universität innsbruck

Ein statistisches Intermezzo

Grundlagen der Ökonometrie

herbert.stocker@uibk.ac.at

www.hsto.info/econometrics

Statistisches Intermezzo

- Deskriptiv: wir beschreiben, was wir *‘sehen’*.
- Stochastik: wir versuchen zu beschreiben, *‘was sein könnte’*.

- Erfordert abstraktes Modell mit Annahmen!

Statistisches Intermezzo

- Deskriptiv: wir beschreiben, was wir *‘sehen’*.
- Stochastik: wir versuchen zu beschreiben, *‘was sein könnte’*.

- Erfordert abstraktes Modell mit Annahmen!
- Vorgangsweise:
  - Unsicherheit $\rightarrow$ *Zufallsexperiment* (Mengen) $\rightarrow$ mathematische Formulierung: *Wahrscheinlichkeitsraum*
  - Abbildung eines interessierenden Aspekts des Zufallsexperiments in die reellen Zahlen $\rightarrow$ *Zufallsvariable*
  - Darstellung von *Zufallsvariablen* durch (Zähl-)*Dichtefunktionen* (relative Häufigkeiten $\leftrightarrow$ Wahrscheinlichkeiten)
  - Beschreibung von (Zähl-)*Dichtefunktionen* durch $\rightarrow$ *Kennwerte der Verteilung*, z.B. (bedingte) Erwartungswerte, (bedingte) Varianzen, . . .

Statistisches Intermezzo

Zufallsexperiment:

1. alle möglichen Versuchsausgänge, d.h. die Menge aller möglichen *Elementarereignisse* (Ergebnisse) des Experiments sind a priori bekannt;

Statistisches Intermezzo

Zufallsexperiment:

1. alle möglichen Versuchsausgänge, d.h. die Menge aller möglichen *Elementarereignisse* (Ergebnisse) des Experiments sind a priori bekannt;
2. das Ergebnis einer einzelnen Durchführung des Experiments kann nicht mit Sicherheit vorhergesagt werden, aber es gibt eine bestimmte Regelmäßigkeit bei wiederholten Durchführungen; und

Statistisches Intermezzo

Zufallsexperiment:

1. alle möglichen Versuchsausgänge, d.h. die Menge aller möglichen *Elementarereignisse* (Ergebnisse) des Experiments sind a priori bekannt;
2. das Ergebnis einer einzelnen Durchführung des Experiments kann nicht mit Sicherheit vorhergesagt werden, aber es gibt eine bestimmte Regelmäßigkeit bei wiederholten Durchführungen; und
3. das Experiment kann unter identischen Bedingungen beliebig oft wiederholt werden.

Statistisches Intermezzo

ad 1: alle möglichen Versuchsausgänge des Zufallsxperiments
sind a priori bekannt:

- keine *“Unknown unkowns”* (D. Rumsfeld)
- auch bekannt als *Knightsche Unsicherheit* (nach Frank Knight, 1885 – 1972, ein Begründer der Chicagoer Schule der Ökonomie)
- *‘fundamental uncertainty’* kann in diesem Modellrahmen nicht vernünftig behandelt werden! $\rightarrow$ Im Folgenden geht es nur um *kalkulierbares Risiko*!

Statistisches Intermezzo

ad 1: alle möglichen Versuchsausgänge des Zufallsxperiments sind a priori bekannt:

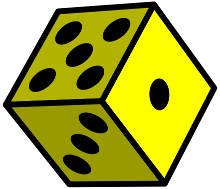
- keine *“Unknown unkowns”* (D. Rumsfeld)
- auch bekannt als *Knightsche Unsicherheit* (nach Frank Knight, 1885 – 1972, ein Begründer der Chicagoer Schule der Ökonomie)
- *‘fundamental uncertainty’* kann in diesem Modellrahmen nicht vernünftig behandelt werden! $\rightarrow$ Im Folgenden geht es nur um *kalkulierbares Risiko*!

ad 3: Experiment kann unter identischen Bedingungen beliebig oft wiederholt werden:

- ebenfalls selten erfüllt, . . . aber oft reicht *Gedankenexperiment*.
- siehe *frequentistische Wahrscheinlichkeitsdefinition*: Wahrscheinlichkeit als relative Häufigkeit einer großen Anzahl gleicher, wiederholter, voneinander unabhängiger Zufallsexperimente.

Grundlegende Konzepte

- **Ergebnisraum eines Zufallsexperiments (bzw. Menge der Elementarereignisse):** Menge aller möglichen Ausgänge eines Zufallsexperiments. Die Menge aller Elementarereignisse wird in der Literatur häufig mit Ω bezeichnet. So ist z.B. für das Werfen eines Würfels $\Omega = \{1, 2, 3, 4, 5, 6\}$, und die einzelnen Elementarereignisse sind $\{1\}$, $\{2\}$, $\{3\}$, . . . , $\{6\}$.



Grundlegende Konzepte

- **Ereignis**: eine beliebige Teilmenge des Ergebnisraums. Ein Ereignis setzt sich aus einem oder mehreren Elementarereignissen zusammen. Beispielsweise setzt sich beim Würfeln das Ereignis “Werfen einer geraden Augenzahl” $A = \{2, 4, 6\}$ aus den Elementarereignissen $\{2\}$, $\{4\}$ und $\{6\}$ zusammen.

Grundlegende Konzepte

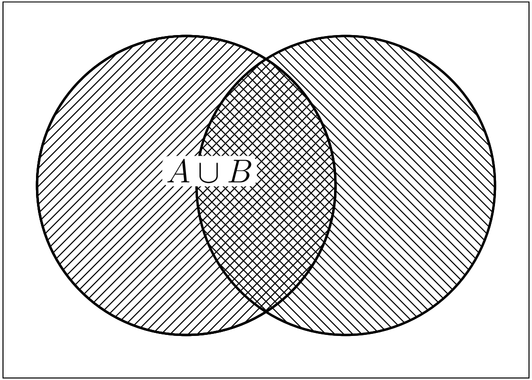
- **Ereignis**: eine beliebige Teilmenge des Ergebnisraums. Ein Ereignis setzt sich aus einem oder mehreren Elementarereignissen zusammen. Beispielsweise setzt sich beim Würfeln das Ereignis “Werfen einer geraden Augenzahl” $A = \{2, 4, 6\}$ aus den Elementarereignissen $\{2\}$, $\{4\}$ und $\{6\}$ zusammen.

- Damit ist es z.B. möglich die Vereinigung zweier Ereignisse A und B ($A \cup B$), den Durchschnitt zweier Ereignisse ($A \cap B$) oder die Komplementärmenge zu definieren.

Grundlegende Konzepte

Vereinigungsmenge:

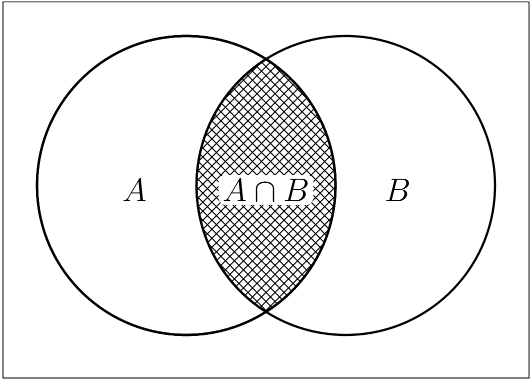
$$A \cup B := \{x \colon x \in A \text{ oder } x \in B\}$$



Grundlegende Konzepte

Durchschnittsmenge:

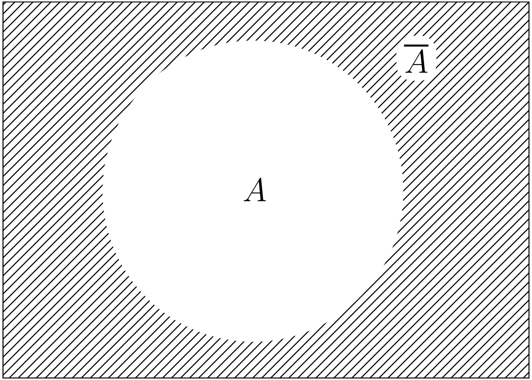
$$A \cap B := \{x \colon x \in A \text{ und } x \in B\}$$



Grundlegende Konzepte

Komplementärmenge:

$$\overline{A} := \{x : x \in \Omega \text{ und } x \notin A\}$$



Grundlegende Konzepte

- **Wahrscheinlichkeit**: ein Maß zur Quantifizierung der Sicherheit bzw. Unsicherheit eines Zufallsexperiments.
- Frequentistischer Wahrscheinlichkeitbegriff: Grenzwert der relativen Häufigkeit, wenn die Anzahl der ZF-Experimente gegen Unendlich geht

$$P(A) = \lim_{n \to \infty} \left(\frac{n_A}{n} \right)$$

- Bayesscher Wahrscheinlichkeitbegriff: *“probability is viewed as representing a degree of reasonable belief with the limiting values of zero being complete disbelief or disproof and of one being complete belief or proof.”* (Zellner 1984)

Grundlegende Konzepte

Axiomatische Wahrscheinlichkeitsdefinition (*A.N. Kolmogorov*) versucht nicht das ‘Wesen’ von Wahrscheinlichkeit zu ‘erklären’, sondern definiert deren mathematische Eigenschaften.
Ermöglicht *‘Rechnen’* mit Wahrscheinlichkeiten!

Drei Axiome:

1. Die Wahrscheinlichkeit $P(A)$ des Ereignisses A ist eine reelle, nichtnegative Zahl, die die Bedingung $0 \leq P(A) \leq 1$ erfüllt.
2. Wenn ein Ereignis Ω alle Elementarereignisse eines Zufallsexperiments enthält ist Ω ein sicheres Ereignis mit $P(\Omega) = 1$.
3. Für sich paarweise gegenseitig ausschließende Ereignisse gilt $P(A_1 \cup A_2 \cup \ldots \cup A_N) = P(A_1) + P(A_2) + \cdots + P(A_N)$.

Grundlegende Konzepte

- **Additionssatz**

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

- **Bedingte Wahrscheinlichkeit**: Die Wahrscheinlichkeit für das Eintreten von A unter der Bedingung, dass Ereignis B bereits eingetreten ist oder gleichzeitig eintritt, wird *bedingte Wahrscheinlichkeit* $P(A|B)$ genannt. Sie ist (für $P(B) > 0$) definiert als

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Der *Multiplikationssatz* $P(A \cap B) = P(B) \cdot P(A|B)$ ermöglicht die Berechnung der Wahrscheinlichkeit für das Eintreten von A *und* B ($P(A \cap B)$)

Grundlegende Konzepte

- **Multiplikationssatz**

  $$P(A \cap B) = P(B) \cdot P(A|B)$$

  ermöglicht die Berechnung der Wahrscheinlichkeit für das Eintreten von A *und* B ($P(A \cap B)$)

- Zwei Ereignisse A und B heißen **stochastisch unabhängig**, wenn das Eintreten des Ereignisses A nicht vom Eintreten oder Nichteintreten des Ereignisses B abhängt, d.h. wenn $P(A|B) = P(A)$.
  z.B.: Die Wahrscheinlichkeit bei zweimaligen Würfeln zwei Sechsen zu erhalten ist $\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$

Grundlegende Konzepte

Wahrscheinlichkeitsraum: Das Triple

$$[\Omega, \mathcal{A}, P(\cdot)]$$

bildet einen *Wahrscheinlichkeitsraum* (*'probability space'*), d.h. die mathematische Beschreibung des Zufallsexperiments ($\rightarrow$ Messbarkeit, Maßraum).

- Ω: *Ergebnismenge* oder Menge aller möglichen *Elementarereignisse*
- $\mathcal{A}$: **Ereignisraum**, enthält alle Mengen, denen man eine Wahrscheinlichkeit zuweisen will. Enthält neben interessierenden Ereignisse auch relevante mit den interessierenden Ereignissen über Mengenoperationen verknüpfte Mengen sowie Ω und die Nullmenge; auch σ-Algebra genannt (σ: sigma).
- $P(\cdot)$: Eine Mengenfunktion, die jeder Teilmenge von $\mathcal{A}$ eine Zahl zwischen Null und Eins zuordnet.

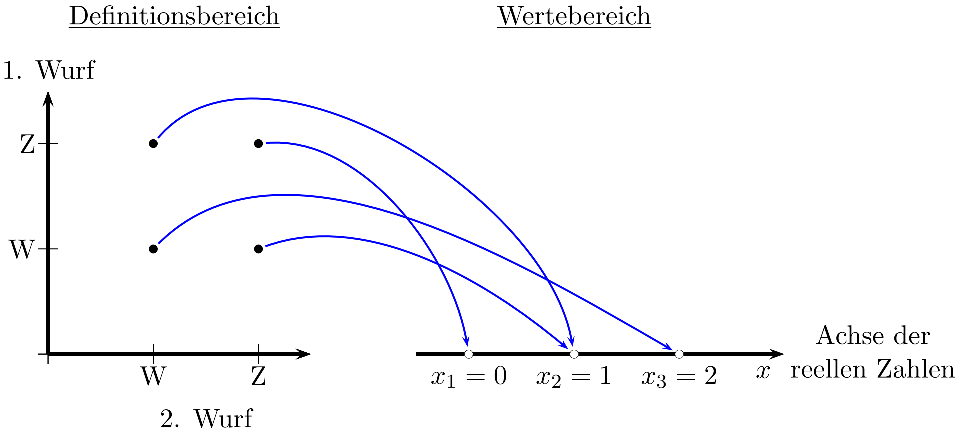
Axiomatische Wahrscheinlichkeitstheorie erlaubt Abbildung der interessierenden Ereignisse in $\mathbb{R}$ $\rightarrow$ *Zufallsvariable*

Grundlegende Konzepte

- Eine **Zufallsvariable (ZV)** ist eine Funktion, die den Ergebnissen eines Zufallsexperimentes (d.h. den Elementarereignissen oder Ereignissen) reelle Zahlen zuordnet.
- Eine ZV bildet *alle möglichen* Ergebnisse eines Zufallsexperimentes in die reellen Zahlen ab!
- Zum Beispiel: Münzwurf mit Elementarereignissen {Zahl, Wappen}, d.h. $\Omega = \{Z, W\}$
  Die Zufallsvariable könnte dem Ereignis $\{Z\}$ den Wert 0 mit Wahrscheinlichkeit 0.5 zuordnen, und dem Ereignis $\{W\}$ den Wert 1 ebenfalls mit Wahrscheinlichkeit 0.5.
- Abbildung *aller möglichen* Ergebnisse in *(Zähl-)Dichtefunktionen*.

Grundlegende Konzepte

Zufallsexperiment “Zweimaliges Werfen einer Münze”; **Zufallsvariable** X = “Anzahl der Wappen” kann die Werte $x_1 = 0$, $x_2 = 1$, oder $x_3 = 2$ annehmen



Grundlegende Konzepte

- **Wahrscheinlichkeitsfunktion** einer *diskreten ZV* X: die Funktion $f(x_j)$, die für jede Ausprägung der ZV die Wahrscheinlichkeit ihres Auftretens angibt

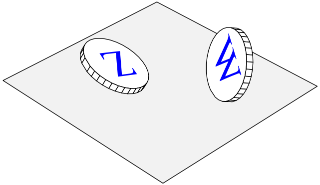
$$f(x_j) = \Pr(X = x_j) \qquad \text{für } j = 1, 2, \ldots k$$

- Das Analogon für *stetige ZV* heißt **Dichtefunktion**
Achtung: für stetige ZV ist $\Pr(X = x) = 0 \neq f(x)$

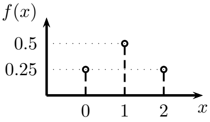
$$\Pr(a \leq X \leq b) = \int_a^b f(x)\,dx$$

Zufallsexperiment: 2-facher Münzwurf

Zufallsvariable: Anzahl Wappen



x	Elemente im Ergebnisraum	$f(x)$
0	(ZZ)	0.25
1	(ZW), (WZ)	0.5
2	(WW)	0.25



$$f(x) = \begin{cases} 0.25 & \text{für } X = 0 \\ 0.5 & \text{für } X = 1 \\ 0.25 & \text{für } X = 2 \\ 0 & \text{sonst} \end{cases}$$

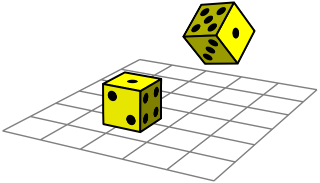
Grundlegende Konzepte

Jede Wahrscheinlichkeits- bzw. Dichtefunktion muss folgende beiden Eigenschaften erfüllen:

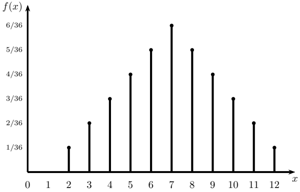
- **Diskrete ZV:**
  1. $f(x_j) = \Pr(X = x_j) \geq 0$ für $j = 1, 2, \dots, k$ (mit k Ausprägungen)
  2. $\sum_{j=1}^{k} f(x_k) = 1$

- **Stetige ZV:**
  1. $f(x) \geq 0, \qquad \Pr(a \leq X \leq b) = f(a \leq X \leq b) = \int_a^b f(x)\, dx$
  2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$

Beispiel: ZV: "Augensumme beim zweimaligen Würfeln"



x	Elemente im Ergebnisraum	$f(x)$
2	1 1	1/36
3	1 2, 2 1	2/36
4	1 3, 3 1, 2 2	3/36
5	1 4, 4 1, 2 3, 3 2	4/36
6	1 5, 5 1, 2 4, 4 2, 3 3	5/36
7	1 6, 6 1, 2 5, 5 2, 3 4, 4 3	6/36
8	2 6, 6 2, 3 5, 5 3, 4 4	5/36
9	3 6, 6 3, 4 5, 5 4	4/36
10	4 6, 6 4, 5 5	3/36
11	5 6, 6 5	2/36
12	6 6	1/36

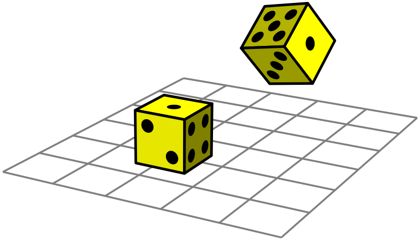


$$f(x) = \Pr(x) = \begin{cases} 1/36 & \text{für } X = 2 \\ 2/36 & \text{für } X = 3 \\ \vdots & \vdots \\ 1/36 & \text{für } X = 12 \\ 0 & \text{sonst.} \end{cases}$$

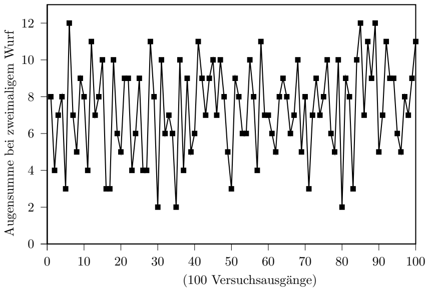
ZV beschreibt alle möglichen Ausgänge der Zufallsexperiments!!!

Deskriptiv: Empirische Verteilung

Beispiel: **“Augensumme beim zweimaligen Würfeln”**

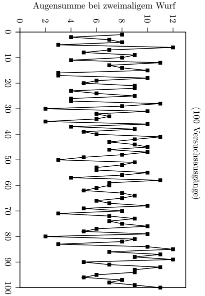


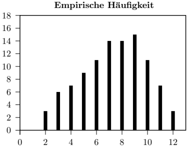
Deskriptive Statistik:
100 **Realisationen**



Empirische vs. theoretische Verteilung

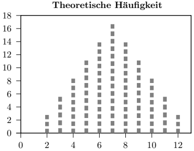
Deskriptiv: 100 Versuche





Zufallsvariable:

x	Elemente im Ergebnisraum	$f(x)$
2	1 1	1/36
3	1 2, 2 1	2/36
4	1 3, 3 1, 2 2	3/36
5	1 4, 4 1, 2 3, 3 2	4/36
6	1 5, 5 1, 2 4, 4 2, 3 3	5/36
7	1 6, 6 1, 2 5, 5 2, 3 4, 4 3	6/36
8	2 6, 6 2, 3 5, 5 3, 4 4	5/36
9	3 6, 6 3, 4 5, 5 4	4/36
10	4 6, 6 4, 5 5	3/36
11	5 6, 6 5	2/36
12	6 6	1/36



$$f(x) = \Pr(x) = \begin{cases} 1/36 & \text{für } X = 2 \\ 2/36 & \text{für } X = 3 \\ \vdots & \vdots \\ 1/36 & \text{für } X = 12 \\ 0 & \text{sonst.} \end{cases}$$

Deskriptive vs. Induktive Statistik

Deskriptiv	*Stochastisch*
Histogramm	(Zähl-)Dichtefunktion
Emp. Verteilungsfunktion	Verteilungsfunktion
(bedingter) Mittelwert	(bedingter) Erwartungswert
(bedingte) emp. Varianz	(bedingte) Varianz einer ZV

Grundlegende Konzepte

- Eine (kumulative) **Verteilungsfunktion** $F(x_l)$ (*cumulative distribution function*) gibt die Wahrscheinlichkeit dafür an, dass eine Zufallsvariable X *höchstens* den Wert x_l annimmt.
- Für diskrtete ZV: Wenn die x_j aufsteigend nach ihrem Wert geordnet sind gilt

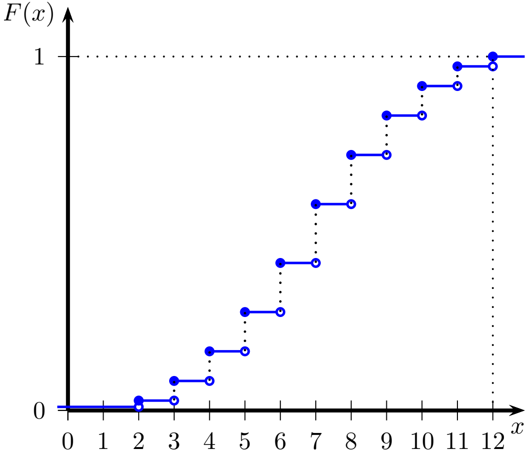
$$F(x_l) = \Pr(X \leq x_l) \ = \sum_{j=1}^{l} f(x_j)$$

- Für stetige ZV:

$$F(x) = \int_{-\infty}^{x} f(v)dv$$

Beispiel: Verteilungsfunktion der diskreten ZV *“Augensumme beim zweimaligen Würfeln”*:

x	$f(x)$	$F(x)$
2	1/36	1/36
3	2/36	3/36
4	3/36	6/36
5	4/36	10/36
6	5/36	15/36
7	6/36	21/36
8	5/36	26/36
9	4/36	30/36
10	3/36	33/36
11	2/36	35/36
12	1/36	36/36

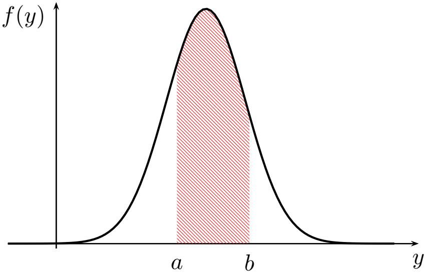


Grundlegende Konzepte

Dichtefunktionen stetiger Zufallsvariablen:

Wenn $f(x)$ eine Dichtefunktion ist, dann ist die Wahrscheinlichkeit dafür, dass X einen Wert in einem beliebigen Intervall $[a, b]$ (mit $a < b$ und $a, b \in \mathbb{R}$) annimmt, gleich

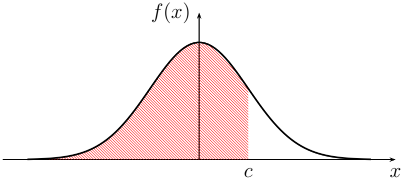
$$\Pr(a < X < b) = \int_a^b f(x)dx$$



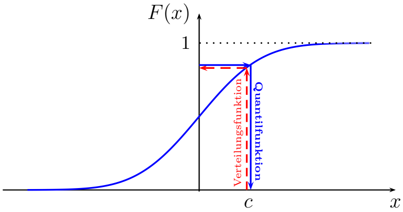
Grundlegende Konzepte

Verteilungs- und Quantilfunktion einer stetigen ZV:

Dichtefunktion:

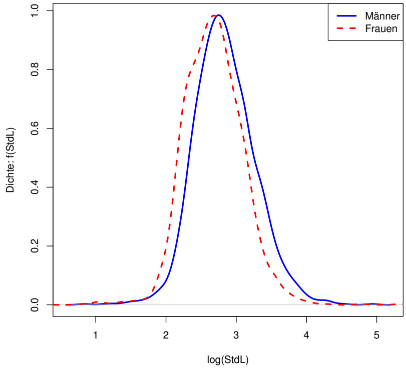


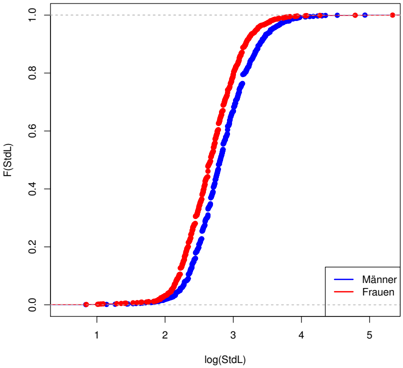
Verteilungsfunktion: (kumulierte Dichte)



Grundlegende Konzepte

Beispiel: Dichte- & Verteilungsfunktion der $\log$(Stundenlöhne) für Österreich 2018 nach Geschlecht





Datenquelle: EU-Silc, Statistik Austria

Grundlegende Konzepte

- Häufig interessieren wir uns dafür, ob bzw. wie zwei ZV zusammenhängen (z.B. Körpergröße X und Gewicht Y einer zufällig gewählten Person.)

- **Gemeinsame Wahrscheinlichkeitsfunktion (WF):** Die gemeinsame Wahrscheinlichkeitsfunktion zweier diskreten ZV X und Y ist

  $$f(x, y) = \Pr(X = x, Y = y)$$

  gibt die Wahrscheinlichkeit dafür an, dass die diskrete Zufallsvariable X den Wert x_i und die diskrete Zufallsvariable Y gleichzeitig den Wert y_j annimmt.

Grundlegende Konzepte

Für zwei diskrete ZV X und Y ist die gemeinsame Wahrscheinlichkeitsfunktion (WF)

$$f(x, y) = \Pr(X = x, Y = y)$$

Grundlegende Konzepte

Für zwei diskrete ZV X und Y ist die gemeinsame Wahrscheinlichkeitsfunktion (WF)

$$f(x,y) = \Pr(X = x, Y = y)$$

	y_1	y_2	$\cdots$	y_M
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	$\cdots$	$f(x_1, y_M)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	$\cdots$	$f(x_2, y_M)$
x_3	$f(x_3, y_1)$	$f(x_3, y_2)$	$\cdots$	$f(x_3, y_M)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	$f(x_n, y_1)$	$f(x_n, y_2)$	$\cdots$	$f(x_n, y_M)$

Grundlegende Konzepte

Beispiel für eine Gemeinsame WF:

		Werte von Y		
		0	1	$f_X(x)$
	1	0	1/3	1/3
Werte von X	2	1/3	0	1/3
	3	0	1/3	1/3
	$f_Y(y)$	1/3	2/3	1

Randverteilungen:

$$f_X(X) = \sum_y f(x,y) \quad \text{Randverteilung von } X$$

$$f_Y(Y) = \sum_x f(x,y) \quad \text{Randverteilung von of } Y$$

Grundlegende Konzepte

Beispiel für eine Gemeinsame WF:

		Werte von Y		
		0	1	$f_X(x)$
	1	0	1/3	1/3
Werte von X	2	1/3	0	1/3
	3	0	1/3	1/3
	$f_Y(y)$	1/3	2/3	1

Randverteilungen:

$$f_X(X) = \sum_y f(x,y) \qquad \text{Randverteilung von } X$$

$$f_Y(Y) = \sum_x f(x,y) \qquad \text{Randverteilung von of } Y$$

Grundlegende Konzepte

Beispiel für eine Gemeinsame WF:

		Werte von Y		
		0	1	$f_X(x)$
	1	0	1/3	1/3
Werte von X	2	1/3	0	1/3
	3	0	1/3	1/3
	$f_Y(y)$	1/3	2/3	1

Randverteilungen:

$$\begin{aligned} f_X(x=3) &= \sum_y f(x,y) = 0 + 1/3 = 1/3 \\ f_Y(y=1) &= \sum_x f(x,y) = 1/3 + 0 + 1/3 = 2/3 \end{aligned}$$

Grundlegende Konzepte

Bedingte Wahrscheinlichkeitsfunktionen:

- Bedingte Wahrscheinlichkeitsfunktionen von X:

$$f(x|y) = \Pr(X = x|Y = y) = \frac{f(x,y)}{f_Y(y)}$$

- Bedingte Wahrscheinlichkeitsfunktionen von Y:

$$f(y|x) = \Pr(Y = y|X = x) = \frac{f(x,y)}{f_X(x)}$$

Grundlegende Konzepte

Beispiel:

		Werte von Y		
		0	1	$f_X(x)$
	1	0	1/3	1/3
Werte von X	2	1/3	0	1/3
	3	0	1/3	1/3
	$f_Y(y)$	1/3	2/3	1

Bedingte Wahrscheinlichkeiten: (Zähldichten)

$$f(x=3|y=1) \quad = \quad \frac{f(x=3, y=1)}{f_Y(y=1)} = \frac{1/3}{2/3} = \frac{1}{2}$$

$$f(y=0|x=2) \quad = \quad \frac{f(y=0, x=2)}{f_X(x=2)} = \frac{1/3}{1/3} = 1$$

Grundlegende Konzepte

Beispiel:

		Werte von Y		
		0	1	$f_X(x)$
	1	0	1/3	1/3
Werte von X	2	1/3	0	1/3
	3	0	1/3	1/3
	$f_Y(y)$	1/3	2/3	1

Bedingte Wahrscheinlichkeiten: (Zähldichten)

$$f(x = 3|y = 1) = \frac{f(x = 3, y = 1)}{f_Y(y = 1)} = \frac{1/3}{2/3} = \frac{1}{2}$$

$$f(y = 0|x = 2) = \frac{f(y = 0, x = 2)}{f_X(x = 2)} = \frac{1/3}{1/3} = 1$$

Grundlegende Konzepte

Statistische (oder stochastische) Unabhängigkeit von zwei ZV:

Zwei Zufallsvariablen X und Y sind statistisch unabhängig, wenn und nur wenn

$$f(x, y) = f_X(x) f_Y(y)$$

bzw. für diskrete ZV

$$\Pr(X = x_i, Y = y_j) = \Pr(X = x_i) \Pr(Y = y_j)$$

(für alle X and Y)

Grundlegende Konzepte

Beispiel: Bedingte WF:

		Werte von Y		
		0	1	$f_X(x)$
	1	0	1/3	1/3
Werte von X	2	1/3	0	1/3
	3	0	1/3	1/3
	$f_Y(y)$	1/3	2/3	1

X und Y sind **nicht statistisch unabhängig**, da z.B.

$$f(X=1 \text{ und } Y=0) = 0 \neq f_X(X=1)f_Y(Y=0) = \frac{1}{3} \times \frac{1}{3}$$

(wir werden später zeigen, dass die Korrelation zwischen X und Y trotzdem Null ist, d.h. $\mathrm{cov}(X,Y)=0$!!!)

Grundlegende Konzepte

- Wahrscheinlichkeitsfunktionen sind wie normale Häufigkeitsverteilungen durch bestimmte Parameter charakterisiert.
- Der **Erwartungswert** ($\mathrm{E}(X)$ oder μ) einer Zufallsvariablen X ist jener Wert, der sich (in der Regel) bei oftmaligem Wiederholen des zugrunde liegenden Experiments als Mittelwert der Ergebnisse ergibt.
- Er bestimmt die Lage einer Verteilung und ist das Analogon zum arithmetischen Mittel einer empirischen Häufigkeitsverteilung in der deskriptiven Statistik.
- μ ist keine ZV, sondern eine fixe Zahl.

Grundlegende Konzepte

Erwartungswert

Der **Erwartungswert** ist die *mit den Eintrittswahrscheinlichkeiten gewichtete* Summe aller möglichen Ausprägungen einer Zufallsvariable.

$$\begin{aligned} \mathrm{E}(X) &= \sum_{j=1}^{k} x_j f(x_j) \quad \text{für diskrete ZV} \\ &= \sum_{j=1}^{k} x_j \Pr(X = x_j) \\ \mathrm{E}(X) &= \int_{-\infty}^{\infty} x f(x)\, dx \quad \text{für stetige ZV} \end{aligned}$$

Grundlegende Konzepte

Beispiel für diskrete ZV: Erwartungswert der Augenzahl beim Würfeln:

$$\begin{aligned} \mathrm{E}(X) &\equiv \mu_X = x_1 f(x_1) + x_2 f(x_2) + \ldots + x_6 f(x_6) \\ &= 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6} = 3.5 \end{aligned}$$

Grundlegende Konzepte

Rechnen mit Erwartungswerten:

Mit Erwartungswerten kann 'sehr ähnlich' gerechnet werden wie mit dem Summenzeichen. So gilt z.B. für konstante a und b

$$\begin{aligned}
\mathrm{E}(a) &= a \qquad \text{für a = const.} \\
\mathrm{E}(aX) &= a\,\mathrm{E}(X) \qquad \text{für a const.} \\
\mathrm{E}\left[\mathrm{E}(X)\right] &= \mathrm{E}(X) \quad (\text{bzw. } \mathrm{E}(\mu) = \mu \\
\mathrm{E}\left[g(X)\right] &= \int_{-\infty}^{\infty} g(x) f(x)\, dx \quad \text{für eine Funktion } g(\cdot) \\
\mathrm{E}(a + bX) &= a + b\,\mathrm{E}(X)
\end{aligned}$$

Grundlegende Konzepte

Beispiel: Wenn X die Augenzahl eines fairen Würfels ist, wie groß ist der Erwartungswert von $g(X) = X^2$?

$$\mathrm{E}(X^2) = 1^2\,\frac{1}{6} + 2^2\,\frac{1}{6} + 3^2\,\frac{1}{6} + 4^2\,\frac{1}{6} + 5^2\,\frac{1}{6} + 6^2\,\frac{1}{6} = 15.1\dot{6}$$

Achtung: $\mathrm{E}(X^2) = 15.1\dot{6} \;\neq\; [\mathrm{E}(X)]^2 = 3.5^2 = 12.25$.

Grundlegende Konzepte

Rechnen mit Erwartungswerten:

- Der Erwartungswert einer Summe von ZV ist gleich der Summe der Erwartungswerte, d.h.
  $\mathrm{E}(X + Y + Z + \cdots) = \mathrm{E}(X) + \mathrm{E}(Y) + \mathrm{E}(Z) + \cdots$
- Wenn – *und nur wenn* – X und Y zwei *stochastisch unabhängige* ZV sind, dann ist der Erwartungswert ihres Produktes gleich dem Produkt der Erwartungswerte, d.h.
  $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$.

Grundlegende Konzepte

Varianz

Die **Varianz** σ_X^2 der Grundgesamtheit ist definiert als

$$\mathrm{var}(X) \equiv \sigma_X^2 = \mathrm{E}\left[X - \mathrm{E}(X)\right]^2$$

Für diskrete Zufallsvariablen

$$\sigma_X^2 = \sum_i (x_i - \mu)^2 f(x_i)$$

Grundlegende Konzepte

Beispiel: Varianz der Augenzahl bei einem einmaligen Wurf mit einem fairen Würfel:

$$
\begin{aligned}
\operatorname{var}(X) &= \mathrm{E}\left[X - \mathrm{E}(X)\right]^2 \\
&= \mathrm{E}\left[x_i - \mu\right]^2 \\
&= \sum_{i=1}^{6} (x_i - \mu)^2 \, f(x) \\
&= (1 - 3.5)^2 \, \frac{1}{6} + (2 - 3.5)^2 \, \frac{1}{6} + \cdots + (6 - 3.5)^2 \, \frac{1}{6} \\
&= 2.91667
\end{aligned}
$$

Grundlegende Konzepte

Die Varianz kann auch folgendermaßen berechnet werden:

$$\begin{aligned}
\sigma_X^2 &= \mathrm{E}\left[X - \mathrm{E}(X)\right]^2 = \mathrm{E}(X-\mu)^2 = \mathrm{E}(X^2 - 2\mu X + \mu^2) \\
&= \mathrm{E}(X^2) - 2\mu\,\mathrm{E}(X) + \mu^2 = \mathrm{E}(X^2) - 2\mu\mu + \mu^2 \\
&= \mathrm{E}(X^2) - \mu^2 \\
&= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2
\end{aligned}$$

da $\mathrm{E}(X) := \mu$

Beispiel: Varianz der Augenzahl bei einmaligem Wurf:

$$\mathrm{var}(X) = (1^2 + 2^2 + \cdots + 6^2)/6 - 3.5^2 = 91/6 - 3.5^2 = 2.91667$$

Grundlegende Konzepte

Rechnen mit Varianzen:

- $\operatorname{var}(a) = 0$
- $\operatorname{var}(a + bX) = b^2 \operatorname{var}(X)$
- $\operatorname{var}(X + Y) = \operatorname{var}(X) + \operatorname{var}(Y) + 2\operatorname{cov}(X, Y)$ bzw.,

  $\operatorname{var}(X - Y) = \operatorname{var}(X) + \operatorname{var}(Y) - 2\operatorname{cov}(X, Y)$

Grundlegende Konzepte

Kovarianz

Die **Kovarianz** ist eine (nicht standardisierte) Maßzahl für den (linearen) Zusammenhang zweier Zufallsvariablen mit gemeinsamer Wahrscheinlichkeitsverteilung

$$\operatorname{cov}(X,Y) = \mathrm{E}\left[X - \mathrm{E}(X)\right]\left[Y - \mathrm{E}(Y)\right]$$

Alternativ:

$$\operatorname{cov}(X,Y) = \mathrm{E}[X \times Y] - \mathrm{E}[X] \times \mathrm{E}[Y]$$

weil:

$$\operatorname{cov}(X,Y) = \mathrm{E}[XY - Y\,\mathrm{E}(X) - X\,\mathrm{E}(Y) + \mathrm{E}(X)\,\mathrm{E}(Y)] =$$
$$\mathrm{E}(XY) - 2\,\mathrm{E}(X)\,\mathrm{E}(Y) + \mathrm{E}(X)\,\mathrm{E}(Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$$

Grundlegende Konzepte

Beispiel:

Zeigen Sie, dass für konstante a und b gilt

$$\operatorname{cov}[X, (a + bX)] = b \operatorname{var}(X)$$

Beispiel:

Zeigen Sie, dass für konstante a und b gilt

$$\operatorname{cov}[X,(a+bX)] = b\operatorname{var}(X)$$

Lösung:

$$\begin{aligned}
\operatorname{cov}[X,(a+bX)] &:= \mathrm{E}[X-\mathrm{E}(X)][(a+bX)-\mathrm{E}(a+bX)] \\
&= \mathrm{E}[X-\mathrm{E}(X)][(a+bX)-\mathrm{E}(a)-\mathrm{E}(bX)] \\
&= \mathrm{E}[X-\mathrm{E}(X)][(a-a)+b(X-\mathrm{E}(X)] \\
&= \mathrm{E}[X-\mathrm{E}(X)]b[(X-\mathrm{E}(X)] \\
&= b\,\mathrm{E}[X-\mathrm{E}(X)][(X-\mathrm{E}(X)] \\
&= b\,\mathrm{E}[X-\mathrm{E}(X)]^2 \\
&= b\operatorname{var}(X)
\end{aligned}$$

Grundlegende Konzepte

Beispiel: Zeigen Sie

$$\mathrm{cov}[X, (Y + Z)] = \mathrm{cov}(X, Y) + \mathrm{cov}(X, Z)$$

Grundlegende Konzepte

Beispiel: Zeigen Sie

$$\mathrm{cov}[X,(Y+Z)] = \mathrm{cov}(X,Y) + \mathrm{cov}(X,Z)$$

Lösung:

$$\begin{aligned}
\mathrm{cov}[X,(Y+Z)] &:= \mathrm{E}[X-\mathrm{E}(X)][(Y+Z)-\mathrm{E}(Y+Z)] \\
&= \mathrm{E}[X-\mathrm{E}(X)][(Y+Z)-\mathrm{E}(Y)-\mathrm{E}(Z)] \\
&= \mathrm{E}[X-\mathrm{E}(X)][(Y-\mathrm{E}(Y)+(Z-\mathrm{E}(Z)] \\
&= \mathrm{E}[X-\mathrm{E}(X)][(Y-\mathrm{E}(Y)]+ \\
&\quad\ \mathrm{E}[X-\mathrm{E}(X)][(Z-\mathrm{E}(Z)] \\
&= \mathrm{cov}(X,Y) + \mathrm{cov}(X,Z)
\end{aligned}$$

Beispiel

Gegeben sei folgende diskrete Wahrscheinlichkeitsverteilung:

		Werte von Y			
		-1	0	+1	
	1	3/24	1/24	2/24	
Werte von X	2	6/24	6/24	0	
	3	3/24	1/24	2/24	

Beispiel

Erwartungswert und Varianz von X:

		Werte von Y			
		-1	0	+1	$f_X(x)$
	1	3/24	1/24	2/24	6/24
Werte von X	2	6/24	6/24	0	12/24
	3	3/24	1/24	2/24	6/24
	$f_Y(y)$	12/24	8/24	4/24	1

$$
\begin{aligned}
\mathrm{E}(X) &= 1 \times 6/24 + 2 \times 12/24 + 3 \times 6/24 = 2 \\
\mathrm{E}(X^2) &= 1^2 \times 6/24 + 2^2 \times 12/24 + 3^2 \times 6/24 = 108/24 = 4.5 \\
\mathrm{var}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = 4.5 - 2^2 = 0.5 \\
\mathrm{var}(Y) &= 5/9 \qquad \text{(warum?)}
\end{aligned}
$$

Beispiel

		Werte von Y			
		-1	0	+1	$f_X(x)$
	1	3/24	1/24	2/24	6/24
Werte von X	2	6/24	6/24	0	12/24
	3	3/24	1/24	2/24	6/24
	$f_Y(y)$	12/24	8/24	4/24	1

$$
\begin{aligned}
\mathrm{E}(XY) &= 1 \times (-1) \times 3/24 + 1 \times 0 \times 1/24 + 1 \times 1 \times 2/24 + \\
&\quad 2 \times (-1) \times 6/24 + 2 \times 0 \times 6/24 + 2 \times 1 \times 0 + \\
&\quad 3 \times (-1) \times 3/24 + 3 \times 0 \times 1/24 + 3 \times 1 \times 2/24 \\
&= -16/24
\end{aligned}
$$

$$\mathrm{cov}(XY) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = -16/24 - 2 \times (-8/24) = 0 \;!$$

Beispiel

		Werte von Y			
		-1	0	+1	$f_X(x)$
Werte von X	1	3/24	1/24	2/24	6/24
	2	6/24	6/24	0	12/24
	3	3/24	1/24	2/24	6/24
	$f_Y(y)$	12/24	8/24	4/24	1

Bedingte Erwartungswertfunktion (CEF): $\mathrm{E}(Y|X)$

$$\mathrm{E}(Y|X) = \begin{cases} -\frac{1}{6} & = -1 \times \frac{3}{6} + 0 \times \frac{1}{6} + 1 \times \frac{2}{6} & \text{für } X = 1 \\ -\frac{1}{2} & = -1 \times \frac{6}{12} + 0 \times \frac{6}{12} + 1 \times \frac{0}{12} & \text{für } X = 2 \\ -\frac{1}{6} & = -1 \times \frac{3}{6} + 0 \times \frac{1}{6} + 1 \times \frac{2}{6} & \text{für } X = 3 \end{cases}$$

Beispiel

		Werte von Y			
		-1	0	+1	$f_X(x)$
	1	3/24	1/24	2/24	6/24
Werte von X	2	6/24	6/24	0	12/24
	3	3/24	1/24	2/24	6/24
	$f_Y(y)$	12/24	8/24	4/24	1

Lineare bedingte Erwartungswertfunktion: $\mathrm{E}(Y|X) = \beta_1 + \beta_2 x$

$$\beta_2 = \frac{\mathrm{E}[x - \mathrm{E}(x)][y - \mathrm{E}(y)]}{\mathrm{E}[x - \mathrm{E}(x)]^2} = \frac{\mathrm{cov}(y, x)}{\mathrm{var}(x)} = \frac{0}{0.5} = 0$$

$$\beta_1 = \mathrm{E}(y) - \beta_2\,\mathrm{E}(x) = -\frac{1}{3} + 0 \times 2 = -\frac{1}{3}$$

Beispiel

		Werte von Y			
		-1	0	+1	$f_X(x)$
	1	3/24	1/24	2/24	6/24
Werte von X	2	6/24	6/24	0	12/24
	3	3/24	1/24	2/24	6/24
	$f_Y(y)$	12/24	8/24	4/24	1

Bedingte Varianz: $\mathrm{var}(Y|X)$

$$\mathrm{var}(Y|X) = \begin{cases} \mathrm{E}(y^2|X=1) - [\mathrm{E}(y|X=1)]^2 & = \frac{5}{6} - [-\frac{1}{6}]^2 & = 0.58 \\ \mathrm{E}(y^2|X=2) - [\mathrm{E}(y|X=2)]^2 & = \frac{6}{12} - [-\frac{1}{2}]^2 & = 0.25 \\ \mathrm{E}(y^2|X=3) - [\mathrm{E}(y|X=3)]^2 & = \frac{5}{6} - [-\frac{1}{6}]^2 & = 0.58 \end{cases}$$

$\Rightarrow$ heteroskedastisch!

Beispiel

		Werte von Y			
		-1	0	+1	$f_X(x)$
	1	3/24	1/24	2/24	6/24
Werte von X	2	6/24	6/24	0	12/24
	3	3/24	1/24	2/24	6/24
	$f_Y(y)$	12/24	8/24	4/24	1

Bedingte Varianz: $\mathrm{var}(Y|X)$

$$\mathrm{var}(Y|X) = \begin{cases} \mathrm{E}(y^2|X=1) - [\mathrm{E}(y|X=1)]^2 & = \frac{5}{6} - [-\frac{1}{6}]^2 & = 0.58 \\ \mathrm{E}(y^2|X=2) - [\mathrm{E}(y|X=2)]^2 & = \frac{6}{12} - [-\frac{1}{2}]^2 & = 0.25 \\ \mathrm{E}(y^2|X=3) - [\mathrm{E}(y|X=3)]^2 & = \frac{5}{6} - [-\frac{1}{6}]^2 & = 0.58 \end{cases}$$

$\Rightarrow$ heteroskedastisch!

Thanx . . .