

# Deep Dive (Anne)

## Univariat - eine Variable

### Modus

Bei Daten, die mindestens nominalskaliert sind (also kategorial), kann man den Modus berechnen. Der Modus als Maß der zentralen Tendenz ist der Wert, den die Variable am häufigsten annimmt (z.B. das lokale Maximum einer Normalverteilung). Er wird, im Gegensatz zum Mittelwert (bei metrischen Daten), nicht durch extreme Werte verzerrt (Kaliyadan & Kulkarni, 2019).

comment.an.mich: 3 Bilder von skewed distr. raussuchen, wo man Einfluss auf mode, median, mean sieht

### Median

Bei Daten, die mindestens ordinalskaliert (kategorial) sind, kann man den Median berechnen. Der Median als Maß der zentralen Tendenz ist die Stelle der Verteilung, über bzw. unter der je 50% der Daten liegen.

Der Median wird, im Gegensatz zum Mittelwert, nicht durch extreme Werte (outliers) verzerrt (Crump et al., 2018).

### Mittelwert

Bei Daten, die mindestens intervallskaliert sind (also metrisch), kann man den Mittelwert (Durchschnitt) berechnen. Der Mittelwert als Maß der zentralen Tendenz ist die Summe aller Werte, die die Variable angenommen hat, geteilt durch die Anzahl dieser Werte.

Achtung: Der Mittelwert wird durch extreme Werte (outliers) verzerrt. Bei sehr asymmetrischen Verteilungen ist ggf. der Median ein besseres Maß der zentralen Tendenz (Crump et al., 2018).

### Standardabweichung

Die Standardabweichung ist ein Streuungsmaß, gibt also an, wie stark die Daten um den Mittelwert streuen. Je verschiedener die Werte sind, desto größer die Standardabweichung. Sie ist die Wurzel aus der Varianz einer Variablen und liegt somit auf demselben Skalenniveau wie der Mittelwert.

Innerhalb der ersten Standardabweichungen über und unter dem Mittelwert einer Normalverteilung liegen ca. 68% der Daten. Innerhalb der ersten zwei Standardabweichungen über und unter dem Mittelwert einer Normalverteilung liegen mehr als 95% der Daten.

comment.an.mich: NormalVT; Binomial;  $X^2$  etc. einfügen?

## Multivariat - mehrere Variablen

### t-Test

Allgemein vergleichen t-Tests die Mittelwerte zweier Gruppen mithilfe einer t-verteilten Statistik. Man hat also zwei mindestens intervallskalierte Variablen, die man miteinander vergleichen möchte. Bei einer gerichteten Hypothese erfolgt die Testung einseitig, bei einer ungerichteten Hypothese testet man zweiseitig. Ist man sich nicht sicher über die Richtung des erwarteten Effekts, lohnt es sich, zweiseitig zu testen. In theoretisch gut begründeten kann man auch einseitig testen.

Da t-Tests zu den parametrischen Verfahren gehören, müssen Annahmen erfüllt sein, um diese berechnen zu dürfen:

- ausreichend große Stichprobe (Faustregel  $n=30$ )
- normalverteilte Daten

Sind die Annahmen verletzt, kann man auf nicht-parametrische Alternativen des jeweiligen t-Tests ausweichen. Die Unterarten des t-Tests haben z.T. zusätzlich zu den oben genannten noch weitere Annahmen, für die bei Verletzung ein Alternativtest am Ende des Abschnitts aufgelistet wird. (Crump et al., 2018)

**one-sample t-Test** Der one-sample t-Test vergleicht einen Stichprobenmittelwert mit einem (geschätzten) Populationsmittelwert, um zu schauen, ob die Stichprobe mit ausreichender Wahrscheinlichkeit aus dieser Population stammt oder aus einer anderen. Dazu braucht man Mittelwert und Standardabweichung der Stichprobe und der Population und den Standardfehler der Mittelwertsverteilung. Da so gut wie immer die Statistiken der Population unbekannt sind, muss man diese schätzen.

Ist der t-Test signifikant, gibt es eine Abweichung zwischen der Stichprobe und der Population, die mit ausreichender Wahrscheinlichkeit nicht zufällig zustande gekommen ist. Ist er nicht signifikant, geht man davon aus, dass die Stichprobe aus der Population stammt.

Nicht-parametrische Alternative: Wilcoxon Signed-Rank Test (Crump et al., 2018)

**correlated-samples t-Test** Der correlated-samples t-Test (auch paired-samples t-Test) ist dem one-sample t-Test sehr ähnlich. Er wird häufig für within-subjects Experimente genutzt (z.B. die Gedächtnisleistung einer Person vor einem Training wird mit ihrer Leistung nach einem Training verglichen) oder wenn Versuchspersonen aus zwei Gruppen miteinander verbunden sind (z.B. Zwillinge, Paare etc.). Die Berechnung ist ähnlich wie beim one-sample t-Test. (Crump et al., 2018)

Nicht-parametrische Alternative: Wilcoxon Signed-Rang Test (wenn die Differenzen nicht normalverteilt sind; Variable darf ordinalskaliert sein)

**independent-samples t-Test** Der independent-samples t-Test wird für between-subjects Experimente (z.B. der Mittelwertsvergleich zweier Gruppen, die verschiedene Treatments bekamen) genutzt, ähnelt aber ansonsten stark den bereits genannten t-Tests. (Crump et al., 2018)

Nicht-parametrische Alternative: Welch-Test (wenn die Annahme der Varianzhomogenität der beiden unabhängigen Gruppen verletzt ist); Mann-Whitney U-Test (wenn die Annahme der Normalverteilung verletzt ist; Variable darf ordinalskaliert sein)

## ANOVA

Eine Varianzanalyse (ANOVA) ähnelt strukturell den t-Tests. Die F-Statistik, die bei einer ANOVA verwendet wird, ist eine quadrierte t-Statistik. Man nutzt Varianzanalysen, um herauszufinden, ob gefundene Mittelwertsunterschiede zwischen Gruppen überzufällig sind oder nur durch Messfehler zustande kamen. Im Gegensatz zu den t-Tests, kann man mit einer ANOVA mehr als zwei Gruppen miteinander vergleichen. Das ist hilfreich, wenn es z.B. mehr als nur zwei Experimentalbedingungen gab, die miteinander verglichen werden sollen. Einfach mehrere t-Tests zu berechnen, würde das die Wahrscheinlichkeit eines Alpha-Fehlers erhöhen und ist daher keine sinnvolle Alternative.

Annahmen der ANOVA: - mindestens intervallskalierte Daten - unabhängige und zufällige Ziehung von k Stichproben - gleiche Größe der k Stichproben (oder Normalverteilung der Daten und Varianzhomogenität der k samples müssen gelten)(Crump et al., 2018)

**One-factor ANOVA** Eine einfaktorielle ANOVA benutzt man, wenn man eine unabhängige Variable (UV; Faktor) mit mindestens zwei (sinnvollerweise mindestens drei, sonst ginge auch ein t-Test) Faktorstufen hat. Man vergleicht dann im Prinzip auch die Mittelwerte der Faktorstufen miteinander, geht aber einen "Umweg" über die Varianzen. Die Versuchspersonen der einzelnen Faktorstufen sind dabei unkorreliert, wie beim independent-samples t-Test. Die resultierende F-Statistik ergibt das Verhältnis aus erklärbarer Varianz durch die experimentelle Manipulation und der Fehlervarianz.

Mit einer ANOVA kann man nur herausfinden, ob mindestens eine Faktorstufe sich signifikant von mindestens einer weiteren unterscheidet. Um herauszufinden, welche Faktorstufen sich unterscheiden, muss man post-hoc Tests durchführen. Man kann nur Haupteffekte finden, da es nur einen Faktor gibt. Um mögliche Interaktionen aufzudecken, braucht man Daten von mindestens zwei unabhängigen Variablen. (Crump et al., 2018)

Gängige post-hoc Tests (Auswahl): - Tukey HSD Test - Least Significant Difference (LSD) - Bonferroni

**Repeated-measures ANOVA** Eine repeated-measures ANOVA nutzt man für within-subjects Designs. Man hat eine unabhängige Variable (UV; Faktor) mit mindestens zwei bzw. drei Faktorstufen. Im Gegensatz zur one-factor ANOVA sind aber die Versuchspersonen nicht unabhängig voneinander, sondern man erhebt z.B. Daten derselben Personen zu drei Messzeitpunkten und vergleicht dann sinngemäß jede Person mit sich selbst zu verschiedenen Zeitpunkten. Nach wie vor werden aber nur Daten einer unabhängigen Variable erhoben, weshalb es auch hier nur Haupteffekte geben kann und noch keine Interaktionen. (Crump et al., 2018)

**Factorial ANOVA** Um Interaktionen zweier (oder mehrerer) Faktoren finden zu können, braucht man eine faktorielle ANOVA, d.h. man hat nun mehr als nur eine unabhängige Variable mit mehreren Faktorstufen. Beliebte Forschungsdesigns, wie das 2x2 factorial design, können mithilfe einer faktoriellen ANOVA berechnet werden. Faktorielle ANOVAs erlauben es, Haupteffekte und Interaktionseffekte der Faktoren auf die abhängige Variable zu messen und können sowohl für within- als auch für between-subjects Experimente genutzt werden. (Crump et al., 2018)

## Lineare Regression

Bei der linearen Regression möchte man mithilfe einer oder mehrerer unabhängigen Variablen (Prädiktoren) eine abhängige Variable vorhersagen. Die unabhängigen Variablen wie auch die abhängige Variable sind metrisch. Grafisch dargestellt ist die Regressionsgerade die Linie, die die Daten am besten beschreibt, d.h. zu der die Abstände von jedem Datenpunkt eines Scatterplots aus minimal sind. Diese Abstände zeigen den Messfehler an. Gäbe es keinen Messfehler, würden alle Datenpunkte auf der Regressionsgerade liegen (Crump et al., 2018; UZH, 2023).

## Chi2

Der Chi2 Test (Kontingenzanalyse) gehört zu den nicht-parametrischen Verfahren, es wird also keine Annahme über die Verteilung der zugrundeliegenden Daten gemacht. Er untersucht den Zusammenhang zweier nominal- bis ordinalskalierten Variablen, die in einer "Kreuztabelle" gegenübergestellt werden, indem beobachtete Häufigkeiten der Daten mit den erwarteten Häufigkeiten verglichen werden (nicht Mittelwerte!, s. t-Tests). Der Chi2 Test nutzt die Chi2 Statistik, ansonsten funktioniert das Signifikanztesten analog zu den bereits genannten Verfahren. Chi2 Tests können auf ein- und mehrdimensionale Zusammenhänge angewandt werden. (Lowry, 1998; UZH, 2023)

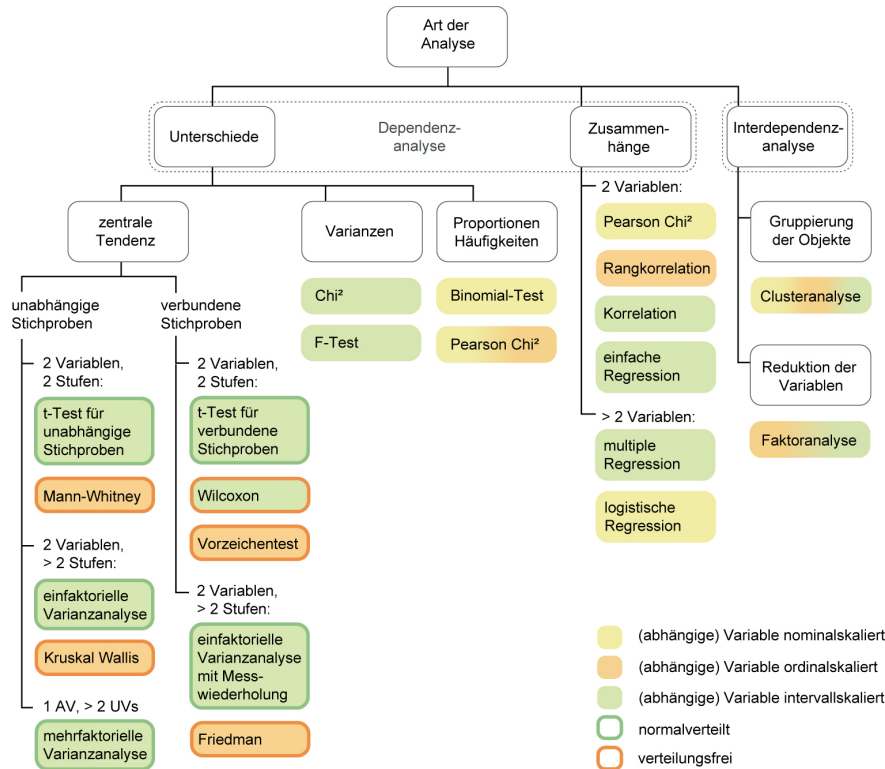
## Faktorenanalyse

Faktorenanalysen gehören zu den Interdependenzanalysen. Sie werden genutzt, um Strukturen in den Daten zu entdecken (explorative Faktorenanalyse) oder erwartete Strukturen zu bestätigen (konfirmatorische Faktorenanalyse). Dabei ist das Ziel, hoch korrelierende Variablen zu übergeordneten Faktoren zusammenzufassen. Gefundene Faktoren sollten möglichst gering mit anderen Faktoren korrelieren.

Für eine Faktorenanalyse braucht man: - eine ausreichend große Stichprobe - ausreichend viele Variablen - intervallskalierte Variablen (Häufig werden dennoch ordinalskalierte Variablen verwendet.)

Bei explorativen Faktorenanalysen (EFA) hat man keine Hypothesen über die Struktur, die geprüft werden soll, wie bei der konfirmatorischen Faktorenanalyse (CFA), die ein strukturüberprüfendes Verfahren darstellt. (UZH, 2023)

## Übersicht der Universität Zürich (UZH, 2023)



## Literatur

- Crumplab, M. J. C., Navarro, D. J., & Suzuki, J. (2018). *Answering questions with data*. <https://www.crumplab.com/statistics/>.
- Kaliyadan, F., & Kulkarni, V. (2019). Types of Variables, Descriptive Statistics, and Sample Size. *Indian Dermatology Online Journal*, 10(1), 82–86. [https://doi.org/10.4103/idoj.IDOJ\\_468\\_18](https://doi.org/10.4103/idoj.IDOJ_468_18)
- Lowry, R. (1998). *Concepts and Applications of Inferential Statistics*.
- UZH. (2023). Datenanalyse mit SPSS. In *Universität Zürich: Methodenberatung*. <http://www.methodenberatung.uzh.ch/de/data>