

Beschreibende Statistik

Beispiel 1)

Um welchen Skalentyp handelt es sich bei folgenden Merkmalen:

- Körpergröße
- Grad Celsius
- Mann = 0, Frau = 1
- Anzahl Mitbewohner WG
- Kleidergröße M
- IQ-Wert
- Monatseinkommen
- Eiergewichtsklassen

Wie kann man eine Skalentransformation des metrischen Wert „Alter“ zu einem ordinalen Wert durchführen?

Beispiel 2)

Warum sind folgende Merkmale stetig/diskret?

Postleitzahl, Körpergröße, Monatseinkommen

Beispiel 3)

Welche Lageparameter gibt es? Was ist der Zusammenhang zwischen Lageparametern und Skalenniveaus?

Gegeben ist folgende Datenreihe: 9 4 7 9 8 10 11 9 4 7

Bestimmen Sie Modalwert, Median, arithmetisches Mittel

Beispiel 4)

In der nachstehenden Tabelle sind die Mietpreise (in €/m²) von Zwei-Zimmer-Mietwohnungen in vergleichbarer Wohnlage, die im ersten Quartal auf dem Wohnungsmarkt einer Kleinstadt angeboten wurden, angegeben:

Mietpreis (in €/m ²) [von, bis[[4,6[[6,9[[9,12[[12,18]
Anzahl der Wohnungen	8	30	42	24

- Zeichnen Sie das zugehörige Histogramm!
- Zeichnen Sie die approximierende Verteilungsfunktion und markieren Sie in der Zeichnung den Median.
- Berechnen und interpretieren Sie den Median. Hinterfragen Sie ihn kritisch.

Beispiel 5)

In einer kleinen Goldgräberstadt wurde erhoben, wie der Besitz von Goldbarren verteilt ist. Dabei ergaben sich die Besitzverhältnisse wie folgt:

Anzahl Goldbarren	0	1	2	5	20
Anzahl Personen	150	40	30	10	2

- Zeichnen Sie die Lorenzkurve.
- Bestimmen Sie den Gini-Koeffizienten.

Beispiel 6 - EXCEL)

Um eine möglicherweise vorhandene Abhängigkeit zwischen höchstem Schulabschluss und regionaler Herkunft aufzudecken, wurden dreißigjährige Personen befragt und man erhielt folgende Ergebnisse in Form einer Tabelle:

Herkunft	Schulabschluss		
	Universität/FH	Matura	Hauptschule
Wien, Linz, Graz	46	78	53
Andere Landeshauptstädte	40	39	79
Sonstige Gemeinden	24	72	68

Berechnen Sie eine geeignete Kennzahl, um eine Aussage über die Abhängigkeit treffen zu können!

Beispiel 7)

Für acht Unternehmen desselben Wirtschaftszweiges soll untersucht werden, welcher Zusammenhang zwischen Umsatz und Beschäftigtenzahl besteht. Im Jahr 2018 wurden folgende Zahlen festgestellt:

Beschäftigte (in Tsd.)	0,5	0,7	1,3	1,8	1,6	1,4	1,8	2,2
Umsatz (in Mio. Euro)	50	120	260	220	180	160	210	280

- Bestimmen Sie den (Pearson-) Korrelationskoeffizienten und interpretieren Sie ihn.
- Bestimmen und zeichnen Sie die Regressionsgerade zur Beschreibung der Abhängigkeit des Umsatzes von der Anzahl der Beschäftigten.

Beispiel 8)

Ein Luxusrestaurant hat durch einen neuen Geschäftsführer einen beträchtlichen Gästerückgang erlebt. Die folgenden Werte wurden ermittelt:

Monat	1	2	3	4	5	6
Gäste	200	150	150	160	100	80

- Bestimmen Sie die Regressionsgerade.
- Prognostizieren Sie damit die Gästezahlen für den 8. und 9. Monat.

Wahrscheinlichkeitsrechnung

Beispiel 1)

Bei einem Brettspiel darf man, wenn man eine Glückskarte zieht, mit einem Würfel einmal Würfeln und dreimal so viele Felder vorziehen wie die Augensumme anzeigt.

Erstellen Sie eine Tabelle der Wahrscheinlichkeitsfunktion und bestimmen Sie Erwartungswert und Varianz.

Beispiel 2)

Wenn Sie nun mit dem Begriff Wahrscheinlichkeit vertraut sind, sollten die folgenden Fragen für Sie – zumindest prinzipiell – ganz einfach sein: Sie werfen einen Würfel sechsmal.

- a) Mit welcher Wahrscheinlichkeit ist der sechste Wurf eine Sechs?
- b) Mit welcher Wahrscheinlichkeit erhalten Sie bei diesen sechs Würfeln genau eine 6?
- c) Mit welcher Wahrscheinlichkeit erhalten sie beim ersten und beim sechsten Wurf jeweils eine 6?
- d) Mit welcher Wahrscheinlichkeit erhalten Sie beim sechsten Wurf zum ersten Mal eine Sechs?
- e) Mit welcher Wahrscheinlichkeit erhalten Sie nur beim ersten und beim letzten Wurf jeweils eine Sechs?
- f) Mit welcher Wahrscheinlichkeit erhalten Sie beim sechsten Wurf schon das sechste Mal die Sechs?
- g) Mit welcher Wahrscheinlichkeit ist mindestens eine Sechs dabei?
- h) Mit welcher Wahrscheinlichkeit ist die Augensumme aller Würfe größer als sechs?
- i) Mit welcher Wahrscheinlichkeit ist das Produkt aller erzielten Zahlen kleiner als 6?
- j) Mit welcher Wahrscheinlichkeit haben Sie sechs Mal gewürfelt?

Beispiel 3)

Das Finanzamt überprüft auf 12 Baustellen die Arbeitsgenehmigungen der Beschäftigten. Es werden jeweils drei Arbeiter zufällig ausgewählt. Unabhängig von der Baustelle betrage der Anteil der illegal Beschäftigten (Schwarzarbeiter) zwanzig Prozent.

- a) Wie groß ist die Wahrscheinlichkeit dafür, auf einer Baustelle keinen (genau zwei) Schwarzarbeiter zu finden?
- b) Mit welcher Wahrscheinlichkeit werden insgesamt höchstens drei illegal Beschäftigte gefunden?
- c) Mit welcher Wahrscheinlichkeit werden genau zwei von den 12 Baustellen beanstandet (d. h. dort wird mindestens ein Illegaler entdeckt)?

Beispiel 4)

Der Gourmethändler Exquisit prüft die Angebote von zwei Kartoffelbauern: Beide Kartoffelbauern liefern Trüffelkartoffel, aber mit unterschiedlichem Ausschussanteil: Bauer X: 4 % Ausschuss, Bauer Y: 10 % Ausschuss

Um zu einer Entscheidung über einen möglichen Liefervertrag zu kommen, wird jeweils eine Stichprobe von $n = 40$ Stück Trüffelkartoffel verkostet. Da dem Gourmethändler ein Ausschussanteil von 5 % gerade noch akzeptabel erscheint, wird die Lieferung akzeptiert, wenn bei der Verkostung höchstens 2 Kartoffeln ungenießbar sind.

- Mit welcher Wahrscheinlichkeit wird die Lieferung vom Bauern X **nicht** angenommen?
- Mit welcher Wahrscheinlichkeit wird eine Lieferung vom Bauern Y **doch** angenommen?

Beispiel 5)

Bestimmen Sie die Wahrscheinlichkeit dafür, beim Lotto „6 aus 45“

- einen Sechser (Fünfer)
- mindestens eine richtige Zahl

zu tippen.

Beispiel 6)

Um teure Rückrufaktionen von produzierten Mobiltelefonen zu vermeiden, werden von einem Hersteller Akkus, die von einem Zulieferer hergestellt werden, einer Abnahmekontrolle unterzogen.

Dabei werden zehn Akkus aus einer eingegangenen Lieferung entnommen und auf Funktionsfähigkeit geprüft. Sind alle zehn Akkus in Ordnung, wird die Lieferung angenommen. Sind jedoch zwei oder mehr defekt, wird die gesamte Lieferung zurück geschickt. Ist jedoch nur ein Stück fehlerhaft, wird eine weitere Stichprobe von vier Stück entnommen. Nur dann, wenn in dieser zweiten Stichprobe alle Akkus einwandfrei sind, wird die Lieferung angenommen, andernfalls zurückgeschickt.

Wie groß ist die Wahrscheinlichkeit, dass eine Lieferung angenommen wird, wenn 25 % der Akkus fehlerhaft sind?

Beispiel 7)

Eine Prüfung wird als Multiple-Choice-Test abgehalten. Es werden fünf Multiple-Choice-Fragen mit jeweils sechs möglichen Antworten vorgelegt, von denen jeweils genau eine Antwort richtig ist.

- Ein schlecht vorbereiteter Student hat keine Ahnung von dem abgefragten Stoff und kreuzt bei jeder Frage jeweils eine Antwort rein zufällig an. Mit welcher Wahrscheinlichkeit hat er danach mindestens zwei Fragen richtig beantwortet?
- Mit welcher Wahrscheinlichkeit beantwortet eine besser vorbereitete Kollegin mindestens drei Fragen richtig, wenn sie ganz sicher bei jeder Frage vier Antwortmöglichkeiten ausschließen kann und zwischen den verbleibenden beiden Antwortmöglichkeiten zufällig wählt?

Beispiel 8)

Eine Zufallsgröße X ist stetig gleichverteilt über dem Intervall $[a, b]$.
Man weiß: $P(X < 45) = 0,4$ und $P(X > 60) = 0,3$. Bestimmen Sie a und b !

Beispiel 9)

Gegeben sei eine $N(0; 1)$ -verteilte Zufallsvariable. Gesucht sind unter Verwendung der Tabellen im Anhang:

$$P(-1 \leq X \leq 2)$$

$$P(X \geq 1.5)$$

$$P(X \leq -2)$$

$$P(-2 \leq X \leq -1)$$

$$\approx 0.95$$

$$\approx 0.01$$

Beispiel 10)

Gegeben sei eine $N(5; 2)$ -verteilte Zufallsvariable. Gesucht sind unter Verwendung der Tabellen im Anhang:

$$P(3 \leq X \leq 6)$$

$$P(X \geq 3)$$

$$P(X \leq 1)$$

$$P(1 \leq X \leq 3)$$

$$\approx 0.95$$

Beispiel 11)

- Bestimmen Sie μ und σ jener Normalverteilung, für die gilt: $P(X < 30) = 0,4$ und $P(X > 50) = 0,1$.
- Wie ist die Summe S von zehn derartigen unabhängigen Zufallsgrößen verteilt? Mit welcher Wahrscheinlichkeit liegt diese Summe über 500?
- Wie ist der Mittelwert \bar{X} von zehn derartigen unabhängigen Zufallsgrößen verteilt? Mit welcher Wahrscheinlichkeit liegt dieser Mittelwert über 40?

Schließende Statistik

Beispiel 1)

Zur Beschreibung der wirtschaftlichen und sozialen Lage von BWL-Studierenden einer Universität wurden 201 Studenten zufällig ausgewählt und befragt. Die befragten Studenten gaben ihre zeitliche Gesamtbelastung durch Studium und Erwerbstätigkeit während der Vorlesungszeit mit durchschnittlich 42,8 Stunden pro Woche an; die Standardabweichung der erhobenen Daten betrug dabei 11,35 Stunden. Die zeitliche Belastung wird als normalverteilt angenommen.

- a) Bestimmen Sie das nach oben begrenzte 90-%-Konfidenzintervall für μ .
- b) Zu welchem Konfidenzniveau gehört das mit 40 Stunden und 43 Minuten nach unten begrenzte Intervall für μ ?

Beispiel 2)

In einem Seminar soll im Rahmen eines Projektes eine Studie über das Wahlverhalten der Steirer erarbeitet werden. 350 zufällig ausgewählte Steirer wurden unter anderem danach befragt, ob sie mit den landespolitischen Entscheidungen der Landesregierung zufrieden seien. 84 Befragte beantworteten diese Frage mit einem: „Ja“.

- a) Welchen Stichprobenumfang würden Sie in der Vorbereitungsphase der Erhebung empfehlen, wenn noch keine Informationen über den Stichprobenanteil vorliegen und die Forderung gestellt ist, dass das Konfidenzintervall höchstens die Länge 0,08 bei einem Konfidenzniveau von 0,99 haben soll?
- b) Welchen Mindeststichprobenumfang empfehlen Sie, wenn Sie die von den Studenten bereits durchgeführte Erhebung als Vorabinformation nutzen?

Beispiel 3)

- a) Von 320 befragten Personen gaben 104 an, im Vorjahr eine Auslandsreise gemacht zu haben. Bestimmen Sie ein zweiseitiges Konfidenzintervall für den Anteilswert p der Auslandsreisenden des Jahres 2004. Niveau 90 % (Niveau 99 %)
- b) Geben Sie ein nach oben abgeschlossenes einseitiges 90-Prozent-Konfidenzintervall an und interpretieren Sie ihr Ergebnis!
- c) Führen Sie dieselbe Rechnung wie in a) durch, wenn von 32.000 Befragten 10.400 angaben, im Vorjahr eine Auslandsreise unternommen zu haben. Wie ändert sich das Konfidenzintervall?

Beispiel 4)

Wie viele Daten muss man (im ungünstigsten Fall) erheben, um einen Anteilswert eines Merkmals mit einer Sicherheit von 90 Prozent auf plus/minus einen Prozentpunkt bestimmen zu können?

Beispiel 5)

Die Zeitschrift „NEUES“ schreibt: „Aufgrund einer Befragung von 800 Studierenden konnten wir feststellen, dass 23,7 Prozent (plus/minus 3 Prozentpunkte) der Studenten ihr Studium vorwiegend selbst finanzieren.“ Welche Treffsicherheit hat (zu welchem Niveau gehört) dieses Konfidenzintervall?

Beispiel 6)

Eine Zufallsgröße unterliege einer Normalverteilung $N(\mu; 6)$. Eine Stichprobe vom Umfang $n = 49$ ergab den Mittelwert 56,9.

- Testen Sie zum Niveau $\alpha = 0,1$ die Nullhypothese $H_0 : \mu = 55$ gegen $H_1 : \mu \neq 55$
- Innerhalb welcher Grenzen muss der beobachtete Mittelwert liegen, damit die Nullhypothese akzeptiert wird?

Beispiel 7)

Eine Zufallsgröße X sei $N(\mu; 3)$ -verteilt.

- Testen Sie zum Niveau $\alpha = 0,01$ die Hypothese $H_0 : \mu \leq 5$ aufgrund einer Stichprobe vom Umfang $n = 400$, die einen Mittelwert von 5,19 ergab.
- Wie ist zu den Niveaus $\alpha = 0,25$ (bzw. $\alpha = 0,001$) zu entscheiden?
- Zu welchem Signifikanzniveau gehört der realisierte Mittelwert von 5,19, d. h. für welches α ist dieser die Grenze des kritischen Bereiches für den Mittelwert?
- Der wahre Erwartungswert μ_{wahr} sei nun 5,2, die Nullhypothese ist also offensichtlich falsch. Mit welcher Wahrscheinlichkeit wird dann die H_0 abgelehnt, d. h. erkennt der Test die H_0 als falsch?

Beispiel 8)

Ein Test zur Messung der sozialen Anpassungsfähigkeit von Schulkindern ist genormt auf Mittelwert $\mu = 50$ und Varianz $\sigma_0^2 = 36$. Ein Soziologe glaubt, eine Möglichkeit zur Organisation des Unterrichts gefunden zu haben, die den Umgang der Schüler miteinander u. a. durch vermehrte Teamarbeit fördert und damit die soziale Anpassungsfähigkeit erhöht. Aus der Grundgesamtheit aller Schülerinnen und Schüler werden 60 zufällig ausgewählt und entsprechend dieses neuen Konzepts unterrichtet. Nach Ablauf eines zuvor festgelegten Zeitraums wird bei diesen Kindern ein mittlerer Testwert für die soziale Anpassungsfähigkeit von 52 beobachtet.

- Lässt sich damit die Beobachtung des Soziologen stützen? D.h. entscheiden Sie über die Behauptung des Soziologen anhand eines geeigneten statistischen Tests zum Niveau $\alpha = 0,05$. Formulieren Sie zunächst die Fragestellung als statistisches Testproblem.
- Was ändert sich in a), wenn
 - der Stichprobenumfang $n = 20$,
 - der beobachtete Mittelwert $\bar{x} = 51$,
 - die Standardabweichung $\sigma = 10$ beträgt?

Beispiel 9)

Eine Verbraucherzentrale möchte überprüfen, ob ein bestimmtes Medikament Übelkeit bei den Konsumenten auslöst. In einer Studie mit zehn Personen wird bei sieben Personen nach dem Genuss dieses Medikaments eine auftretende Übelkeit registriert.

- a) Überprüfen Sie zum Signifikanzniveau $\alpha = 0,05$ die statistische Nullhypothese, dass der Anteil der Personen mit Übelkeitssymptomen nach dem Genuss dieses Produkts in der Grundgesamtheit höchstens 50 % beträgt.
- b) Wie groß ist die Wahrscheinlichkeit für die ungerechtfertigte Akzeptanz (Beibehalten) der H_0 , d. h. die Wahrscheinlichkeit für einen Fehler 2. Art, wenn der tatsächliche Anteil siebzig Prozent beträgt?

Beispiel 10)

In einem Betrieb wurden für acht Lehrlinge jeweils der Notenschnitt des Abschlusszeugnisses aus der Schule und eine im Betrieb ermittelte Leistungskennzahl miteinander verglichen. Für die Leistungskennzahl wurden dabei Bewertungen zwischen 1 (= „sehr schlecht“) und 6 (= „ausgezeichnet“) vergeben.

Lehrling	1	2	3	4	5	6	7	8
Notenschnitt	1,5	2,6	2,1	1,8	2,5	3,2	1,3	1,4
L-Kennzahl	2	3	4	5	4	1	6	5

Lässt sich mit diesen Daten folgende Behauptung „bestätigen“: „Ein guter Notenschnitt in der Schule geht mit einer hohen Leistungskennzahl im Betrieb einher?“ Testen Sie unter Verwendung der Rangkorrelation zum Testniveau $\alpha = 0,1!$