

Klausurfragen Evaluation

Die gängigen Formeln zu SD, SE, MW, z-Standardisierung, Konfidenzintervallen, T-Prüfgröße und Cohens D, Kombinatorik, Quadratsummenzerlegung, Satz von Bayes, Additions und Multiplikationssatz sollte man in Kopf haben

Psychologie als empirische Wissenschaft

→ Kommt so nicht vor

Hypothesenprüfen und Testen

1. Wie lautet die allgemeine Form prüfbarer Hypothesen, und wie wird sie in statistische Hypothesen übersetzt?

Hypothese: Aussagen, die aus Theorien abgeleitet sind, empirisch testbar und somit falsifizierbar.

Allgemeine Form prüfbarer Hypothesen:

-wenn.. dann Form (wenn Prädiktor... dann Kriterium)

-eine gültige Hypothese (Istzustand)

-eine Alternativ Hypothese (angenommener Zustand/ erwarteter Zustand)

-statistisch übersetzt in H_0 und H_1

Wichtig ist, theoretische Hypothese in statistische zu übersetzen d.h. operationalisierbar zu machen

→ wir wollen H_0 widerlegen, nicht H_1 annehmen: es könnte auch eine andere H_1 geben.

Bsp: Depressive sind nach Therapie weniger depressiv

AV: Depression

UV Therapie (ja/nein)

2. Was ist der Unterschied zwischen Unterschiedshypothese und Zusammenhangshypothese ? (Beispiele):

Zusammenhangshypothese:

Trifft Aussage über Zusammenhang zwischen Variablen bzw. Ausprägungen von Variablen

Bsp.: Mit zunehmender Intelligenzminderung ist die Reaktionszeit länger → es geht um Korrelationen oder Regressionsanalysen

Unterschiedshypothese:

Es geht um Differenz zwischen zwei Gruppen oder Veränderung über die Zeit → wird inferenzstatistisch gerechnet z.B. über T-test oder Varianzanalysen

Bsp. 1: Wirksamkeit einer Maßnahme

Bsp. 2: Reaktionszeit: je älter desto langsamer und Unterschied: Alte langsamer als junge

Inwiefern lassen sich Unterschieds- und Zusammenhangshypothese auf das gleiche Thema beziehen?

Unterschiedshypothese: Mädchen lesen mehr als Jungen

Zusammenhangshypothese: Je nach Geschlecht wird mehr gelesen. bzw. Geschlecht steht in Zusammenhang mit Häufigkeit des Lesens/Leseverhaltens.

3. Welche Information liefert der Ergebnissatz der Prüfung einer statistischen Unterschiedshypothese?

Er liefert den Unterschied zwischen zwei Variablen: ob es Mittelwertsunterschiede oder Unterschiede über die Zeit gibt

→ Es gibt aber gewisse Irrtumswahrscheinlichkeit (Alpha Fehler)

Keine Aussage über Wahrheit möglich, da immer Alpha Fehler möglich

Bei gerichteter Testung kann man Richtung des Unterschieds angeben, aber nicht die Größe des Unterschieds.

Nichtsignifikantes Ergebnis: heißt nicht, dass H_0 richtig ist, sondern H_0 konnte durch vorliegende Untersuchung nicht verworfen werden. Über die Richtigkeit der Aussagen kann keine Angabe gemacht werden.

4. Welche Voraussetzungen müssen im Allgemeinen für die Angabe von Auftretenswahrscheinlichkeiten beim statistischen Testen erfüllt sein? (Geben Sie ein Beispiel und erläutern Sie die zugrundeliegenden Annahmen über die Testsituation)

Voraussetzungen:

- 1) Variable muss multinomial sein → mehrere Ausprägungen einer Variable, unterscheidbares Merkmal (Fehleranzahl)
- 2) Unabhängigkeit der Stichprobe bzw. Vergleichsgruppe (Vergleich zwischen Personen oder zeitlicher Vergleich)

→ Bsp.: Wählen Frauen Parteien in anderen Häufigkeiten als Männer?

5. Erläutern Sie die Logik der Bestimmung von Konfidenzintervallen

Voraussetzung: Normalverteilung

Konfidenzintervall: Bezeichnung des Bereichs, der bei unendlicher Wiederholung des Zufallsexperiments mit einer gewissen zuvor festgestellten Irrtumswahrscheinlichkeit den wahren Wert einschließt.

Durch allgemeinere Intervallschätzung ist richtigere Aussage möglich als durch eine Punktschätzung (geschätzter Mittelwert der Population)

- Wie verteilen sich Daten um wahren Wert/Mittelwert?
- Formel: $CI = \text{Mittelwert von } x \pm Z(\alpha/2) * SE$

Intervalle bestimmen, wo liegen z.B. 99% der Daten? (1SD, 2SD und 2,5 SD)

6. Erläutern Sie die Logik der Prüfung von Mittelwertsunterschieden

Mittelwertsunterschiede gelten als Vergleichsmaß von Personen (unabhängig) oder von Werten innerhalb einer Stichprobe (abhängig). Ob die Unterschiede zwischen oder innerhalb von Personen nur auf Zufall beruhen oder ob der Mittelwertsunterschied wirklich signifikant ist, kann nur mit statistischen Verfahren geprüft werden.

a) für abhängige Stichproben:

Man beobachtet nur eine Stichprobe zu mehrere Zeitpunkten und bestimmt die Wahrscheinlichkeit mit der sich die MW zu den Zeitpunkten unterscheiden. Wenn nicht: H_0 verwerfen

Voraussetzung:

für Stichprobe unter 30 normalverteilte Daten.

Intervallskaliert

abhängige Stichprobe

b) für unabhängige Stichproben

bestimmte Wahrscheinlichkeit, dass sich MW von zwei oder mehr Stichproben z.B. Männer und Frauen (Stichproben), unterscheiden.

Voraussetzungen: intervallskaliert, unabhängige Daten, Homogenität der Populationsvarianzen. Normalverteilt für N_1 und N_2 unter 50.

7. Was ist die Logik des Signifikanztests allgemein?

1. Man bestimmt Auftretens-Wahrscheinlichkeit für bestimmte (Mittelwerts-) Differenz oder Wahrscheinlichkeit eines bestimmten Wertes. Dies geschieht mithilfe einer Prüfgröße z.B. F-Wert oder T-Wert)
→ Diese Wahrscheinlichkeit ist bedingte Wahrscheinlichkeit (Wahrscheinlichkeit des Wertes, gegeben dass H_0 gilt.
2. Wenn die Wahrscheinlichkeit zu klein (0,05 oder 0,01) ist, dann wird Test signifikant
Ist diese bedingte Wahrscheinlichkeit zu klein, dann kann die H_0 nicht gelten → ist zu unwahrscheinlich. Andernfalls beruht der Unterschied auf Zufall.
3. Wir nehmen H_1 an und verwerfen H_0 als zu unwahrscheinlich (unter einer gewissen Irrtumswahrscheinlichkeit!)
4. Allgemein gilt: Es wird signifikant: Wahrscheinlichkeit für H_0 wird zu klein.
p-Wert: Wahrscheinlichkeit für Mittelwertsunterschied, gegeben dass die H_0 gilt.

Was kann man aus signifikantem Ergebnis nicht ablesen?

Man kann nicht mit 100%iger Sicherheit sagen, dass H_1 zutrifft, es bleibt immer noch Irrtumswahrscheinlichkeit von 5 oder 1 Prozent.

Fehler 1. Art: richtige Hypothese wird zugunsten der Alternativhypothese abgelehnt

Fehler 2. Art: wenn eine falsche Nullhypothese beibehalten wird.

Man kann an Signifikanz bzw. Nicht-Signifikanz, ablesen wie groß Mittelwerts Unterschiede sind und in welche Richtung der Unterschied geht (nicht an p Wert, sondern an Prüfgröße kann man Richtung ablesen, aber nicht bei χ^2)

Wahrscheinlichkeitsbegriff

1. Für welche Arten von Ereignissen kann man Wahrscheinlichkeiten bestimmen?

Wahrscheinlichkeiten beziehen sich immer auf abstrakte Ereignisse (zb Wahrscheinlichkeit eine 1 zu würfeln), nicht auf konkrete Ereignisse (Ich habe eine 1 gewürfelt).

Man kann also nur für Zufallsereignisse eine Wahrscheinlichkeit bestimmen In der Psychologie sind das alle Experimente/Zufallsexperimente. Doch diese müssen **wiederholbar sein und mit stabiler Häufigkeit auftreten (d.h. bei sehr häufiger Wiederholung ungefähr gleich stabil auftreten)**

Elementar-Ereignis: Ereignisse, die nur ein mögliches Ereignis enthalten

Gegenereignis (zu A): enthält alle Elemente, die nicht Teil von A sind.

2. Vergleichen Sie die Wahrscheinlichkeits-Definitionen von Laplace, von Mises und Kolmogoroff. Warum gibt es drei Ansätze zum Wahrscheinlichkeitsbegriff? Welches grundlegende Problem hat die Wahrscheinlichkeitsdefinition nach LaPlace?

Es gibt mehrere, weil es ja nicht immer die gleiche Wahrscheinlichkeiten gibt.

Laplace	Mises	Kolmogoroff
<p>Hat ein Zufallsexperiment nur <u>endlich viele</u> Ergebnisse und haben diese alle die <u>gleiche Wahrscheinlichkeit</u>, so gilt für $P(A)$ eines Ereignisses:</p> <p>$P(A)$: Anzahl der Ereignisse, bei denen A eintritt</p> <p>Geteilt durch:</p> <p>Anzahl aller möglichen Ereignisse</p> <p>Bsp.: bei zweimaligen Münzwurf gibt es 36 mögliche Ereignisse für die Augenzahlkombinationen. Die Augensumme 9 kann bei vier Ereignissen auftreten (6,3; 3,6; 5,4;4,5) Wahrscheinlichkeit für Summe 9 ist also: $4/36 = 1/9$</p>	<p>Geht von relativer Häufigkeit aus bzw. Wahrscheinlichkeiten aufgrund von Beobachtungen</p> <p>Wahre Wahrscheinlichkeit ergibt sich Anzahl Stichproben: je mehr Durchläufe umso eher nähert man sich dem wahren Wert.</p> <p>Nicht alle haben die gleiche Wahrscheinlichkeit! $P(A) = \lim (n \text{ Anzahl der Versuche gegen unendlich}): nA/n$</p> <p>n: Gesamtzahl der Versuche nA: Versuche bei denen A beobachtet wurde.</p> <p>Wird n größer so wächst nA entsprechend an.</p> <p>Der Quotient nA/n strebt dabei gegen einen Grenzwert $p(A)$</p> <p>Man nennt diese Definition auch die statische Wahrscheinlichkeitsdefinition oder a posteriori Definition, da keine apriori Annahmen über die Ereignisse gemacht wurden, die Wahrscheinlichkeit wird nur induktiv über den Weg der Beobachtung gewonnen. Dabei muss angenommen werden, dass ein Grenzwert</p>	<p>Macht dasselbe wie Mises, nur K. hat noch Axiome.</p> <p>Stellt drei Axiome auf, die fundamentalen Eigenschaften von Wahrscheinlichkeiten beschreiben</p> <p>0. Unmögliches Ereignis = 0</p> <p>1. Für die Wahrscheinlichkeit eines zufälligen Ereignisses gilt: $0 < p(A) < 1$</p> <p>2. Für das sichere Ereignis gilt: $P(\Omega/\text{Stichproben-raum}) = 1$</p> <p>3. Disjunkte Ereignisse dürfen addiert werden.</p> <p>Wenn Ereignis E in disjunkte Teilereignisse zerlegbar ist, kann man die Wahrscheinlichkeit berechnen, indem man die Wahrscheinlichkeiten der disjunkten Ereignisse addiert: $P(A \text{ und } B) = P(A) + p(B)$</p>

	überhaupt existiert, dem die relative Häufigkeit zustrebt.	
Problem an der Laplace Definition:	Man kann nicht für jedes Zufallsexperiment annehmen, dass alle möglichen Ereignisse gleich wahrscheinlich sind. Die Definition von Laplace lässt sich also nicht immer anwenden.	Bsp.: Reißnagel: Spitze kann nach Wurf nach oben oder unten zeigen. Wahrscheinlichkeit ist nicht bekannt, aber eher nicht gleich

3. Erläutern Sie die Begriffe Zufallsexperiment, Stichprobenraum und Ereignisalgebra

Zufallsexperiment: Satz von Regeln, unter denen eine Beobachtung stattfindet. Es ist ein Vorgang, dessen Ergebnis in der Weise vom Zufall abhängt, dass man vor dem Experiment nicht weiß, zu welchem der möglichen Ereignisse das Experiment führen wird

1. **Es ist unter gleichen Bedingungen beliebig oft wiederholbar**
2. **Es besitzt mehrere, sich gegenseitig ausschließende Ereignisse**
3. **Ergebnisse sind zufallsbedingt**

Stichprobenraum=Merkmalsraum=Ergebnisraum. Dies ist die Menge aller möglichen Ereignisse eines Zufallsexperiments

Bsp. Stichprobenraum beim einmaligen Würfelwurf (1,2,3,4,5,6)

Ereignisalgebra=Sigma Algebra=die Potenzmenge zum Stichprobenraum: enthält alle Kombinationen von Ereignissen eines Zufallsexperimentes, auf die man wetten könnte und das unmögliche Ereignis (leere Menge)

Bei Ereignisalgebra geht es nicht um die wiederholte Anwendung eines Ereignisses.

4. Was sind disjunkte, was unabhängige Ereignisse?

Disjunkte Ereignisse: wenn das gleichzeitige Auftreten von 2 Ereignissen A und B unmöglich ist, schließen sich diese aus (sind disjunkt) → haben keine Schnittmenge

A und B= unmögliches Ereignis

Bsp: dass Kopf und Zahl gleichzeitig kommt bei einem Wurf oder man kann nicht gleichzeitig tot oder lebendig sein

Unabhängige Ereignisse:

A und B sind voneinander unabhängige Ereignisse, wenn das Auftreten von B die Wahrscheinlichkeit, dass A eintritt, nicht beeinflusst.

Bsp: Karten ziehen mit Zurücklegen. Zweiter Zug einer Karte wird vom ersten nicht beeinflusst

Achtung: disjunkte Ereignisse sind nicht unabhängig!

4. Welche wichtigen Sätze der Wahrscheinlichkeitslehre folgen aus Disjunktheit und Unabhängigkeit und was besagen sie?

Disjunktheit (wenn es keine Schnittmenge gibt, nicht gleichzeitig aufzutreten):

daraus folgt der Additionssatz: Ereignisse ohne Schnittmenge dürfen addiert werden um die Wahrscheinlichkeit für eines der Ereignisse zu erhalten

$$P(A \text{ UND/ODER } B) = P(A) + P(B) - P(A \text{ UND } B)$$

Bei Nicht-Disjunktheit: $P(A \text{ oder } B) = P(A) + P(B) - P(A \text{ UND } B)$

Bei Disjunktheit: $P(A \text{ oder } B) = P(A) + P(B)$

Stochastische Unabhängigkeit:

daraus folgt der Multiplikationssatz:

Multiplikation der Grundwahrscheinlichkeit mit einer bedingten Wahrscheinlichkeit die Verbundwahrscheinlichkeit

Bei Unabhängigkeit: $P(A \text{ +} B) = P(A) * P(B)$

Bei Abhngigkeit: $p(A+B) = p(A) * P(B/A) \text{ oder } P(B) * P(A/B)$

6. Was versteht man unter einer bedingten Wahrscheinlichkeit (Beispiel)?

Dies meint die Wahrscheinlichkeit eines Ereignisses B, gegeben das Ereignis A schon eingetreten ist.

$P(B/A) = p(A \text{ und } B) \text{ Verbundwahrscheinlichkeit (Schnittmenge A und B)} / p(A) \text{ (Grundwahrscheinlichkeit A)}$

Bsp: Ereignis A Zahl größer 3 und B „ungerade Zahl“

Nach dem Würfeln hat man die Information, dass A eingetreten ist ($p=0,5$). Nun ist nur eine Option bezüglich B günstig, also $p(B/A) = 1/6$ durch $1/2$

7. Was besagt der Satz von Bayes?

Der Satz beschreibt, wie sich die Wahrscheinlichkeit des Ereignisses A durch die Berücksichtigung des Ereignisses B verändert. Für 2 Ereignisse A und B (für B ungleich 0) lautet der Satz von Bayes:

$p(A/B)$ ist die Wahrscheinlichkeit des Ereignisses A unter der Bedingung, dass B eingetreten ist.

Der Satz sagt uns, wie wir „A gegeben B“ $p(A/B)$ ausrechnen können, wenn wir nur „B gegeben A“ haben und keine Verbundwahrscheinlichkeiten.

Die bedingte Wahrscheinlichkeit $p(A/B)$ wird mit der unbedingten Wahrscheinlichkeit verknüpft.

Satz von Bayes: $p(A/B) = p(A/B) * p(A) / p(B)$ oder

$$P(B/A) = p(B/A) * p(B) / p(A)$$

Exkurs: Weitere Formeln

Verbundwahrscheinlichkeiten: Grundwahrscheinlichkeit * bedingte Wahrscheinlichkeit
(Multiplikationssatz)

Satz der totalen Wahrscheinlichkeit: alle Verbundwahrscheinlichkeiten der Ereignisse, die ich suche, addieren

8. Was versteht man unter einer Kombination, was unter einer Permutation? (Beispiele)

Permutation: ist eine Menge mit n Elementen gegeben, so bezeichnet man die möglichen Anordnungen von allen n Elementen als Permutation.

Bei Permutationen wird **immer die Reihenfolge** berücksichtigt! Dies ist immer der Fall, wenn Dinge unterscheidbar sind (zb verschiedenfarbige Kugeln oder Kindern beim Geburtstag)

Bsp.: „Permutation ohne Wiederholung“: eine Anordnung von n Objekten, die alle unterscheidbar sind. Nachdem es für das erste Objekt n Platzierungsmöglichkeiten gibt, kommen für das folgende Objekt nur noch n-1 Optionen in Betracht usw.

Die Anzahl der möglichen Permutationen von n Objekten wird demnach durch die Fakultät angegeben

Fakultät: $n! = n * (n-1) * (n-2) * \dots * 1$

Es gibt z.B. $4! = 4 * 3 * 2 * 1$ mögliche Anordnungen von 4 verschiedenen farbigen Kugeln in einer Reihe ($4! = 24$)

Wenn Reihenfolge wichtig ist (bei Permutation) gibt es mehr Möglichkeiten/Ereignisse als bei Kombination.

Ohne Wiederholung: $n! / (n-k)!$

Mit Wiederholung: $n! / k!(n-k)!$

Dabei ist N Anzahl aller Elemente und k entnommene Elemente

Kombination: (auch ungeordnete Stichprobe genannt):

Bezeichnet eine Auswahl von Objekten aus einer Grundmenge, die jedoch

- 1.) **nicht alle Objekte** der Grundmenge/Stichprobe enthalten **muss** und
- 2.) bei der die Reihenfolge unberücksichtigt bleibt.

-> beides steht im Gegensatz zur Permutation

Können Objekte nur einmal gewählt werden: Kombination ohne Wiederholung

Können Objekte mehrfach gewählt werden: Kombination mit Wiederholung

Kombination mit Wdh.: $(n+k-1)! / k! * (n-1)!$

Kombination ohne Wdh: $N! / K!(n-k)!$

Vierfelderschema der Formeln

	Mit Wiederholung	Ohne Wiederholung
Reihenfolge wichtig/Permutation	N hoch k	$N! / (n-k)!$
Reihenfolge egal/Kombination	$(n+k-1)! / k! * (n-1)!$	$N! / K!(n-k)!$

9. Was versteht man unter einer Wahrscheinlichkeitsverteilung und einer Verteilungsfunktion?

Wahrscheinlichkeitsfunktion:

man kann direkt Wahrscheinlichkeit ablesen

Gibt an, wie sich Auftretenswahrscheinlichkeiten für exakte Trefferzahlen verteilen (beispielsweise: die 3 würfeln), also $1/3$

Verteilungsfunktion: Kumulierung einer diskreten Wahrscheinlichkeitsfunktion: liest Wahrscheinlichkeit für einen Wert ab und für alle Möglichkeiten darunter

Gibt an, wie wahrscheinlich es ist eine bestimmte Würfelzahl oder eine darunter zu würfeln (beispielsweise höchstens die 3 zu würfeln) also $3/6$

10. Was versteht man unter einer Binomialverteilung und in welcher Art Zufallsexperimenten tritt sie auf? Welche Parameter hat sie?

Die Binomialverteilung ist eine diskrete Wahrscheinlichkeitsverteilung, die die Anzahl von Erfolgen in einer Serie gleichartiger und unabhängiger Versuche beschreibt, die jeweils zwei mögliche Ergebnisse haben (Erfolg und Misserfolg).

Also tritt sie in Zufallsexperimenten auf, die voneinander unabhängig sind und jeweils zwei mögliche Ergebnisse haben => Bernoulli Experimente

(Beim Bernoulli Experiment muss die Wahrscheinlichkeit nicht gleich sein, aber die Summe der Wahrscheinlichkeiten muss 1 sein.)

Parameter der Bernoulli Verteilung:

N: Anzahl der Versuche/Wiederholungen

Mü bzw. P=Erfolgswahrscheinlichkeit/Trefferwahrscheinlichkeit

X: tatsächliche Treffer

- ➔ Dichotomer Ausgang
- ➔ Diskrete Wahrscheinlichkeitsverteilung

11. Was ist eine Normalverteilung, in welcher Art Zufallsexperimenten tritt sie auf, und welche Parameter hat sie? Wie hängt sie mit der Binomialverteilung zusammen?

Die Normalverteilung ist eine symmetrische, eingipflige und glockenförmige **stetige** Wahrscheinlichkeits- und **Zufallsverteilung**

Bei richtiger Skalierung des Koordinatensystems: Modus, Median und Mittelwert fallen auf einen Punkt.

Die Normalverteilung kann als das Ergebnis des Zusammenwirkens vieler gleichartiger Zufallsprozesse bezeichnet werden (sogar Messfehler sind normalverteilt)

Wenn wir eine ganz große Zahl/Wiederholungen des Experiments mit identischer Wahrscheinlichkeitsverteilung aufsummieren, dann ist diese Summe normalverteilt:

- ⇒ zentraler Grenzwertsatz: Summe einer großen Zahl identisch verteilter Zufallsvariablen (z.B. millionen mal einen Würfel würfeln und die Summe der Augenzahlen jedes Mal aufaddieren: es entsteht eine Normalverteilungskurve

- ⇒ polytomer Ausgang (Gegenteil von dichotom): es gibt mehr als zwei Möglichkeiten
- ⇒ stetig: ohne Unterbrechung und Sprünge

Parameter der Normalverteilung:

Erwartungswert μ : legt fest, an welcher Stelle die Normalverteilung ihr Maximum haben wird

Standardabweichung: je höher, desto breiter die Normalverteilung

Der Zusammenhang von Normalverteilung und Binomialverteilung:

Die Binomialverteilung kann zur Annäherung an die Normalverteilung verwendet werden, wenn n hinreichend groß ist.

Wenn n unendlich ist, dann sind Binomial- und Normalverteilung identisch

Exkurs: wie prüfe ich auf Normalverteilung?

1) mit Chi² Test: wenn nicht signifikant: Normalverteilung gegeben

2.) mit QQ Plot: wenn wir eine Gerade sehen, ist Normalverteilung gegeben: Je stärker gekrümmt die Linie ist, umso mehr ist Normalverteilung verletzt.

Normalverteilung, Skalen und Transformation

1. Was sind Skalenniveaus? Welche Skalen-Niveaus gibt es?

Ein Skalenniveau drückt aus, inwieweit sinnvolle Rechenoperationen angewendet werden können und ist eine wichtige Eigenschaft von Merkmalen bzw. von Variablen. Je nach der Art eines Merkmals bzw. je nachdem, welche Vorschriften bei seiner Messung eingehalten werden können, lassen sich verschiedene Stufen unterscheiden (s.u.). Skalen beschreiben die Qualität der Daten!

Skalen: Nominal, Ordinal, Intervall, Verhältnis, Absolut

Welche Kennwerte kann man ablesen?

Nominal: reine Unterschiede z.B. Geschlecht → keine Reihenfolge

Ordinal: natürliche Rangordnung zwischen den Ausprägungen (größer/kleiner) Bsp.: Schulnote oder Schulabschlüsse → keine Differenzen

Intervallskala: Rangunterschiede und Differenzen/gleiche Abstände zwischen den Werten z.B. Temperatur, Lautstärke in Dezibel → keine Verhältnisse

(Bsp. 30 Grad ist nicht doppelt so heiß wie 15 Grad, aber 5 Grad ist 5 Grad mehr als 0 Grad.)

Verhältnis: absoluter Nullpunkt existiert und Verhältnisse zwischen den Werten. Bsp.: Körpergröße in cm

→ Keine natürliche Einheit (es könnten statt cm auch Feet sein etc..)

Absolutsskala: natürliche Einheit Bsp.: Kinderzahl

Skalen:

		Nominalskala	Ordinalskala	Intervallskala	Verhältnisskala	Absolutskala
Mathematisch logische Operationen	x /			(x)	x	x
	+ -			x	x	x
	< >		x	x	x	x
	= ≠	x	x	x	x	x
Natürlicher Nullpunkt						x
Berechenbare Kennwerte		Modus	Median Extrema Quantile	Mean Varianz SD	Alle vorherigen	Alle vorherigen
Qualitativ/Quantitativ		Qualitativ	Qualitativ	Quantitativ	Quantitativ	Quantitativ
Beispiel		Geschlecht	Schulnoten	Temperatur in Celcius	Geld/Einkommen	Größe in cm

Adaptiert von: matheguru.com

Achtung:

In der Tabelle ist das Bsp. Größe in cm meines Erachtens falsch, da es nicht natürlich ist.

Bei Intervallskala ist Multiplikation bzw. Division nur bedingt sinnvoll weil ein Nullpunkt fehlt: 10 Grad ist eben NICHT doppelt so warm wie 20 Grad.

2. Warum hat die Intervallskala in der Statistik eine besondere Bedeutung?

-Intervallskalen Niveau ist eine wichtige Voraussetzung für viele Tests → höchstmögliches Skalenniveau in der Psychologie

-Psychologen „machen“ Dinge intervallskaliert (zb Likertskala). Dies kann man machen, weil sich Unterschiede rausmitteln.

Inwiefern treten sie natürlicherweise auf?

Intervallskalen treten natürlicherweise auf, da wir oft annehmen, dass ein Merkmal in der Population normalverteilt ist. Dieses muss intervallskaliert sein, da erst ab dann Mittelwert und Standardabweichung ausgerechnet bzw. geschätzt werden können.

Probleme: ungeprüfte Annahme der Intervallskala in psychologischen Untersuchungen → oft ist Intervallskala nicht gegeben.

3. Welche Skalentransformationen sind auf Intervallskalen zulässig? Was bedeutet zulässige Transformation auf einem Skalenniveau?

Eine Skalentransformation wird als zulässig bezeichnet, wenn das Skalenniveau der ursprünglichen Skala erhalten bleibt. (Eine Umwandlung in eine andere Skala ist nur zu einer niedrigeren möglich und dabei auch immer mit Informationsverlust verbunden.)

Lineare Transformation sind bei Intervallskala zulässig: alle Werte werden in gleichem Maße verändert.

Lineare Transformationen: addieren, subtrahieren, multiplizieren und dividieren

5. Was sind die wichtigsten und geläufigsten Transformationen?

- z-Standardisierung → Mittelwert 0, Standardabweichung 1 → damit könne Zufallsvariablen mit beliebigen MW und SD in solche transformiert werden, die einen definierten Mittelwert und Standardabweichung aufweisen.
- Nullpunkt-Transformation/Zentrierung (in Regressionen um Interkorrelationen zu vermeiden)
- Einheitstransformation

Wie geht z-Standardisierung?

z-Standardisierung jedes Datenpunktes, z-Werte mit der neuen SD multiplizieren und dem neuen MW addieren

Bsp: neuer MW wäre bei IQ: 100, SD wäre 15

Formel: $z = (x - \text{MW von } x) / \text{SD}$

6. Inwieweit verändert sich die Wahrscheinlichkeitsverteilung bei linearen Funktionen im Hinblick auf die folgenden Punkte?

	Verteilungsform:	Varianz	Mittelwert	Produkt Moment- Korrelation	Kovarianz: Maß für den linearen Zusammenhang zwischen zwei Variablen
addieren	Nein, nur nach rechts verschoben	Bleibt gleich	Vergrößert sich um addierten Wert	Bleibt gleich	Bleibt gleich
subtrahieren	Nein, nur nach links	Bleibt gleich	Verringert sich um subtrahierten Wert	Bleibt gleich	Bleibt gleich
multiplizieren	Ja, wird breiter, Form gleich	Vergrößert sich um das multiplizierte ins Quadrat	Vergrößert um das multiplizierte	Bleibt gleich	Vergrößert sich um das multiplizierte ins Quadrat
dividieren	Ja, wird enger	Verkleinert sich um das Dividierte zum Quadrat	Verringert sich um das Dividierte	Bleibt gleich	Verkleinert sich um das Dividierte zum Quadrat

7. Was hat die z-Transformation mit der Normalverteilung zu tun? Wie verändert sich die Normalverteilung von Daten, wenn man die Daten einer z-Transformation unterzieht?

Normalverteilung ist die Voraussetzung für z-Standardisierung.

Form der Verteilung bleibt gleich, aber es gibt eine andere Skala (von minus 3 SD bis plus 3SD). Man kann so berechnen, wo genau 68% (1SD), 95% (2SD) und 99% (2,5SD) der Daten liegen.

8. Warum ist die Normalverteilung eine so wichtige Verteilung? In welchem Zusammenhang tritt sie beim Testen von Hypothesen auf?

Viele Variablen in der Natur sind normalverteilt bzw. bei großen Datenmengen greift der zentrale Grenzwertsatz. Sie hilft, weil wir wissen, was der Normbereich ist. Dann können wir genau sagen, ab wann jemand nicht mehr „normal“ ist.

Die Normalverteilung hat große Bedeutung, weil sämtliche normalverteilte Variablen durch eine einfache Transformation in die Standardnormalverteilung überführbar sind. Auch Stichprobenmittelwerte sind normalverteilt (Zentraler Grenzwertsatz)→ Grundstein um Skalen zu konstruieren und Werte einzuordnen.

Für das Testen von Hypothesen ist die Normalverteilung der Werte Voraussetzung:

Für die Annahme oder Ablehnung einer Hypothese wird ein gewisser Bereich unter der Gauß'schen Normalverteilungskurve festgelegt -> Ablehnungsbereich der H_0 → Signifikanzberechnung

Standardnormalverteilung ist Voraussetzung für z-Transformation

9. Wie kann die Annahme der Normalverteilung bei der Konstruktion von Skalen benutzt werden? Welche bekannten Beispiele für Skalenkonstruktion mithilfe der Normalverteilung gibt es in der Psychologie?

Man kann Personen in über-, unter- und normaldurchschnittlich klassifizieren und somit vergleichen. Um die Testergebnisse eines Probanden in verschiedenen Tests oder die Ergebnisse unterschiedlicher Probanden in einem Test miteinander vergleichen zu können, können unterschiedlichen Normskalen konstruiert werden, wenn die Normalverteilung angenommen wird.

Man geht davon aus, dass einige human- und sozialwissenschaftlich relevante Merkmale zumindest annähernd normalverteilt sind, weil dass die Grundlage für die Konstruktion von entsprechenden Skalen ist.

(Man erstellt innerhalb von Skalen Items, um Merkmalsausprägungen zu erfassen. Die Daten werden dann anhand einer repräsentativen Stichprobe in Bezug auf die Population normiert, um Vergleiche zwischen den Probanden anstellen zu können. (Falls Testwertverteilung nicht normalverteilt ist, könnte es sein, dass auch das erhobene Merkmal in der Population nicht normalverteilt ist.)

Bsp:

Neo FFI : faktorenanalytische Skalenkonstruktion: Beginn mit möglichst umfangreichem, repräsentativen Stichproben von Items und Personen

HAWIE: kriterienorientierte Skalenkonstruktion: VPs unterscheiden sich hinsichtlich eines Kriteriums (z.B. Intelligenz): Items sollen die VP je nach IQ gut differenzieren
Viele Fragen müssen auf den mittleren Bereich abzielen, wo Großteil der Personen liegt (höchste Schwierigkeit der Fragen in der Mitte) und nur einige an den Rändern.

9. Was ist die Thurstone Skala?

Eine Skalierungsmethode, die z.B. für Einstellungsmessungen verwendet wird. Einer großen Zahl an Beurteilern wird eine Sammlung von Items vorgelegt, die neutrale wie auch extreme Ausprägungen über ein Einstellungsobjekt enthalten. Beurteiler sollen dabei nicht ihre eigene Einstellung bewerten, sondern die Items an sich. Die endgültige Thurstone Skala besteht nur aus Items, bei denen sich die Beurteiler über die Einstellung einig sind und die kann den Testpersonen vorgelegt werden.

→ Kurz gesagt: Sie fragen Leute, was für sie z.B. Gewissheit ist und machen daraus einen Test. Dann lassen sie andere raten, wie gut sie die Fragen finden d.h. wie gut typische Merkmale des Konstrukts erfasst werden und schmeißen sie dann rein oder raus.

10. Inwiefern sind psychometrische Persönlichkeitstest Skalen?

Eine Skala erfasst die Ausprägungen eines Merkmals - genauso sollen bei Persönlichkeitstests die individuellen Merkmale erfasst werden. Außerdem werden die Eigenschaften bei Persönlichkeitstests zumeist auch auf einer Skala eingestuft.

Persönlichkeitstests sind Skalen weil sie Likert Skalen nutzen, d.h. man kann die Konstrukte in Zahlen übersetzen.

Statistische Prüfverfahren

1a) Welche Voraussetzungen hat der T-test für unabhängige Stichproben?

- nicht zweimal die gleiche Person: unabhängige Ziehung
- Varianzhomogenität (→ wird geprüft über Fmax Test)
- mindestens intervallskaliert
- Normalverteilung oder beide Stichproben Größe zusammen größer als 50.

1b) Welche Voraussetzungen hat der T-test für abhängige Stichproben?

- Normalverteilung oder Stichprobe größer 30
- mindestens intervallskaliert
- abhängige Ziehungen

2. Warum ist der T-test für abhängige Stichproben im Allgemeinen teststärker als der T-test für unabhängige Stichproben? In welchem Fall ist er es nicht?

Hohe Korrelation der Datenreihen bei gleichen Personen gilt als Annahme. Je ähnlicher sich die Messwertpaare sind, umso kleiner wird der Standard Fehler. Dadurch wird der T-wert größer und dieser wird dann eher signifikant.

Es gilt nicht, wenn beide Zeitpunkte nicht stark miteinander korrelieren bzw. wenn die Messwerte beim Test der abhängigen Stichproben negativ korrelieren, dann verliert der Test an Teststärke.

Warum wird die Prüfgröße größer?

Das kann man mit der Berechnung des T-Werts erklären:

$T = (\text{MW von } X - \mu) / SE$: wenn der Standardfehler kleiner wird, teilt man durch kleinere Zahl und erhält eine größere Prüfgröße

3. Wenn die Voraussetzungen des T-Werts grob verletzt sind, welche Möglichkeiten bestehen dann noch, die Unterschiedshypothese zu prüfen? Beschreiben Sie die dann in Frage kommenden Verfahren und möglichen Voraussetzungen von diesen.

Wenn Varianzhomogenität verletzt wird:

Welchtest: testet das gleich wie T-test, hat nur nicht Voraussetzung der Varianzhomogenität

Wenn intervallskaliert verletzt

Für unabhängige Stichproben:

U-Test: testet, ob sich zwei unabhängige Stichproben in ihrer Ausprägung auf einem ordinalskalierten Merkmal unterscheiden

Voraussetzung: Stichproben müssen unabhängig sein, Stichprobe größer als 20, Daten um Median symmetrisch.

Meist ist Symmetrie um Median nicht erfüllt → dann: Welch auf den Rängen.

Keine Alternative wenn Stichprobe zu klein, hoffen auf Normal -Verteilung

Für abhängige Stichproben:

Wilcoxon Test: testet, ob sich 2 abhängige Stichproben in ihrer Ausprägung auf einem ordinal-skalierten Merkmal unterscheiden.

Voraussetzungen: abhängige Stichprobe, die dem Merkmal tatsächlich zugrunde liegende Verteilungsfunktion sollte stetig sein (Idee: für jede Beobachtungseinheit können Differenzen zwischen den beiden Stichproben berechnet werden. Der absolute Betrag dieser Differenzen ist nicht interpretierbar, doch sie sind ordinalskaliert: größere Differenzen bedeuten also größere Veränderungen zwischen den Stichproben.)

4. Warum heißt die Varianzanalyse (ANOVA) wie sie heißt, obwohl sie doch die Unterschiedlichkeit von Mittelwerten zwischen mehreren Bedingungen prüft?

Anova: Analysis of Variance

Ziel: Analyse des Einflusses der UVN auf eine abhängige Variable

(Bsp. Wie gut wirken verschiedene Therapieformen (UV) auf spezifische Phobien (UV) bei verschiedenen Geschlechtern? (UV) → Therapieerfolg (AV))

Es werden zwar Mittelwertsvergleiche untersucht, jedoch geht die Anova davon aus, dass beim Zustandekommen eines jeden Messwerts verschiedene Komponenten beteiligt sind:

-Populationswert: den Wert, den sich alle Gruppen teilen

-Effekt des Treatments A

-Effekts des Treatments B

-Interaktionseffekt A mal B

-Zufallsfehler

Die Anova versucht Anteile der Komponenten an der gesamten Streuung/Varianz herauszufinden.

Diese Komponenten werden in ihren Streuungen (um den Stufenmittelwert) auf die Messwerte wieder gegeben und so kann durch die Varianzanalyse untersucht werden, wie genau eine abhängige Variable durch die verschiedenen Komponenten erklärt werden kann.

→ Grundfrage: ist es eine überzufällige Streuung/größer als Zufallsfehler?

H_0 ist immer: Fehlervarianz=Treatmentvarianz

→ Anova leistet also das Auffinden von Unterschieden zwischen Gruppenmittelwerten, indem sie Streuungen und Treatments miteinander vergleicht

→ sind diese Unterschiede überzufällig?

Die Anova arbeitet mit Quadratsummen und das ist eine Vorform der Varianz und Quadratsummen berechnen sich aus Mittelwerten. Unterschiede in Mittelwerten drücken sich auch in Quadratsummen aus. Varianzen geben dann auch Aufschluss über mögliche Mittelwerte.

5. Was sind die zentralen Annahmen des Prüfmodells der Varianzanalyse? Wie wird in der Regel getestet bei der Anova:

Zentrale Annahmen:

-Messwert einer Person besteht aus verschiedenen Komponenten

-Ziel: Analyse des Einflusses der UVS auf eine AV.

-Anova versucht generell Unterschiede zwischen Gruppen auf der Höhe der abhängigen Variable aufzuzeigen.

-Voraussetzung: AV muss stetig (intervallskaliert) sein, UVs sind in der Regel nominalskaliert
-Nach Anzahl der Treatments unterscheidet man einfaktorielle, zweifaktorielle oder mehrfaktorielle Anovas.

Zweifaktoriell=zwei Prädiktoren

A) Für unabhängige Stichproben:

0. Mittelwerte berechnen
 1. Quadratsummenzerlegung (Summe aller quadrierten Abweichungen aller Einzelwerte vom Gesamtmittelwert)
 2. Berechnung der Varianz für jeweils Treatment und Fehler
 3. F-Test: Berechnung der Prüfgröße F und Vergleich mit F-Verteilung:
 $F = \frac{QS \text{ Treatment}}{QS \text{ Fehler}}$
 4. Signifikanz
 5. Kontrastanalyse um zu wissen, wo genau die Unterschiede liegen
- Mehrere unabhängige Variablen (z.B. Brille, Tattoo, Piercing) und schaut wie die sich hinsichtlich wahrgenommene Attraktivität unterscheiden → mehrere unabhängige Variablen

B) für abhängige Stichproben:

Gleiche Abfolge wie bei abhängiger Stichprobe, aber anderer Vergleich: statt Unterschiede zwischen Gruppen nun z.B. Interessenunterschiede innerhalb derselben Person

6. Welche Bedeutung hat der F-Test? Was prüft er eigentlich?

Test, der prüft, ob sich Quadratsummen/Varianzen signifikant unterscheiden, kann man auch für ANOVA nutzen, nicht nur als Voraussetzungstest für T-Test

→ Testet Varianzhomogenität als Voraussetzung für Anova

F-Prüfgröße: damit lässt sich Wahrscheinlichkeit ableiten mit der eine solche Varianz oder eine größere auftritt.

Voraussetzungen für F-Test sind Intervallskalierung und Normalverteilung

Beispiel: Gibt es signifikante Abweichung der Körpergröße der Streuungen der Körpergröße in Mainz und FFM? Wenn signifikant, ist es sinnvoll Größe für beide Städte getrennt zu halten.

7. Wie ist die Quadratsummenzerlegung in der ANOVA?

A) für unabhängige Stichproben:

QS total setzt sich zusammen aus:

QS Fehlervarianz (Messschwankungen etc.) + QS Treatments (Treatment A und Treatment B etc. + Interaktion von beiden)

B) für abhängige Stichproben:

QS total= Between/Zwischen QS (Summe zwischen den Personen, z.B. wie sehr sich Leute generell in Fahrleistung unterscheiden, aber interessiert uns gerade nicht) +

QS Within/Innerhalb (setzt sich zusammen aus: QSTreatment + QS Fehlervarianz)

8. Erläutern Sie die Prüfgröße Eta Quadrat und Partial Eta Quadrat.

Eta Quadrat: =Effektstärke

sagt aus, wie groß die Varianzaufklärung (ohne Fehlervarianz) bzw. die Effektstärke ist.

Allerdings hat Eta Quadrat den Nachteil eines positiven Bias, das heißt die aufgeklärte Varianz wird immer überschätzt.

Formel: QS Zwischen/QS Gesamt bzw. QS Effekt/QS Gesamt

Partielles ETA Quadrat: ist identisch mit Eta Quadrat im Fall der einfaktoriellen Anova ohne Messwiederholung, bei mehreren Faktoren oder bei Messwiederholungen wird jedoch das partielle Eta Quadrat verwendet. Während beim Eta Quadrat im Nenner die gesamte Varianz steht, steht beim partiellen Eta Quadrat die ungeklärte Varianz von Y plus die aufgeklärte Varianz einer Variablen X. Dies erlaubt es die Effekte über verschiedene Studien hinweg zu vergleichen.

Formel: QS Treatment/ (QS Effekt von X + QS Fehler von Y)

9. Welche Voraussetzungen gibt es für die ANOVA?

a) für unabhängige Stichproben:

-abhängige Variable sollte normalverteilt sein und intervallskaliert

-unabhängige Stichprobe: d.h. von verschiedenen Merkmalsträgern

-keine Ausreißer

-Homoskedastizität: Gleichheit der Fehler- Varianzen der verschiedenen Gruppen (→ prüft man über Levene Test, Bartlett Test oder F-Test)

-Treatment und Fehler nicht korreliert

-Normalverteilte Fehler → Prüfung über Kolmogoroff-Smirnoff Test

b) für abhängige Stichproben:

-Abhängigkeit der Ziehungen

- normalverteilte Fehler

-Sphärizität:

Def. Sphärizität: ist eine der wichtigsten Voraussetzungen der ANOVA mit Messwiederholung. Sphärizität ist eine Annahme, die bei allen Verfahren mit Messwiederholung gemacht wird, die mehr als zwei Stufen haben. Bei vorliegender Sphärizität sind die Differenzen aller Faktor/Treatmentstufen der UV gleich. Man sagt auch, das Homoskedastizität zwischen den Stufen vorliegt.

Sphärizität wird über den Mauchly Test geprüft.

Welche Voraussetzungen sind kritisch für die Gültigkeit des F-Tests?

Verletzungen der Sphärizitätsannahme führen zu progressiven Entscheidungen bei der F-Prüfgröße, das heißt es stellen sich unerwünscht schneller Signifikanzen ein.

Ist die Sphärizitätsannahme verletzt, so ist eine Korrektur der Freiheitsgrade vorzunehmen (z.B. Greenhouse Geisser). Es empfiehlt sich allerdings, diese Korrektur immer durchzuführen.

10. Was versteht man unter einer statistischen Interaktion? Geben Sie ein Beispiel, erläutern Sie mithilfe einer Zeichnung.

Immer wenn die Option besteht, dass mindestens zwei Prädiktoren (UVs) ein Kriterium vorhersagen, kann eine Wechselwirkung auftreten.

Beispiel:

Prädiktor 1: Hitze

Prädiktor 2: frustrierende Situation

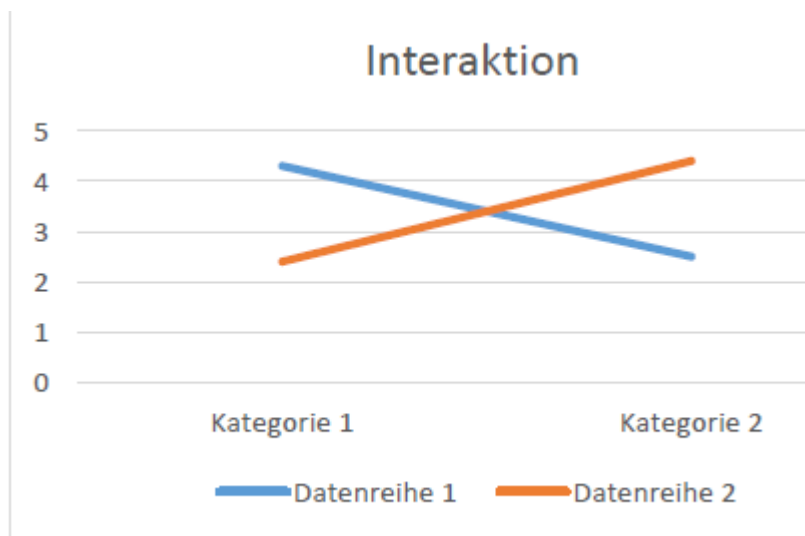
Kriterium: Aggressivität

Wenn Prädiktor 1 Hitze und Prädiktor 2 Situation, alleine nicht viel ausmachen, doch zusammen zu Aggressivität führen, liegt ein Interaktionseffekt vor.

Man sieht: Haupteffekt für Frustration, denn Frustration führt immer zu mehr Aggression, Aber man sieht auch Interaktionseffekt denn Hitze und Frustration zusammen führen zu besonders viel Aggression.

Achtung: Zeichnung passt nicht zu obiger Interaktion.

Disordinale Interaktion: keine Haupteffekte von A und B, aber Interaktionseffekt



- 11. Welche Probleme tauchen auf, wenn man mit paarweisen Mittelwertvergleichen innerhalb eines signifikant gewordenen Faktors (Intelligenz) oder eine signifikant gewordenen Interaktion prüfen möchte, welche Bedingungen (unterschiedliche Messzeitpunkte) sich von welchen anderen signifikant unterscheiden (Kontrastanalyse)? Welche Lösungsmöglichkeiten gibt es für diese Probleme und welche Eigenschaften haben die verschiedenen Lösungen?**
- Warum sollte man in einer Kontrastanalyse möglichst sparsam testen?**

Probleme: je mehr Einzelvergleiche durchgeführt werden, also Hypothesen getestet werden, desto größer wird bei festgelegtem Alpha Niveau die Wahrscheinlichkeit einer zufälligen Signifikanz.

→ Alpha Fehler Kumulierung.

Lösung:

Apriori: Bonferonni: Alpha-Korrektur: Alpha-Niveau/ Anzahl der Tests → Signifikanz wird viel kleiner

Apriori: Dunn-Sidak Korrektur: ist nicht so hart bzw. konservativ wie Bonferroni, komplexere Formel, man verkleinert Alpha nicht ganz so hart.

Apriori: Scheffe Test: sichert die Gesamtheit aller mit den Einzelvergleichen verbundenen Entscheidungen für eine bestimmte Wahrscheinlichkeit ab – sehr konservativ
Man sollte in der Kontrastanalyse möglichst sparsam testen, damit die Alpha Fehler Kumulierung nicht zu groß wird.

12. Varianzanalysen ermöglichen auch die Verwendung von Kovariaten. Welches Problem löst man hierdurch? Welche Annahmen werden hierbei gemacht, und müssen sie erfüllt sein? Gibt es andere Möglichkeiten zur Kontrolle des Effekts von Drittvariablen?

Kovariaten: ein in einem konkreten Experiment nicht als relevant angesehener Faktor (z.B. Geschlecht), der herausgerechnet werden soll, um so einen möglichen Effekt einer interessierenden UV auf die AV statistisch nachweisen zu können.

Falls Kovariaten nicht herausgerechnet werden können, können sie unerwünscht die Teststärke erhöhen.

Mithilfe der Kovarianzanalyse wird der Einfluss einer Kovariaten auf die abhängige Variable „neutralisiert“. Nun kann eine Anova über den Einfluss der interessierenden UVS auf die AV gerechnet werden.

Annahmen:

- Kovariate sollte mit der AV korrelieren, nicht jedoch mit der UV.
- Korrelation zwischen UV und Kovariate könnte den eigentlichen Effekt reduzieren.

Voraussetzungen u.a.

- Normalverteilung der Residuen/Fehler
- Varianzhomogenität der Residuen/Fehler
- AV: kontinuierlich, intervallskaliert

Andere Möglichkeiten zur Kontrolle des Effekts von Drittvariablen:

1. Blockbildungsverfahren: VPN zb nach Berufserfahrung in Blöcke teilen (homogene Teilgruppen) und Berufserfahrung als Kontrollvariable miteinfassen.
2. Untersuchungsbedingte Störvariablen: versuchen alle Untersuchungsbedingungen und somit auch alle Störvariablen konstant zu halten (Art und Intensität der

Störvariablen möglichst genau zu protokollieren, ggf. nachträglich die Ergebnisse bzgl. des Einflusses der Störvariablen korrigieren.)

Versuchsplanung und Güteaspekte im Kontext von Evaluationsuntersuchungen

1. Was versteht man unter Teststärke, was unter Effektstärke? Wie hängen beide Größen mit der Stichprobengröße zusammen?

Teststärke/Poweranalyse: mit welcher Wahrscheinlichkeit ein Test signifikant wird, wenn es tatsächlichen Effekt gibt

--> Stichprobengröße sehr wichtig, um zu wissen ob ich noch VP brauche.

Effektstärke: gibt die tatsächliche praktische Bedeutsamkeit eines signifikanten Effekts an
Wie groß der Effekt ist, wenn Test signifikant ist → Stichproben Größe ist egal

Zusammenhang zwischen beiden Größen:

Effektstärke ist nicht von Anzahl der VP abhängig, darum ist es immer gut sie zusätzlich zu betrachten. Teststärke allein kann Bewertung einer Studie verzerren. (Man sieht einen hohe Signifikanz, weil ich 100 000 Vp habe, aber der Effekt (zb Cohens D) ist dann sehr klein.)

2. Wie bestimmt man Effektstärken/Effektgrößen und welche Werte gelten als klein, mittel, groß?

Sie werden in der Regel mittels der Varianzen bestimmt und drücken aus wieviel Unterschiedlichkeit des Kriteriums aufgeklärt werden konnte.

Bei unabhängigen Daten: $D = (\text{Mittelwert} - \text{Mü}) / \text{geschätzte SD}$

Bei abhängigen Daten: $D = \text{MW von Messzeitpunkt 1} - \text{MW von Messzeitpunkt 2} / \text{geschätzte SD}$

$D \leq 20$: kleiner Effekt

$D \leq 50$: mittlerer Effekt

$D \leq 80$: großer Effekt → Bewertung nach Cohen

3. Wie sollte man die Stichprobengröße seiner Untersuchung planen? Welche Parameter sollte man berücksichtigen?

Die ideale Stichprobe richtet sich

- nach dem Anteil der für die Untersuchung relevanten Personen in der Population
- dem Signifikanzniveau → bei $\alpha = 0,01$ braucht man eine größere Stichprobe

Die Stichprobengröße kann man durch Erfahrung oder bestimmte Normen festlegen.

Wenn man einen wenig sensitiven Test hat, braucht man eine größere Stichprobe
Eine größere Stichprobe braucht man auch, wenn man ohne Messwiederholung /unabhängig testet, da nicht jeder doppelt getestet wird → größere Korrelationen bei abhängigen Tests

4. Wie sieht ein klassisches Evaluationsdesign zum Nachweis der Effektivität einer Intervention aus? Aus welchem Ergebnisschema kann man eindeutig auf die Wirksamkeit der Intervention schließen? (Ausschluss von alternativen Erklärungen)? Welche Ergebnisschemata sind hinsichtlich dieses Schlusses mehrdeutig, und lassen alternative Erklärungen zu?

Klassisches Evaluationsdesign: Experimental und Kontrollgruppe: randomisiertes Kontrollgruppendesign (dadurch Kontrolle von Störvariablen)

→ Kausalität ableitbar

Ergebnisschema:

Prätest: vorher Unterschiede? → sollte es nicht geben

Posttest: hinterher Unterschiede? (am besten Messwiederholung über die Zeit) (repeated Measurement)

Mehrdeutige sind Ergebnisschema bei:

z.B. Ex post facto Studien (wenn weder Experiment noch Quasiexperiment möglich ist)

Dabei werden AV und UV nachträglich bestimmt

Dann besteht keine Möglichkeit zur Kausalinterpretation, nur zur Zusammenhangsinterpretation (→ denn fehlende Kontrolle der Störvariablen)

Bsp.: Folgen von Rauchen erforschen und Daten einer Lungenklinik ansehen: schauen, was Raucher und Nichtraucher trennt (UV) und gucken, wer Lungenkrebs hatte (AV).

→ Doch diese Gruppen unterscheiden sich wahrscheinlich nicht nur in Nikotingebrauch sondern auch z.B. im Lebensraum (Störvariable).

5. In welcher Weise ist das Experiment anderen Beobachtungsformen bezüglich der ermöglichten Schlussweisen überlegen? Warum können in Evaluationsuntersuchungen oftmals nicht echte Experimente durchgeführt werden? Was zeichnet Quasi-Experiment und Ex Post Facto Untersuchungen aus?

Überlegenheit des Experiments

-Kontrolle von UVs und Störfaktoren durch Randomisierung der Gruppen und Prä-Post Design möglich

-Systematische Variation von UVs möglich

-Wiederholungsmessungen möglich

-Verringerung der Messfehler

➔ Folgen: höhere interne Validität (niedrigere externe Validität)

➔ Kausalschlüsse ableitbar

Echte Experimente oft nicht möglich, weil

- Durchführung nicht wiederholbar ist
- es unethisch ist, eine Kontrollgruppe zu erstellen (z.B. suizidgefährdete Personen ohne Behandlung)
- eine UV nicht variierbar oder künstlich manipulierbar ist (wie Geschlecht, Nationalität, Krankheit)
- es zu teuer ist, alle Bedingungen zu standardisieren.

Quasi-Experiment:

Natürlich Gruppen ohne randomisierte Zuordnung von Versuchspersonen werden verglichen (z.B. Schulklassen). Persönliche Störvariablen sind daher nicht ausgeschlossen und Kausalinterpretationen sind nicht möglich, da die Rückführung von Unterschieden in der AV nicht eindeutig auf die UV zurück gehen kann.

Ex post facto Experiment: retrospektiv befragen und im Nachhinein UV und AV nachträglich festlegen➔ schlechte Interpretierbarkeit, nur Berechnung korrelativer Zusammenhänge.

6. **Was ist interne Validität, was externe? Wie kann man diese bestimmen bzw. Daten dazu erheben?**

Interne Validität: (Eindeutigkeit)➔ Effektstärke

Untersuchung ist intern valide, wenn ihr Ergebnis eindeutig interpretierbar ist, die Störvariablen also kontrolliert sind. Sie sinkt mit wachsender Anzahl plausibler Alternativerklärungen für das Ergebnis aufgrund nicht kontrollierbarer Störvariablen.

Bestimmbar durch: Laboruntersuchung, Experiment

Externe Validität:

Untersuchung ist extern valide, wenn hohe Generalisierbarkeit vorliegt

➔ wie gut sind Ergebnisse übertragbar auf Population, auf die Lebenswelt?

Sie sinkt mit wachsender Unnatürlichkeit der Umgebung und der Künstlichkeit der Untersuchungsbedingungen bzw. mit abnehmender Repräsentativität der Stichprobe.

Bestimbar durch Felduntersuchung, (systematische wissenschaftliche Beobachtung unter natürlichen Bedingungen, also außerhalb des Labors)

7. Was ist Kriteriums-, was ist Konstruktvalidität? Mit welchen Methoden kann man diese Validitätsarten bestimmen, und warum haben sie in der Evaluationsforschung eine zentrale Wichtigkeit?

Kriteriumsvalidität: Zusammenhang zwischen den Ergebnissen des Instruments und einem empirischen bzw- äußeren Kriterium: Wie gut ich Kriterium wirklich messe z.B. wie gut ein IQ- Test die Intelligenz eines Menschen erfasst

Bestimbar durch Übereinstimmungsvalidität und prognostische Validität

Konstruktvalidität: ein empirischer Beleg dafür, dass der Test das Konstrukt misst, das er zu messen vorgibt und nicht ein anderes. Z.B. ob alle Teilaspekte des Tests tatsächlich das zu messende Gesamtmerkmal repräsentieren z.B. sprachliche Intelligenz wird in einem IQ Test nicht getestet → mangelnde Konstruktvalidität

Bestimbar durch konvergente Validität → Korrelation (sollte hoch sein), diskriminante Validität → Korrelation (sollte gering sein)

Methoden zur Bestimmung für beide: Multitrait(versch. Skalen) Multimethods

Wichtig in der Evaluationsforschung:

Validität ist das wichtigste Gütekriterium. Für Untersuchungen ist es am wichtigsten, dass der Test auch genau in der Lage ist, das zu messen, was er zu messen vorgibt. Dies kann man nur über Kriteriums- und Konstruktvalidität herausbekommen. Darauf aufbauend können dann Hypothesen geprüft und Entscheidungen getroffen werden.

8. Erläutern Sie Methoden der Kriteriumsvalidität, die auf der Bestimmung einer biserialen oder punktbiserialen Korrelation beruhen.

Punktbiserialer Korrelation:

Ein dichotomes und ein intervallskaliertes Merkmal z.B. Übereinstimmungsvalidität (Zusammenhang) von Geschlecht (Mann/Frau) und IQ (intervallskaliert)

Biserialer Korrelation:

Ein künstlich dichotomisierendes (Z.B. dumm/intelligent) und ein intervallskaliertes Merkmal (ausgehend von ursprünglicher Normalverteilung beider Merkmale) z.B.

prognostische Validität im Alter im Format alt/jung und Berufserfolg, operationalisiert durch Gehalt.

9. Erläutern Sie den Multitrait-MultiMethod Ansatz zur Bestimmung der Konstruktvalidität.

Multitrait Multimethod bezeichnet eine Gruppe statistischer Verfahren, die zum Nachweis der Konstruktvalidität eines psychologischen Tests dienen. Es wird untersucht, in welchem Grad ein Instrument zwischen den verschiedenen Konstrukten differenziert (diskriminante Validität) und auch wie gut ein Konstrukt durch verschiedene Instrumente erfasst werden kann (konvergente Validität).

- ➔ konvergente Validität liegt vor, wenn Messungen eines Konstrukts, das mit verschiedenen Methoden erfasst wird, hoch miteinander korrelieren.
- ➔ Diskriminante Validität liegt vor, wenn verschiedene Tests verschiedene Konstrukte erfassen oder ein Test verschiedene Konstrukte
- ➔ Konvergente sollte immer höher als diskriminante sein

Kritik an MTMM

Keine exakten Entscheidungsregeln

Reliabilitäten der Messungen bleiben unberücksichtigt

Weitere Prüfungsfragen/Alte Prüfungsfragen (WS 2018)

1. Wie ist der Satz der totalen Wahrscheinlichkeit?

Wahrscheinlichkeit eines Ereignisses ist gleich Summe der Wahrscheinlichkeiten aller Pfade die zu diesem Ereignis führen.

Satz der totalen Wahrscheinlichkeit für zwei Ereignisse A und B

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap \bar{B}) \\ &= P(B) \cdot P_B(A) + P(\bar{B}) \cdot P_{\bar{B}}(A) \end{aligned}$$

BStrich=B2

Bsp:

Eine Gemeinde wird zur Bürgermeisterwahl in zwei Wahlbezirke (B1 und B2) eingeteilt.

60% der Wähler kommen aus B1, 40% aus B2.

In B1 erhält der Kandidat Albrecht 30% der Stimmen, in B2

dagegen 80%.

Wie viel Prozent der Stimmen hat der Kandidat Albrecht insgesamt bekommen?

$$P(A) = P(A \cap B_1) + P(A \cap B_2) \\ = 0,6 \cdot 0,3 + 0,4 \cdot 0,8 = 0,5$$

Antwort:

Genau 50% aller Wähler haben für den Kandidaten Albrecht gestimmt.

3. Wie sind die Lageregeln bei Verteilungen?

Symmetrisch wenn gilt: Modus gleich Median gleich Mittelwert

Linksschief/Rechtssteil, wenn gilt: Modus größer Median größer Mittelwert

Rechtsschief/linkssteil, wenn gilt: Modus \leq Median \leq Mittelwert

Den Median würde man angeben, wenn man eine realistische Darstellung haben möchte. Er sagt genau aus, an welchen Datenpunkt 50% drunter und drüber liegen.

Der Mittelwert ist im Gegensatz zum Median anfällig für Ausreißer, er wird stark durch hohe Werte beeinflusst.

4. Man geht in die Schule und würfelt zwei Klassen zusammen, mit diesen macht man einen IQ Test. In einer anderen Schule testet man zwei Klassen so wie sie sind. Was ist der Unterschied?

Das zweite wäre ein Quasi Experiment (ohne Randomisierung), das erste ein Experiment (mit Randomisierung). In der zweiten Gruppe sind es natürliche Gruppen, da können Störvariablen mit einwirken. Eine Kausalinterpretation ist nicht möglich.

Allerdings haben wir eine Randomisierung in der ersten Bedingung auch nur für Schüler, Das Experiment wäre also auch nur für Schüler gültig.

5. Wenn man aus Versehen Konzentration mittesten würde, muss man die Daten dann entsorgen oder ist noch etwas zu retten?

Man kann die Konzentration (durch eine Kovarianzanalyse) rausrechnen oder versuchen für alle konstant zu halten.

6. Führen Sie eine Permutation mit dem Wort „Statistik“ durch.

=nicht wohl verschiedene Permutation: es gibt gleiche Elemente (hier z.B. T und S)

STATISTIK

123456789

S=2mal T=3mal I=2mal

Formel: $n!$ (gesamt) / $(n_1! \cdot n_2! \cdot \dots)$ n_1 etc. = einzelne Untergruppen wie hier S oder T

$9! / (3! \cdot 2! \cdot 2! \cdot 1! \cdot 1!)$

$9! = 362.880$

$3! = 6$ $2! = 2$ $\rightarrow 362880 / (6 \cdot 2 \cdot 2) = 15120$

7. Nennen Sie ein Beispiel für Permutation und Kombination aus dem Alltag.

Permutation: Auf wieviele Weisen kann ich sechs verschiedene Schokoladen auf sechs Kinder verteilen?

- Mit Reihenfolge und ohne Wiederholung
- $N! / (n! - k!) = 6! / (6! - 6!) = 6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 = 720$
- Achtung: Wenn $n! - k! = 0$ ist das Ergebnis immer 1!

Kombination: Auf wieviele Weisen können 5 Spieler einer Fussballmannschaft fürs Elfmeterschießen ausgewählt werden?

Ohne Reihenfolge (Mannschaft wird als Ganzes betrachtet)

Ohne Wiederholung (Jeder schießt nur ein Mal)

- $N = 11$ $k = 5$ $n! / k!(n-k)! = 11! / 5!(11-5)! = 462$

8. Wenn ich beim T-Test 30 Versuchspersonen habe und er nicht signifikant wird, warum bilde ich dann nicht einfach ganz viele Gruppen?

Es kommt bei multiplen Testen zur Zufallssignifikanz (Alpha Fehler Kumulierung) → man muss Alpha anpassen z.B. durch Bonferonni Korrektur

Wenn ich noch 120 VP dazu nehme, hätte ich ein Problem, wenn es dann signifikant wäre?

Es kann gut signifikant werden, aber es heißt nicht, dass es einen Unterschied gibt. Je mehr Leute man testet, je wahrscheinlicher wird es signifikant und die von der Stichprobengröße unabhängige Effektstärke wäre nach wie vor klein.

10. Was ist das Problem der Bonferroni Korrektur?

Diese ist recht konservativ und irgendwann macht sie auch keinen Sinn mehr, weil das sehr konservative Signifikanzniveau einfach nicht erreicht werden kann.

9. Wo liegen Modus, Median und Mittelwert in der Normalverteilung?

Modus, Median und Mittelwert sind in der symmetrischen Normalverteilung identisch und stellen den höchsten Punkt der Normalverteilung dar.

10. Wenn man 1 2 3 3 3 4 555 hat, was ist der Modus, was der Median?

Der Modus wäre 3 und 5, weil beide gleich häufig vorkommen – bimodale Verteilung.
Median wäre an fünfter Stelle und wäre 3.

11. Wird das Konfidenzintervall größer oder kleiner, wenn sich Alpha ändert?

$$CI = x \pm z(\alpha/2) * SE$$

Beim 95% Konfidenzintervall wäre Alpha 1,96, bei 99% wäre es 2,57.

Wenn Alpha kleiner wird, also z.B. vom 5% aufs 1% Nivea geht, wird unser Konfidenzintervall größer. → Denn wenn wir eine sichere Aussage machen wollen, schließen wird mehr Daten ein.

12. Wann ist der abhängige und wann der unabhängige t-Test teststärker und wie berechnet man jeweils die Freiheitsgrade?

In der Regel ist der abhängige T-test teststärker, da aus der stärkeren Korrelation der Messwertpaare ein kleinerer SE hervorgeht und wenn durch etwas Kleineres geteilt wird, eine größere Zahl rauskommt.

(Formel: $SE = SD / \sqrt{N}$)

Wenn die Messwertpaare allerdings negativ miteinander korrelieren, ist der unabhängige T-Test teststärker.

Für abhängigen Test: $df = n - 1$

Für unabhängigen Test: $df = n_1 - 1 + n_2 - 1 = n - 2$

13. Was passiert mit der externen Validität, wenn ich die interne möglichst hoch halten möchte?

Eine möglichst hohe interne Validität wird durch maximale Kontrolle aller Störvariablen erreicht. Ein hohes Ausmaß an Kontrolle kann jedoch dazu führen, dass die experimentelle Situation so künstlich wird, dass sie nicht mehr realitätsnah ist und die externe Validität sehr niedrig wird (Wenn man externe erhöhen will, sinkt die interne dagegen.).

14. Was ist progressiver: die Bonferroni oder die Dunn-Skidat Korrektur?

Die Dunn-Skidat Korrektur, da die Bonferroni Korrektur sehr konservativ ist. (Die Dunn-Skidat Korrektur geht von unabhängigen Stichproben aus und berechnet das neue Alpha Niveau über die komplementäre Wahrscheinlichkeit.

15. Wie funktioniert die QS Zerlegung bei der abhängigen Anova und was interessiert uns davon?

Abhängig=Rm Anova: repeated measurement

QS Total = QS Zwischen + QS Innerhalb (Die QS Zwischen ist uns egal, uns interessiert vor allem die QS innerhalb, da diese Aussagen über das Treatment macht.)

QS Innerhalb = QS Treatment (a, b und Interaktion → meint Wirkung der Faktorstufen) + QS Fehler (Residual → bezeichnet Streuung innerhalb der Person)

16. Wie berechnet man Median, Modus und Mittelwert?

Modus: häufigster Wert → einfach ablesen

Mittelwert: Summe aller Zahlen durch Anzahl Zahlen

Median: Mitte der Datenreihe, wenn sie nach der Größe geordnet werden. Bei einer ungeraden Datenanzahl ist der Median der Wert in der Mitte (rechts und links von ihm sind gleich viele Zahlen.)

Bei einer geraden Datenanzahl ist es das arithmetische Mittel der beiden mittleren Werte z.B. mittlere Werte = 8 und 9 → $(8+9)/2 = 8,5$

17. Zeichnen und beurteilen Sie einen Wahrscheinlichkeitsbaum und zwar ob 3 Würfel, bei denen zweimal die 6 kommt als typisches LaPlace Experiment gilt.

Zeichnung

Achtung hier gibt es nur zwei Äste jeweils: einer ist 5/6, einer ist 1/6. Zeichnet man sechs bekommt man Punktabzug!

Obwohl wir es mit einem Würfelwurf zu tun haben, handelt es sich **nicht** um ein LaPlace Experiment, da die Wahrscheinlichkeit für Erfolg (6 würfeln) 1/6 beträgt und die Wahrscheinlichkeit für keinen Erfolg (1-5 würfeln) 5/6. Damit haben wir keine gleiche Wahrscheinlichkeit und LaPlace fällt raus. Stattdessen: Kolmogoroff (bzw. Bernoulli dem aber eine Kolmogoroff Wahrscheinlichkeit zugrunde liegt.)

18. Berechnen Sie Cohens D und interpretieren Sie es.

Cohens D ist die Effektgröße für Mittelwertsunterschiede zwischen zwei Gruppen.

$D = \frac{MW - M\ddot{U}}{SD}$ → bei gleichen Gruppenvarianzen

$D = \frac{\text{Mittelwert } x_1 - \text{Mittelwert } x_2}{\sqrt{\text{geschätzte Stabw}_1 \text{ zum Quadrat} + \text{geschätzte Stabw}_2 \text{ zum Quadrat}}}$ durch 2

→ bei unterschiedlichen Gruppenvarianzen

$D > 0,2$ kleiner Effekt

$D > 0.5$ mittlerer Effekt

$D > 0.8$ großer Effekt

Was ist besser zur Bewertung: Cohens D oder der T-Wert?

Cohens D ist aussagefähiger als T-Wert, weil durch SD geteilt wird und nicht wie beim T-Wert durch SE. SE ist verzerrt von der Stichprobengröße

19. Wie berechnet man aus der Varianz die Standardabweichung?

Die Standardabweichung ist die Wurzel aus der Varianz zB SD, wenn Varianz 112: $SD=11$

20. Was besagt der zentrale Grenzwertsatz?

Die Summe einer großen Anzahl von unabhängigen Zufallsvariablen folgt asymptotisch einer stabilen Verteilung. Bei endlicher und positiver Varianz der Zufallsvariablen ist die Summe annähernd normalverteilt, was die Sonderstellung der Normalverteilung erklärt. (Je größer die Stichprobe wird, desto näher wird die Stichprobenverteilung normalverteilt sein.) Dank des zentralen Grenzwertsatzes können wir Hypothesentests durchführen, auch wenn die Grundgesamtheit nicht normalverteilt ist, vorausgesetzt die Stichprobe ist hinreichend groß. Die meisten Statistikbücher geben eine Empfehlung von $N=30$ an, ab der wir von einer Normalverteilung ausgehen dürfen.

21. Was ist eine Ringpermutation?

Allgemeine Formel: $(n-1)!$

Das sind die Möglichkeiten verschiedener Anordnungen z.B. von Personen, die um einen runden Tisch sitzen.

Erklärung: Person 1 wählt beliebig einen Platz, die anderen haben dann die freie Wahl. Damit ist klar, dass es $n-1$ Ringpermutationen gibt, denn die nächste Person hat immer einen Platz weniger, den sie wählen kann zB für 5 Personen um einen Tisch würden wir $4!$ Berechnen.

Was ändert sich wenn zusätzlich zwei bestimmte Personen nicht nebeneinander sitzen sollen?

Person 1 sitzt immer auf Platz 1. Dann darf Person 2 nicht auf Platz 2 und nicht auf Platz N , also bleiben $N-3$ Plätze für Person 2. Die restlichen Personen können sich beliebig auf die $N-2$ Plätze verteilen.

22. Was bedeutet ein Z-Wert von 0,5?

Das bedeutet dass unser Datenpunkt 0,5 Standardabweichung vom Mittelwert entfernt ist (größer als der MW). Wäre der Z-Wert negativ, wäre er kleiner als der

MW. Man kann auch daraus ableiten, wieviel Prozent bei der Standardnormalverteilung über und unter dem Datenpunkt liegen.

23. Was wird mit dem MauchlyTest getestet?

Dies ist der bekannteste Test, um Daten auf Sphärizität zu prüfen. Wenn der P-Wert des Mauchly Test größer oder gleich des festgelegten Alpha Niveau ist, können wir davon ausgehen, dass die Sphärizität der Daten gegeben ist.

24. Geben Sie Mittelwert und Standardabweichung der Zahlenreihe 5,5,5,5,5 an.

Mittelwert: 5 Stabw.:0

Berechnung Standardabweichung:

1. MW berechnen, dann Varianz, dann daraus Wurzel

-Varianz Formel: $(\text{Wert1}-\text{MW})^2+(\text{Wert2}-\text{MW})^2$ etc./Anzahl aller Werte n

25. Berechnen Sie die Bonferonni Korrektur und erklären Sie, was sie ist.

Die Bonferonni Korrektur ist eine von mehreren Methoden, die verwendet werden, um der Alphafehler Kumulierung entgegenzuwirken.

Sie wird immer bei mehreren **Mittelwertvergleichen** verwendet. Sie gehört zu den einfachsten und konservativsten Verfahren. Sie begrenzt die Wahrscheinlichkeit einen Fehler 1. Art zu begehen.

Berechnung= Signifikanzniveau von Alpha/Anzahl der Tests

26. Ein König hat vier Töchter. Wie hoch ist die Wahrscheinlichkeit für einen Sohn?

Die Wahrscheinlichkeit beträgt immer 50%

27. Warum werden die Freiheitsgrade vom T-test mit n-1 berechnet?

Die t Verteilung hat einen Parameter, der ihre Varianz und damit auch ihre Form verändert: Die Freiheitsgrade. Diese kann man sich als Anzahl von Möglichkeiten vorstellen, die man in einem System bei einem feststehenden Wert, unabhängig voneinander variieren kann. Wenn wir beispielsweise drei Zahlen a, b und c haben und ihre Summe 100 kennen, dann wissen wir, dass wenn wir zwei Zahlen kennen (zb a=20 und b=30, c=50sein muss). Diese Gleichung hätte dann 2 Freiheitsgrade, also n-1 (Da der Letzte Wert immer festgelegt ist.)

28. Wovon sind Freiheitsgrade abhängig?

Von der Stichprobengröße

Die Verteilungsform ist abhängig von den Freiheitsgraden:

Je weniger Freiheitsgrade desto flacher ist die Verteilungsform in der Mitte und desto höher am Rand und desto später signifikant.

29. Was ist eine Greenhouse Geisser Korrektur und bei welchem Versuchsdesign kann man davon ausgehen, dass diese verletzt ist?

Wenn die Annahme der Sphärizität verletzt ist (P-wert beim Mauchly Test kleiner als das Alpha Niveau), so müssen die Freiheitsgrade nach unten korrigiert werden (d.h. die Freiheitsgrade werden mit einem Korrekturfaktor multipliziert.). Dann kann trotzdem eine rm Anova durchgeführt werden, auch wenn die Varianzen unähnlich sind.

Bei einem Design mit vielen Treatmentstufen kann man davon ausgehen, dass sich die Varianzen wahrscheinlich stark unterscheiden, da sich die VP über die Zeit/Messwiederholungen stärker verändern, z.B. weil sie mehr beansprucht sind.

30. Wie ist die Formel für die T-Prüfgröße?

$T = (\text{MW von allen } x - \text{Erwartungswert}) / SE$

$SE = SD / \sqrt{N}$

31. Wie werden Kovariaten herausgerechnet?

Das Ziel bei der Kovarianzanalyse besteht in der statistischen Kontrolle einer Störvariablen, die möglicherweise die Daten der Untersuchung beeinflusst haben könnte (z.B. vor dem Vergleich der Gedächtnisleitung zwischen zwei Lernbedingungen den Einfluss des Alters herausrechnen). Man möchte wissen, wie das Ergebnis aussähe, wenn die Kovariate (Störvariable) aus allen Gruppen gleich gewesen wäre.

Dazu wird die Störvariable zusätzlich erhoben und ihr Einfluss mit einer Kovarianzanalyse neutralisiert. Genauer: Die Kovarianzanalyse entfernt die Varianz der Kovariate aus der AV, indem eine Regression der AV auf die Kovariate berechnet wird. Die Regressionresiduen beschreiben den Anteil der AV, der nicht durch die Kovariate erklärt werden kann. Diese Residuen werden als neue AV in eine Varianzanalyse gegeben. Die nach der Regressionsanalyse verbleibende, nicht erklärbare Varianz wird mithilfe der Anova erklärt.

32. Wenn man Störvariablen hat: Ist es wichtig die vorher zu berechnen oder geht das auch im Nachhinein?

Wenn eine Konfundierung zweier Variablen erst im Nachhinein festgestellt wird und die Störvariablen im Experiment nicht erhoben werden, wird das ganze Experiment unbrauchbar, da nicht mehr deutlich von der UV auf die AV geschlossen werden kann.

(Im Vorhin kann man durch Randomisierung Konfundierung verhindern.)

33. Was ist der Unterschied von Verhältnis- und Absolutskala?

Absolutsskala hat eine natürliche Einheit (auch bei Verhältnis!)

Bsp.: Kinder oder Menge eines Artikels, Einwohner eines Landes

Verhältnisskala hat keine natürliche Einheit.

34. Was gibt man an, wenn eine Firma seine Gehälter offen legen will?

Arbeitgeber will angeben: Mittelwert: verzerrt nach oben: er steht besser da

Arbeitnehmer: Median: stabiler Wert

Modus: gibt reales Gehalt wieder

Rechenbeispiele

1. 6 Personen sollen sich auf 6 Stühle setzen. Wieviele Optionen gibt es?

Reihenfolge wichtig, ohne Wdh.

$$6!/(6-6)!$$

2. 6 Leute sollen sich auf 8 Stühle setzen

Reihenfolge wichtig, ohne Wdh

Aus Sicht der Stühle betrachten

$$N!/(n-k)! = 8!/(8-6)! = 20160$$

3. Acht Personen gehen in ein Haus mit sechs Räumen. Wieviele Optionen sich zu verteilen gibt es?

Reihenfolge wichtig, mit Wdh. → es können auch mehrere Leute in einem Raum sein.

$$N \text{ hoch } k = 6 \text{ hoch } 8 = 1679616 \text{ Optionen}$$

Bei 9 Touristen auf 12 Stockwerken, wären es $12 \text{ hoch } 9$

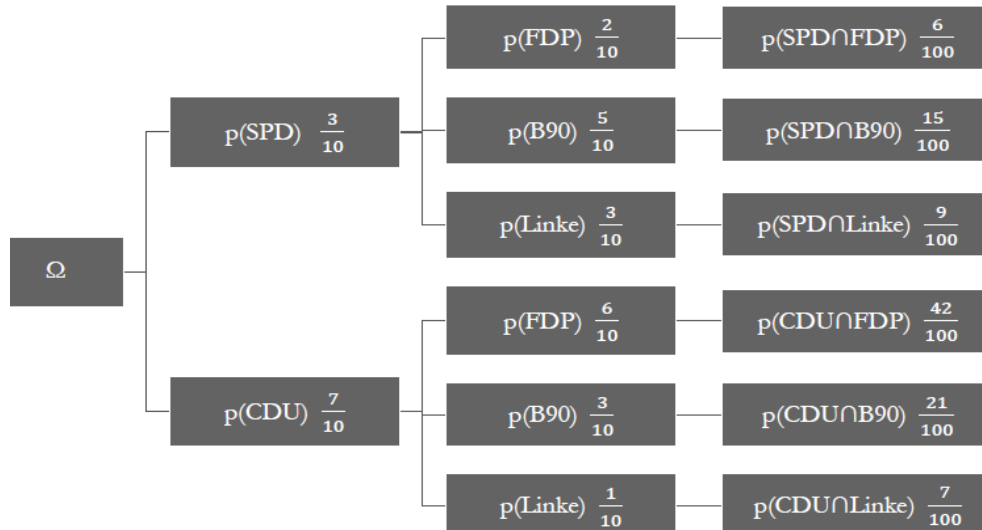
Bei einer Münze, die 15 mal geworfen wird und nur Kopf und Zahl gilt, wäre es $2 \text{ hoch } 15$

Beispiele, wo die Reihenfolge egal ist

Wieviele Optionen gibt es acht rote und fünf blaue in einer Reihe hintereinander aufzustellen?

→ Nicht wohl verschiedene Permutation

→ $13!/(8!*5!) = 1287$



Wie hoch ist die Wahrscheinlichkeit ist eine Koalition aus SPD und FDP: $\frac{6}{100}$

Wie wahrscheinlich ist, dass die SPD regiert unter der Bedingung, dass die Grünen in der Koalition sind (Antwort s.u.)

- Wir suchen zunächst die Verbundwahrscheinlichkeit aus $p(\text{SPD} \cap \text{B90})$

$$p(\text{SPD} \cap \text{B90}) = \frac{15}{100}$$

- Als nächstes benötigen wir die Grundwahrscheinlichkeit für $p(\text{B90})$

Satz der totalen Wahrscheinlichkeit!

Der Satz der totalen Wahrscheinlichkeit:

- Durch Addition der Verbundwahrscheinlichkeiten mit dem gleichen Ereignis, erhält man die Grundwahrscheinlichkeit für genau dieses

$$p(\text{SPD} \cap \text{B90}) \frac{15}{100} + p(\text{CDU} \cap \text{B90}) \frac{21}{100} = p(\text{B90}) \frac{36}{100}$$

Bedingte Wahrscheinlichkeit rechnen:

- Nun können wir die Formel der bedingten Wahrscheinlichkeit anwenden

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

$$p(\text{SPD}|\text{B90}) = \frac{\frac{15}{100}}{\frac{36}{100}} = \frac{15}{100} * \frac{100}{36} = \frac{1500}{3600} = \frac{15}{36} = \frac{5}{12}$$

Prüfung stochastischer Unabhängigkeit

- Wir wollen wissen, ob die Ereignisse „SPD“ und „Grüne“ stochastisch unabhängig sind
- Hierzu muss folgendes gelten:

$$p(A|B) = p(A)$$

Wir haben:

$$p(\text{SPD}|\text{B90}) \neq p(\text{SPD})$$

$$\frac{5}{12} \neq \frac{3}{10}$$

→ Die beiden Ereignisse sind nicht stochastisch unabhängig

Beispiel: Satz von Bayes

- Man kennt nicht den Wahrscheinlichkeitsbaum, sondern lediglich folgende Wahrscheinlichkeiten:

$$p(\text{Linke}) = \frac{16}{100}$$

$$p(\text{CDU}) = \frac{70}{100}$$

$$p(\text{Linke}|\text{CDU}) = \frac{10}{100}$$

- Nun möchte man die Wahrscheinlichkeit $p(\text{CDU} | \text{Linke})$ bestimmen, gemäß folgender Formel:

$$p(A|B) = \frac{p(B|A) * p(A)}{p(B)}$$

$$p(\text{CDU}|\text{Linke}) = \frac{\frac{10}{100} * \frac{70}{100}}{\frac{16}{100}} = \frac{7}{100} * \frac{100}{16} = \frac{7}{16}$$

Wichtige Tipps für die Prüfung

Variation: Auswahl in einer bestimmten Reihenfolge

Permutation: bestimmte Reihenfolge

Kombination: ohne Reihenfolge

Größere Zahl ist immer N

Reihenfolge: wenn Elemente unterscheidbar sind

QS Residium=QS Fehler!

Mit Zurücklegen: gilt nur, wenn alle Elemente zurück gelegt werden, nicht wenn nicht alle zurückgelegt werden

Nicht wohl verschieden= gleiche Elemente

Normale Permutation= nur verschieden Elemente kommen vor

Wenn nicht alles zurückgelegt wird: Ohne Zurücklegen

„alle möglichen Kombinationen angeben“= mit Zurücklegen