FERNSTUDIUM GUIDE *Zukunft beginnen.*

Grundlagen der Statistik

Teil 6

Demo - Version

Dieses Werk ist urheberrechtlich geschützt. Alle Rechte vorbehalten. **FSGU® AKADEMIE**
Staatlich geprüft und zugelassen unter den Zulassungsnummern 7272514c sowie 7272614c sowie 7272714c

Inhalt

- Wahrscheinlichkeitsverteilungen
  - Bedingte Verteilungen
  - Wahrscheinlichkeits- und Dichtefunktion
  - Verteilungsfunktion
  - Erwartungswert
  - Varianz, Standardabweichung
  - Gleichverteilung
  - Normalverteilung
  - Binomialverteilung

Analogie deskriptive Statistik und Wahrscheinlichkeitsverteilung

Deskriptive Statistik	Wahrscheinlichkeitsverteilung
Merkmal	Zufallsvariable X
Mittelwert $\overline{x}$	Erwartungswert E(X)
Relative Häufigkeiten $f(x_i)$	Wahrscheinlichkeitsfunktion $f_X(x_i)$ (Dichtefunktion bei stetigen Zufallsvariablen)
Summenhäufigkeit $F(x_i)$	Verteilungsfunktion $F_X(x_i)$
Bedingte Verteilung $f(x_j \mid y_k)$	Bedingte Verteilung $f_X(x_i \mid y_j)$
Abweichung vom Mittelwert $x_i - \overline{x}$	Abweichung vom Erwartungswert $x_i - E(X)$
Varianz $\frac{1}{n} \cdot \sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2$	Varianz $\frac{1}{n} \cdot \sum_{i=1}^{n} \left(x_i - E(X)\right)^2 \cdot f_X(x_i)$

Formeln: bei **diskreten** Zufallsvariablen **Summen**, bei stetigen Zufallsvariablen **Integrale**.

In einer früheren Lerneinheit haben wir in dieser Tabelle **relative Häufigkeiten** der Merkmale $\mathbf{x_i}$ (Einkommen) und $\mathbf{y_i}$ (höchster Bildungsabschluss nach ISCED) festgehalten.

Hier ist die **relative Häufigkei**t nichts anderes als die **Wahrscheinlichkeit**, dass ein bestimmtes Merkmal $\mathbf{x_i}$ **zusammen** mit einem bestimmten Merkmal $\mathbf{y_i}$ **auftrit**t. So ist beispielsweise die Wahrscheinlichkeit gleich **25%,** dass eine beliebig ausgewählte Person über **1500 bis 2000** € verdient und **ISCED III** als Bildungsabschluss besitzt.

Einkommen (€) x_i; über; bis	**y_i; ISCED**				
	y_1 **II**	y_2 **III**	y_3 **IV**	y_4 **V**	Σ
x_1 **1000; 1500**	0,195	0,055	0,008	0,000	h_1. **0,258**
x_2 **1500; 2000**	0,141	0,250	0,070	0,000	h_2. **0,461**
x_3 **2000; 2500**	0,047	0,109	0,094	0,031	h_3. **0,281**
Σ	$h._1$ **0,383**	$h._2$ **0,414**	$h._3$ **0,172**	$h._4$ **0,031**	1,000

Schreibweise zur Wahrscheinlichkeitsverteilung:

$P(Y = III) = f_Y(Y = III)$

$P(Y = III, X = 1500\text{-}2000) = f_{XY}(X = 1500\text{-}2000, Y=III)$

Einkommen (€) x_i; über; bis	**y_i; ISCED**				
	y_1 **II**	y_2 **III**	y_3 **IV**	y_4 **V**	Σ
x_1 **1000; 1500**	0,195	0,055	0,008	0,000	h_1. **0,258**
x_2 **1500; 2000**	0,141	0,250	0,070	0,000	h_2. **0,461**
x_3 **2000; 2500**	0,047	0,109	0,094	0,031	h_3. **0,281**
Σ	$h_{.1}$ **0,383**	$h_{.2}$ **0,414**	$h_{.3}$ **0,172**	$h_{.4}$ **0,031**	1,000

$f_Y(Y = III) =$ **0,414**

$f_{XY}(X = 1500\text{-}2000, Y=III) =$ **0,250**

Schreibweise zur bedingten Wahrscheinlichkeitsverteilung:

$P(Y = III \mid X = 1000\text{-}1500) = f_{XY}(Y=III|X = 1000\text{-}1500)$

Die Zufallsvariable Y wird **bedingt** für jedes $X = x_1$ verteilt.

Einkommen (€) x_i; über; bis	**y_i; ISCED**				
	y_1 **II**	y_2 **III**	y_3 **IV**	y_4 **V**	**Σ**
x_1 **1000; 1500**	**0,195**	**0,055**	**0,008**	**0,000**	h_1. **0,258**
x_2 **1500; 2000**	**0,141**	**0,250**	**0,070**	**0,000**	h_2. **0,461**
x_3 **2000; 2500**	**0,047**	**0,109**	**0,094**	**0,031**	h_3. **0,281**
Σ	$h._1$ **0,383**	$h._2$ **0,414**	$h._3$ **0,172**	$h._4$ **0,031**	**1,000**

Dazu erweitern wir die Zeilenwerte von x_1 mit (/0,258) für eine Zeilensumme von 1.

Einkommen (€) x; über; bis	**y; ISCED**				
	y_1 **II**	y_2 **III**	y_3 **IV**	y_4 **V**	**Σ**
x_1 **1000; 1500**	**0,756**	**0,213**	**0,031**	**0,000**	h_1. **1,000**

Gesucht ist die **bedingte** Verteilung, dass ein zufällig ausgewählter Mitarbeiter in der Einkommensklasse € 1000-€ 1500 ist und als höchsten Bildungsabschluss ISCED Niveau III besitzt.

Einkommen (€) x; über; bis	**y; ISCED**				
	y_1 **II**	y_2 **III**	y_3 **IV**	y_4 **V**	**Σ**
x_1 **1000; 1500**	**0,756**	**0,213**	**0,031**	**0,000**	h_1. **1,000**

Wir erhalten für $\mathbf{f_Y(Y=III|X=1000\text{-}1500)}$ = **0,213.**

Für die **bedingte** Verteilung $\mathbf{f_x(x_i\ |y_j)}$ werden die entsprechende **Spaltenwerte** durch die zugehörige **Spaltensumme** geteilt, für eine Spaltensumme von 1.

Z.B. $\mathbf{f_x(X=1000\text{-}1500\ |Y=III)}$ = **0,132.**

Einkommen (€) x; über; bis	**y; ISCED**
	y_2 **III**
x_1 **1000; 1500**	**0,132**
x_2 **1500; 2000**	**0,604**
x_3 **2000; 2500**	**0,264**
Σ	h.2 **1,000**

Herausgeber:
FSGU® AKADEMIE - Ein Unternehmen der FSGU® GmbH
Erlenweg 1
D-77948 Friesenheim
kontakt@fsgu-akademie.de | www.fsgu-akademie.de
info@fernstudium-guide.de | www.fernstudium-guide.de

Alle Rechte vorbehalten