

Statistische Grundlagen I

UE-KFOR SS 2005
Mag. Gabriele Tatzl

Arten der Statistik

Deskriptive Statistik

→ Zusammenfassung und Darstellung von Daten

Beschäftigt sich mit der Untersuchung u. Beschreibung von Gesamtheiten oder Teilmengen von Gesamtheiten durch z.B. Mittelwerte, Prozentsätze, Streuungsmaße, Korrelationskoeffizienten etc.

UE-KFOR SS 2005
Mag. Gabriele Tatzl

Arten der Statistik

Inferenzstatistik bzw. schließende Statistik

→ Ermöglicht empirischen Wissenschaften objektive Entscheidungen über die Brauchbarkeit problemorientierter, prüfbarer Hypothesen. Untersucht eine repräsentative Teilmasse der Grundgesamtheit (Population) u. schließt von dieser Teilmasse auf die Charakteristika der Grundgesamtheit.

Deskriptive Statistik Skalenniveau

➤ Nominalskala

Setzt Gleichheit oder Ungleichheit von Eigenschaften (einfache Klassifikation diskreter Merkmale wie z. B. Geschlecht) bzw. Möglichkeit mehrklassiger Einteilungen (z.B. in Berufe, Haarfarbe, Studienrichtung etc.) in Kategorien voraus. Kategorien müssen exakt definiert, sich gegenseitig ausschließend und erschöpfend sein.

Statistische Operationen:

- ✓ Häufigkeitsverteilung
 - ✓ Modalwert
-

Deskriptive Statistik Skalenniveau

➤ Ordinalskala (Rangskala)

Aussage über Beziehung zwischen den Messwerten, wobei Messwerte in einem Über- bzw. Unterordnungsverhältnis stehen. Die statistische Analyse läuft auf die Auswertung von Ranginformationen hinaus (z.B. Präferenzen für Automarken). Größe des Unterschieds zwischen 2 Objekten ist hier allerdings nicht definiert!

Statistische Operationen:

- ✓ Auswertung v. Ranginformationen,
 - ✓ Median
-

Deskriptive Statistik Skalenniveau

➤ Intervallskala (metrisch)

Über den Unterschied zweier Messwerte kann ausgesagt werden, ob er größer, gleich oder kleiner ist als der Unterschied zweier anderer Messwerte. Die Differenz der Messwerte ist genau quantitativ definiert. Skalenwerte einer Intervallskala können also bezüglich ihrer Differenzen (und Summen) verglichen werden wie z.B. Längenmaße, IQ-Skala.

Statistische Operationen:

- ✓ Differenzen und Summen
 - ✓ Arithmetischer Mittelwert
-

Deskriptive Statistik Skalenniveau

Nominal-, Ordinal- u. Intervallskalen sind abwärts kompatibel, d.h. eine Intervallskala kann auch als Ordinal- oder Nominalskala behandelt werden, eine Ordinal- auch als Nominalskala.

Bestimmung des Skalenniveaus in d. Sozialwissenschaften:

„Per-Fiat Messungen“ (d.h. Messungen „durch Vertrauen“)
Viele Messungen in der Forschungspraxis sind „Per-fiat“-Messungen für die Erhebungsinstrumente (Fragebögen, Tests etc.) konstruiert werden, von denen man annimmt, sie würden das jeweilige Merkmal auf einer Intervallskala messen, damit die gesamten statistischen Operationen eingesetzt werden können.

Deskriptive Statistik Statistische Kennwerte

Die statistischen Kennwerte haben die Funktion, über spezielle Eigenschaften einer Merkmalsverteilung summarisch Auskunft zu geben. Statistische Kennwerte wie z.B. das arithmetische Mittel oder die Standardabweichung werden berechnet, um eine Stichprobe hinsichtlich der zentralen Tendenz bzw. der Dispersion ihrer Messwerte zu beschreiben.

Deskriptive Statistik

Maße der zentralen Tendenz

➤ Modus (Modalwert)

Mo ist der am häufigsten in einer Verteilung vorkommende Messwert bzw. in der grafischen Darstellung einer Verteilung der Wert, bei dem die Verteilung ihr Maximum hat.

Unterscheidung: bimodale und breitgipflige Verteilung

Der Modus kann für Daten jeden Skalenniveaus bestimmt werden.

Deskriptive Statistik

Maße der zentralen Tendenz

➤ Median

Der Median (Zentralwert einer Verteilung) ist der Wert, der eine nach ihrer Größe geordnete Rangreihe halbiert. Ist der Wert, von dem alle übrigen Werte so abweichen, dass die Summe der Absolutbeträge ein Minimum ergibt.

Haben bsp. 9 Vpn die Messwerte 3,5,6,7,9,11,15,16,19 erhalten, so lautet der Median $Md=9$

Der Median setzt mindestens Ordinalskalenniveau voraus.

Deskriptive Statistik

Maße der zentralen Tendenz

➤ **Arithmetischer Mittelwert**

Das arithmetische Mittel (AM) ist das gebräuchlichste und am häufigsten verwendete Maß zur Kennzeichnung der zentralen Tendenz einer Verteilung.

Es wird berechnet, indem die Summe aller Werte durch die Anzahl aller Werte dividiert wird.

Die Berechnung des AM setzt voraus, dass das untersuchte Merkmal intervallskaliert ist.

Deskriptive Statistik

Dispersionsmaße/Streuungsmaße

Ähneln sich zwei Verteilungen hinsichtlich ihrer zentralen Tendenz, können sie wegen unterschiedlicher Streuungen (Dispersionen) der einzelnen Werte stark voneinander abweichen. Während Maße der zentralen Tendenz angeben, durch welchen Wert eine Verteilung am besten repräsentiert ist, informieren **Dispersionsmaße** über die **Unterschiedlichkeit der Werte**.

Deskriptive Statistik

Dispersionsmaße/Streuungsmaße

➤ **Variationsbreite (Spannweite)**

Das einfachste Dispersionsmaß ist die Variationsbreite („range“), die angibt, in welchem Bereich sich die Messwerte befinden.

Sie wird ermittelt, indem man die Differenz aus dem größten und kleinsten Wert bildet.

Deskriptive Statistik

Dispersionsmaße/Streuungsmaße

➤ **AD-Streuung**

Informationsreicher als die Streubreiten aller od. einiger Werte ist die AD-Streuung („average-deviation“), die den Durchschnitt der in Absolutbeträgen gemessenen Abweichungen aller Messwerte vom AM angibt.

$$\text{AD} = \frac{\text{Summer der Abweichungen (Absolutbeträge)}}{n}$$

Deskriptive Statistik

Dispersionsmaße/Streuungsmaße

➤ Varianz (s^2) u. Standardabweichung (s)

Sind die gebräuchlichsten Maße zur Kennzeichnung der Variabilität. Berechnung: Die Summe der quadrierten Abweichungen aller Messwerte vom arithmetischen Mittel, dividiert durch die Anzahl aller Messwerte. Da ein solches Maß aber nur schwer interpretierbar ist, wird die Quadrierung rückgängig gemacht, indem man die Wurzel aus der Varianz berechnet. Der positive Wert dieser Wurzel $\rightarrow s$ (od. kurz Streuung)

s^2 u. s setzen intervallskalierte Merkmale voraus

Deskriptive Statistik

Dispersionsmaße/Streuungsmaße

➤ Variationskoeffizient (Variabilitätskoeffizient)

Variationskoeffizient relativiert die Standardabweichung am Mittelwert und drückt damit die Standardabweichung in Mittelwertseinheiten aus. Maß wird eingesetzt, wenn Streuungen von Verteilungen mit unterschiedlichen Mittelwerten zu vergleichen sind u. Mittelwert u. Streuung voneinander abhängen.

$$VK = \frac{\text{Standardabweichung}}{\text{Mittelwert}} * 100\%$$

Deskriptive Statistik

Zusammenfassung

Statistische Methoden zur Beschreibung der Daten in Form von Grafiken, Tabellen oder einzelnen Kennwerten
= deskriptive Statistik.

1. Tabellen u. Grafiken informieren über die gesamte Verteilungsform.
2. Maße der zentralen Tendenz (z.B. Mittelwert) geben beispielsweise an, welches Alter alle Probanden am besten charakterisiert.
3. Dispersionsmaße (z.B. Streuung) kennzeichnen beispielsweise die Unterschiedlichkeit der Befragten in bezug auf das Alter.