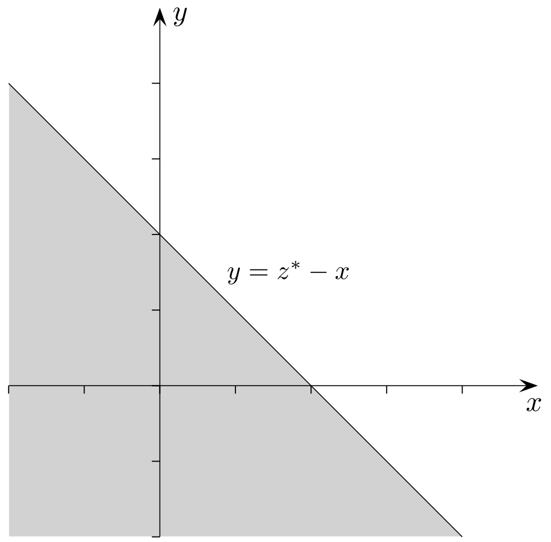
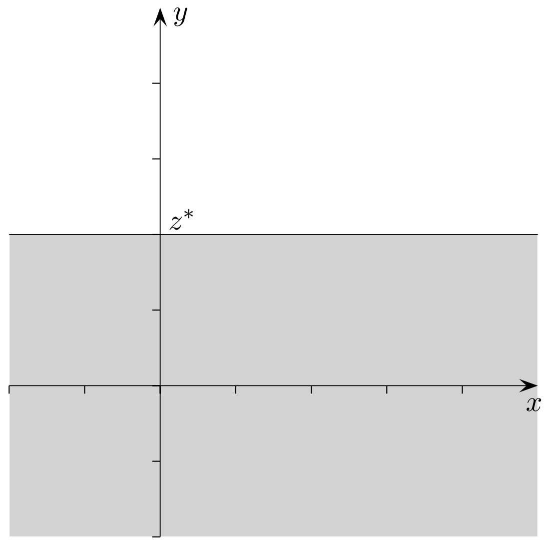
Summe, Produkt, Quotient von Zufallsvariablen

Die Zufallsvariablen X und Y haben die gemeinsame Dichte $f(x,y)$.
Falls sie unabhängig sind, gilt $f(x,y) = f_X(x) \cdot f_Y(y)$.

Untersuchen wir die Verteilungsfunktion der Summe $X+Y$:

$$P(X+Y \leq z^*) \;=\; \int\!\!\int\limits_{x+y\,\leq\,z^*} f(x,y)\;dx\,dy$$





Mit der Transformation $\qquad (x,\, z-x) \quad \longleftarrow \quad (x,\, z) \qquad$ Probe: $x + (z-x) = z$
kann die Integration vereinfacht werden:

$$P(X+Y \leq z^*) \;=\; \int_{-\infty}^{z^*} \underbrace{\int_{-\infty}^{\infty} f(x, z-x)\;dx}_{\text{Dichte von } X+Y}\,dz$$

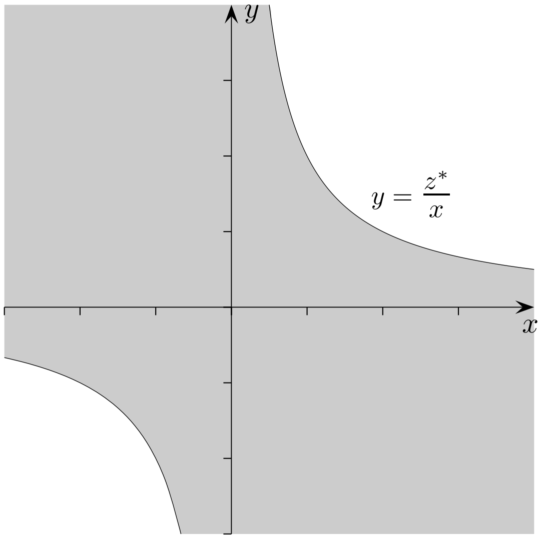
Allgemein gilt für eine Transformation (siehe Verschiedenes, krummlinige Koordinaten)

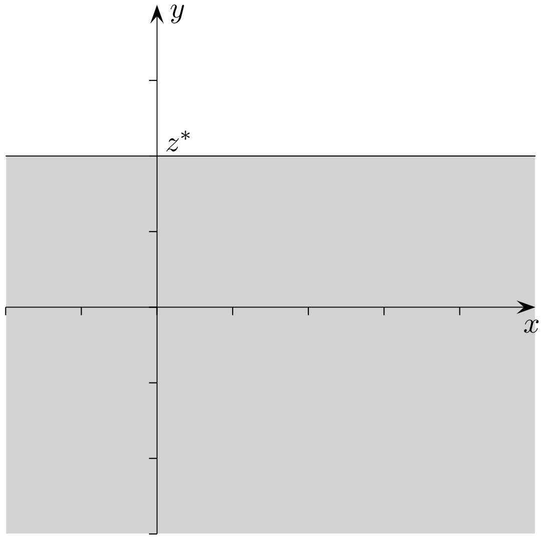
$$(u,v) \longrightarrow \big(\,x(u,v),\; y(u,v)\,\big)$$

$$dA^* = \left|\,\begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}\,\right|\,du\,dv \qquad\qquad \text{und in diesem Fall} \qquad \left|\,\begin{vmatrix} 1 & 0 \\ \dots & 1 \end{vmatrix}\,\right| = 1.$$

Produkt von Zufallsvariablen

$$P(X \cdot Y \leq z^*) \ = \iint\limits_{x \cdot y \leq z^*} f(x,y) \ dx\, dy$$





Die Transformation lautet: $(x, \frac{z}{x}) \quad \longleftarrow \quad (x, z)$ Probe: $x \cdot \frac{z}{x} = z$

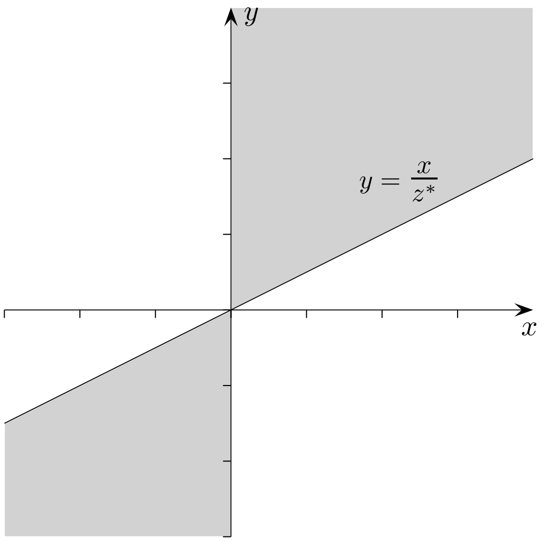
$$P(X \cdot Y \leq z^*) \ = \int_{-\infty}^{z^*} \underbrace{\int_{-\infty}^{\infty} f(x, \frac{z}{x}) \, | \, \frac{1}{x} \, | \ dx}_{\text{Dichte von } X \cdot Y} \, dz$$

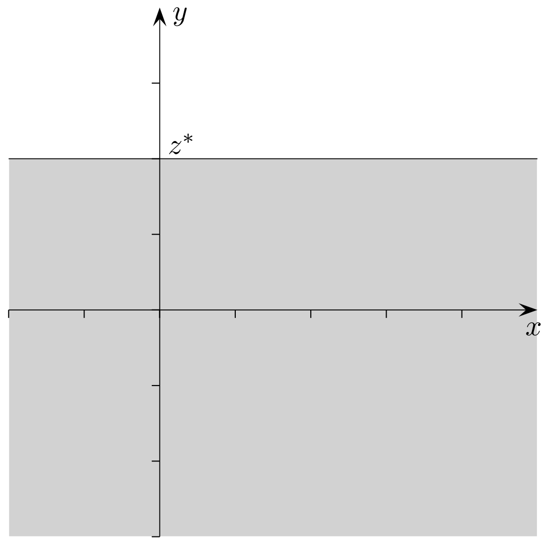
$$(u, v) \longrightarrow \big(\, x(u, v), \ y(u, v)\big)$$

$dA^* = \left| \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} \right| du\, dv$ in diesem Fall $\left| \begin{vmatrix} 1 & 0 \\ \dots & \frac{1}{x} \end{vmatrix} \right| = \left| \, \frac{1}{x} \, \right|$

Quotient von Zufallsvariablen

$$P(\frac{X}{Y} \le z^*) \;=\; \underset{\frac{x}{y}\,\le\, z^*}{\int\int} f(x,y)\;dx\,dy$$





Eine Transformation lautet: $\quad (x, \frac{x}{z}) \;\longleftarrow\; (x, z)$ $\qquad$ Probe: $x : \frac{x}{z} = z$

$$P(\frac{X}{Y} \le z^*) \;=\; \int_{-\infty}^{z^*} \underbrace{\int_{-\infty}^{\infty} f(x, \frac{x}{z})\,|\,\frac{x}{z^2}\,|\;dx}_{\text{Dichte von } \frac{X}{Y}}\;dz$$

$$(u, v) \longrightarrow (\,x(u, v),\; y(u, v))$$

$$dA^* = |\begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}|\;du\,dv \qquad \text{in diesem Fall} \qquad |\begin{vmatrix} 1 & 0 \\ \dots & -\frac{x}{z^2} \end{vmatrix}| = |\,-\frac{x}{z^2}\,|$$

und eine weitere: $\quad (z \cdot x, x) \;\longleftarrow\; (x, z)$

$$P(\frac{X}{Y} \le z^*) \;=\; \int_{-\infty}^{z^*} \underbrace{\int_{-\infty}^{\infty} f(z\cdot x, x)\,|\,x\,|\;dx}_{\text{Dichte von } \frac{X}{Y}}\;dz$$

$$|\begin{vmatrix} z & x \\ 1 & 0 \end{vmatrix}| = |\,x\,|$$

Differenz von Zufallsvariablen

Für die Verteilungsfunktion der Zufallsvariablen $X - Y$ gilt:

$$P(X - Y \leq z^*) \ = \iint\limits_{x-y\,\leq\, z^*} f(x,y)\ dx\,dy$$

Mit der Transformation $\qquad (x, x-z) \quad \longleftarrow \quad (x, z) \qquad$ Probe: $x - (x - z) = z$
kann die Integration vereinfacht werden:

$$P(X - Y \leq z^*) \ = \int_{-\infty}^{z^*} \underbrace{\int_{-\infty}^{\infty} f(x, x-z)\ dx}_{\text{Dichte von } X-Y}\ dz$$

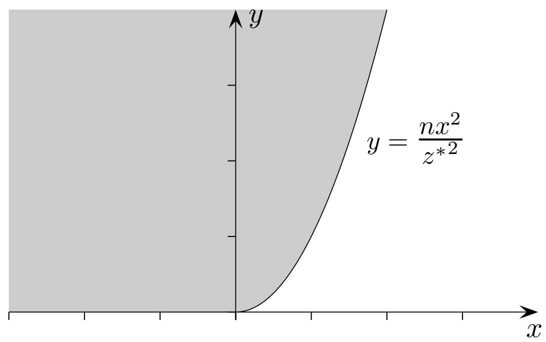
$$(u, v) \longrightarrow (\,x(u,v),\ y(u,v)\,)$$

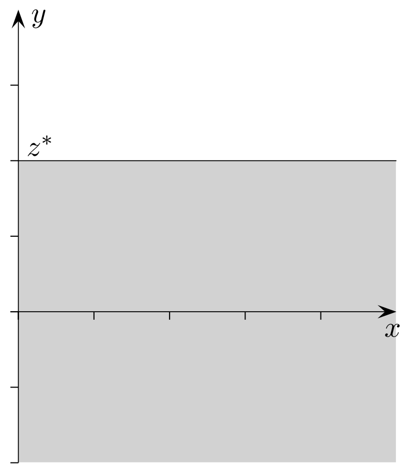
$$dA^* = \left|\,\begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}\,\right| du\,dv \qquad\qquad \left|\,\begin{vmatrix} 1 & 0 \\ \dots & -1 \end{vmatrix}\,\right| = 1.$$

t-Verteilung

X sei $N(0,1)$-verteilt und Y χ^2_n-verteilt.

$$P(\frac{X}{\sqrt{\frac{Y}{n}}} \leq z^*) \;=\; \iint\limits_{\frac{x}{\sqrt{\frac{y}{n}}} \leq z^*} f(x,y)\; dx\, dy$$





Transformation: $\qquad (z\sqrt{\frac{x}{n}}, x) \quad \longleftarrow \quad (x,\, z)$

$$P(\frac{X}{\sqrt{\frac{Y}{n}}} \leq z^*) \;=\; \int_{-\infty}^{z^*} \underbrace{\int_0^{\infty} f(z\sqrt{\frac{x}{n}}, x) \cdot \sqrt{\frac{x}{n}}\; dx}_{\text{Dichte von } \frac{X}{\sqrt{\frac{Y}{n}}}}\; dz$$

$$(u,v) \longrightarrow \big(x(u,v),\; y(u,v)\big)$$

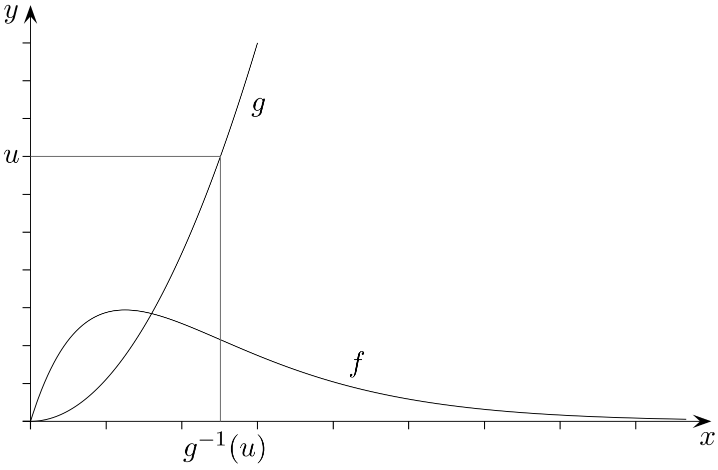
$$dA^* = |\begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}|\; du\, dv \qquad\qquad \text{in diesem Fall} \quad |\begin{vmatrix} \cdots & \sqrt{\frac{x}{n}} \\ 1 & 0 \end{vmatrix}| = \sqrt{\frac{x}{n}}$$

Die Dichte hat die Form: $g_n(x) = \ldots \dfrac{1}{\left(1+\frac{x^2}{n}\right)^{\frac{n+1}{2}}}$ (Konstanten werden mit der Γ-Funktion formuliert).

Verteilung von $g(X)$

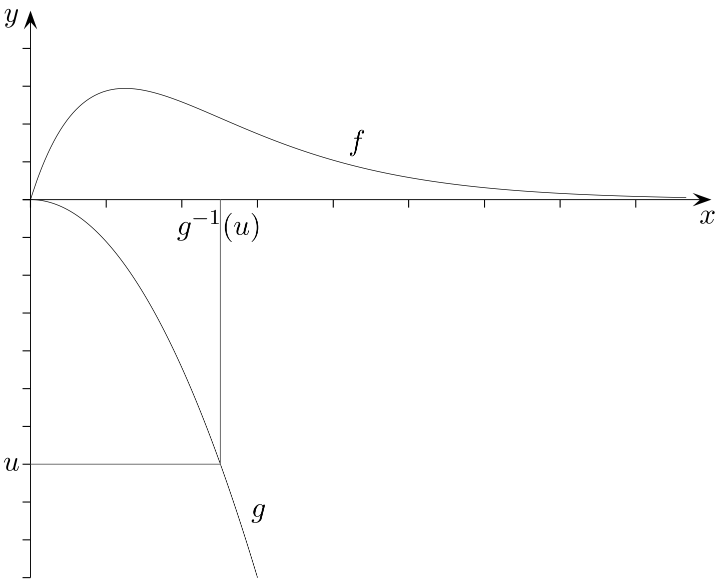
Sei die Dichte von X f, die Verteilungsfunktion F, g umkehrbar.
Wie lautet die Dichte von $g(X)$?

$$\begin{aligned} P(g(X) \le u) &= P(X \le g^{-1}(u)) \\ &= F(g^{-1}(u)) \\ \frac{d}{du} F(g^{-1}(u)) &= f(g^{-1}(u)) \cdot \left(g^{-1}(u)\right)' \end{aligned}$$



Sei nun g monoton fallend.

$$\begin{aligned} P(g(X) \le u) &= 1 - F(g^{-1}(u)) \\ \frac{d}{du}\left(1 - F(g^{-1}(u))\right) &= -f(g^{-1}(u)) \cdot \left(g^{-1}(u)\right)' \\ &= f(g^{-1}(u)) \cdot |\left(g^{-1}(u)\right)'| \end{aligned}$$



Für eine $N(0,1)$-verteilte Zufallsvariable X und $g(x) = \sigma x + \mu$ erhalten wir die Dichte

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

Mit $g(x) = \sqrt{\frac{x}{n}}$ und der Quotientenbildung für Zufallsvariablen ist eine weitere Herleitung der t-Verteilung möglich.

Verteilung der Zufallsvariablen $\sum_{i=1}^{n}(\frac{X_i-\overline{X}}{\sigma})^2$

X sei normalverteilt und $n=2$.

$$\begin{aligned}
\sum_{i=1}^{n}\Big(\frac{X_i-\overline{X}}{\sigma}\Big)^2 &= \tfrac{1}{\sigma^2}\big[(X_1-\overline{X})^2+(X_2-\overline{X})^2\big] \\
&= \tfrac{1}{\sigma^2}\Big[\big(X_1-\tfrac{X_1+X_2}{2}\big)^2+\big(X_2-\tfrac{X_1+X_2}{2}\big)^2\Big] \\
&= \tfrac{1}{\sigma^2}\Big[\big(\tfrac{X_1-X_2}{2}\big)^2+\big(\tfrac{X_1-X_2}{2}\big)^2\Big] \\
&= \tfrac{1}{\sigma^2}\,2\big(\tfrac{X_1-X_2}{2}\big)^2 \\
&= \big(\tfrac{X_1-X_2}{\sqrt{2}\,\sigma}\big)^2
\end{aligned}$$

X_1-X_2 ist $N(0,\sqrt{2}\sigma)$ verteilt (Varianzen addieren sich).

Daher sind $\frac{X_1-X_2}{\sqrt{2}\,\sigma}$ $N(0,1)$- und $\big(\frac{X_1-X_2}{\sqrt{2}\,\sigma}\big)^2$ χ_1-verteilt.

Allgemein gilt:

$$\sum_{i=1}^{n}\Big(\frac{X_i-\overline{X}}{\sigma}\Big)^2 \quad \text{ist } \chi_{n-1}\text{-verteilt.}$$

Der Übergang von μ zu $\overline{X}$ verringert also den Freiheitsgrad um 1.

Beweisskizze:

$$\begin{aligned}
\sum_{i=1}^{n}(X_i-\mu)^2 &= \sum_{i=1}^{n}\big((X_i-\overline{X})+(\overline{X}-\mu)\big)^2 \\
&= \sum_{i=1}^{n}(X_i-\overline{X})^2+2(\overline{X}-\mu)\cdot\underbrace{\sum_{i=1}^{n}(X_i-\overline{X})}_{=0}+n(\overline{X}-\mu)^2
\end{aligned}$$

$$\underbrace{\sum_{i=1}^{n}\Big(\frac{X_i-\mu}{\sigma}\Big)^2}_{\chi^2_n\text{-verteilt}} = \sum_{i=1}^{n}\Big(\frac{X_i-\overline{X}}{\sigma}\Big)^2+\underbrace{\Big(\frac{\sqrt{n}\,(\overline{X}-\mu)}{\sigma}\Big)^2}_{\chi^2_1\text{-verteilt}}$$

Aus einer Unabhängigkeitsüberlegung folgt, dass die fehlende Verteilung eindeutig bestimmt ist.
Daher muss es die χ^2_{n-1}-Verteilung sein.

t-Test

Um für eine normalverteilte Zufallsvariable X ein Schwankungsintervall für den Mittelwert μ zu ermitteln, wird die $N(0,1)$-Verteilung von $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ herangezogen.
Die Standardabweichung σ muss bekannt sein.
Falls dies nicht der Fall ist, kann σ mit der Stichprobenvarianz

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$$

geschätzt werden.

s für σ eingesetzt, ergibt die Zufallsvariable:

$$T = \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} = \frac{(\overline{X}-\mu)\sqrt{n}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}}$$

Für einen kleinen Stichprobenumfang n ist T nicht normalverteilt,
sondern unterliegt der sogenannten t-Verteilung. Dies konnte Gosset 1908 nachweisen.

Hierzu wird der Bruch mit $\frac{1}{\sigma}$ erweitert und etwas umgeformt:

$$T = \frac{\frac{(\overline{X}-\mu)}{\sigma}\sqrt{n}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\frac{(X_i - \overline{X})^2}{\sigma^2}}} = \frac{\frac{\frac{(\overline{X}-\mu)}{\sigma}}{\sqrt{n}}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma}\right)^2}}$$

Nun ist zu erkennen, dass T vom Typ $\frac{Y}{\sqrt{\frac{Z}{n-1}}}$ ist, wobei Y $N(0,1)$- und Z χ^2_{n-1}-verteilt ist.

T ist daher t-verteilt mit $n-1$ Freiheitsgraden.

Es liege das Stichprobenergebnis X_1, X_2, …, X_n vor.
Dann gilt (Sicherheitswahrscheinlichkeit α):

$$-c \leq \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \leq c$$

$$\iff \quad \overline{X} - c\cdot\frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + c\cdot\frac{s}{\sqrt{n}}$$

c wird mit $P(T \leq c) = \frac{1+\alpha}{2}$ bestimmt (analog zur Normalverteilung), $n-1$ Freiheitsgrade.