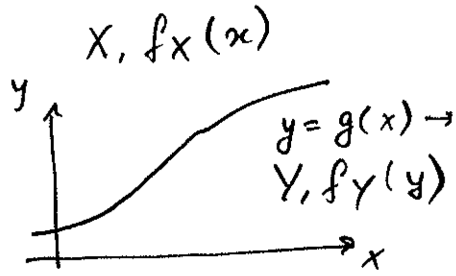
Kapitel 4 Funktionen von Zufallsvariablen

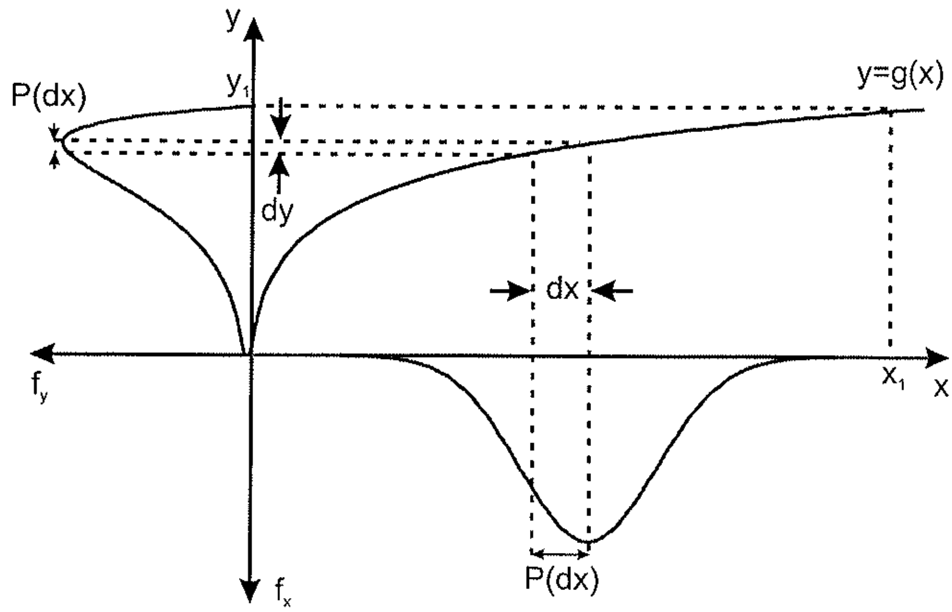


In vielen praktischen Anwendungen treten Situationen auf, in denen eine Zufallsvariable X mit bekanntem wahrscheinlichkeitstheoretischem Verhalten als Argument einer monoton wachsenden oder fallenden Funktion $g(X)$ auftritt. Hierdurch entsteht eine neue Zufallsvariable $Y = g(X)$. Ein einfaches Beispiel für eine solche Situation in einer technischen Anwendung ist in der Beschreibung einer monoton wachsenden Verstärkerkennlinie gegeben, die durch die Funktion $g(.)$ analytisch beschrieben wird. Eine zufällig gewählte Amplitude, mit bekannter Verteilungsfunktion, wird auf diese Verstärkerkennlinie angewandt. Dabei wird die Frage nach dem wahrscheinlichkeitstheoretischen Gesetz der neuen Zufallsvariablen Y gestellt. Diese Frage kann mit Angabe der Verteilungs- oder Wahrscheinlichkeitsdichtefunktion für die Zufallsvariable Y direkt und vollständig beantwortet werden.

Bestimmung der Wahrscheinlichkeitsdichtefunktion durch Transformation

Gegeben sei eine beispielsweise monoton wachsende Funktion $y = g(x)$. Die Wahrscheinlichkeit, mit der die Zufallsvariable X in einem beliebigen Intervall der Breite dx auftritt, muß gleich der Wahrscheinlichkeit sein, mit der die neue Zufallsvariable Y im zugehörigen Intervall dy Zufallswerte annimmt

$$f_X(x)|dx| = f_Y(y)|dy|.$$



Transf. der Dichtefunk.

Wir betrachten die zur Funktion $y = g(x)$ zugehörige eindeutig bekannte Umkehrfunktion $x = h(y)$. Es gilt

$$h(y) = g^{-1}(y) = x$$

und

$$\frac{\mathrm{d}x}{\mathrm{d}y} = \frac{\mathrm{d}h(y)}{\mathrm{d}y}$$

entsteht folgende mathematische Beziehung zwischen den Wahrscheinlichkeitsdichtefunktionen der Zufallsvariablen X und Y:

$$\boxed{f_Y(y) = f_X(x = h(y))\left|\frac{\mathrm{d}h(y)}{\mathrm{d}y}\right|.}$$

Aus der Sicht der monoton wachsenden Funktion $g(x)$ erhält man alternativ:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}g(x)}{\mathrm{d}x} = g'(x)$$

$$\boxed{f_Y(y) = \left.\frac{f_X(x)}{|g'(x)|}\right|_{x=h(y)}.}$$

Durch die Angabe des Betragszeichens wird gleichzeitig sowohl der Fall monoton steigender als auch monoton fallender Funktionen $g(x)$ berücksichtigt. Das Vorzeichen der Steigung ist für die Angabe der Wahrscheinlichkeit unbedeutend.

Beispiel 4.1 (Funktionen von Zufallsvariablen)
Sei X eine gleichverteilte Zufallsvariable im Intervall $[0, 1]$

$$f_X(x) = \begin{cases} 0 & \text{für } x < 0 \vee x > 1 \\ 1 & \text{für } 0 < x < 1 \end{cases}$$

und $g(x) = \sinh(x)$ die betrachtete monoton wachsende Funktion. Dann ist die Ableitung dieser Funktion $g(x)$

$$g'(x) = \cosh(x).$$

Unter Berücksichtigung der allgemein gültigen Beziehung

$$\cosh^2(x) - \sinh^2(x) = 1$$

kann die gesuchte Wahrscheinlichkeitsdichtefunktion $f_Y(y)$ der neuen Zufallsvariablen $Y = \sinh(x)$ wie folgt berechnet werden:

$$f_Y(y) = \begin{cases} 0 & \text{für } y < 0 \vee y > \sinh(1) \\ \frac{1}{\cosh(x)} = \frac{1}{\sqrt{1+y^2}} & \text{für } 0 < y < \sinh(1) \end{cases}$$

Verteilungsfunktion der Zufallsvariablen $Y = g(X)$

Bisher haben wir uns auf die Berechnung der Wahrscheinlichkeitsdichtefunktion $f_Y(y)$ der neuen Zufallsvariablen konzentriert. In diesem Abschnitt wollen wir zusätzlich, oder alternativ, die zugehörige Verteilungsfunktion berechnen. Für eine beispielsweise *monoton wachsende* Funktionen $g(x)$ folgt aus $Y \leq y$ und damit $X \leq h(y) = g^{-1}(y)$ für die Verteilungsfunktion von Y

$$F_Y(y) = F_X(h(y)).$$

Für eine *monoton fallende* Funktionen $g(x)$ folgt aus $Y \leq y$ und damit $X > h(y) = g^{-1}(y)$ für die Verteilungsfunktion

$$F_Y(y) = 1 - F_X(h(y)).$$

Beispiel 4.3 (Monoton fallende Transformation)

Sei die Transformation der Zufallsvariablen

$$y = g(x) = ax$$

und die zugehörige Umkehrfunktion $h(y)$

$$h(y) = \frac{1}{a}y,$$

dann ist die Wahrscheinlichkeitsdichtefunktion der Zufallsvariablen Y

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right).$$

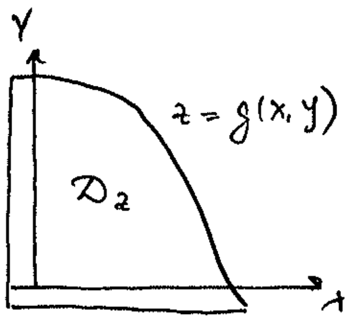
Die zugehörige Verteilungsfunktion der neuen Zufallsvariablen Y kann mit den obigen Vorbereitungen relativ einfach ermittelt werden:

$$F_Y(y) = \begin{cases} F_X\left(\frac{y}{a}\right) & \text{für } a > 0 \\ 1 - F_X\left(\frac{y}{a}\right) & \text{für } a < 0 \end{cases}.$$

Funktionen angewandt auf zwei Zufallsvariable

Um die Wahrscheinlichkeitsdichtefunktion einer Zufallsvariablen Z mit

$$Z = g(X, Y)$$



bestimmen zu können, ermittelt man die Verteilungsfunktion $F_Z(z)$ der Zufallsvariablen Z

$$F_Z(z) = P(Z \leq z) = P(g(X, Y) \leq z).$$

Hierdurch wird eine Fläche in der (x, y)-Ebene repräsentiert, welche durch die Kurve $g(x, y) = z$ begrenzt wird

Man erhält $F_Z(z)$ durch Integration der Verbundwahrscheinlichkeitsdichte von X und Y über dieser Fläche R

$$F_Z(z) = \iint_R f_{(X,Y)}(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

und die Wahrscheinlichkeitsdichtefunktion von Z per Definition durch Differentiation von $F_Z(z)$

$$f_Z(z) = \frac{\mathrm{d}F_Z(z)}{\mathrm{d}z}.$$

Beispiel 4.4 (Verteilungsdichte von $Z = X^2 + Y^2$)

Die Zufallsvariablen X und Y seien statistisch unabhängig und identisch normalverteilt, so dass die gemeinsame Wahrscheinlichkeitsdichte durch die folgende Funktion angegeben wird.

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

Durch den Wechsel des Koordinatensystems zu Polarkoordinaten mit

$$x = r\cos(\theta), \quad y = r\sin(\theta), \quad \mathrm{d}x\,\mathrm{d}y = r\mathrm{d}r\,\mathrm{d}\theta$$

erhält man für die Verteilungsfunktion der Zufallsvariablen Z:

$$\begin{aligned} F_Z(z) &= \frac{1}{2\pi\sigma^2} \iint_R \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \mathrm{d}x\,\mathrm{d}y \\ &= \frac{1}{2\pi\sigma^2} \int_0^{\sqrt{z}} \int_0^{2\pi} \exp\left(-\frac{r^2}{2\sigma^2}\right) r\,\mathrm{d}\theta\,\mathrm{d}r \\ &= 1 - \exp\left(-\frac{z}{2\sigma^2}\right) \quad \forall z \geq 0 \end{aligned}$$

$f_Z(z) = F_Z'(z)$
$= \frac{1}{2\sigma^2} e^{-\frac{z}{2\sigma^2}}, \; z \geq 0$
Exponentialverteilung

Festhalten einer Zufallsvariablen.

Eine weitere und alternative Möglichkeit zur Ermittelung der Wahrscheinlichkeitsdichte der Zufallsvariablen Z erhält man durch „Festhalten" einer der beiden Zufallsvariablen X oder Y. Im folgenden Beispiel wird die Zufallsvariable Y beispielsweise als Konstante betrachtet. Es wird die zweidimensionale, von den Parametern x und y abhängige Funktion $Z = g(x, y)$ betrachtet. Wenn der Wert y zunächst als konstant betrachtet wird, dann ist die Funktion $Z = g(x, y)$ nur noch von x aber nicht mehr vom Parameter y abhängig. Dieser Sachverhalt kann analytisch wie folgt ausgenutzt werden:

$$f_Z(z|y) = \left.\frac{f_X(x|y)}{\left|\frac{\partial g(x,y)}{\partial x}\right|}\right|_{x=g^{-1}(z,y)}$$

Aus der gemeinsamen Wahrscheinlichkeitsdichte von Y und Z kann man mit der bedingten Wahrscheinlichkeitsdichte dann $f_Z(z)$ berechnen:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_{(Z,Y)}(z, y)\,\mathrm{d}y \\ &= \int_{-\infty}^{\infty} f_Z(z|y) f_Y(y)\,\mathrm{d}y \end{aligned}$$

Beispiel 4.5 (Produkt von Zufallsvariablen)

Durch Festhalten von Y in der Beziehung

$$Z = g(x, y) = x \cdot y$$

wird die durch Y bedingte Wahrscheinlichkeitsdichte von Z

$$f_Z(z|y) = \frac{f_X\left(\frac{z}{y}|y\right)}{|y|}$$

und die Wahrscheinlichkeitsdichte von Z

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} \frac{f_X\left(\frac{z}{y}|y\right)}{|y|} f_Y(y)\,\mathrm{d}y \\ &= \int_{-\infty}^{\infty} \frac{f_{(X,Y)}\left(\frac{z}{y}, y\right)}{|y|}\,\mathrm{d}y \end{aligned}$$

Transformation von Zufallsvektoren.

Insbesondere bei einer Variablentransformation tritt das Problem auf, zwei Zufallsvariablen X_1 und X_2 in zwei neue Zufallsvariablen Z_1 und Z_2 zu transformieren (z.B. Real- und Imaginärteil in Betrag und Phase).

$$\begin{aligned} z_1 &= g_1(x_1, x_2) \\ z_2 &= g_2(x_1, x_2) \end{aligned}$$

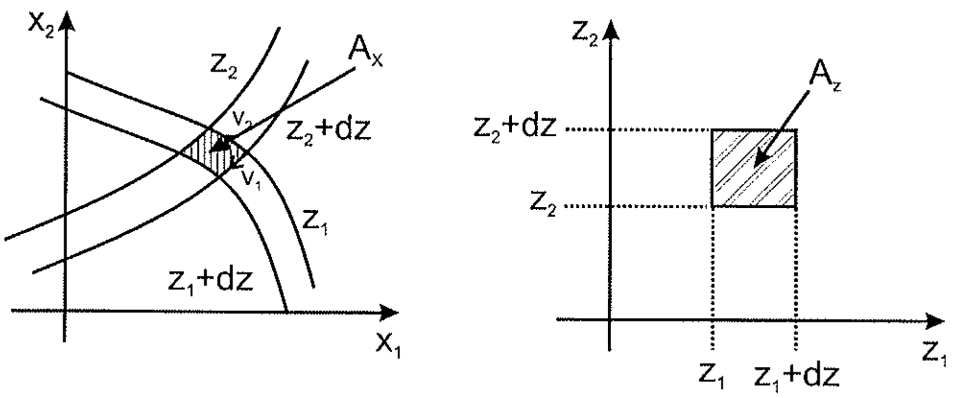
Mit dem gleichen Ansatz wie im Falle einer Zufallsvariablen, nach dem die Wahrscheinlichkeit einer transformierten Fläche gleich der Wahrscheinlichkeit der Originalfläche sein muss

$$f_{(Z_1,Z_2)}(z_1, z_2)\, dz_1 dz_2 = f_{(X_1,X_2)}(x_1, x_2)\, \mathrm{d}x_1 \mathrm{d}x_2$$

bzw.

$$f_{(Z_1,Z_2)}(z_1, z_2)\, |A_Z| = f_{(X_1,X_2)}(x_1, x_2)\, |A_X|$$

ist also lediglich das Verhältnis der Flächen A_X und A_Z zu bestimmen. (Zur Herleitung geht man zunächst von der vereinfachenden Annahme aus, dass eindeutige Umkehrfunktionen $g_1^{-1}(z_1, z_2)$ und $g_2^{-1}(z_1, z_2)$ existieren. Falls diese Voraussetzung nicht erfüllt ist, so sind – wie aus den Betrachtungen für eine einzelne Zufallsvariable bekannt – die einzelnen Teilwahrscheinlichkeiten zu addieren.)



. . Variablentransformation

Wird die Fläche A_X durch die zwei Vektoren v_1 und v_2 aufgespannt und ist β der eingeschlossene Winkel, so ist die Fläche des Parallelograms

$$\begin{aligned} |v_1||v_2|\sin(\beta) &= \left(|v_1|^2|v_2|^2(1-\cos^2(\beta))\right)^{1/2} \\ &= \left(v_1^2 v_2^2 - v_1^2 v_2^2 \cos^2(\beta)\right)^{1/2} \\ &= \left(v_1^2 v_2^2 - (v_1 \cdot v_2)^2\right)^{1/2} \\ &= \left((v_{11}^2 + v_{12}^2)(v_{21}^2 + v_{22}^2) - (v_{11}v_{12} + v_{21}v_{22})^2\right)^{1/2} \\ &= |v_{11}v_{22} - v_{12}v_{21}| = \left|\det \begin{pmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{pmatrix}\right| \end{aligned}$$

Die Länge der Vektoren ermittelt man aus dem Anfangspunkt

$$(x_1, x_2) = (g_1^{-1}(z_1, z_2), g_2^{-1}(z_1, z_2))$$

und dem Endpunkt

$$(g_1^{-1}(z_1 + dz_1, z_2), g_2^{-1}(z_1 + dz_1, z_2)) = \left(x_1 + \frac{\partial g_1^{-1}}{\partial z_1} dz_1, x_2 + \frac{\partial g_2^{-1}}{\partial z_1} dz_1\right)$$

Die Vektoren können also ausgedrückt werden als

$$v_1 = \left(\frac{\partial g_1^{-1}}{\partial z_1} dz_1, \frac{\partial g_2^{-1}}{\partial z_1} dz_1\right)$$

und

$$v_2 = \left(\frac{\partial g_1^{-1}}{\partial z_2} dz_2, \frac{\partial g_2^{-1}}{\partial z_2} dz_2\right).$$

Die Fläche A_X ist damit

$$\begin{aligned} A_X &= \left|\det \begin{pmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{pmatrix}\right| \\ &= \left|\det \begin{pmatrix} \frac{\partial g_1^{-1}}{\partial z_1} dz_1 & \frac{\partial g_1^{-1}}{\partial z_2} dz_2 \\ \frac{\partial g_2^{-1}}{\partial z_1} dz_1 & \frac{\partial g_2^{-1}}{\partial z_2} dz_2 \end{pmatrix}\right| \\ &= \left|J \begin{pmatrix} x_1 & x_2 \\ z_1 & z_2 \end{pmatrix}\right| |dz_1\, dz_2| \end{aligned}$$

Mit

$$|A_Z| = |dz_1\, dz_2|$$

kann man schließlich die gemeinsame Wahrscheinlichkeitsdichtefunktion von Z_1 und Z_2 durch die von X_1 und X_2 gegebene ausdrücken:

$$\begin{aligned} f_{(Z_1,Z_2)}(z_1, z_2) &= f_{(X_1,X_2)}(x_1, x_2) \left|J \begin{pmatrix} x_1 & x_2 \\ z_1 & z_2 \end{pmatrix}\right| \\ &= f_{(X_1,X_2)}(x_1, x_2) \cdot \frac{1}{\left|J \begin{pmatrix} z_1 & z_2 \\ x_1 & x_2 \end{pmatrix}\right|} \end{aligned}$$

an der Stelle $x_1 = g_1^{-1}(z_1, z_2), x_2 = g_2^{-1}(z_1, z_2)$.

J - wird als Jacobi-Determinante bezeichnet

Beispiel 4.6 (Einfache Transformation)

Es sei die Transformation

$$\begin{aligned} z_1 &= a_{11}x_1 + a_{12}x_2 \\ z_2 &= a_{21}x_1 + a_{22}x_2 \end{aligned}$$

gegeben. Die Umkehrabbildung $(x_1, x_2) = g^{-1}(z_1, z_2)$ ist

$$\begin{aligned} x_1 &= b_{11}z_1 + b_{12}z_2 \\ x_2 &= b_{21}z_1 + b_{22}z_2 \end{aligned}$$

Die Jacobi-Determinante ist

$$J \begin{pmatrix} x_1 & x_2 \\ z_1 & z_2 \end{pmatrix} = \det \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

so dass die gemeinsame Verteilungsdichte von Z_1 und Z_2 durch

$$f_{(Z_1,Z_2)}(z_1, z_2) = f_{(X_1,X_2)}(b_{11}z_1 + b_{12}z_2, b_{21}z_1 + b_{22}z_2) J \begin{pmatrix} x_1 & x_2 \\ z_1 & z_2 \end{pmatrix}$$

gegeben ist.

Beispiel 4.7 (Etwas schwierigere Transformation)

Die Transformation der Zufallsvariablen X und Y in die Zufallsvariablen Z und W sei gegeben durch

$$z = xy, \quad w = \frac{y}{x}.$$

Die Umkehrabbildung ist definiert durch

$$x = \pm\sqrt{\frac{z}{w}}, \quad y = \pm\sqrt{zw}.$$

Die Jacobi-Determinante ist

$$J\begin{pmatrix} z & w \\ x & y \end{pmatrix} = \det\begin{pmatrix} y & x \\ -\frac{y}{x^2} & \frac{1}{x} \end{pmatrix} = \frac{2y}{x} = 2w.$$

Die Punkte $(\sqrt{z/w}, \sqrt{zw})$ und $(-\sqrt{z/w}, -\sqrt{zw})$ werden beide in den Punkt (z, w) transformiert, d.h. die entsprechenden Wahrscheinlichkeiten sind zu addieren!

Außerdem sind Z und W beide positiv (negativ) falls X und Y das gleiche (verschiedene) Vorzeichen besitzen, so dass

$$f_{(Z,W)}(z, w) = 0 \quad \forall (z > 0, w < 0) \,\text{oder}\, (z < 0, w > 0).$$

Die transformierte Wahrscheinlichkeitsdichtefunktion ist hiermit (z und w haben gleiches Vorzeichen)

$$f_{(Z,W)}(z, w) = \frac{1}{2|w|}\left[f_{(X,Y)}\left(\sqrt{\frac{z}{w}}, \epsilon\sqrt{zw}\right) + f_{(X,Y)}\left(-\sqrt{\frac{z}{w}}, -\epsilon\sqrt{zw}\right)\right]$$

mit

$$\epsilon = \begin{cases} 1 & \text{für } z > 0 \land w > 0 \\ -1 & \text{für } z < 0 \land w < 0 \end{cases}$$

Geordnete Statistik.

In vielen Anwendungsfällen wird die Frage nach einer aufsteigenden oder absteigenden Reihenfolge von Zufallsvariablen gestellt. Zu dieser Rangfolge von Zufallsvariablen, die beispielsweise aus einer Menge von statistisch unabhängig und identisch verteilten Zufallsvariablen gebildet wurde, soll die Wahrscheinlichkeitsdichtefunktion eines definierten Ranges[1] gebildet werden.

Als Rang wird in diesem Zusammenhang der Wert der Zufallsvariablen an einer definierten Position innerhalb der Reihenfolge verstanden.

Als Ausgangspunkt der Analyse wird eine Situation mit insgesamt N statistisch unabhängigen identisch verteilten reellwertigen Zufallsvariablen X_i betrachtet, $i = 1, 2, \ldots, N$. Die Wahrscheinlichkeitsdichtefunktion der Zufallsvariablen sei $f(x)$ und die Verteilungsfunktion sei $F(x)$. Diese Zufallsvariablen werden jetzt der Größe nach aufsteigend sortiert. (Die in Klammern gesetzten Indizes weisen auf die Reihenfolge hin.)

$$X_{(1)} \leq X_{(2)} \leq X_{(3)} \leq \cdots \leq X_{(N)}$$

. Die WDF für die an der k-ten Position beobachtete Zufallsvariable $X_{(k)}$ berechnet sich wie folgt:

$$f_k(x) = k\binom{N}{k}(1 - F(x))^{N-k}F(x)^{k-1}f(x).$$

Die Verteilungsfunktion ist in der folgenden Gleichung für jeden individuellen Rang k angegeben.

$$F_k(x) = \sum_{j=k}^{N} \binom{N}{j} \cdot [F(x)]^j \cdot [1 - F(x)]^{N-j}$$

Insbesondere gilt für die WDF bzw. für die Verteilungsfunktion des Minimums

$$\begin{aligned} f_1(x) &= N \cdot [1 - F(x)]^{N-1} f(x) \\ F_1(x) &= 1 - [1 - F(x)]^N \end{aligned}$$

und des Maximums

$$\begin{aligned} f_N(x) &= NF(x)^{N-1} f(x) \\ F_N(x) &= F(x)^N. \end{aligned}$$

Beispiel 4.8 (Minimum von N exponentialverteilten Zufallsvariablen)

Gegeben seien N statistisch unabhängige identisch verteilte Realisierungen einer exponentialverteilten Zufallsvariable X. Es soll die Wahrscheinlichkeitsdichtefunktion des Minimums dieser N Zufallsvariablen angegeben werden.

Für die einzelnen Zufallsvariablen gilt

$$\begin{aligned} f(x) &= \lambda e^{-\lambda x} \\ F(x) &= 1 - e^{-\lambda x}. \end{aligned}$$

Allgemein gilt in diesem Fall für die WDF des k-ten Ranges nach Gleichung

$$f_k(x) = k\lambda \binom{N}{k} e^{-(N-k+1)\lambda x} \left(1 - e^{-\lambda x}\right)^{k-1}.$$

Die WDF des Minimums, also der Zufallsvariablen des Ranges 1 der geordneten Statistik, berechnet sich nach Gleichung zu

$$\begin{aligned} f_1(x) &= N\left[1 - (1 - e^{-\lambda x})\right]^{N-1} \cdot \lambda e^{-\lambda x} \\ &= N\lambda e^{-N\lambda x}. \end{aligned}$$

Das Minimum gehorcht also einer Exponentialverteilung mit dem Parameter $N\lambda$.

Geordnete Statistiken werden beispielsweise in den Fällen eingesetzt, in denen der Erwartungswert einer Zufallsvariablen geschätzt werden soll, aber in den gemessenen Werten mit einigen Ausreißern gerechnet werden muss. Zur Schätzung des Erwartungswertes wird

Beispiel 4.9 (Robustheit des Medians gegenüber Ausreißern)

Gegeben seien die folgenden 11 Messwerte:

5.0	4.2	4.6	4.5	4.9	4.8	4.5	4.0	4.8	4.4	4.6

Der arithmetische Mittelwert berechnet sich in diesem Fall zu 4.6 und der Median lässt sich mit Hilfe der geordneten Stichprobe

4.0	4.2	4.4	4.5	4.5	4.6	4.6	4.8	4.8	4.9	5.0

ebenfalls zu 4.6 bestimmen (Wert auf Rang 6). Nimmt man jetzt an, dass als dritter Messwert statt des Wertes 4.6 durch eine Störung ein Wert von 20.0 gemessen wurde, so ergibt sich für das arithmetische Mittel der Wert 6.0. Es wird durch den Ausreißer also stark verfälscht. Für den mit Hilfe der geordneten Statistik berechneten Median ergibt sich jedoch immer noch der Wert 4.6 – der Median ist *robust* gegenüber Ausreißern.