Wahrscheinlichkeitstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2022/23

Prof. Dr. Dirk Ostwald

(4) Zufallsvariablen

Repräsentation zentraler Eigenschaften

Modellierung

Modell

Wahrscheinlichkeitstheorie

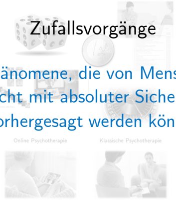
Probabilistisches Modell

$(\Omega, \mathcal{A}, \mathbb{P}), \xi : \Omega \to \mathbb{R}$

Wahrscheinlichkeitsrechnung

$\mathbb{P}_\xi(S) = \mathbb{P}(\xi^{-1}(S))$

Realität



Zufallsvorgänge

Phänomene, die von Menschen nicht mit absoluter Sicherheit vorhergesagt werden können.

Vorhersagen

Quantifizierung von Unsicherheit

Jedes Augenzahlpaar kommt im Mittel gleich häufig vor.

Basierend auf der Physik sollte jedes Augenzahlpaar die gleiche Wahrscheinlichkeit haben.

Modellierung

Modell

Wahrscheinlichkeitstheorie

$\Omega \coloneqq \{(r,b) | r \in \mathbb{N}_6, b \in \mathbb{N}_6\}$

$\mathcal{A} \coloneqq \mathcal{P}(\Omega)$ $\mathbb{P}(\{\omega\}) \coloneqq 1/36$

$\xi : \Omega \to \mathbb{R}, \xi(r,b) \coloneqq r + b$

Realität

Zufallsvorgang

Vorhersagen

Die Summe der Augenzahlen ist eine Zufallsvariable mit Verteilung $\mathbb{P}_\xi$.

Konstruktion, Definition, Notation, Intuition

Wahrscheinlichkeitsmassefunktionen

Wahrscheinlichkeitsdichtefunktionen

Kumulative Verteilungsfunktionen

Selbstkontrollfragen

Konstruktion, Definition, Notation, Intuition

Wahrscheinlichkeitsmassefunktionen

Wahrscheinlichkeitsdichtefunktionen

Kumulative Verteilungsfunktionen

Selbstkontrollfragen

Konstruktion, Definition, Notation, Intuition

Konstruktion von Zufallsvariablen und Verteilungen

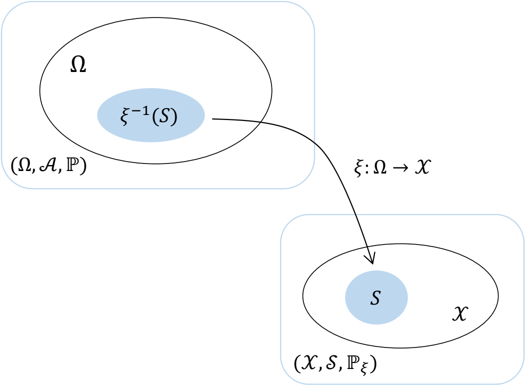
- Es seien $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $\xi : \Omega \to \mathcal{X}$ eine Abbildung.
- ξ ist das kleine griechische Xi.
- Es sei $\mathcal{S}$ eine σ-Algebra auf $\mathcal{X}$.
- Für jedes $S \in \mathcal{S}$ sei das *Urbild von* S definiert als

$$\xi^{-1}(S) := \{\omega \in \Omega | \xi(\omega) \in S\}. \tag{1}$$

- Wenn $\xi^{-1}(S) \in \mathcal{A}$ für alle $S \in \mathcal{S}$ gilt, dann heißt ξ *messbar*.
- $\xi : \Omega \to \mathcal{X}$ sei messbar. Allen $S \in \mathcal{S}$ kann die Wahrscheinlichkeit

$$\mathbb{P}_\xi : \mathcal{S} \to [0,1], S \mapsto \mathbb{P}_\xi(S) := \mathbb{P}\left(\xi^{-1}(S)\right) = \mathbb{P}\left(\{\omega \in \Omega | \xi(\omega) \in S\}\right) \tag{2}$$

  zugeordnet werden.
- ξ heißt nun *Zufallsvariable* und $\mathbb{P}_\xi$ heißt *Bildmaß* oder *Verteilung von* ξ.
- $(\mathcal{X}, \mathcal{S}, \mathbb{P}_\xi)$ ist ein Wahrscheinlichkeitsraum.
- Mit $\mathcal{X} := \mathbb{R}$ und $\mathcal{S} := \mathcal{B}(\mathbb{R})$ rückt der Wahrscheinlichkeitsraum $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_\xi)$ ins Zentrum.



$$\mathbb{P}\left(\xi^{-1}(S)\right) = \mathbb{P}(\{\omega \in \Omega | \xi(\omega) \in S\}) =: \mathbb{P}_\xi(S)$$

Definition (Zufallsvariable)

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $(\mathcal{X}, \mathcal{S})$ ein *Messraum*. Dann ist eine *Zufallsvariable (ZV)* definiert als eine Abbildung $\xi : \Omega \to \mathcal{X}$ mit der *Messbarkeitseigenschaft*

$$\{\omega \in \Omega | \xi(\omega) \in S\} \in \mathcal{A} \text{ für alle } S \in \mathcal{S}. \tag{3}$$

Bemerkungen

- ZVen sind weder "zufällig" noch "Variablen".
- Intuitiv wird $\omega \in \Omega$ "zufällig" anhand von $\mathbb{P}$ gezogen und $\xi(\omega)$ realisiert.
- Wir nennen $\mathcal{X}$ den *Ergebnisraum der ZV* ξ.
- Die Verteilungen (Bildmaße) von ZVen sind in der Statistik zentral.
- Der Begriff der Verteilung wird oft auch für W-Maße und Dichten verwendet.

Konstruktion, Definition, Notation, Intuition

Beispiel (Summe eines roten und eines blauen Würfels)

- Für das Werfen zweier Würfel ist ein sinnvolles Wahrscheinlichkeitsraum-Modell
  - $\Omega := \{(r,b) | r \in \mathbb{N}_6, b \in \mathbb{N}_6\}$
  - $\mathcal{A} := \mathcal{P}(\Omega)$.
  - $\mathbb{P} : \mathcal{A} \to [0,1]$ mit $\mathbb{P}(\{(r,b)\}) = 1/36$ für alle $(r,b) \in \Omega$.
- Die Augenzahl-Summenbildung wird dann sinnvoller Weise durch die Zufallsvariable

$$\xi : \Omega \to \mathcal{X}, (r,b) \mapsto \xi((r,b)) := r + b. \tag{4}$$

  beschrieben, wobei $\mathcal{X} := \{2, 3, ..., 12\}$.
- $\mathcal{S} := \mathcal{P}(\mathcal{X})$ ist eine sinnvolle σ-Algebra auf $\mathcal{X}$.
- Mithilfe der σ-Addivität von $\mathbb{P}$ können wir die Verteilung $\mathbb{P}_\xi$ von ξ für alle Elementarereignisse $\{x\} \in \mathcal{S}$ berechnen, wie auf der nächsten Folie gezeigt.
- Wir haben damit ein weiteres Wahrscheinlichkeitsraum-Modell $(\mathcal{X}, \mathcal{S}, \mathbb{P}_\xi)$ konstruiert.

Beispiel (Summe zweier Würfel)

Bestimmung der Verteilung $\mathbb{P}_\xi$ von ξ

$$
\begin{aligned}
\mathbb{P}_\xi(\{2\}) &= \mathbb{P}\left(\xi^{-1}(\{2\})\right) &&= \mathbb{P}\left(\{(1,1)\}\right) &&= \tfrac{1}{36} \\
\mathbb{P}_\xi(\{3\}) &= \mathbb{P}\left(\xi^{-1}(\{3\})\right) &&= \mathbb{P}\left(\{(1,2),(2,1)\}\right) &&= \tfrac{2}{36} \\
\mathbb{P}_\xi(\{4\}) &= \mathbb{P}\left(\xi^{-1}(\{4\})\right) &&= \mathbb{P}\left(\{(1,3),(3,1),(2,2)\}\right) &&= \tfrac{3}{36} \\
\mathbb{P}_\xi(\{5\}) &= \mathbb{P}\left(\xi^{-1}(\{5\})\right) &&= \mathbb{P}\left(\{(1,4),(4,1),(2,3),(3,2)\}\right) &&= \tfrac{4}{36} \\
\mathbb{P}_\xi(\{6\}) &= \mathbb{P}\left(\xi^{-1}(\{6\})\right) &&= \mathbb{P}\left(\{(1,5),(5,1),(2,4),(4,2),(3,3)\}\right) &&= \tfrac{5}{36} \\
\mathbb{P}_\xi(\{7\}) &= \mathbb{P}\left(\xi^{-1}(\{7\})\right) &&= \mathbb{P}\left(\{(1,6),(6,1),(2,5),(5,2),(3,4),(4,3)\}\right) &&= \tfrac{6}{36} \\
\mathbb{P}_\xi(\{8\}) &= \mathbb{P}\left(\xi^{-1}(\{8\})\right) &&= \mathbb{P}\left(\{(2,6),(6,2),(3,5),(5,3),(4,4)\}\right) &&= \tfrac{5}{36} \\
\mathbb{P}_\xi(\{9\}) &= \mathbb{P}\left(\xi^{-1}(\{9\})\right) &&= \mathbb{P}\left(\{(3,6),(6,3),(4,5),(5,4)\}\right) &&= \tfrac{4}{36} \\
\mathbb{P}_\xi(\{10\}) &= \mathbb{P}\left(\xi^{-1}(\{10\})\right) &&= \mathbb{P}\left(\{(4,6),(6,4),(5,5)\}\right) &&= \tfrac{3}{36} \\
\mathbb{P}_\xi(\{11\}) &= \mathbb{P}\left(\xi^{-1}(\{11\})\right) &&= \mathbb{P}\left(\{(5,6),(6,5)\}\right) &&= \tfrac{2}{36} \\
\mathbb{P}_\xi(\{12\}) &= \mathbb{P}\left(\xi^{-1}(\{12\})\right) &&= \mathbb{P}\left(\{(6,6)\}\right) &&= \tfrac{1}{36}
\end{aligned}
$$

Definition (Notation für Zufallsvariablen)

Es seien $(\Omega, \mathcal{A}, \mathbb{P})$ und $(\mathcal{X}, \mathcal{S}, \mathbb{P}_\xi)$ W-Räume und $\xi : \Omega \to \mathcal{X}$ sei eine ZV. Dann gelten folgende Konventionen:

$$
\begin{aligned}
\{\xi \in S\} &:= \{\omega \in \Omega | \xi(\omega) \in S\}, S \subset \mathcal{X}, \\
\{\xi = x\} &:= \{\omega \in \Omega | \xi(\omega) = x\}, x \in \mathcal{X}, \\
\{\xi \le x\} &:= \{\omega \in \Omega | \xi(\omega) \le x\}, x \in \mathcal{X}, \\
\{\xi < x\} &:= \{\omega \in \Omega | \xi(\omega) < x\}, x \in \mathcal{X}.
\end{aligned}
$$

Aus diesen Konventionen folgen exemplarisch die folgenden Konventionen für Verteilungen:

$$
\begin{aligned}
\mathbb{P}_\xi\left(\xi \in S\right) = \mathbb{P}\left(\{\xi \in S\}\right) = \mathbb{P}\left(\{\omega \in \Omega | \xi(\omega) \in S\}\right), S \subset \mathcal{X} \\
\mathbb{P}_\xi\left(\xi \le x\right) = \mathbb{P}\left(\{\xi \le x\}\right) = \mathbb{P}\left(\{\omega \in \Omega | \xi(\omega) \le x\}\right), x \in \mathcal{X}.
\end{aligned}
$$

Oft wird zudem auf das ZV Subskript bei Verteilungssymbolen verzichtet, z.B.

$$
\begin{aligned}
\mathbb{P}\left(\xi \in S\right) = \mathbb{P}_\xi\left(\xi \in S\right), S \subset \mathcal{X}, \\
\mathbb{P}\left(\xi \le x\right) = \mathbb{P}_\xi\left(\xi \le S\right), x \in \mathcal{X}.
\end{aligned}
$$

Definition (Realisierung einer Zufallsvariable)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum $(\mathcal{X}, \mathcal{S})$ sei ein Messraum und $\xi : \Omega \to \mathcal{X}$ sei eine Zufallsvariable. Dann heißt $\xi(\omega) \in \mathcal{X}$ auch *Realisierung der Zufallvariable*.

- In der Datenanalyse werden Daten typischerweise als Realisierungen von Zufallsvariablen modelliert.
- Da die Auswahl eines $\omega \in \Omega$ in einem Zufallsvorgang zufällig ist, erscheint $\xi(\omega)$ zufällig.

Simulation von Zufallsvariablenrealisierungen (Summe zweier Würfel)

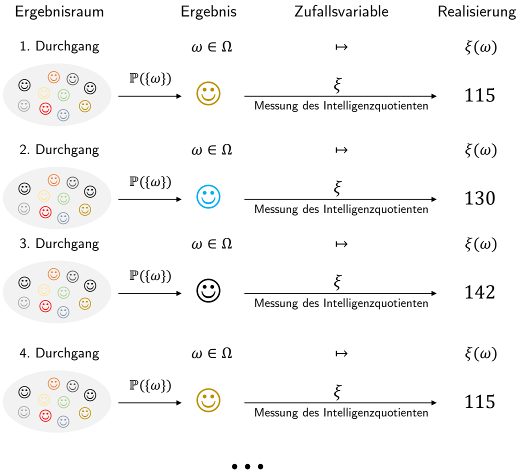
```
# Wahrscheinlichkeitsraummodell
Omega    = list()                             # Ergebnisrauminitialisierung
idx      = 1                                  # Ergebnisindexinitialisierung
for(r in 1:6){                                # Ergebnisse roter  Würfel
    for(b in 1:6){                            # Ergebnisse blauer Würfel
        Omega[[idx]] = c(r,b)                 # \omega \in \Omega
        idx          = idx + 1 }}             # Ergebnisindexupdate
K        = length(Omega)                      # Kardinalität von \Omega
pi       = rep(1/K,1,K)                       # Wahrscheinlichkeitsfunktion \pi

# Zufallsvorgang
omega    = Omega[[which(rmultinom(1,1,pi) == 1)]] # Auswahl von \omega anhand \mathbb{P}({\omega})

# Auswertung der Zufallsvariable
xi_omega = sum(omega)                         # \xi(\omega)
```

```
> omega     :  4 2
> xi(omega) :  6
```

Zufallsvariablen als Modelle von Messvorgängen



Konstruktion, Definition, Notation, Intuition

Theorem (Arithmetik reeller Zufallsvariablen)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum, $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ sei der reelle Messraum, $\xi : \Omega \to \mathbb{R}$, $\upsilon : \Omega \to \mathbb{R}$ seien reellwertige Zufallsvariablen und $c \in \mathbb{R}$ sei eine Konstante. Weiterhin seien

$$
\begin{aligned}
\xi + c : \Omega \to \mathbb{R}, \omega \mapsto (\xi + c)(\omega) &:= \xi(\omega) + c \text{ für } c \in \mathbb{R} \\
c\xi : \Omega \to \mathbb{R}, \omega \mapsto (c\xi)(\omega) &:= c\xi(\omega) \text{ für } c \in \mathbb{R} \\
\xi + \upsilon : \Omega \to \mathbb{R}, \omega \mapsto (\xi + \upsilon)(\omega) &:= \xi(\omega) + \upsilon(\omega) \\
\xi\upsilon : \Omega \to \mathbb{R}, \omega \mapsto (\xi\upsilon)(\omega) &:= \xi(\omega)\upsilon(\omega)
\end{aligned}
\tag{5}
$$

die Addition einer Konstante zu einer reellwertigen Zufallsvariable, die Multiplikation einer reellwertigen Zufallsvariable mit einer Konstante, die Addition zweier reellwertiger Zufallsvariablen und die Multiplikation zweier reellwertigen Zufallsvariablen, respektive. Dann sind auch $\xi + c$, $c\xi$, $\xi + \upsilon$ und $\xi\upsilon$ reellwertige Zufallsvariablen.

Bemerkungen

- Intuitiv ergibt die Addition einer zufälligen Größe zu einer konstanten Größe eine zufällige Größe.
- Intuitiv ergibt die Multiplikation einer zufälligen Größe mit einer Konstante eine zufällige Größe.
- Intuitiv ergibt die Addition zweier zufälliger Größen wiederum eine zufällige Größe.
- Intuitiv ergibt die Multiplikation zweier zufälliger Größen wiederum eine zufällige Größe.
- Für einen Beweis, siehe Hesse (2009), Seite 33-34.

Konstruktion, Definition, Notation, Intuition

Wahrscheinlichkeitsmassefunktionen

Wahrscheinlichkeitsdichtefunktionen

Kumulative Verteilungsfunktionen

Selbstkontrollfragen

Definition (Diskrete ZV, Wahrscheinlichkeitsmassefunktion)

Eine Zufallsvariable ξ heißt *diskret*, wenn ihr Ergebnisraum $\mathcal{X}$ endlich oder abzählbar ist und eine Funktion der Form

$$p : \mathcal{X} \to \mathbb{R}_{\geq 0}, x \mapsto p(x) \tag{6}$$

existiert, für die gilt

(1) $\sum_{x \in \mathcal{X}} p(x) = 1$ und

(2) $\mathbb{P}_\xi(\xi = x) = p(x)$ für alle $x \in \mathcal{X}$.

Eine entsprechende Funktion p heißt *Wahrscheinlichkeitsmassefunktion (WMF)* von ξ.

Bemerkungen

- Eine Menge heißt abzählbar, wenn sie bijektiv auf $\mathbb{N}$ abgebildet werden kann.
- WMFen heißen im Deutschen auch *W-Funktionen* oder *Zähldichten*.
- WMFen heißen auf Englisch *probability mass functions (PMFs)*.
- Die Eigenschaft $\sum_{x \in \mathcal{X}} p(x) = 1$ nennt man auch *Normiertheit* von p.
- Zur Parallelität mit PMFs und WDFs bevorzugen wird den Begriff WMF.

Definition (Bernoulli-Zufallsvariable)

Es sei ξ eine Zufallsvariable mit Ergebnisraum $\mathcal{X} = \{0, 1\}$ und WMF

$$p : \mathcal{X} \to [0, 1], x \mapsto p(x) := \mu^x (1-\mu)^{1-x} \text{ mit } \mu \in [0, 1]. \tag{7}$$

Dann sagen wir, dass ξ einer *Bernoulli-Verteilung mit Parameter* $\mu \in [0, 1]$ unterliegt und nennen ξ eine *Bernoulli-Zufallsvariable*. Wir kürzen dies mit $\xi \sim \text{Bern}(\mu)$ ab. Die WMF einer Bernoulli-Zufallsvariable bezeichnen wir mit

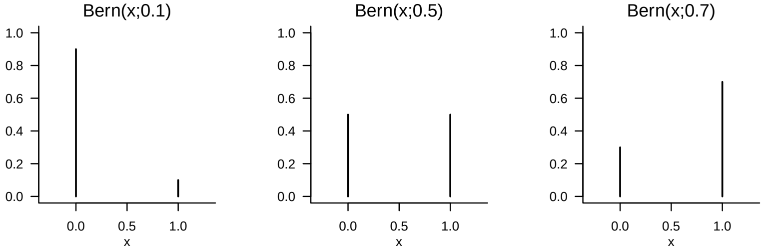
$$\text{Bern}(x; \mu) := \mu^x (1-\mu)^{1-x}. \tag{8}$$

Bemerkungen

- Eine Bernoulli-Zufallsvariable kann als Modell eines Münzwurfs dienen.
- μ ist die Wahrscheinlichkeit dafür, dass ξ den Wert 1 annimmt,

$$\mathbb{P}(\xi = 1) = \mu^1 (1-\mu)^{1-1} = \mu. \tag{9}$$

Bernoulli-Zufallsvariable



Definition (Binomial-Zufallsvariable)

Es sei ξ eine Zufallsvariable mit Ergebnisraum $\mathcal{X} := \mathbb{N}_n^0$ und WMF

$$p : \mathcal{X} \to [0,1], x \mapsto p(x) := \binom{n}{x} \mu^x (1-\mu)^{n-x} \text{ für } \mu \in [0,1]. \tag{10}$$

Dann sagen wir, dass ξ einer *Binomialverteilung mit Parametern* $\mu \in [0,1]$ *und* $n \in \mathbb{N}$ unterliegt und nennen ξ eine Binomial-Zufallsvariable. Wir kürzen dies mit $\xi \sim \text{Bin}(\mu, n)$ ab. Die WMF einer Binomial-Zufallsvariable bezeichnen wir mit

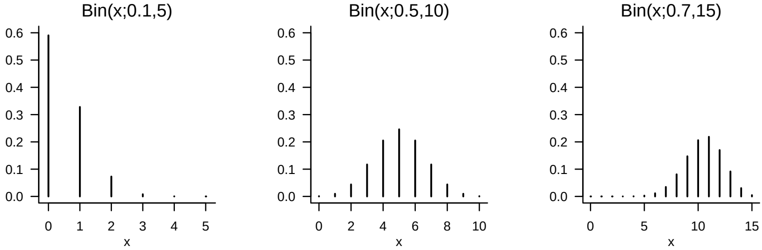
$$\text{Bin}(x; \mu, n) := \binom{n}{x} \mu^x (1-\mu)^{n-x}. \tag{11}$$

Bemerkung

- Es gilt $\text{Bin}(x; \mu, 1) = \text{Bern}(x; \mu)$.

Wahrscheinlichkeitsmassefunktionen

Binomial-Zufallsvariable



Definition (Diskret-gleichverteilte Zufallsvariable)

Es sei ξ eine diskrete Zufallsvariable mit endlichem Ergebnisraum $\mathcal{X}$ und WMF

$$p : \mathcal{X} \to \mathbb{R}_{\geq 0}, x \mapsto p(x) := \frac{1}{|\mathcal{X}|}. \tag{12}$$

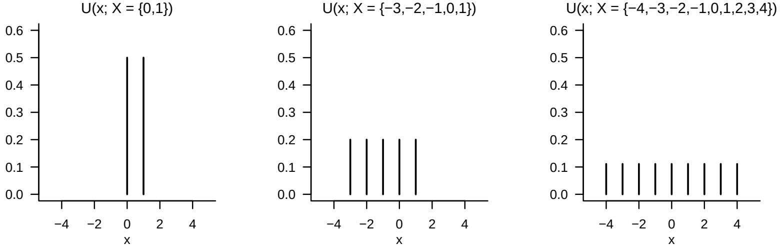
Dann sagen wir, dass ξ einer *diskreten Gleichverteilung* unterliegt und nennen ξ eine *diskret-gleichverteilte Zufallsvariable*. Wir kürzen dies mit $\xi \sim U(|\mathcal{X}|)$ ab. Die WMF einer diskret-gleichverteilten Zufallsvariable bezeichnen wir mit

$$U(x; |\mathcal{X}|) := \frac{1}{|\mathcal{X}|}. \tag{13}$$

Bemerkungen

- Bern$(x; 0.5) = U(x; |\mathcal{X}|)$ für $\mathcal{X} = \{0, 1\}$.

Diskret-gleichverteilte Zufallsvariable



Konstruktion, Definition, Notation, Intuition

Wahrscheinlichkeitsmassefunktionen

Wahrscheinlichkeitsdichtefunktionen

Kumulative Verteilungsfunktionen

Selbstkontrollfragen

Definition (Kontinuierliche ZV, Wahrscheinlichkeitsdichtefunktion)

Eine Zufallsvariable ξ heißt *kontinuierlich*, wenn $\mathbb{R}$ der Ergebnisraum von ξ ist und eine Funktion

$$p : \mathbb{R} \to \mathbb{R}_{\geq 0}, x \mapsto p(x) \tag{14}$$

existiert, für die gilt

(1) $\int_{-\infty}^{\infty} p(x)dx = 1$ und

(2) $\mathbb{P}_\xi(\xi \in [a, b]) = \int_a^b p(x)\,dx$ für alle $a, b \in \mathbb{R}$ mit $a \leq b$.

Eine entsprechende Funktion p heißt *Wahrscheinlichkeitsdichtefunktion (WDF)* von ξ.

Bemerkungen

- WDFen können Werte größer als 1 annehmen.
- Es gilt $\mathbb{P}_\xi(\xi = a) = \int_a^a p(x)\,dx = 0$.
- Wahrscheinlichkeiten werden aus WDFen durch Integration berechnet.
- (Wahrscheinlichkeits)Masse $=$ (Wahrscheinlichkeits)Dichte $\times$ (Mengen)Volumen.

Definition (Normalverteilte und standardnormalverteilte Zufallsvariablen)

Es sei ξ eine Zufallsvariable mit Ergebnisraum $\mathbb{R}$ und WDF

$$p : \mathbb{R} \to \mathbb{R}_{>0}, x \mapsto p(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right). \tag{15}$$

Dann sagen wir, dass ξ einer *Normalverteilung (oder Gauß-Verteilung) mit Parametern $\mu \in \mathbb{R}$ und $\sigma^2 > 0$* unterliegt und nennen ξ eine *normalverteilte Zufallsvariable*. Wir kürzen dies mit $\xi \sim N\left(\mu, \sigma^2\right)$ ab. Die WDF einer normalverteilten Zufallsvariable bezeichnen wir mit

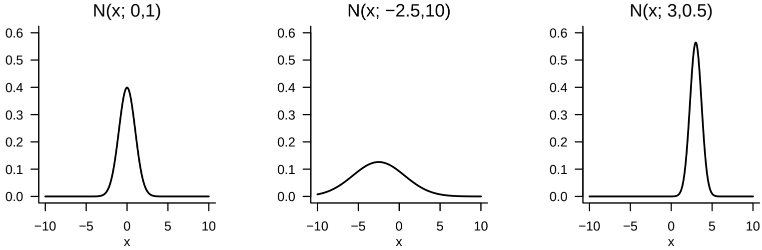
$$N\left(x; \mu, \sigma^2\right) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right). \tag{16}$$

Eine normalverteilte Zufallsvariable mit $\mu = 0$ und $\sigma^2 = 1$ heißt *standardnormalverteilte Zufallsvariable* und wird oft als *Z-Zufallsvariable* bezeichnet.

Bemerkungen

- Der Parameter μ entspricht dem Wert höchster Wahrscheinlichkeitsdichte.
- Der Parameter σ^2 spezifiziert die Breite der WDF.

Normalverteilte Zufallsvariablen



Definition (Gamma-Zufallsvariable)

Es sei ξ eine Zufallsvariable mit Ergebnisraum $\mathcal{X} := \mathbb{R}_{>0}$ und WDF

$$p : \mathbb{R}_{>0} \to \mathbb{R}_{>0}, x \mapsto p(x) := \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), \tag{17}$$

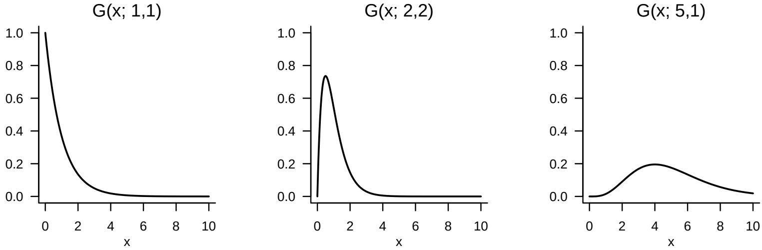
wobei Γ die Gammafunktion bezeichne. Dann sagen wir, dass ξ einer *Gammaverteilung mit Formparameter* $\alpha > 0$ *und Skalenparameter* $\beta > 0$ unterliegt und nennen ξ eine *gammaverteilte Zufallsvariable*. Wir kürzen dies mit $\xi \sim G(\alpha, \beta)$ ab. Die WDF einer gammaverteilen Zufallsvariable bezeichnen wir mit

$$G(x; \alpha, \beta) := \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right). \tag{18}$$

Bemerkung

- $G\left(\frac{n}{2}, 2\right)$ heißt auch *Chi-Quadrat (χ^2) Verteilung mit n Freiheitsgraden*.

Gamma-Zufallsvariablen



Definition (Beta-Zufallsvariable)

Es sei ξ eine Zufallsvariable mit Ergebnisraum $\mathcal{X} := [0,1]$ und WDF

$$p : \mathcal{X} \to [0,1], x \mapsto p(x) := \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \text{ mit } \alpha, \beta \in \mathbb{R}_{>0}, \tag{19}$$

wobei Γ die Gammafunktion bezeichne. Dann sagen wir, dass ξ einer *Beta-Verteilung* mit Parametern $\alpha > 0$ und $\beta > 0$ unterliegt, und nennen ξ eine beta-verteilte Zufallsvariable. Wir kürzen dies mit $\xi \sim \text{Beta}(\alpha, \beta)$ ab. Die WDF einer beta-verteilten Zufallsvariable bezeichnen wir mit

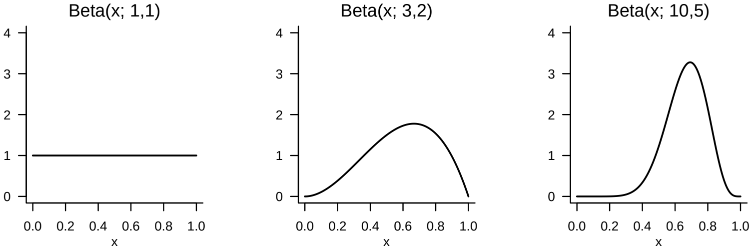
$$\text{Beta}(x; \alpha, \beta) := \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}. \tag{20}$$

Bemerkung

- Für $\alpha < 1, \beta < 1$ ist der Ergebnisraum $\mathcal{X} :=]0,1[$.

Wahrscheinlichkeitsdichtefunktionen

Beta-Zufallsvariable



Definition (Gleichverteilte Zufallsvariable)

Es sei ξ eine kontinuierliche Zufallsvariable mit Ergebnisraum $\mathbb{R}$ und WDF

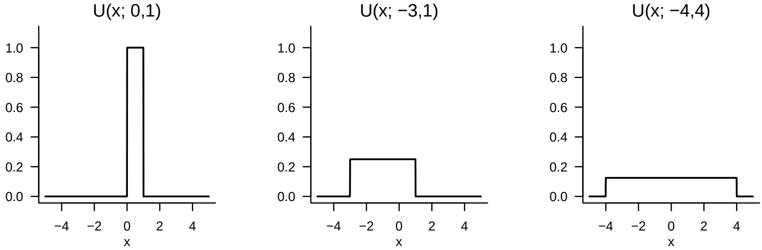
$$p : \mathbb{R} \to \mathbb{R}_{\geq 0}, x \mapsto p(x) := \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & x \notin [a,b] \end{cases}. \tag{21}$$

Dann sagen wir, dass ξ einer *Gleichverteilung mit Parametern a und b* unterliegt und nennen ξ eine *gleichverteilte Zufallsvariable*. Wir kürzen dies mit $\xi \sim U(a,b)$ ab. Die WDF einer gleichverteilten Zufallsvariable bezeichnen wir mit

$$U(x; a, b) := \frac{1}{b-a}. \tag{22}$$

Wahrscheinlichkeitsdichtefunktionen

Gleichverteilte Zufallsvariablen



Konstruktion, Definition, Notation, Intuition

Wahrscheinlichkeitsmassefunktionen

Wahrscheinlichkeitsdichtefunktionen

Kumulative Verteilungsfunktionen

Selbstkontrollfragen

Kumulative Verteilungsfunktionen

Definition (Kumulative Verteilungsfunktion)

Die *kumulative Verteilungsfunktion (KVF)* einer Zufallsvariable ξ ist definiert als

$$P : \mathbb{R} \to [0,1], x \mapsto P(x) := \mathbb{P}(\xi \le x). \tag{23}$$

Bemerkungen

- KVFen sind sowohl für diskrete als auch kontinuierliche ZVen definiert.
- $P(x)$ ist für jedes $x \in \mathbb{R}$ definiert, auch wenn $x \notin \mathcal{X}$.
- Mithilfe von KVFen können Intervallwahrscheinlichkeiten angegeben werden

Theorem (Überschreitungswahrscheinlichkeit)

Es sei ξ eine Zufallsvariable mit Ereignisraum $\mathcal{X}$ und P ihre kumulative Verteilungsfunktion. Dann gilt für die *Überschreitungswahrscheinlichkeit* $\mathbb{P}(\xi > x)$, dass

$$\mathbb{P}(\xi > x) = 1 - P(x) \text{ für alle } x \in \mathcal{X}. \tag{24}$$

Beweis

Die Ereignisse $\{\xi > x\}$ und $\{\xi \le x\}$ sind disjunkt und

$$\Omega = \{\omega \in \Omega | \xi(\omega) > x\} \cup \{\omega \in \Omega | \xi(\omega) \le x\} = \{\xi > x\} \cup \{\xi \le x\}. \tag{25}$$

Mit der σ-Additivität von $\mathbb{P}$ folgt dann

$$\begin{aligned}
\mathbb{P}(\Omega) &= 1 \\
\Leftrightarrow \mathbb{P}(\{\xi > x\} \cup \{\xi \le x\}) &= 1 \\
\Leftrightarrow \mathbb{P}(\{\xi > x\}) + \mathbb{P}(\{\xi \le x\}) &= 1 \\
\Leftrightarrow \mathbb{P}(\{\xi > x\}) &= 1 - \mathbb{P}(\{\xi \le x\}) \\
\Leftrightarrow \mathbb{P}(\{\xi > x\}) &= 1 - P(x).
\end{aligned} \tag{26}$$

$\square$

Theorem (Intervallwahrscheinlichkeiten)

Es sei ξ eine Zufallsvariable mit Ereignisraum $\mathcal{X}$ und P ihre kumulative Verteilungsfunktion. Dann gilt für die *Intervallwahrscheinlichkeit* $\mathbb{P}(\xi \in]x_1, x_2])$, dass

$$\mathbb{P}(\xi \in]x_1, x_2]) = P(x_2) - P(x_1) \text{ für alle } x_1, x_2 \in \mathcal{X} \text{ mit } x_1 < x_2. \tag{27}$$

Beweis

Wir betrachten die Ereignisse $\{\xi \le x_1\}$,$\{x_1 < \xi \le x_2\}$ und $\{\xi \le x_2\}$, wobei

$$\{\xi \le x_1\} \cap \{x_1 < \xi \le x_2\} = \emptyset \text{ und } \{\xi \le x_1\} \cup \{x_1 < \xi \le x_2\} = \{\xi \le x_2\}. \tag{28}$$

gelten. Mit der σ-Additivität von $\mathbb{P}$ gilt dann

$$\begin{aligned}
\mathbb{P}(\{\xi \le x_1\} \cup \{x_1 < \xi \le x_2\}) &= \mathbb{P}(\{\xi \le x_2\}) \\
\Leftrightarrow \mathbb{P}(\{\xi \le x_1\}) + \mathbb{P}(\{x_1 < \xi \le x_2\}) &= \mathbb{P}(\{\xi \le x_2\}) \\
\Leftrightarrow \mathbb{P}(\{x_1 < \xi \le x_2\}) &= \mathbb{P}(\{\xi \le x_2\}) - \mathbb{P}(\{\xi \le x_1\}) \\
\Leftrightarrow \mathbb{P}(\{x_1 < \xi \le x_2\}) &= P(x_2) - P(x_1) \\
\Leftrightarrow \mathbb{P}(\xi \in]x_1, x_2]) &= P(x_2) - P(x_1).
\end{aligned} \tag{29}$$

$\square$

Theorem (Eigenschaften von kumulative Verteilungsfunktionen)

Es sei ξ eine Zufallsvariable und P ihre kumulative Verteilungsfunktion. Dann hat P die folgenden Eigenschaften

(1) P ist *monoton steigend*, i.e., wenn $x_1 < x_2$, dann gilt $P(x_1) \le P(x_2)$.

(2) $\lim_{x \to -\infty} P(x) = 0$ und $\lim_{x \to \infty} P(x) = 1$.

(3) P ist *rechtsseitig stetig*, d.h., $P(x) = P(x^+) = \lim_{y \to x, y > x} P(y)$ für alle $x \in \mathbb{R}$

Bemerkungen

- Die genannten Eigenschaften können auch zur Definition einer KVF genutzt werden.
- (3) $\Leftrightarrow$ Eine KVF hat keine Sprünge, wenn man sich Grenzpunkten von rechts nähert.

Kumulative Verteilungsfunktionen

Beweis

(1) Wir halten zunächst fest, dass für Ereignisse $A \subset B$ gilt, dass $\mathbb{P}(A) \leq \mathbb{P}(B)$. Wie halten dann fest, dass für $x_1 < x_2$,

$$\{\xi \leq x_1\} = \{\omega \in \Omega | \xi(\omega) \leq x_1\} \subset \{\omega \in \Omega | \xi(\omega) \leq x_2\} = \{\xi \leq x_2\} \tag{30}$$

Also gilt

$$\mathbb{P}(\{\xi \leq x_1\}) \leq \mathbb{P}\{\xi \leq x_2\} \Rightarrow P(x_1) \leq P(x_2). \tag{31}$$

(2) Wir verzichten auf einen Beweis

(3) Wir definieren

$$P(x^+) = \lim_{y \to x, y > x} P(y). \tag{32}$$

Seien nun $y_1 > y_2 > \cdots$ so, dass $\lim_{n \to \infty} y_n = x$. Dann gilt

$$\{\xi \leq x\} = \cap_{n=1}^{\infty} \{\xi \leq y_n\}. \tag{33}$$

Es gilt also

$$P(x) = \mathbb{P}(\{\xi \leq x\}) = \mathbb{P}(\cap_{n=1}^{\infty} \{\xi \leq y_n\}) = \lim_{n \to \infty} \mathbb{P}(\{\xi \leq y_n\}) = P(x^+), \tag{34}$$

wobei wir die dritte Gleichung unbegründet stehen lassen.

Kumulative Verteilungsfunktionen von diskreten Zufallsvariablen

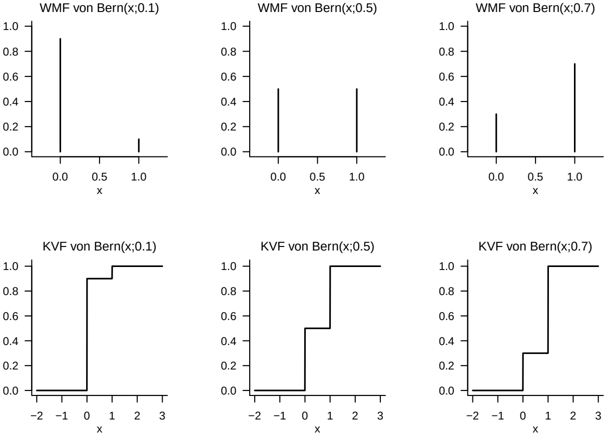
- Wenn $a < b$ und $\mathbb{P}(a < \xi < b) = 0$, dann ist P konstant horizontal auf $]a, b[$.
- An jedem Punkt x mit $\mathbb{P}(\xi = x) > 0$ springt die KVF um den Betrag $\mathbb{P}(\xi = x)$.
- $\Leftrightarrow$ An jedem Punkt x mit $p(x) > 0$ springt die KVF um den Betrag $p(x) > 0$.
- Generell ist die KVF einer diskreten Zufallsvariable mit Ergebnisraum $\mathbb{N}_0$ durch

$$P : \mathbb{R} \to [0, 1], x \mapsto P(x) := \sum_{k=0}^{\lfloor x \rfloor} \mathbb{P}(\xi = k) \tag{35}$$

gegeben, wobei $\lfloor x \rfloor$ die Abrundungsfunktion bezeichnet.

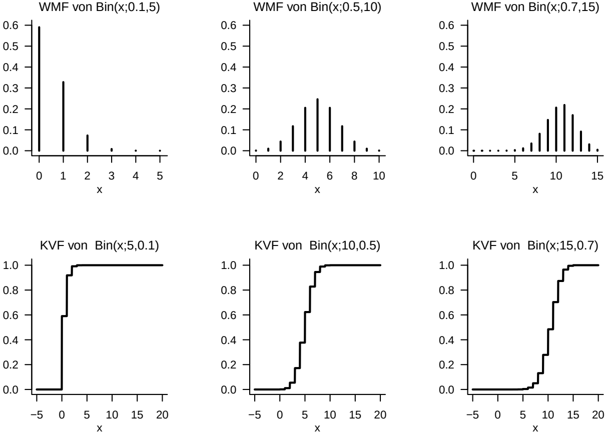
Kumulative Verteilungsfunktionen

Bernoulli-Zufallsvariablen



Kumulative Verteilungsfunktionen

Binomial-Zufallsvariablen



Theorem (Kumulative Verteilungsfunktionen von kontinuierlichen ZVen)

ξ sei eine kontinuierliche Zufallsvariable mit WDF p und KVF P. Dann gilt

$$P(x) = \int_{-\infty}^{x} p(t)\,dt \text{ und } p(x) = \frac{d}{dx}P(x). \tag{36}$$

Beweis

Wir halten zunächst fest, dass weil $\mathbb{P}(\xi = x) = 0$ für alle $x \in \mathbb{R}$ gilt, die KVF von ξ keine Sprünge hat, d.h. P ist stetig. Mit der Definitionen von WDF und KVF, folgt, dass P die Form einer Stammfunktion von p hat. Dass p die Ableitung von P ist, folgt dann unmittelbar aus dem Fundamentalsatz der Analysis.

□

Bemerkungen

- Die KVF ist eine Stammfunktion der WDF, die WDF ist die Ableitung der KVF.
- Das *Theorem von Radon-Nikodym* ist eine generalisierte Variante dieser Einsicht.
- KVFen von kontinuierlichen ZV heißen auch kumulative Dichtefunktionen (KDFen).

Beispiel (Normalverteilung)

Es sei $X \sim N(\mu, \sigma^2)$.

- Die WDF von ξ ist

$$p : \mathbb{R} \to \mathbb{R}_{>0}, x \mapsto p(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right).$$

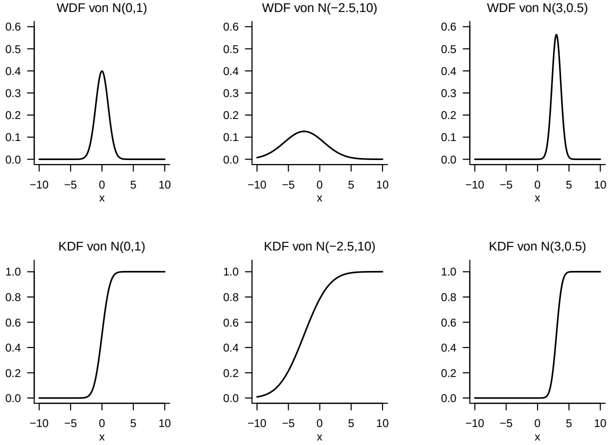
- Die KVF von ξ ist

$$P : \mathbb{R} \to]0,1[, x \mapsto P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \exp\left(-\frac{1}{2\sigma^2}(\xi-\mu)^2\right) d\xi.$$

- Die KVF von $\xi \sim N(\mu, \sigma^2)$ kann nur numerisch, nicht analytisch, berechnet werden.
- Für $\mu = 1, \sigma^2 = 1$, gilt zum Beispiel $p(2) = 0.24$ und $P(2) = 0.84$.
- Die WDF und KVF von $Z \sim N(0,1)$ werden oft mit ϕ und Φ, respektive, bezeichnet.

Kumulative Verteilungsfunktionen

Normalverteilte Zufallsvariablen



Definition (Inverse Kumulative Verteilungsfunktion)

ξ sei eine kontinuierliche Zufallsvariable mit KVF P. Dann heißt die Funktion

$$P^{-1} :]0,1[\to \mathbb{R}, q \mapsto P^{-1}(q) := \{x \in \mathbb{R} | P(x) = q\} \tag{37}$$

die *inverse kumulative Verteilungsfunktion von* ξ.

Bemerkungen

- P^{-1} ist die Inverse von P, d.h. $P^{-1}(P(x)) = x$.
- Offenbar gilt $P(x) = q \Leftrightarrow \mathbb{P}(\xi \le x) = q$.
- Für $q \in]0,1[$ ist also $P^{-1}(q)$ der Wert x von ξ, so dass $\mathbb{P}(\xi \le x) = q$ gilt.
- Wenn $Z \sim N(0,1)$ mit KVF Φ ist, dann gilt zum Beispiel $\Phi^{-1}(0.975) = 1.960$.

Kumulative Verteilungsfunktionen

Beispiel (Normalverteilung)

Es sei $\xi \sim N(\mu, \sigma^2)$.

- Die KVF von ξ ist

$$P : \mathbb{R} \to]0,1[, x \mapsto \mathbb{P}(\xi \le x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \exp\left(-\frac{1}{2\sigma^2}(\xi - \mu)^2\right) d\xi \qquad (38)$$

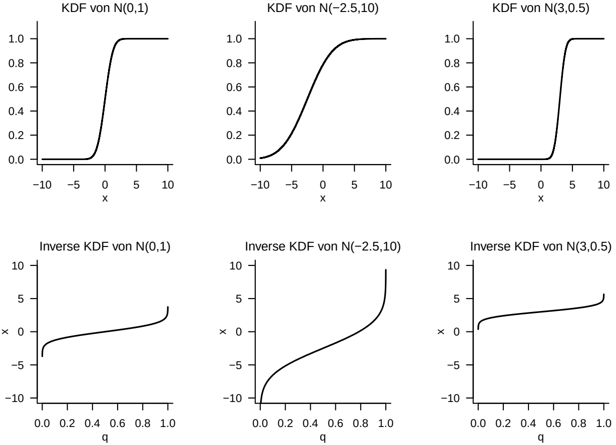
- Die inverse KVF von ξ ist

$$P^{-1} :]0,1[\to \mathbb{R}, q \mapsto P^{-1}(q) = \{x \in \mathbb{R} | P(x) = q\}. \qquad (39)$$

- Für $\mu = 1, \sigma^2 = 1$ gilt z.B., dass $P(2) = 0.84$ und $P^{-1}(0.84) = 2$.
- Die inverse KVF von $\xi \sim N(0,1)$ wird oft mit Φ^{-1} bezeichnet.
- Typische Beispielwerte für die KVF und inverse KVF von $N(0,1)$ sind
  - $\Phi(1.645) = 0.950$, $\Phi^{-1}(0.950) = \Phi^{-1}(1 - 0.050) = 1.645$.
  - $\Phi(1.960) = 0.975$, $\Phi^{-1}(0.975) = \Phi^{-1}(1 - \frac{0.050}{2})$.

Kumulative Verteilungsfunktionen

Normalverteilte Zufallsvariablen



Konstruktion, Definition, Notation, Intuition

Wahrscheinlichkeitsmassefunktionen

Wahrscheinlichkeitsdichtefunktionen

Kumulative Verteilungsfunktionen

Selbstkontrollfragen

Selbstkontrollfragen

1. Definieren Sie den Begriff der Zufallsvariable.
2. Erläutern Sie die Gleichung $\mathbb{P}_\xi(\xi = x) = \mathbb{P}(\{\xi = x\})$.
3. Erläutern Sie die Bedeutung von $\mathbb{P}(\xi = x)$.
4. Definieren Sie den Begriff der Wahrscheinlichkeitsmassefunktion.
5. Definieren Sie die Begriffe der Wahrscheinlichkeitsdichtefunktion.
6. Definieren Sie den Begriff der kumulativen Verteilungsfunktion.
7. Schreiben sie die Intervallwahrscheinlichkeit einer Zufallsvariable mithilfer ihrer KVF.
8. Definieren Sie die WDF und KVF einer normalverteilten Zufallsvariable.
9. Schreiben Sie den Wert $P(x)$ der KVF einer Zufallsvariable mithilfe ihrer WDF.
10. Schreiben Sie den Wert $p(x)$ der WDF einer Zufallsvariable mithilfe ihrer KVF.
11. Definieren Sie den Begriff der inversen Verteilungsfunktion.

Referenzen

Hesse, Christian. 2009. *Wahrscheinlichkeitstheorie*. 2nd ed. Wiesbaden: Vieweg + Teubner.