WDDA FS 2026: Leitfaden für Aufgabenserie 1

2026-02-13

Einleitung

Dieser Leitfaden bietet detaillierte Erklärungen für die Übungen in WDDA FS 2026 Aufgabenserie 1. Für jede Aufgabe werden wir den Denkprozess durchgehen und erklären, wie man zur richtigen Lösung gelangt.

Aufgabe 1: Variablenklassifikation

In dieser Aufgabe müssen wir Variablen als nominal, ordinal oder metrisch (quantitativ) klassifizieren und bestimmen, ob sie diskret oder stetig sind.

Grundlegende Konzepte

Bevor wir beginnen, wiederholen wir die wichtigsten Konzepte:

- **Nominale Variablen**: Kategorien ohne inhärente Reihenfolge (z.B. Farben, Geschlecht)
- **Ordinale Variablen**: Kategorien mit einer sinnvollen Reihenfolge (z.B. Rangfolgen, Noten)
- **Metrische/Quantitative Variablen**: Numerische Messungen, die sein können:
  - **Diskret**: Zählbare Werte (z.B. Anzahl der Kinder), d.h. ganze Zahlen (keine Brüche)
  - **Stetig**: Kann jeden Wert innerhalb eines Bereichs annehmen (z.B. Grösse, Zeit)

Lösungen mit Erklärungen

Gehen wir jede Variable durch:

(a) Anzahl der wöchentlich von einer Fussballmannschaft erzielten Tore

Einteilung: Metrisch und diskret

Erklärung: Dies ist eine Zählvariable, die nur ganzzahlige Werte annehmen kann (0, 1, 2, usw.). Sie ist metrisch, weil wir sinnvolle arithmetische Operationen durchführen können (durchschnittliche Tore pro Spiel, Gesamttore, usw.), und sie ist diskret, weil Tore nur in ganzen Zahlen vorkommen.

(b) Körpergrössen der Spieler eines Footballteams

Einteilung: Metrisch und stetig

Erklärung: Die Körpergrösse wird auf einer kontinuierlichen Skala gemessen (z.B. 175,6 cm). Sie ist metrisch, weil wir arithmetische Operationen durchführen können, und sie ist stetig, weil die Körpergrösse theoretisch jeden Wert innerhalb eines Bereichs annehmen kann, nicht nur ganze Zahlen.

(c) Die beliebtesten Radiosender

Einteilung: Nominal und diskret

Erklärung: Radiosender sind Kategorien ohne inhärente Reihenfolge. Ein Sender ist nicht “mehr” als ein anderer in irgendeinem quantitativen Sinne. Die Variable ist diskret, weil wir verschiedene Einheiten kategorisieren.

(d) Die Anzahl der Kinder in einer japanischen Familie

Einteilung: Metrisch und diskret

Erklärung: Dies ist eine Zählvariable, die nur ganzzahlige Werte annehmen kann. Sie ist metrisch, weil wir aussagekräftige Statistiken wie Durchschnittswerte berechnen können, und sie ist diskret, weil Menschen nur in ganzen Zahlen vorkommen.

(e) Der Dienstgrad eines Soldaten

Einteilung: Ordinal und diskret

Erklärung: Militärische Ränge haben eine klare hierarchische Ordnung (z.B. Gefreiter, Korporal, Feldwebel). Die Variable ist ordinal, weil die Ränge eine sinnvolle Reihenfolge haben, und sie ist diskret, weil es sich um verschiedene Kategorien handelt.

(f) Die Anzahl der Brote, die eine Familie pro Woche kauft

Einteilung: Metrisch und diskret

Erklärung: Dies ist eine Zählvariable, die nur ganzzahlige Werte annehmen kann. Sie ist metrisch, weil wir arithmetische Operationen durchführen können, und sie ist diskret, weil Brote als ganze Einheiten gekauft werden.

(g) Die Art der Haustiere, welche die Schüler einer 8. Klasse besitzen

Einteilung: Nominal und diskret

Erklärung: Haustierarten (Hund, Katze, Fisch, usw.) sind Kategorien ohne inhärente Reihenfolge. Die Variable ist nominal, weil diese Kategorien keine natürliche Rangfolge haben, und sie ist diskret, weil wir mit verschiedenen Kategorien arbeiten.

(h) Die Anzahl der Blätter an Bäumen

Einteilung: Metrisch und diskret

Erklärung: Dies ist eine Zählvariable, die nur ganzzahlige Werte annehmen kann. Sie ist metrisch, weil wir arithmetische Operationen durchführen können, und sie ist diskret, weil Blätter als ganze Einheiten gezählt werden.

(i) Die Anzahl der Sonnenstunden an einem Tag

Einteilung: Metrisch und stetig

Erklärung: Sonnenstunden können mit einer Präzision jenseits ganzer Zahlen gemessen werden (z.B. 5,7 Stunden). Die Variable ist metrisch, weil wir arithmetische Operationen durchführen können, und sie ist stetig, weil Zeit auf einer kontinuierlichen Skala gemessen werden kann.

(j) Die Anzahl der Menschen, die jedes Jahr in den USA an Krebs sterben

Einteilung: Metrisch und diskret

Erklärung: Dies ist eine Zählvariable, die nur ganzzahlige Werte annehmen kann. Sie ist metrisch, weil wir arithmetische Operationen durchführen können, und sie ist diskret, weil wir einzelne Menschen zählen.

(k) Die Herkunftsländer der Einwanderer

Einteilung: Nominal und diskret

Erklärung: Länder sind Kategorien ohne inhärente numerische Reihenfolge. Die Variable ist nominal, weil Länder keine natürliche Rangfolge haben, und sie ist diskret, weil wir mit verschiedenen Kategorien arbeiten.

(l) Die Farben der Autos

Einteilung: Nominal und diskret

Erklärung: Farben sind Kategorien ohne inhärente Reihenfolge. Die Variable ist nominal, weil Farben keine natürliche Rangfolge haben, und sie ist diskret, weil wir mit verschiedenen Kategorien arbeiten.

(m) Das Geschlecht von Schulleitern

Einteilung: Nominal und diskret

Erklärung: Geschlechtskategorien haben keine inhärente Reihenfolge. Die Variable ist nominal, weil Geschlecht keine natürliche Rangfolge hat, und sie ist diskret, weil wir mit verschiedenen Kategorien arbeiten.

(n) Die Zeit, die mit Hausaufgaben verbracht wird

Einteilung: Metrisch und stetig

Erklärung: Zeit kann mit einer Präzision jenseits ganzer Zahlen gemessen werden (z.B. 2,5 Stunden). Die Variable ist metrisch, weil wir arithmetische Operationen durchführen können, und sie ist stetig, weil Zeit auf einer kontinuierlichen Skala gemessen werden kann.

(o) Die Kategorien von Pokerhänden

Einteilung: Ordinal und diskret

Erklärung: Pokerhände haben eine klare Rangfolge (High Card, Paar, zwei Paare, usw.). Die Variable ist ordinal, weil es eine sinnvolle Reihenfolge der Kategorien gibt, und sie ist diskret, weil es verschiedene Kategorien gibt.

(p) Der Betrag, der in Dollar pro Woche für Lebensmittel ausgegeben wird

Einteilung: Metrisch und stetig

Erklärung: Geld kann mit einer Präzision jenseits ganzer Zahlen gemessen werden (z.B. $45,67). Die Variable ist metrisch, weil wir arithmetische Operationen durchführen können, und sie ist stetig, weil Geld auf einer kontinuierlichen Skala gemessen werden kann.

(q) Die in einer Klassenarbeit erzielten Noten

Einteilung: Ordinal und diskret

Erklärung: Noten (A, B, C, usw. oder numerische Noten) haben eine klare Rangfolge. Die Variable ist ordinal, weil es eine sinnvolle Reihenfolge der Kategorien gibt, und sie ist diskret, weil es verschiedene Notenstufen gibt.

(r) Die Anzahl der im Schulbistro verkauften Artikel

Einteilung: Metrisch und diskret

Erklärung: Dies ist eine Zählvariable, die nur ganzzahlige Werte annehmen kann. Sie ist metrisch, weil wir arithmetische Operationen durchführen können, und sie ist diskret, weil Artikel als ganze Einheiten verkauft werden.

(s) Die Anzahl der Streichhölzer in einer Schachtel

Einteilung: Metrisch und diskret

Erklärung: Dies ist eine Zählvariable, die nur ganzzahlige Werte annehmen kann. Sie ist metrisch, weil wir arithmetische Operationen durchführen können, und sie ist diskret, weil Streichhölzer als ganze Einheiten gezählt werden.

(t) Die Endpositionen der Athleten in einem Rennen

Einteilung: Ordinal und diskret

Erklärung: Rennpositionen (1., 2., 3., usw.) haben eine klare Rangfolge. Die Variable ist ordinal, weil es eine sinnvolle Reihenfolge der Positionen gibt, und sie ist diskret, weil es verschiedene Positionen gibt.

(u) Die Endzeiten der Athleten eines Rennens

Einteilung: Metrisch und stetig

Erklärung: Rennzeiten können mit einer Präzision jenseits ganzer Zahlen gemessen werden (z.B. 10,45 Sekunden). Die Variable ist metrisch, weil wir arithmetische Operationen durchführen können, und sie ist stetig, weil Zeit auf einer kontinuierlichen Skala gemessen werden kann.

(v) Die Gründe, warum Menschen Taxis benutzen

Einteilung: Nominal und diskret

Erklärung: Gründe (Bequemlichkeit, kein Auto, usw.) sind Kategorien ohne inhärente Reihenfolge. Die Variable ist nominal, weil diese Gründe keine natürliche Rangfolge haben, und sie ist diskret, weil wir mit verschiedenen Kategorien arbeiten.

(w) Die Sportarten, die von Schülern an Gymnasien gespielt werden

Einteilung: Nominal und diskret

Erklärung: Sportarten (Basketball, Fussball, usw.) sind Kategorien ohne inhärente Reihenfolge. Die Variable ist nominal, weil Sportarten keine natürliche Rangfolge haben, und sie ist diskret, weil wir mit verschiedenen Kategorien arbeiten.

(x) Die Verbraucherpräferenz zwischen drei verschiedenen Produkten

Einteilung: Nominal und diskret

Erklärung: Produktpräferenzen sind Kategorien ohne inhärente Reihenfolge. Die Variable ist nominal, weil Präferenzen nicht unbedingt eine natürliche Rangfolge haben, und sie ist diskret, weil wir mit verschiedenen Kategorien arbeiten.

Nebenbemerkung: Aus mikroökonomischer Sicht nehmen wir an, dass Konsumenten nicht nur zwischen verschiedenen Produktkategorien unterscheiden, sondern diese auch in einer bestimmten Rangfolge ordnen können. Während in der reinen Datenskala die Produktpräferenzen als nominale und diskrete Kategorien betrachtet werden–also als unterschiedliche Gruppen ohne inhärente quantitative Ordnung–, impliziert die Theorie der Verbraucherwahl, dass diese Kategorien tatsächlich in eine ordinale Reihenfolge gebracht werden können.

(y) Die Pulsraten einer Gruppe von Sportlern in Ruhe

Einteilung: Metrisch und stetig

Erklärung: Pulsraten können mit einer Präzision jenseits ganzer Zahlen gemessen werden (z.B. 72,5 Schläge pro Minute). Die Variable ist metrisch, weil wir arithmetische Operationen durchführen können, und sie ist stetig, weil Herzfrequenzen kontinuierlich variieren können.

Aufgabe 2: Bestimmung von Variablenwerten

In dieser Aufgabe sollen wir diskutieren, wie wir Werte für verschiedene Variablen ermitteln könnten.

(a) Wo auf dem politischen Spektrum eine Person steht

Um zu bestimmen, wo eine Person auf dem politischen Spektrum steht, könnten wir:

1. **Ein Umfrageinstrument verwenden**: Einen Fragebogen mit Aussagen zu politischen Themen erstellen und die Befragten bitten, ihren Grad der Zustimmung auf einer Skala anzugeben (z.B. 1-7 Likert-Skala).
2. **Wahlverhalten analysieren**: Vergangene Wahlmuster betrachten, um Personen auf einem Spektrum zu platzieren.
3. **Selbstidentifikation**: Einfach Menschen bitten, sich selbst auf einem politischen Spektrum zu platzieren (z.B. "Auf einer Skala von 1 (sehr liberal) bis 10 (sehr konservativ), wo würden Sie sich einordnen?").
4. **Themenbasierte Zuordnung**: Nach Positionen zu bestimmten Themen fragen und Antworten auf ein politisches Spektrum abbilden.

Die resultierende Variable wäre wahrscheinlich ordinal (wenn Kategorien wie "sehr liberal", "etwas liberal" usw. verwendet werden) oder potenziell metrisch, wenn eine numerische Skala verwendet wird.

(b) Der Blutdruck eines Patienten

Der Blutdruck wird typischerweise bestimmt durch:

1. **Direkte Messung**: Verwendung eines Sphygmomanometers (Blutdruckmanschette) und Stethoskops oder eines automatischen Blutdruckmessgeräts.
2. **Mehrfache Messungen**: Durchführung mehrerer Messungen zu verschiedenen Zeiten, um Schwankungen zu berücksichtigen.
3. **Standardisierte Bedingungen**: Sicherstellen, dass der Patient sitzt, entspannt ist, der Arm auf Herzhöhe ist, usw.

Der Blutdruck wird als zwei Zahlen aufgezeichnet: systolisch (Druck beim Herzschlag) und diastolisch (Druck bei Herzruhe). Zum Beispiel 120/80 mmHg.

Dies ist eine metrische, stetige Variable, da sie jeden Wert innerhalb eines Bereichs annehmen kann und mit hoher Präzision gemessen werden kann.

(c) Die Armutsgrenze eines Landes

Die Armutsgrenze eines Landes kann bestimmt werden durch:

1. **Absoluter Ansatz**: Schätzung des Einkommens, das für grundlegende Bedürfnisse benötigt wird (Nahrung, Unterkunft, Kleidung).
2. **Relativer Ansatz**: Definition von Armut als Prozentsatz des Medianeinkommens (z.B. 60% des Medianeinkommens).
3. **Multidimensionaler Ansatz**: Berücksichtigung von Faktoren jenseits des Einkommens, wie Zugang zu Bildung, Gesundheitsversorgung und grundlegenden Dienstleistungen.
4. **Konsumbasierte Messungen**: Betrachtung dessen, was Haushalte konsumieren, anstatt ihres Einkommens.

Die resultierende Variable ist metrisch und stetig, typischerweise ausgedrückt als Geldwert pro Person oder Haushalt pro Zeitraum.

(d) Die Ausdauerleistung eines Sportlers

Die Ausdauerleistung eines Sportlers kann gemessen werden durch:

1. **Zeit-bis-zur-Erschöpfung-Tests**: Wie lange ein Sportler eine bestimmte Intensität der Übung aufrechterhalten kann.
2. **Distanztests**: Wie weit ein Sportler in einer festgelegten Zeit gehen kann (z.B. 12-Minuten-Lauftest).
3. **VO2-max-Tests**: Messung des maximalen Sauerstoffverbrauchs während inkrementeller Übung.

Dies ist eine metrische Variable, die stetig sein könnte (wenn Zeit oder physiologische Parameter gemessen werden) oder diskret (wenn standardisierte Leistungskategorien verwendet werden).

(e) Die Anzahl, wie oft eine Person an einem Tag lügt

Dies ist schwierig genau zu messen, aber Ansätze könnten umfassen:

1. **Selbstberichterstattung**: Menschen bitten, ein Tagebuch zu führen, wann sie lügen, obwohl dies Ehrlichkeits- und Bewusstseinsproblemen unterliegt.
2. **Beobachtung**: Geschulte Beobachter überwachen Interaktionen, obwohl dies ressourcenintensiv ist und subtile Lügen übersehen könnte.
3. **Experimentelle Einstellungen**: Situationen schaffen, in denen Lügen auftreten könnten, und das Verhalten überwachen.

Dies wäre eine metrische, diskrete Variable, da Lügen als ganze Zahlen gezählt würden.

Aufgabe 3: Grafikanalyse

In dieser Aufgabe analysieren wir eine Grafik, die MSNBC und Fox News vergleicht.

(a) Identifizieren Sie die beiden Variablen

Die beiden in der Grafik gezeigten Variablen sind:

1. **Faktische Berichterstattung**: Dargestellt durch den oberen Balken, der von "Sehr niedrig" bis "Sehr hoch" reicht
2. **Politische Voreingenommenheit**: Dargestellt durch den unteren Pfeil, der von "Extrem links" bis "Extrem rechts" reicht

(b) Erklären Sie, wie die Werte für diese Variablen ermittelt werden könnten

Faktische Berichterstattung

Werte für faktische Berichterstattung könnten ermittelt werden durch:

1. **Inhaltsanalyse**: Untersuchung einer Stichprobe von Nachrichtenberichten und Überprüfung von Behauptungen anhand primärer Quellen oder etablierter Fakten.
2. **Fehlerratenverfolgung**: Zählung von Korrekturen, Rücknahmen und identifizierten falschen Behauptungen im Laufe der Zeit.
3. **Expertenpanel-Bewertung**: Medienexperten oder Faktenprüfer bewerten Inhalte basierend auf etablierten Kriterien.
4. **Quellentransparenz**: Bewertung, ob Behauptungen ordnungsgemäss überprüfbaren Quellen zugeschrieben werden.
5. **Methodologische Offenlegung**: Bewertung, ob das Nachrichtenmedium erklärt, wie Informationen gesammelt wurden.

Politische Voreingenommenheit

Werte für politische Voreingenommenheit könnten ermittelt werden durch:

1. **Inhaltsanalyse**: Untersuchung von Sprache, Framing und Themenauswahl auf parteiische Neigungen.
2. **Quellenauswahl**: Analyse, welche Experten, Politiker oder Organisationen zitiert oder vorgestellt werden.
3. **Themenabdeckung**: Messung der Zeit/des Raums, die/der verschiedenen politischen Themen gewidmet wird.
4. **Publikumsbefragungen**: Befragung von Zuschauern/Lesern zu ihren eigenen politischen Neigungen, um Muster zu identifizieren.
5. **Vergleichende Analyse**: Vergleich der Berichterstattung über dieselben Ereignisse in verschiedenen Medien.

(c) Von welchem Typ sind die Variablen?

Beide Variablen werden in der Grafik als **ordinal** und **diskret** dargestellt.

Erklärung: - Sie sind ordinal, weil die Kategorien eine sinnvolle Reihenfolge haben (von "Sehr niedrig" bis "Sehr hoch" für faktische Berichterstattung und von "Extrem links" bis "Extrem rechts" für politische Voreingenommenheit). - Sie werden als diskret dargestellt, weil sie verschiedene Kategorien anstelle einer kontinuierlichen Skala verwenden.

Es ist jedoch erwähnenswert, dass die zugrunde liegende Datenerhebung möglicherweise kontinuierlichere Messungen verwendet hat, die dann für die Präsentation in diese ordinalen Kategorien umgewandelt wurden.

Aufgaben 4-7: Arbeiten mit dem BFH-Datensatz

Diese Aufgaben beinhalten das Arbeiten mit Variablen aus dem BFH-Datensatz und das Durchführen von Operationen mit ihnen.

Aufgabe 4: sum(cash) vs sum(gender)

```r
# Dies ist konzeptionell - wir haben den tatsächlichen BFH-Datensatz nicht geladen
# Gesamtbetrag des von allen Personen mitgeführten Bargelds
sum(cash)
# Würde einen Fehler zurückgeben, weil gender eine nominale Variable ist
sum(gender)
```

Erklärung: - `sum(cash)` funktioniert, weil `cash` eine metrische Variable ist (Betrag an CHF, der mitgeführt wird). Die Addition dieser Werte ergibt den Gesamtbetrag des von allen Personen im Datensatz mitgeführten Bargelds. - `sum(gender)` erzeugt einen Fehler, weil `gender` eine nominale Variable ist (Kategorien ohne inhärenten numerischen Wert). R kann keine arithmetischen Operationen mit nicht-numerischen Daten durchführen.

Aufgabe 5: height/100

```r
# Dies ist konzeptionell - wir haben den tatsächlichen BFH-Datensatz nicht geladen
height/100  # Konvertiert Körpergrösse von Zentimetern in Meter
```

Erklärung: Diese Operation konvertiert die Körpergrössenmessungen von Zentimetern in Meter. Da die ursprüngliche Variable `height` in Zentimetern ist, ergibt die Division durch 100 die entsprechende Körpergrösse in Metern.

Zum Beispiel, wenn jemand 175 cm gross ist, würde height/100 1,75 m ergeben.

Aufgabe 6: reaction1-reaction2

```
# Dies ist konzeptionell - wir haben den tatsächlichen BFH-Datensatz nicht geladen
# Berechnet die Differenz der Reaktionszeiten zwischen den Versuchen
reaction1-reaction2
```

Erklärung: Diese Operation berechnet die Differenz zwischen der ersten und zweiten Reaktionszeitmessung.

- Wenn das Ergebnis positiv ist (reaction1 > reaction2), deutet dies auf eine Verbesserung der Reaktionszeit vom ersten zum zweiten Versuch hin.
- Wenn das Ergebnis negativ ist (reaction1 < reaction2), deutet dies auf eine langsamere Reaktionszeit im zweiten Versuch hin.
- Wenn das Ergebnis null ist, deutet dies auf keine Veränderung der Reaktionszeit hin.

Diese neue Variable repräsentiert die Veränderung oder Verbesserung der Reaktionszeit zwischen den Versuchen.

Aufgabe 7: siblings*present

```
# Dies ist konzeptionell - wir haben den tatsächlichen BFH-Datensatz nicht geladen
# Berechnet die Gesamtausgaben für Geschenke pro Person
siblings*present
# Würde die Gesamtausgaben für Geschenke über alle Personen berechnen
sum(siblings*present)
```

Erklärung: Die Operation `siblings*present` multipliziert die Anzahl der Geschwister, die eine Person hat, mit dem Betrag, der für Geschenke pro Geschwister ausgegeben wird. Dies ergibt den Gesamtbetrag, den jede Person für Geschenke für alle ihre Geschwister ausgibt.

Die Summe dieser Variable über alle Personen (`sum(siblings*present)`) würde den Gesamtbetrag darstellen, der von allen im Datensatz für Geschwistergeschenke ausgegeben wird.

Aufgaben 8-16: Arbeiten mit Summationsnotation

Diese Aufgaben konzentrieren sich auf das Verständnis und die Manipulation der Summationsnotation.

Aufgabe 8: Ausschreiben von Summen

(a) $\sum_{r=1}^{4} r^3$

Dies bedeutet: $1^3 + 2^3 + 3^3 + 4^3$

Berechnung: $1 + 8 + 27 + 64 = 100$

(b) $\sum_{r=1}^{6}(2r + 1)$

Dies bedeutet: $(2 \times 1 + 1) + (2 \times 2 + 1) + (2 \times 3 + 1) + (2 \times 4 + 1) + (2 \times 5 + 1) + (2 \times 6 + 1)$

Berechnung: $3 + 5 + 7 + 9 + 11 + 13 = 48$

(c) $\sum_{n=1}^{5} 3^{2n}$

Dies bedeutet: $3^{2\times1} + 3^{2\times2} + 3^{2\times3} + 3^{2\times4} + 3^{2\times5}$

Berechnung: $3^2 + 3^4 + 3^6 + 3^8 + 3^{10}$

Berechnung: $9 + 81 + 729 + 6'561 + 59'049 = 66'429$

(d) $\sum_{n=1}^{7} 4$

Dies bedeutet: $4 + 4 + 4 + 4 + 4 + 4 + 4$

Berechnung: $7 \times 4 = 28$

Aufgabe 9: Verwendung der Sigma-Notation für Summen

(a) $1^2 + 2^2 + 3^2 + ... + 7^2$

Dies kann geschrieben werden als: $\sum_{n=1}^{7} n^2$

(b) $2^1 + 2^2 + 2^3 + ... + 2^{10}$

Dies kann geschrieben werden als: $\sum_{n=1}^{10} 2^n$

(c) $\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{77}$

Dies kann geschrieben werden als: $\sum_{n=1}^{77} \frac{1}{n}$

(d) $\frac{1\times3}{2\times4} + \frac{2\times4}{3\times5} + ... + \frac{19\times21}{20\times22}$

Dies kann geschrieben werden als: $\sum_{n=1}^{19} \frac{n\times(n+2)}{(n+1)\times(n+3)}$

Aufgabe 10: Auswerten von Summen

(a) $\sum_{r=1}^{6} r$

Dies bedeutet: $1 + 2 + 3 + 4 + 5 + 6 = 21$

(b) $\sum_{r=1}^{6} 3r$

Dies bedeutet: $3 \times 1 + 3 \times 2 + 3 \times 3 + 3 \times 4 + 3 \times 5 + 3 \times 6$

Berechnung: $3 + 6 + 9 + 12 + 15 + 18 = 63$

(c) $3 \times \sum_{r=1}^{6} r$

Dies bedeutet: $3 \times (1 + 2 + 3 + 4 + 5 + 6) = 3 \times 21 = 63$

Beachten Sie, dass (b) und (c) das gleiche Ergebnis liefern, was die Eigenschaft demonstriert, dass eine Konstante aus einer Summe herausgezogen werden kann.

Aufgabe 11: Ausschreiben allgemeiner Summen

(a) $\sum_{r=1}^{n} x_r$

Dies bedeutet: $x_1 + x_2 + x_3 + ... + x_n$

(b) $\sum_{r=1}^{n} x_r^2$

Dies bedeutet: $x_1^2 + x_2^2 + x_3^2 + ... + x_n^2$

(c) $\sum_{r=1}^{n} x_r^r$

Dies bedeutet: $x_1^1 + x_2^2 + x_3^3 + ... + x_n^n$

Aufgabe 12: Richtig oder falsch Aussagen

(a) $\sum_{r=1}^{n} x_r + \sum_{r=1}^{n} y_r = \sum_{r=1}^{n}(x_r + y_r)$

Richtig

Erklärung: Diese Eigenschaft besagt, dass die Summe zweier separater Summationen gleich der Summation der Summe entsprechender Terme ist. Dies ist wahr, weil Addition kommutativ und assoziativ ist.

Zum Beispiel, mit n=3: $(x_1 + x_2 + x_3) + (y_1 + y_2 + y_3) = (x_1 + y_1) + (x_2 + y_2) + (x_3 + y_3)$

(b) $\sum_{r=1}^{n} x_r \cdot \sum_{r=1}^{n} y_r = \sum_{r=1}^{n}(x_r \cdot y_r)$

Falsch

Erklärung: Die linke Seite multipliziert jeden Term in der ersten Summe mit jedem Term in der zweiten Summe, während die rechte Seite nur entsprechende Terme multipliziert.

Zum Beispiel, mit n=2: $(x_1 + x_2) \cdot (y_1 + y_2) = x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$ Aber $\sum_{r=1}^{2}(x_r \cdot y_r) = x_1y_1 + x_2y_2$

Diese sind eindeutig nicht gleich.

(c) $\sum_{r=1}^{n}(c \cdot x_r) = c \cdot \sum_{r=1}^{n} x_r$

Richtig

Erklärung: Diese Eigenschaft besagt, dass eine Konstante aus einer Summation herausgezogen werden kann. Dies ist wahr, weil Multiplikation über Addition verteilt.

Zum Beispiel, mit n=3: $c \cdot x_1 + c \cdot x_2 + c \cdot x_3 = c \cdot (x_1 + x_2 + x_3)$

Aufgabe 13: Ausschreiben von $\sum_{m=1}^{4}(-1)^m x$

Dies bedeutet: $(-1)^1 \cdot x + (-1)^2 \cdot x + (-1)^3 \cdot x + (-1)^4 \cdot x$

Berechnung: $-x + x - x + x = 0$

Aufgabe 14: Finden von $\sum_{r=1}^{7}(x_r - 2)^2$ bei gegebenen anderen Summen

Gegeben: - $\sum_{r=1}^{7} x_r^2 = 100$ - $\sum_{r=1}^{7} x_r = 10$

Wir müssen finden: $\sum_{r=1}^{7}(x_r - 2)^2$

Erweitern wir: $\sum_{r=1}^{7}(x_r - 2)^2 = \sum_{r=1}^{7}(x_r^2 - 4x_r + 4)$

$= \sum_{r=1}^{7} x_r^2 - 4\sum_{r=1}^{7} x_r + \sum_{r=1}^{7} 4$

$= \sum_{r=1}^{7} x_r^2 - 4\sum_{r=1}^{7} x_r + 4 \cdot 7$

$= 100 - 4 \cdot 10 + 28$

$= 100 - 40 + 28$

$= 88$

Aufgabe 15: Ausschreiben verschiedener statistischer Notationen

(a) $\frac{1}{N} \cdot \sum_{i=1}^{N} x_i$

Dies ist die Formel für das arithmetische Mittel (Durchschnitt) einer Menge von Werten: $\frac{x_1+x_2+x_3+\ldots+x_N}{N}$

(b) $\sum_{r=1}^{n}(x_r - \bar{x})^2$

Dies ist die Summe der quadrierten Abweichungen vom Mittelwert, verwendet bei der Berechnung der Varianz: $(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + ... + (x_n - \bar{x})^2$

(c) $\sum_{i=1}^{n} w_i x_i$

Dies ist eine gewichtete Summe, bei der jeder Wert x_i mit einem Gewicht w_i multipliziert wird: $w_1x_1 + w_2x_2 + w_3x_3 + ... + w_nx_n$

(d) $\sum_{i=1}^{n} P(E_i)$

Dies ist die Summe der Wahrscheinlichkeiten für Ereignisse E_1 bis E_n: $P(E_1) + P(E_2) + P(E_3) + ... + P(E_n)$

(e) $\sum_{i=1}^{3}\sum_{j=1}^{2} P_{ij}$

Dies ist eine Doppelsummation, die Werte in einer 3×2-Matrix summiert: $P_{11} + P_{12} + P_{21} + P_{22} + P_{31} + P_{32}$

(f) $\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$

Dies ist die Summe der Produkte von Abweichungen, verwendet bei der Berechnung der Kovarianz: $(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + ... + (x_n - \bar{x})(y_n - \bar{y})$

Aufgabe 16: Auswerten tabellenbasierter Summationen

Gegeben die Tabelle:

```
 1  2  3  4
 5  6  7  8
 9 10 11 12
13 14 15 16
```

Wobei a(r,c) der Wert in Zeile r, Spalte c ist.

(a) $\sum_{n=1}^{4} a(n,2)$

Dies bedeutet: a(1,2) + a(2,2) + a(3,2) + a(4,2)

Betrachten wir Spalte 2: 2 + 6 + 10 + 14 = 32

(b) $\sum_{n=1}^{4} a(1,n)$

Dies bedeutet: a(1,1) + a(1,2) + a(1,3) + a(1,4)

Betrachten wir Zeile 1: 1 + 2 + 3 + 4 = 10

(c) $\sum_{n=1}^{4} a(n,n)$

Dies bedeutet: a(1,1) + a(2,2) + a(3,3) + a(4,4)

Betrachten wir die Diagonale: 1 + 6 + 11 + 16 = 34

(d) $\sum_{n=1}^{4}(a(n,1) \times a(n,3))$

Dies bedeutet: a(1,1) × a(1,3) + a(2,1) × a(2,3) + a(3,1) × a(3,3) + a(4,1) × a(4,3)

Berechnung: 1 × 3 + 5 × 7 + 9 × 11 + 13 × 15 = 3 + 35 + 99 + 195 = 332

(e) $\sum_{n=1}^{4} a(n,1) \times \sum_{n=1}^{4} a(n,3)$

Dies bedeutet: (a(1,1) + a(2,1) + a(3,1) + a(4,1)) × (a(1,3) + a(2,3) + a(3,3) + a(4,3))

Berechnung: (1 + 5 + 9 + 13) × (3 + 7 + 11 + 15) = 28 × 36 = 1008

(f) $\sum_{n=1}^{4} a(n,3) + \sum_{m=1}^{4} a(4,m)$

Dies bedeutet: (a(1,3) + a(2,3) + a(3,3) + a(4,3)) + (a(4,1) + a(4,2) + a(4,3) + a(4,4))

Berechnung: (3 + 7 + 11 + 15) + (13 + 14 + 15 + 16) = 36 + 58 = 94

(g) $\sum_{n=1}^{4} \sum_{m=1}^{4} a(n,m)$

Dies bedeutet die Summe aller Werte in der Tabelle: 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 13 + 14 + 15 + 16 = 136

(h) $\sum_{n=2}^{4} \sum_{m=1}^{n-1} a(n,m)$

Dies ist etwas komplexer:

Für n=2: m geht von 1 bis 1, also haben wir a(2,1) = 5 Für n=3: m geht von 1 bis 2, also haben wir a(3,1) + a(3,2) = 9 + 10 = 19 Für n=4: m geht von 1 bis 3, also haben wir a(4,1) + a(4,2) + a(4,3) = 13 + 14 + 15 = 42

Gesamt: 5 + 19 + 42 = 66