamma]], p.\Gamma] - \text{CDF}[\text{NormalDistribution}[\mu.\Gamma, \text{Sqrt}[\Gamma.\Sigma.\Gamma]], q.\Gamma]$   
 $\text{Clear}[\mu, \Sigma, p, q, \Gamma]$ 
```

```
0.531899
```