

Grundlagen der Mathematischen Statistik

Diskrete und kontinuierliche Zufallsvariablen

Uwe Menzel, 2018
 uwe.menzel@matstat.org
www.matstat.org

Diskrete und kontinuierliche Zufallsvariablen

Zufallsvariable (Z.v.):

- Variable welche zufällige Werte annimmt; beschreibt einen Zufallsversuch
- große Buchstaben: X, Y, L usw.
- unbestimmt vor Zufallsversuch

Observation (für eine Z.v.):

- Resultat (Realisation) eines Zufallsversuches
- kleine Buchstaben: k, x, l usw.

Diskrete Z.v.:

- kann nur eine endliche (oder abzählbar unendliche) Anzahl Werte annehmen
- Augenzahl beim Würfel; Anzahl Würfe bevor zweimal hintereinander dieselbe Augenzahl kommt; Anzahl von Kunden in einer Warteschlange; Anzahl der Anwender auf einem Server; Anzahl von Autounfällen pro Jahr usw.

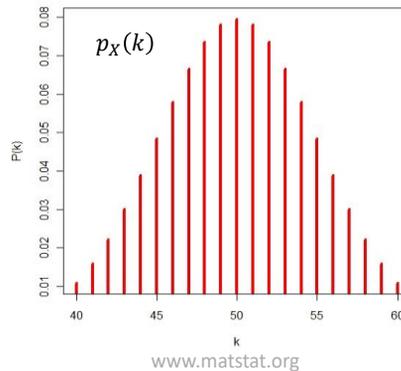
Kontinuierliche Z.v.:

- nicht diskret !
- nimmt reelle Werte an ("unendlich dicht": zwischen zwei beliebigen reellen Werten gibt es unendlich viele reelle Zahlen)
- Lebensdauer einer Glühbirne ; Gewicht ; Windgeschwindigkeit ; elektrische Spannung ; Konzentration einer Chemikalie ; ...

Diskrete Zufallsvariablen

Diskrete Z.v.:

- kann nur eine endliche (oder abzählbar unendliche) Anzahl Werte annehmen
- Augenzahl beim Würfel; Anzahl Würfe bevor zweimal hintereinander dieselbe Augenzahl kommt; Anzahl Kunden in einer Warteschlange; Anzahl User auf einem Server; Anzahl Autounfälle pro Jahr usw.



Wahrscheinlichkeitsfunktionen

- X : diskrete Zufallsvariable (Z.v.)
- k : Realisation (Observation)

Bezeichnung für die Wahrscheinlichkeit, dass die Zufallsvariable X den Wert k annimmt:

$$P(X = k) \text{ oder } p_X(k)$$

Wahrscheinlichkeitsfunktion: $P(X = k)$ für jedes k aus Ω_X

Beispiel: Würfel, Zufallsvariable X : Augenzahl für einen Wurf

$$P(X = 4) = p_X(4) = \frac{1}{6}$$

www.matstat.org

Eine Wahrscheinlichkeitsfunktion

... kann auf drei Arten beschrieben werden:

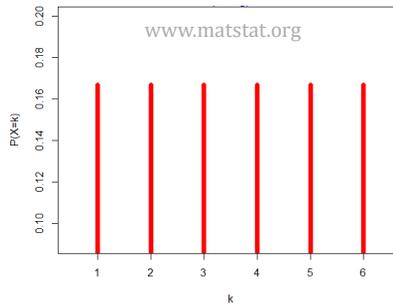
Beispiel: Würfel: $\Omega_X = \{1, 2, 3, 4, 5, 6\}$

1. **Tabelle:**

k	1	2	3	4	5	6
$p_X(k)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

2. **Formel:** $p_X(k) = 1/6 \quad k = 1, 2, \dots, 6$

3. **Säulendiagramm:**



Eigenschaften von Wahrscheinlichkeitsfunktionen

$0 \leq p_X(k) \leq 1$ für alle k (folgt aus **Axiom 1**)

$\sum_{\text{alle } k} p_X(k) = 1$ **Axiom 2** $P(\Omega) = 1$

$P(A) = \sum_{k \in A} p_X(k)$ **Axiom 3** (die Elementarereignisse für verschiedene k müssen unvereinbar sein)

es wird über alle Elementarereignisse, die zum Ereignis A gehören, summiert

Beispiel: Würfel $A = \{2, 4, 6\}$ (gerade Zahl wird gewürfelt)

$$P(A) = \sum_{k=2,4,6} p_X(k) = p_X(2) + p_X(4) + p_X(6) = \frac{3}{6}$$

www.matstat.org

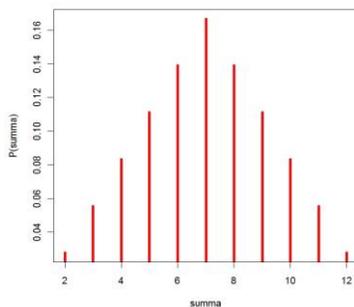
Eigenschaften von Wahrscheinlichkeitsfunktionen

Beispiel : zwei Würfel ; Zufallsvariable $Z = \text{Augensumme}$ $\Omega_Z = \{2, 3, \dots, 12\}$

k	2	3	4	5	6	7	8	9	10	11	12
$p_Z(k)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

In Vorlesung 1 wurde die
Tabelle schon verwendet

	1	2	3	4	5	6
1	②	③	④	5	6	7
2	③	④	5	6	7	8
3	④	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	⑫



Haben wir das richtig gemacht ? :

$$0 \leq p_Z(k) \leq 1 \quad \checkmark$$

$$\sum_{k=2}^{12} p_Z(k) = 1 \quad \checkmark$$

www.matstat.org

Eigenschaften von Wahrscheinlichkeitsfunktionen

Fortsetzung, Beispiel : zwei Würfel

Ereignis A : Summe der Augenzahlen ist eine gerade Zahl

$$P(A) = \sum_{k \in A} p_Z(k) = p_Z(2) + p_Z(4) + \dots + p_Z(12) = \underline{\underline{\frac{1}{2}}}$$

Ereignis B : Summe der Augenzahlen ≤ 4

$$P(B) = P(Z \leq 4) = p_Z(2) + p_Z(3) + p_Z(4) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} = \underline{\underline{\frac{1}{6}}}$$

	1	2	3	4	5	6
1	②	③	④	5	6	7
2	③	④	5	6	7	8
3	④	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	⑫

www.matstat.org

Eigenschaften von Wahrscheinlichkeitsfunktionen

Beispiel 1, diskrete Wahrscheinlichkeitsfunktion

k	1	2	3	4
$p_X(k)$	0.2	0.3	0.4	?

$$\Omega = \{1, 2, 3, 4\}$$

$$\Rightarrow p_X(4) = 0.1 \quad \text{Axiom 2}$$

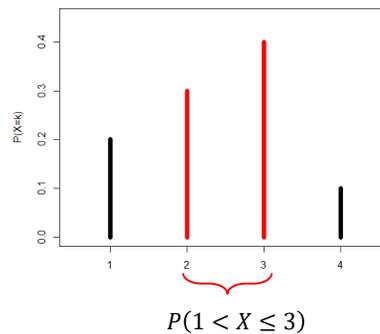
$$P(X \leq 1) = p_X(1) = 0.2$$

$$P(X > 1) = p_X(2) + p_X(3) + p_X(4) = 0.8$$

$$P(1 < X \leq 3) = p_X(2) + p_X(3) = 0.7$$

Vorsicht! Die Eins ist **nicht** einbegriffen!
Für diskrete Z.v. ist es **sehr wichtig**,
zwischen "<" und "≤" zu unterscheiden!

Diskrete Verteilung



www.matstat.org

Eigenschaften von Wahrscheinlichkeitsfunktionen

Beispiel 2, diskrete Wahrscheinlichkeitsfunktion

k	1	2	3	4
$p_X(k)$	$0.2 \cdot c$	$0.3 \cdot c$	$0.4 \cdot c$	$0.5 \cdot c$

$$\Omega = \{1, 2, 3, 4\}$$

$$c = ?$$

$$\sum_{k=1}^4 p_X(k) = 1.4 \cdot c = 1 \quad \Rightarrow \quad c = \frac{10}{14} = \frac{5}{7} \quad (\text{Axiom 2})$$

Zusammenfassung: Eine Wahrscheinlichkeitsfunktion

- gibt die Wahrscheinlichkeit für jedes Elementarereignis einer diskreten Zufallsvariable an
- kann durch Formel, Tabelle oder Säulendiagramm repräsentiert werden

www.matstat.org

Verteilungsfunktionen für diskrete Zufallsvariablen

Definition Verteilungsfunktion:

- X : Zufallsvariable
- a : reelle Zahl

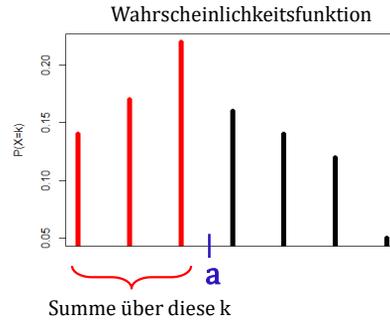
$$F_X(a) = P(X \leq a)$$

Zeichen " \leq " wichtig: "kleiner oder gleich a " !

Verteilungsfunktionen können mit Hilfe der Wahrscheinlichkeitsfunktion berechnet werden:

$$F_X(a) = P(X \leq a) = \sum_{k \leq a} p_X(k)$$

- Summation über alle $p_X(k)$ deren Argument k kleiner oder gleich der reellen Zahl a ist
- (Anm.: die Zahl a muss also **nicht** zum Ereignisraum Ω gehören!)
- die Verteilungsfunktion ist eine kumulative Summe



www.matstat.org

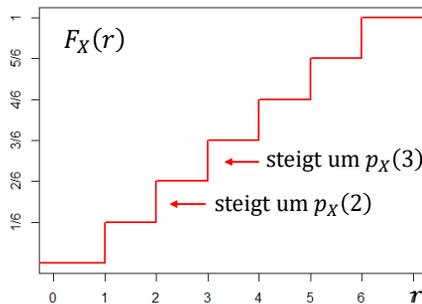
Verteilungsfunktionen für diskrete Zufallsvariablen

Beispiel :
Würfel



Elem.ereignis	k	1	2	3	4	5	6
Wahrscheinl.fkt.	$p_X(k)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$
Verteilungsfkt.	$F_X(k)$	$1/6$	$2/6$	$3/6$	$4/6$	$5/6$	1

Verteilungsfunktionen (engl. **CDF: Cumulative Distribution Functions**) sind **kontinuierliche** Funktionen (definiert für reelle Zahlen). [in der obigen Tabelle nur für ausgewählte Punkte notiert]



www.matstat.org

$$F_X(0.3) = 0$$

rechtskontinuierlich:

$$F_X(1) = P(X \leq 1) = 1/6$$

$$F_X(4.5) = 2/3$$

$$F_X(1000) = 1$$

$F_X(a)$ ist 1 für alle Zahlen a , die größer oder gleich dem größten Elementarereignis sind

Verteilungsfunktion, Eigenschaften

$$0 \leq F_X(t) \leq 1 \quad F_X(t) \text{ ist ja selbst eine Wahrsch.: } F_X(a) = P(X \leq a); \\ F_X(t) \text{ muss also zwischen 0 und 1 sein (Axiom 1)}$$

$$\lim_{t \rightarrow -\infty} F_X(t) = 0 \quad "F_X(-\infty) = 0": F_X \text{ ist Null "links" vom kleinsten Wert in } \Omega$$

$$\lim_{t \rightarrow +\infty} F_X(t) = 1 \quad \text{alle Elementarereignisse müssen kleiner als } \infty \text{ sein, also } "P(X \leq \infty) = 1".$$

$$t_2 > t_1 \rightarrow F_X(t_2) \geq F_X(t_1) \quad F \text{ wächst monoton (nimmt nicht ab).} \\ \text{Es werden ja nur positive Zahlen } p_X(t) \text{ addiert.}$$

Berechnung der Wahrscheinlichkeitsfunktion mit Hilfe der Verteilungsfunktion
(Beispiel Würfel):

$$F_X(3) = p_X(1) + p_X(2) + p_X(3)$$

$$F_X(4) = \underbrace{F_X(3)} + p_X(4)$$

$$\Rightarrow p_X(4) = F_X(4) - F_X(3)$$

$$p_X(k) = F_X(k) - F_X(k-1)$$

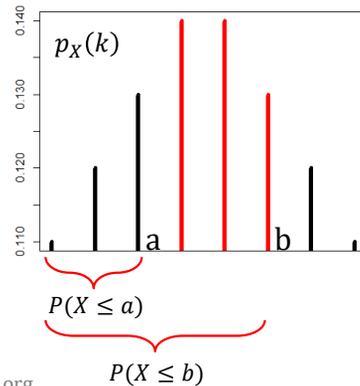
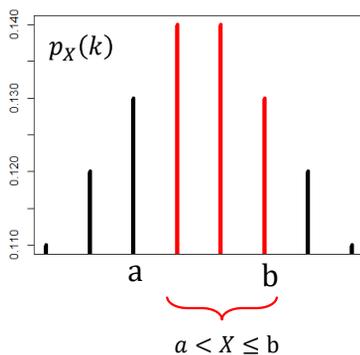
$F_X(k)$ steigt bei k um $p_X(k)$, siehe Bild oben
 F_X ist oft tabelliert \rightarrow Möglichkeit zur Berechnung von $p_X(k)$

Verteilungsfunktion und Wahrscheinlichkeit für Ereignisse

X : diskrete Zufallsvariable; F_X : dessen Verteilungsfunktion korrekte Anwendung der Zeichen \leq sehr wichtig!

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

$$\text{also } P(a < X \leq b) = P(X \leq b) - P(X \leq a)$$

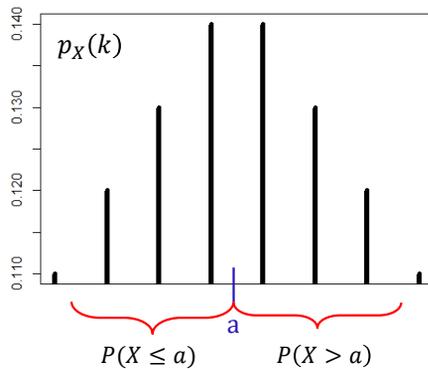


www.matstat.org

Verteilungsfunktion und Komplementsatz

$$P(X > a) = 1 - P(X \leq a)$$

$$P(X > a) = 1 - F_X(a)$$

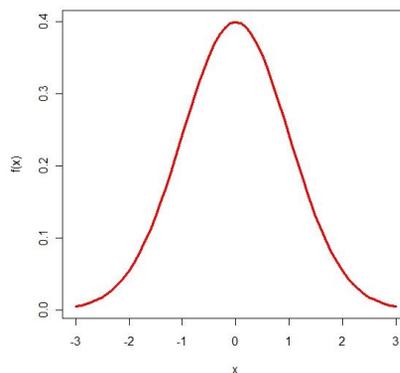


www.matstat.org

Kontinuierliche Zufallsvariablen

Kontinuierliche Zufallsvariable:

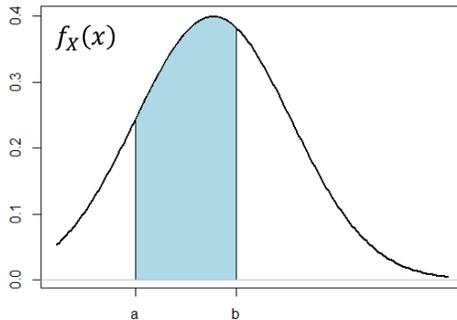
- nicht diskret !
- nimmt reelle Werte an ("unendlich dicht": zwischen zwei beliebigen reellen Werten gibt es unendlich viele reelle Zahlen)
- Lebensdauer einer Glühbirne ; Gewicht ; Windgeschwindigkeit ; elektrische Spannung ; Konzentration einer Chemikalie ; ...



Dichtefunktionen für kontinuierliche Zufallsvariablen

Eine kontinuierliche Zufallsvariable X kann mit Hilfe einer **Dichtefunktion** $f_X(x)$ beschrieben werden.

- X : kontinuierliche Zufallsvariable (Z.v.)
- x : Argument



Definition der Dichtefunktion:

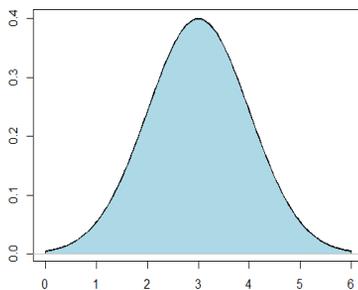
Die Wahrscheinlichkeit, dass die Z. v. X zwischen den reellen Werte a und b landet soll gleich der Fläche unter der Dichtefunktion zwischen a und b sein:

$$P(a < X \leq b) = \int_a^b f_X(x) dx$$

= W., dass X im Intervall (a, b) landet

www.matstat.org

Dichtefunktionen für kontinuierliche Zufallsvariablen



Die Gesamtfläche unter der Dichtefunktion muss 1 sein (**Axiom 2**: $P(\Omega) = 1$)

$$\int_{-\infty}^{+\infty} f_X(x) dx = 1$$

Normierung

dies ist ja die W., dass X zwischen $-\infty$ und $+\infty$ landet - ein Ereignis, welches mit Sicherheit eintreffen muss!

Anmerkungen:

- Wenn der Ereignisraum Ω auf ein gewisses Intervall (u, v) begrenzt ist, können die Integrationsgrenzen in der obigen Formel auf diese Werte gesetzt werden.
- Eine Dichtefunktion kann nur die W. angeben, dass X zwischen zwei Werten landet, dagegen hat eine einzelne reelle Zahl immer die **Wahrscheinlichkeit Null**!

~~$P(X = a)$~~ gibt es nicht für kontinuierliche Zufallsvariablen!

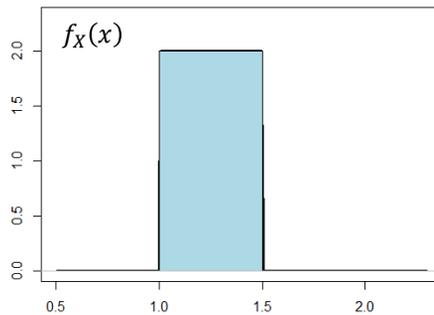
www.matstat.org

Dichtefunktionen für kontinuierliche Zufallsvariablen

Zur Erinnerung: eine Wahrscheinlichkeitsfunktion für eine diskrete Z.v. muss folgende Bedingung erfüllen: $0 \leq p_X(k) \leq 1$

Im Gegensatz dazu kann eine Dichtefunktion Werte annehmen, die größer als 1 sind, solange die Gesamtfläche unter der Kurve Eins ist:

Beispiel: Gleichförmige kontinuierliche Verteilung im Intervall (1, 1.5)



$$\int_{-\infty}^{+\infty} f_X(x) dx = \int_1^{1.5} 2 dx = [2x]_1^{1.5} = \underline{\underline{1}}$$

(viel) einfacher: $2 \cdot 0.5 = 1$

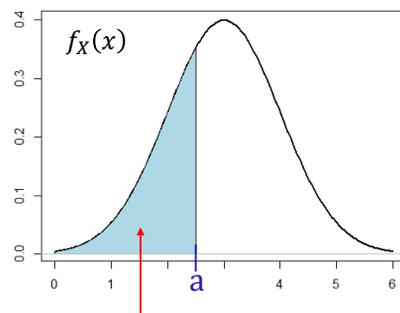
www.matstat.org

Verteilungsfunktionen für kontinuierliche Zufallsvariablen

X : kontinuierliche Zufallsvariable ; F_X : dessen Verteilungsfunktion

$$F_X(a) = P(X \leq a) = \int_{-\infty}^a f_X(x) dx$$

Dadurch wird für jeden Wert von a eine neue Funktion definiert:
Funktionswert = Fläche unter der Dichtefunktion "links" von a .

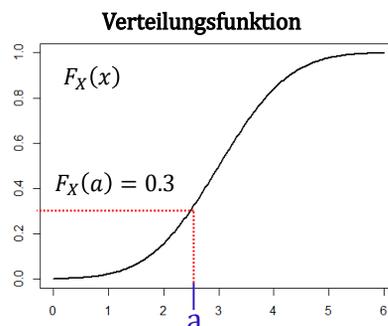
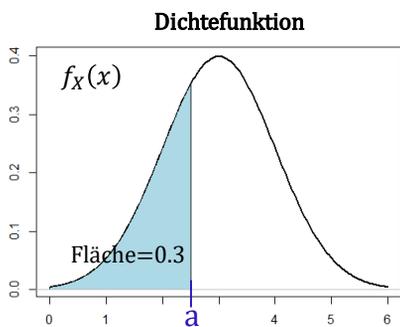


$F_X(a)$ entspricht der blau markierten Fläche

www.matstat.org

Zusammenhang zwischen Dichtefunktion und Verteilungsfunktion

$$F_X(a) = P(X \leq a) = \int_{-\infty}^a f_X(x) dx$$



uwe.menzel@matstat.org

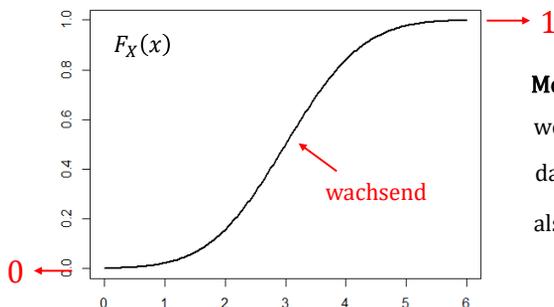
Eigenschaften von Verteilungsfunktionen

X : kontinuierliche Zufallsvariable ; F_X : dessen Verteilungsfunktion

$$\lim_{a \rightarrow -\infty} F_X(a) = 0 \quad "F_X(-\infty) = P(X \leq -\infty)" \quad (\text{letztere Notation ist nicht korrekt})$$

$$\lim_{a \rightarrow +\infty} F_X(a) = 1 \quad "F_X(\infty) = P(X \leq \infty)" \quad (\text{letztere Notation ist nicht korrekt})$$

Die Verteilungsfunktion wächst monoton:



Monotonie:

wenn $a_1 < a_2$

dann $P(X \leq a_1) \leq P(X \leq a_2)$

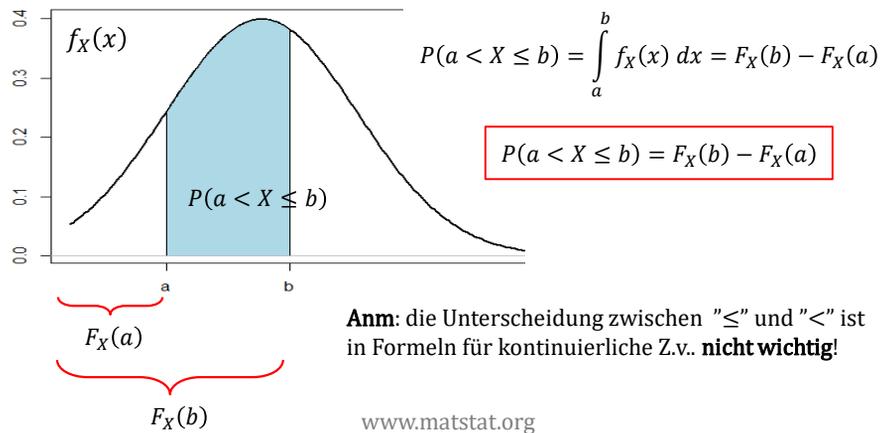
also $F_X(a_1) \leq F_X(a_2)$

www.matstat.org

Berechnung von Wahrscheinlichkeiten mit Hilfe von Verteilungsfunktionen

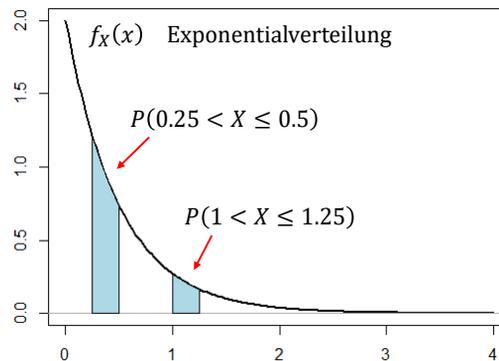
X : kontinuierliche Zufallsvariable ; F_X : dessen Verteilungsfunktion

Zur Erinnerung: die Wahrscheinlichkeit, dass die Z.v. X zwischen den reellen Werten a und b landet ist gleich der Fläche unter der Dichtefunktionen zwischen a und b :



Berechnung von Wahrscheinlichkeiten mit Hilfe von Verteilungsfunktionen

Beispiel: $f_X(x) = \begin{cases} 2 \cdot e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$ Exponentialverteilung



Berechnung von Wahrscheinlichkeiten mit Hilfe von Verteilungsfunktionen

Beispiel, Fortsetzung :

$$f_X(x) = \begin{cases} 2 \cdot e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

a) Stimmt Normierung?: $\int_{-\infty}^{+\infty} f_X(x) dx = 1$??

$$\int_{-\infty}^{+\infty} f_X(x) dx = \int_0^{+\infty} 2 \cdot e^{-2x} dx = \left[2 \cdot \frac{1}{-2} \cdot e^{-2x} \right]_0^{+\infty} = [e^{-2x}]_0^{+\infty} = 1 \quad \checkmark$$

uwe.menzel@matstat.org

Berechnung von Wahrscheinlichkeiten mit Hilfe von Verteilungsfunktionen

Beispiel, Fortsetzung:

$$f_X(x) = \begin{cases} 2 \cdot e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

b) Berechnung der Verteilungsfunktion: $a > 0$ gefordert!

$$F_X(a) = \int_{-\infty}^a f_X(x) dx = \int_0^a 2 \cdot e^{-2x} dx = [-e^{-2x}]_0^a = [e^{-2x}]_a^0 = \underline{\underline{1 - e^{-2a}}}$$

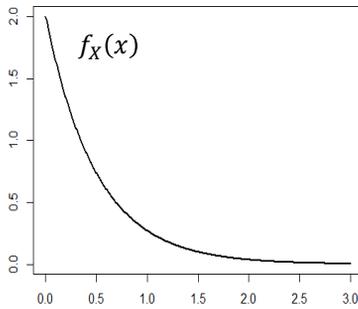
$$F_X(x) = \begin{cases} 1 - e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

uwe.menzel@matstat.org

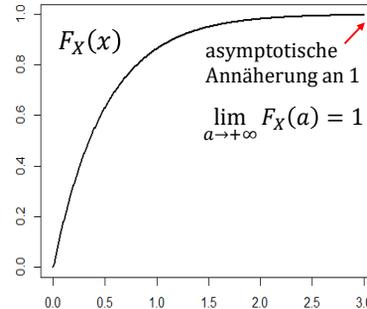
Berechnung von Wahrscheinlichkeiten mit Hilfe von Verteilungsfunktionen

Beispiel, Fortsetzung:

$$f_X(x) = \begin{cases} 2 \cdot e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$



$$F_X(x) = \begin{cases} 1 - e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

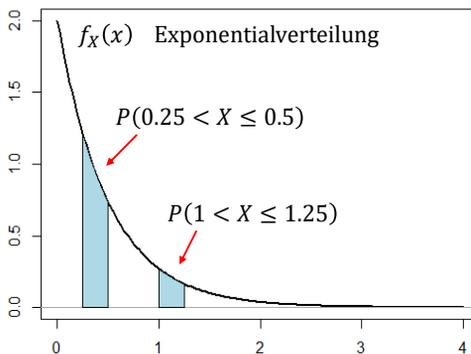


www.matstat.org

Berechnung von Wahrscheinlichkeiten mit Hilfe von Verteilungsfunktionen

Beispiel, Fortsetzung:

$$F_X(x) = \begin{cases} 1 - e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$



$$\begin{aligned} P(0.25 < X \leq 0.5) &= F_X(0.5) - F_X(0.25) \\ &= (1 - e^{-1}) - (1 - e^{-0.5}) \\ &= e^{-0.5} - e^{-1} = \underline{\underline{0.239}} \end{aligned}$$

$$\begin{aligned} P(1 < X \leq 1.25) &= F_X(1.25) - F_X(1) \\ &= (1 - e^{-2.5}) - (1 - e^{-2}) \\ &= e^{-2} - e^{-2.5} = \underline{\underline{0.053}} \end{aligned}$$

uwe.menzel@matstat.org

Der Zusammenhang zwischen Dichtefunktion und Verteilungsfunktion

$$F_X(a) = \int_{-\infty}^a f_X(x) dx \quad \text{Ableitung } \frac{\partial}{\partial a}$$



$$\frac{\partial}{\partial a} F_X(a) = f_X(a)$$

Ableitung eines Parameterintegrals → Leibnizregel:

$$\frac{\partial}{\partial a} \int_{u(a)}^{o(a)} h(x) dx = h[o(a)] \cdot \dot{o}(a) - h[u(a)] \cdot \dot{u}(a)$$

(wenn der Integrand nicht explizit von a abhängt)

Beispiel Exponentialverteilung, Fortsetzung:

$$F_X(a) = 1 - e^{-2a} \quad \text{für } a > 0$$

$$f_X(a) = \frac{\partial}{\partial a} F_X(a) = \frac{\partial}{\partial a} (1 - e^{-2a}) = \underline{\underline{2 \cdot e^{-2a}}} \quad \checkmark \quad (a > 0)$$

www.matstat.org

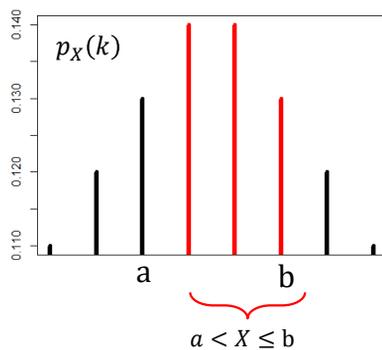
Berechnung von Wahrscheinlichkeiten für diskrete und kontinuierliche Zufallsvariablen

Diskrete Z.v.

$$P(a < X \leq b) = F_X(b) - F_X(a)$$



sehr wichtig!

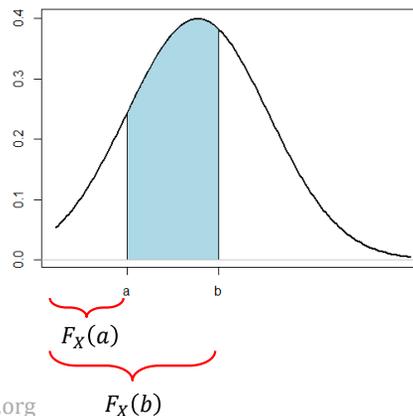


Kontinuierliche Z.v.

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



spielt keine Rolle!



www.matstat.org

Berechnung von Wahrscheinlichkeiten für kontinuierliche Zufallsvariablen

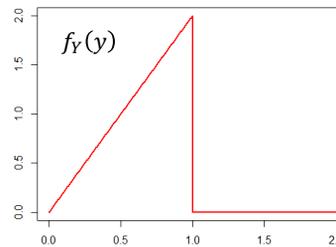
Beispiel: Dichtefunktion $f_Y(y) = \begin{cases} c \cdot y & 0 < y \leq 1 \\ 0 & \text{sonst} \end{cases}$ $c = \text{konstant}$

a) Berechne die Konstante c

$$1 = \int_{-\infty}^{+\infty} f_Y(y) dy = \int_0^1 c \cdot y dy = c \cdot \left[\frac{1}{2} \cdot y^2 \right]_0^1 = c \cdot \frac{1}{2} \quad \Rightarrow \quad \underline{\underline{c = 2}}$$

b) Skizziere die Dichtefunktion

$$f_Y(y) = \begin{cases} 2 \cdot y & 0 < y \leq 1 \\ 0 & \text{sonst} \end{cases}$$



www.matstat.org

Berechnung von Wahrscheinlichkeiten für kontinuierliche Zufallsvariablen

c) Berechne die Verteilungsfunktion für $f_Y(y) = \begin{cases} 2 \cdot y & 0 < y \leq 1 \\ 0 & \text{sonst} \end{cases}$

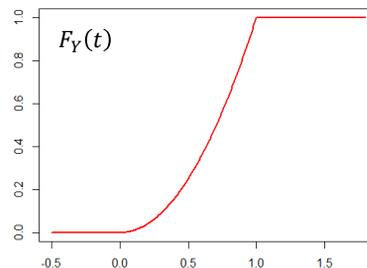
$$F_Y(t) = \int_{-\infty}^t f_Y(y) dy = \int_0^t 2 \cdot y dy = 2 \cdot \left[\frac{1}{2} \cdot y^2 \right]_0^t = t^2 \quad \text{für } 0 < t \leq 1$$

$$F_Y(t) = 0 \quad \text{für } t < 0 \quad (\text{siehe Bild oben})$$

$$F_Y(t) = 1 \quad \text{für } t \geq 1$$



$$F_Y(t) = \begin{cases} 0 & t < 0 \\ t^2 & 0 < t \leq 1 \\ 1 & t > 1 \end{cases}$$



www.matstat.org

Vergleich von diskreten und kontinuierlichen Zufallsvariablen

Diskrete Zufallsvariable	Kontinuierliche Zufallsvariable
Wahrscheinlichkeitsfunktion	Dichtefunktion
$P(X = k) = p_X(k)$	Wahrscheinlichkeit ist Null für einen einzelnen reellen Wert
$0 \leq p_X(k) \leq 1$	$f_X(x)$ kann größer als 1 werden
$\sum_{k \in \Omega} p_X(k) = 1$	$\int_{-\infty}^{+\infty} f_X(x) dx = 1$
$P(A) = \sum_{k \in A} p_X(k)$	$P(a < X \leq b) = \int_a^b f_X(x) dx$

uwe.menzel@matstat.org

Vergleich von diskreten und kontinuierlichen Z.v.

Diskrete Zufallsvariable	Kontinuierliche Zufallsvariable
Verteilungsfunktion	Verteilungsfunktion
$F_X(a) = P(X \leq a)$	$F_X(a) = P(X \leq a)$
$F_X(a) = \sum_{k \leq a} p_X(k)$	$F_X(a) = \int_{-\infty}^a f_X(x) dx$
$P(a < X \leq b) = F_X(b) - F_X(a)$	$P(a < X \leq b) = F_X(b) - F_X(a)$
$p_X(k) = F_X(k) - F_X(k-1)$	$f_X(x) = \frac{d}{dx} F_X(x)$
$P(X > a) = 1 - F_X(a)$	$P(X > a) = 1 - F_X(a)$
$F_X(x)$ monoton wachsend von 0 nach 1	$F_X(x)$ monoton wachsend von 0 nach 1

www.matstat.org