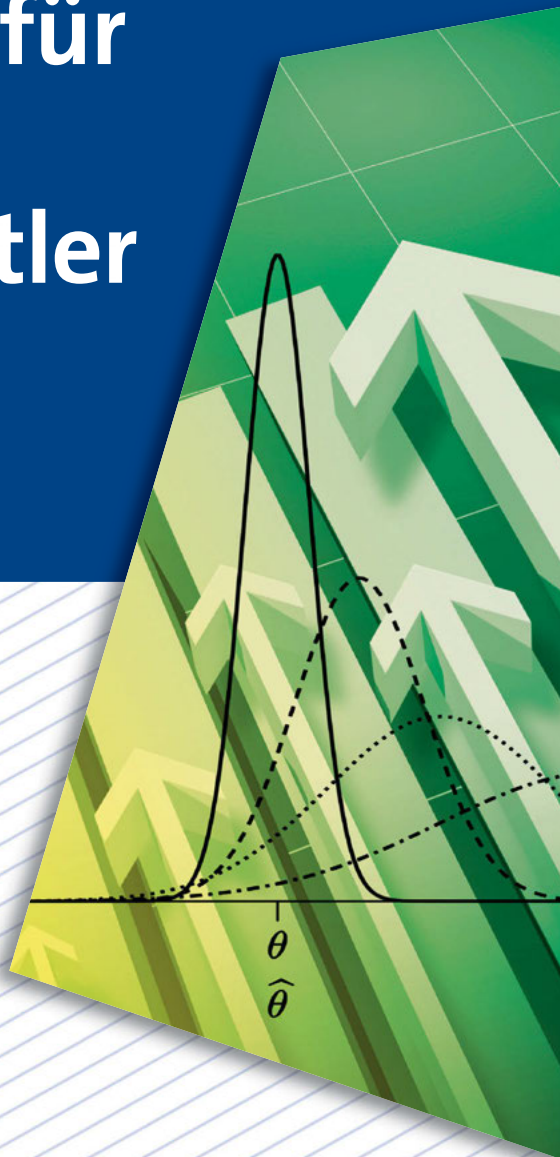


Übungen zur Statistik für Wirtschafts- wissenschaftler

Josef Bleymüller
Rafael Weißbach
Achim Dörre

Vahlen



Entwicklung einer soliden statistischen Methodenkompetenz mit praktischem Bezug

Dieses Übungsbuch ergänzt das Lehrbuch „Statistik für Wirtschaftswissenschaftler“ sowie die dazugehörige Formelsammlung „Statistische Formeln und Tabellen“ von Josef Bleymüller und Rafael Weißbach.

Zu allen 140 Aufgaben sind ausführlich Lösungswege angegeben. Die Konzeption der Übungsaufgaben zielt auf eine umfassende Abdeckung der im Lehrbuch vermittelten Inhalte bei gleichzeitiger Berücksichtigung ansteigender Schwierigkeit und Komplexität.

Prof. Dr. Josef Bleymüller war Direktor des Instituts für Ökonometrie und Wirtschaftsstatistik der Universität Münster. **Prof. Dr. Rafael Weißbach** ist Inhaber des Lehrstuhls für Statistik und Ökonometrie am Institut für Volkswirtschaftslehre der Universität Rostock, an dem auch **Dr. Achim Dörre** als wissenschaftlicher Mitarbeiter tätig ist.



**Das Lehrbuch
zum Übungsbuch:**
Bleymüller/Weißbach,
Statistik für Wirtschafts-
wissenschaftler



**Die Formelsammlung
zum Übungsbuch:**
Bleymüller/Weißbach,
Statistische Formeln
und Tabellen

Übungen zur Statistik für Wirtschafts- wissenschaftler

von

Prof. Dr. Josef Bleymüller

Prof. Dr. Rafael Weißbach

Dr. Achim Dörre

Verlag Franz Vahlen München

Prof. Dr. **Josef Bleymüller** war Direktor des Instituts für Ökonometrie und Wirtschaftsstatistik der Universität Münster. Prof. Dr. **Rafael Weißbach** ist Inhaber des Lehrstuhls für Statistik und Ökonometrie am Institut für Volkswirtschaftslehre der Universität Rostock. Dr. **Achim Dörre** ist wissenschaftlicher Mitarbeiter am Lehrstuhl für Statistik und Ökonometrie am Institut für Volkswirtschaftslehre der Universität Rostock.

ISBN Print: 978 3 8006 5873 2
ISBN E-Book: 978 3 8006 5874 9

© 2018 Verlag Franz Vahlen GmbH, Wilhelmstr. 9, 80801 München
Satz: Fotosatz Buck
Zweikirchener Str. 7, 84036 Kumhausen
Druck und Bindung: Druckhaus Nomos
In den Lissen 12, 76547 Sinzheim
Umschlaggestaltung: Ralph Zimmermann – Bureau Parapluie
Gedruckt auf säurefreiem, alterungsbeständigem Papier
(hergestellt aus chlorfrei gebleichtem Zellstoff)

Vorwort

*Wir lernen durch Irren und Fehlen
und werden Meister durch Übung,
ohne zu merken, wie es zugegangen ist.*
Christoph Martin Wieland (1733–1813)

Dieses Übungsbuch ergänzt das Lehrbuch

Bleymüller, J. und Weißbach, R. (2015): *Statistik für Wirtschaftswissenschaftler*, Vahlen.

und die dazugehörige Formelsammlung

Bleymüller, J. und Weißbach, R. (2015): *Statistische Formeln und Tabellen*, Vahlen.

Die Einteilung der Kapitel erfolgt entsprechend dem Lehrbuch und bietet Stoff, um die behandelten Themen und Konzepte stufenweise zu erarbeiten. Zu allen 140 Aufgaben sind ihre ausführlichen Lösungswege angegeben.

Die Konzeption der Übungsaufgaben zielt auf eine umfassende Abdeckung der im Lehrbuch vermittelten Inhalte bei gleichzeitiger Berücksichtigung ansteigender Schwierigkeit und Komplexität. Angewandte Statistik als Studienfach ist untrennbar mit mathematischen Grundtechniken verknüpft und lässt sich daher ideal durch aktive Übung aneignen.

Es ist beabsichtigt, dass viele Aufgaben einen authentischen Kontext besitzen, aber fiktive Daten enthalten, um einerseits eine potenzielle inhaltliche Motivation für die statistischen Verfahren und Konzepte anzubieten und andererseits Rechnungen an vielen Stellen zu vereinfachen. Im Vordergrund steht daher die Entwicklung eines soliden Verständnisses statistischer Methoden mit praktischem Bezug.

Die Übungsaufgaben basieren zu einem Teil auf dem ursprünglichen Übungsmaterial des Lehrbuchs (bis einschließlich der 17. Auflage), das überarbeitet wurde. Neben vollständig neu erstellten entstammen zahlreiche Aufgaben den Übungsveranstaltungen in den Grundlagenfächern der Statistik, die in den vergangenen Jahren an der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Universität Rostock durchgeführt wurden.

Letztere Übungsaufgaben wurden so kontinuierlich überarbeitet und ergänzt. In diesem Sinne sei den folgenden Personen gedankt: Liane Buchmann, Edelmiro Ricabal Delgado, Alexander Kremer, Benjamin Strohner, Gordon Frank und Lena Ullrich.

Ganz besonderer Dank gilt Ann-Josephine Thieme für ihre wertvollen konstruktiven Ratschläge sowie ihren unermüdlichen Ehrgeiz bei der kritischen Durchsicht des Manuskripts.

Wir danken außerdem Herrn Dennis Brunotte und dem Verlag Franz Vahlen für die langjährige Zusammenarbeit.

Rostock, Dezember 2018

*Josef Bleymüller
Rafael Weißbach
Achim Dörre*

Inhaltsverzeichnis

Vorwort	V
1. Einführung	1
Aufgaben	1
Lösungen	2
2. Empirische Verteilungen	5
Aufgaben	5
Lösungen	7
3. Mittelwerte	16
Aufgaben	16
Lösungen	18
4. Streuungsmaße	23
Aufgaben	23
Lösungen	25
5. Wahrscheinlichkeitsrechnung I	30
Aufgaben	30
Lösungen	31
6. Wahrscheinlichkeitsrechnung II	34
Aufgaben	34
Lösungen	35
7. Zufallsvariablen I (Eindimensionale Zufallsvariablen)	41
Aufgaben	41
Lösungen	43
8. Zufallsvariablen II (Zweidimensionale Zufallsvariablen)	51
Aufgaben	51
Lösungen	52
9. Theoretische Verteilungen I (Diskrete Verteilungen)	58
Aufgaben	58
Lösungen	59
10. Theoretische Verteilungen II (Stetige Verteilungen)	62
Aufgaben	62
Lösungen	63
11. Theoretische Verteilungen III (Approximationen, Reproduktionseigenschaft)	71
Aufgaben	71
Lösungen	72

12. Stichproben und Stichprobenverteilung I	75
Aufgaben	75
Lösungen	76
13. Stichproben und Stichprobenverteilung II	79
Aufgaben	79
Lösungen	80
14. Schätzverfahren I	84
Aufgaben	84
Lösungen	86
15. Schätzverfahren II	91
Aufgaben	91
Lösungen	92
16. Testverfahren I (Parametertests)	95
Aufgaben	95
Lösungen	96
17. Testverfahren II (Parametertests)	101
Aufgaben	101
Lösungen	103
18. Testverfahren III (Varianzanalyse)	108
Aufgaben	108
Lösungen	110
19. Testverfahren IV (Verteilungstests)	115
Aufgaben	115
Lösungen	117
20. Regressionsanalyse I (Lineare Einfachregression – Methode der kleinsten Quadrate)	122
Aufgaben	122
Lösungen	124
21. Regressionsanalyse II (Lineare Einfachregression – Schätz- und Testverfahren)	129
Aufgaben	129
Lösungen	131
22. Regressionsanalyse III (Lineare Einfachregression – Prognosen, Residualanalyse)	137
Aufgaben	137
Lösungen	140
23. Regressionsanalyse IV (Lineare Mehrfachregression – Schätz- und Testverfahren)	148
Aufgaben	148
Lösungen	151

24. Regressionsanalyse V (Lineare und nichtlineare Mehrfachregression)	159
Aufgaben	159
Lösungen	160
25. Indizes	166
Aufgaben	166
Lösungen	168
26. Konzentrationsmessung	172
Aufgaben	172
Lösungen	174

1. Einführung

Aufgabe 1.1

Ermitteln Sie anhand des Statistischen Jahrbuchs 2017 für die Bundesrepublik Deutschland folgende statistische Angaben:

- (a) Prozentanteil der Bevölkerung der Bundesrepublik Deutschland im Jahr 2015 mit einem Mindestalter von 65 Jahren
- (b) Bruttowertschöpfung des Baugewerbes im Jahr 2016
- (c) prozentuale Veränderung des Verbrauchs an Fleisch und Fleischerzeugnissen je Einwohner und Jahr (in kg) zwischen 2001 und 2015
- (d) Produktion von Bier aus Malz (ohne alkoholfreies Bier) und Zahl der produzierten Zigaretten in der Bundesrepublik im Jahr 2016
- (e) Anzahl der zugelassenen ambulanten Pflegedienste im Jahr 2015

Aufgabe 1.2

Handelt es sich bei den folgenden statistischen Gesamtheiten um Bestands- oder Bewegungsmassen?

- (a) Schüler und Schülerinnen einer Gesamtschule
- (b) Geburten in einem Landkreis
- (c) Fahrzeuge eines Unternehmens
- (d) Maschinenausfälle in einer Werkstatt
- (e) Anmeldungen in einem Einwohnermeldeamt
- (f) wartende Kunden an einer Kasse

Aufgabe 1.3

Der Bestand eines Halbfabrikats betrug am Wochenanfang 6318 und am Wochenende 7480 Stück. Für diesen Zeitraum wurde ein Lagerzugang von 3620 Stück festgestellt. Wie hoch war der Lagerabgang in dieser Woche? In der folgenden Tabelle sind die Angaben weiterer Wochen gegeben. Ergänzen Sie alle fehlenden Stückzahlen.

Bestand Wochenbeginn	Lagerzugang	Lagerabgang	Bestand Wochenende
6318	3620		7480
7480		3746	5030
	2750	1920	
	1633	3000	

Aufgabe 1.4

Die folgende Tabelle zeigt einen Ausschnitt aus einer Erhebung im Rahmen einer bundesweiten Studie:

Name	Anzahl Beschäftigte	Jahresumsatz (in Mio. Euro)	Eigenkapital-Anteil (in %)	Bonitätsklasse
GrunwoldPearson GmbH	26	3,56	21,7	ausgezeichnet
TechnikPerfekt AG	17	25,92	11,9	sehr gut
Kimonade UG	3	0,08	70,0	ausreichend
⋮	⋮	⋮	⋮	⋮
TRT Tiefbau GmbH	115	15,07	25,4	gut

- Benennen Sie alle vorliegenden Merkmale und deren Skalenniveau.
- Welches Skalenniveau ergibt sich, wenn die Unternehmen in eine Rangfolge bzgl. ihrer Jahresumsätze gestellt werden und welche Information geht dabei verloren?

Aufgabe 1.5

Geben Sie für die folgenden Merkmale an, ob sie diskret oder stetig sind.

- Kraftstoffverbrauch eines Pkws auf 100 km
- Zeitspanne, die zur Ausführung einer bestimmten Arbeit benötigt wird
- Zahl der pro Stunde in einem Geschäft eintreffenden Kunden
- Grundstücksgröße
- Stimmzahl einer Partei in einem Wahlbüro

Aufgabe 1.6

Auf welche Skalen sind die folgenden Transformationen ohne Informationsverlust anwendbar, wenn $a, b > 0$?

- $x^* = bx^2, x > 0$
- $x^* = a + bx$
- $x^* = bx$

Lösungen**Lösung zu Aufgabe 1.1**

Das Statistische Jahrbuch 2017 enthält folgende Angaben:

- 21,1 %
- 134,6 Mrd. Euro
- $(88,2 - 87,9) / 87,9 \cdot 100\% = 0,34\%$

- (d) Produktion von Bier aus Malz (ohne alkoholfreies Bier): 83139 Tsd. Hektoliter
Zahl der produzierten Zigaretten: 167750 Mio. Stk.
- (e) 13323 zugelassene ambulante Pflegedienste

Lösung zu Aufgabe 1.2

Das entscheidende Kriterium für Bestandsmassen ist, dass sie zu gewissen *Zeitpunkten* erfasst werden. Bewegungsmassen können dagegen nur in *Zeiträumen* sinnvoll gemessen werden. So ist bspw. die Gesamtheit der Schüler und Schülerinnen einer Gesamtschule (Teilaufgabe (a)) zu einem gewissen Zeitpunkt eindeutig definiert, während die Geburten in einem Landkreis (Teilaufgabe (b)) nicht zu einem Zeitpunkt, sondern in einem Zeitraum (z. B. Woche, Monat, Jahr) erfasst werden können. In der Regel sind Bewegungsmassen mit einem bestimmten Ereignis (z. B. Geburt, Maschinenausfall, Anmeldung) verknüpft, während Bestandsmassen sich auf einen Zustand beziehen.

- (a) Bestandsmasse
- (b) Bewegungsmasse
- (c) Bestandsmasse
- (d) Bewegungsmasse
- (e) Bewegungsmasse
- (f) Bestandsmasse

Lösung zu Aufgabe 1.3

Der Lagerabgang x wird mittels der Fortschreibungsformel bestimmt:

$$6318 + 3620 - x = 7480$$

$$\Rightarrow x = 6318 + 3620 - 7480 = 2458$$

Es wird ein Lagerabgang in Höhe von 2458 festgestellt. Alle weiteren fehlenden Werte werden analog bestimmt:

Bestand Wochenbeginn	Lagerzugang	Lagerabgang	Bestand Wochenende
6318	3620	2458	7480
7480	1296	3746	5030
5030	2750	1920	5860
5860	1633	3000	4493

Lösung zu Aufgabe 1.4

- (a) Gegeben sind die folgenden Merkmale:
 - *Name* (Nominalskala)
 - *Anzahl Beschäftigte* (Verhältnisskala)
 - *Jahresumsatz* (Verhältnisskala)
 - *Eigenkapital-Anteil* (Verhältnisskala)
 - *Bonitätsklasse* (Ordinalskala)

- (b) Durch die Bildung einer Rangfolge anhand des Jahresumsatzes entsteht ein ordinalskaliertes Merkmal. Bei einer Rangfolge wird die Information über die absolute Höhe des Umsatzes und das Ausmaß des Abstands zwischen den Unternehmen hinsichtlich des Umsatzes aufgegeben.

Lösung zu Aufgabe 1.5

Entscheidend für die Unterteilung in die Merkmalsarten ist, ob das betrachtete Merkmal nur bestimmte Ausprägungen oder – zumindest in einem bestimmten Intervall – jeden beliebigen Wert annehmen kann. Im ersteren Fall ist das Merkmal diskret, im letzteren stetig.

- (a) stetig
- (b) stetig
- (c) diskret
- (d) stetig
- (e) diskret

Lösung zu Aufgabe 1.6

Informationsverlust tritt auf, wenn nach der Durchführung einer Transformation bestimmte Eigenschaften der ursprünglichen Skala nicht mehr erfüllt sind. Dies wird in allen drei gegebenen Fällen überprüft.

- (a) Ordinalskala
Durch die Quadrierung von x wird die Information der Abstände zwischen den Ausprägungen verzerrt; somit ist diese Transformation auf Intervallskalen (und folglich auf Verhältnisskalen) nicht anwendbar.
Sofern ein ordinalskaliertes Merkmal mit positiven Zahlenwerten angegeben ist, erhält die angegebene Transformation die Information über die Rangfolge.
- (b) Ordinalskala, Intervallskala
Da eine lineare Transformation vorliegt, bleibt die Information über die Abstände zwischen Ausprägungen nach der Transformation unter Beachtung der Skalierung mit b erhalten.
Sofern ein ordinalskaliertes Merkmal mit positiven Zahlenwerten angegeben ist, erhält die angegebene Transformation die Information über die Rangfolge.
- (c) Ordinalskala, Intervallskala, Verhältnisskala
Diese Transformation ist ein Sonderfall der Transformation in Teilaufgabe (b) mit $a = 0$ und entspricht einer Skalierung mit b und ist auf alle gegebenen Skalen ohne Informationsverlust anwendbar.

2. Empirische Verteilungen

Aufgabe 2.1

Für einen Handwerksbetrieb wird im Rahmen einer Beratung die Anzahl der täglich eintreffenden Aufträge erfasst. Für die 20 Werktage eines ausgewählten Monats ergeben sich die folgenden Einzelwerte.

Tag	1	2	3	4	5	6	7	8	9	10
Aufträge	4	1	2	6	4	0	1	0	4	3
Tag	11	12	13	14	15	16	17	18	19	20
Aufträge	5	2	2	1	3	4	1	4	7	5

- Benennen Sie das untersuchte Merkmal, den Merkmalsträger und die Merkmalskategorisierung.
- Erstellen Sie aus den gegebenen Einzelwerten eine Häufigkeitstabelle inkl. der relativen Häufigkeiten.
- Bestimmen und zeichnen Sie die empirische Verteilungsfunktion.

Aufgabe 2.2

Eine Immobilienmaklerin betrachtet für das vergangene Jahr die Anzahl der von ihr erfolgreich abgeschlossenen Vermittlungen je Arbeitswoche.

x_i	h_i
0	20
1	15
2	9
3	5
4	1

- Ordnen Sie die Begriffe Merkmal, Merkmalsträger und Merkmalsausprägung(en) der gegebenen Situation zu.
- Stellen Sie die Häufigkeitsverteilung der erfolgreich abgeschlossenen Vermittlungen inkl. der relativen Häufigkeiten und kumulierten relativen Häufigkeiten in einer Tabelle dar.
- Erstellen Sie ein Stabdiagramm zur Darstellung der relativen Häufigkeitsverteilung.

Aufgabe 2.3

Der Betreiber eines kostenpflichtigen Parkplatzes erstellt monatlich eine Übersicht der täglich ausgestellten Strafzettel. Die folgende Tabelle enthält die entsprechenden Werte für den aktuellen Monat.

Strafzettel	Mo	Di	Mi	Do	Fr	Sa	So
Woche 1	–	–	3	1	6	3	25
Woche 2	7	4	22	4	10	23	22
Woche 3	13	2	23	19	4	12	26
Woche 4	11	9	11	17	5	25	23
Woche 5	20	4	1	8	–	–	–

- Benennen Sie Merkmal, Merkmalsträger und die Merkmalskalierung.
- Prinzipiell könnte aus diesen Einzelwerten eine Tabelle der absoluten und relativen Häufigkeiten erstellt werden. Erläutern Sie, warum dies in der gegebenen Situation zwar technisch möglich, aber inhaltlich nicht sinnvoll ist.
- Unterteilen Sie die Anzahl der ausgestellten Strafzettel in die Klassen 1–5, 6–10, 11–15, ..., 26–30 und erstellen Sie die entsprechende Häufigkeitsverteilung sowie ein Histogramm.

Aufgabe 2.4

Eine Maklerin analysiert den Immobilienmarkt einer bestimmten Stadt und interessiert sich vorrangig für die Verteilung der Kaufpreise. Die folgende Tabelle zeigt die Ergebnisse ihrer Recherche (Kaufpreisangaben in Tsd. Euro).

Kaufpreis		Anzahl
von	bis	
0	200	4
200	300	8
300	400	26
400	500	10
500	1000	2

- Benennen Sie Merkmal, Merkmalsträger und die Merkmalskalierung.
- Ordnen Sie die Klassenunter- und obergrenzen in die formale Notation klassifizierter Daten ein.
- Erstellen Sie ein Histogramm und das zugehörige Häufigkeitspolygon.
- Bestimmen Sie die empirische Verteilungsfunktion. Nutzen Sie die so konstruierte Funktion, um den Anteil der Immobilien mit einem Kaufpreis von höchstens 350000 Euro zu ermitteln.

Aufgabe 2.5

An der Lebensmittelkasse eines Kaufhauses werden die Rechnungsbeträge von den ersten 100 Kunden des Tages erfasst. Es ergibt sich die folgende Häufigkeitsverteilung.

Rechnungsbetrag in Euro	Anzahl der Rechnungen
bis 10	16
über 10 bis 20	48
über 20 bis 40	27
über 40 bis 80	9

- (a) Stellen Sie die empirische Verteilungsfunktion der Rechnungsbeträge grafisch dar.
- (b) Erstellen Sie ein Histogramm für die Verteilung der Rechnungsbeträge.

Aufgabe 2.6

Die Lebensdauern von 1000 Motoren weisen die folgende Verteilung auf.

Lebensdauer in Jahren	Anzahl der Motoren
bis 2	33
über 2 bis 4	276
über 4 bis 6	404
über 6 bis 8	237
über 8 bis 10	50

- (a) Stellen Sie die Häufigkeitsverteilung und die empirische Verteilungsfunktion grafisch dar.
- (b) Bestimmen Sie den Anteil der Motoren mit einer Lebensdauer von mehr als 5 Jahren.

Lösungen

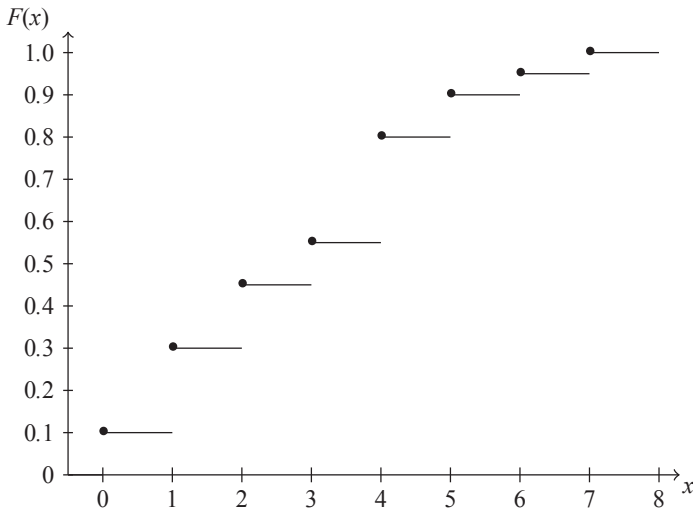
Lösung zu Aufgabe 2.1

- (a) Das untersuchte Merkmal ist die Anzahl der eingehenden Aufträge; der Merkmals-träger in der gegebenen Darstellung ist der Arbeitstag. Das Merkmal besitzt eine Verhältnisskala.
- (b) Zur Erstellung der Häufigkeitstabelle werden aus den gegebenen Einzelwerten die Häufigkeiten ermittelt:

x_i	h_i	f_i	F_i
0	2	0,10	0,10
1	4	0,20	0,30
2	3	0,15	0,45
3	2	0,10	0,55
4	5	0,25	0,80
5	2	0,10	0,90
6	1	0,05	0,95
7	1	0,05	1,00

- (c) Die empirische Verteilungsfunktion wird unter Verwendung der kumulierten relativen Häufigkeiten ermittelt:

$$F(x) = \begin{cases} 0,00 & x < 0 \\ 0,10 & 0 \leq x < 1 \\ 0,30 & 1 \leq x < 2 \\ 0,45 & 2 \leq x < 3 \\ 0,55 & 3 \leq x < 4 \\ 0,80 & 4 \leq x < 5 \\ 0,90 & 5 \leq x < 6 \\ 0,95 & 6 \leq x < 7 \\ 1,00 & 7 \leq x \end{cases}$$

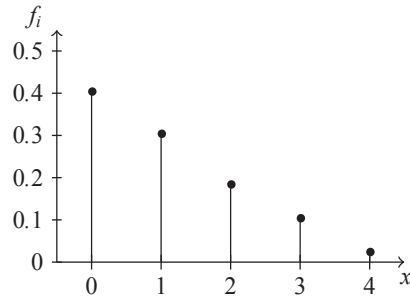


Lösung zu Aufgabe 2.2

- (a) In der gegebenen Situation ist das Merkmal die Anzahl der erfolgreich abgeschlossenen Vermittlungen, die Merkmalsträger sind die Arbeitswochen und die zugehörigen Merkmalsausprägungen sind durch die natürlichen Zahlen einschließlich der Null gegeben.
- (b) Die Häufigkeitsverteilung wird anhand der gegebenen absoluten Häufigkeiten bestimmt:

x_i	h_i	f_i	F_i
0	20	0,40	0,40
1	15	0,30	0,70
2	9	0,18	0,88
3	5	0,10	0,98
4	1	0,02	1,00

(c) Stabdiagramm:

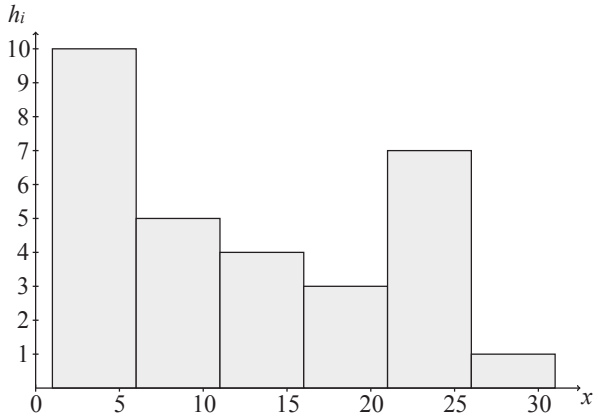


Lösung zu Aufgabe 2.3

- (a) Untersucht wird das Merkmal *ausgestellte Strafzettel*, das am Merkmalsträger *Wochentag* gemessen wird. Das Merkmal besitzt eine Verhältnisskala.
- (b) Die Erstellung einer Häufigkeitsverteilung für die $N = 30$ gegebenen Einzelwerte a_i ist technisch möglich, indem deren Häufigkeiten direkt ausgezählt werden. Auf diesem Weg sind die relativen und kumulierten relativen Häufigkeiten bestimmbar. Allerdings liegen aufgrund der hohen Spannweite der Einzelwerte viele unterschiedliche Ausprägungen vor, weshalb die meisten gebildeten absoluten Häufigkeiten lediglich 1 oder 2 sind. In diesem Sinn würde eine Häufigkeitsverteilung keine wesentliche Zusammenfassung der Einzelwerte bedeuten. Eine strikte Trennung nahe beieinander liegender Merkmalsausprägungen ist zudem im gegebenen Kontext nicht überzeugend.
- (c) Die Häufigkeitsverteilung wird durch Zählen der gegebenen Einzelwerte bestimmt:

x_i^u	x_i^o	Δx_i	h_i	f_i
1	5	4	10	0,333
6	10	4	5	0,167
11	15	4	4	0,133
16	20	4	3	0,100
21	25	4	7	0,233
26	30	4	1	0,033

Da alle Klassenbreiten gleich sind, können die Säulenhöhen des Histogramms mit den Klassenhäufigkeiten gleichgesetzt werden:

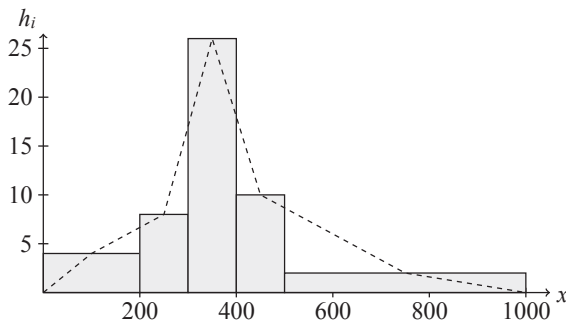


Lösung zu Aufgabe 2.4

- (a) Das untersuchte Merkmal ist der Kaufpreis von Immobilien, wobei die Merkmalsträger die Immobilien sind. Die Immobilienpreise besitzen eine Verhältnisskala und liegen in der gegebenen Situation in klassifizierter (und somit ordinaler) Form vor.
- (b) Es liegen $k = 5$ Klassen vor, deren Unter- und Obergrenzen gegeben sind. Notiert werden diese und die Klassenbreiten Δx_i wie folgt:

$x_1^u = 0$	$x_1^o = 200$	$\Delta x_1 = 200$
$x_2^u = 200$	$x_2^o = 300$	$\Delta x_2 = 100$
$x_3^u = 300$	$x_3^o = 400$	$\Delta x_3 = 100$
$x_4^u = 400$	$x_4^o = 500$	$\Delta x_4 = 100$
$x_5^u = 500$	$x_5^o = 1000$	$\Delta x_5 = 500$

- (c) Histogramm inklusive Häufigkeitspolygon:

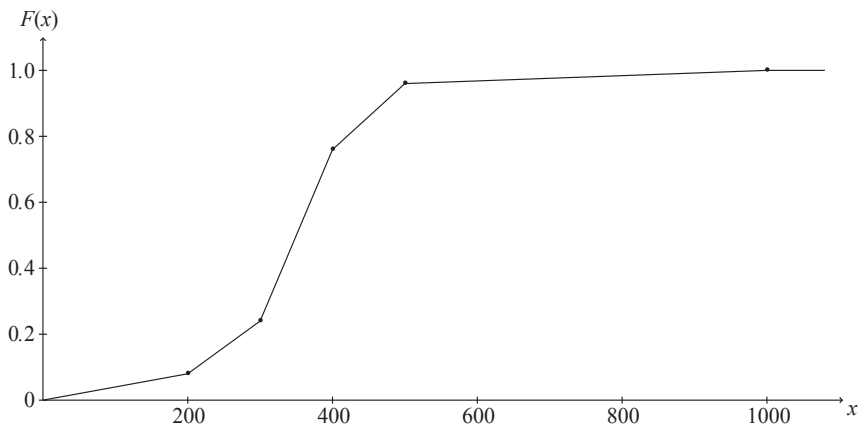


- (d) Die empirische Verteilungsfunktion wird mittels der kumulierten relativen Häufigkeiten ermittelt:

x_i^u	x_i^o	Δx_i	h_i	f_i	F_i
0	200	200	4	0,08	0,08
200	300	100	8	0,16	0,24
300	400	100	26	0,52	0,76
400	500	100	10	0,20	0,96
500	1000	500	2	0,04	1,00

Da klassifizierte Daten vorliegen, wird die Annahme getroffen, dass sich die Merkmalsausprägungen *innerhalb* der Klassen gleichmäßig verteilen. Grafisch entspricht dies der linearen Verbindung der kumulierten relativen Häufigkeiten innerhalb der Klassen, denn so wächst die kumulierte Häufigkeit proportional innerhalb jeder einzelnen Klasse, wie es bei einer perfekten Gleichverteilung der Fall wäre.

Die so gebildete empirische Verteilungsfunktion hat die folgende Form.



Gesucht ist der Anteil der Immobilien mit einem Kaufpreis von höchstens 350000 Euro. Dies entspricht der kumulierten relativen Häufigkeit zur Merkmalsausprägung 350. Da diese innerhalb der dritten Klasse (300, 400) liegt, wird der gesuchte Wert mittels der entsprechenden Teilstrecke der empirischen Verteilungsfunktion bestimmt. Es gilt $F_2 = 0,24$ und $F_3 = 0,76$. Der Wert $x = 350$ befindet sich genau in der Mitte der dritten Klasse, d. h. auf der Hälfte der Strecke zwischen 0,24 und 0,76. Somit ist der approximierte Wert von $F(350)$ gegeben durch

$$F(350) \approx 0,24 + 0,5 \cdot (0,76 - 0,24) = 0,5.$$

Demnach ist der Anteil der Immobilien mit einem Kaufpreis von höchstens 350000 Euro etwa 50%.