

Inhalt:

Teil A - Datenerhebung:

- Methodologie
- Ablaufmodell Empirische Sozialforschung
- Begriffsdefinitionen in der Sozialforschung
- Datenerhebung: Messung & Meßniveaus
- Repräsentativität von Stichproben
- Verfahren (Befragung, Beobachtung, Inhaltsanalyse)

Teil B - Statistik:

- Statistik in P&K-Wissenschaft / Sozialwissenschaften
- Datenformen
- Univariates
- Bivariates
- Wahrscheinlichkeitstheorie
- Stichprobentheorie
- Hypothesenprüfung

Teil A - Datenerhebung:

Methodologie:

Def. einer Wissenschaftsdisziplin durch Erkenntnisgegenstand, -interesse, -verfahren (als Abgrenzung) ;

dabei: a. Theorie (-bildung): zwei unterschiedl Theoriebegriffe: allgem. Th.-bildung <-> konkrete Th.-bildung

b. Empirie: Hypot.-belastung durch system. Beobachtung von Realitätsausschnitten

=> Analyse muß system. & objektivierbar sein (dann sind Ergebnisse wissens. akzeptabel)

c. Verhältnis Empirie / Theorie: widersprüchl. Ergebn. => Theorie falsch oder Vorgehen falsch (s.u.)

d. Verhältnis Wissens. (Erkenntnis) / Praxis (Nutzen): genereller Konflikt (=> unsaubere Methoden)

WICHTIG: Beachtung der mögl. Auswirkungen von wissens. Erkenntnissen

methodologische Grundpositionen in Sozialwissenschaften:

1. Demoskopie (demoskopische Beschreibung sozialer Sachverhalte)
2. emp.- anal. Wissenschaftstheorie [und andere] (emp.- anal. Erklärung sozialer Sachverhalte)
3. Hermeneutik / Phänomenologie (Verstehen von sozialen Sachverhalten, geisteswissens. Tradition)
4. dialektischer Materialismus [+ andere] (normativ-anal./dialektische Kritik/Veränd. gesells. Praxis)

zu 2.: empirisch-analytische Wissenschaftstheorie (als Grundposition hier):

Ziel: Reduktion von Wissens. auf konkrete Th.-bildung & Empirie ; als Einheitslogik für alle Wissens.

Inhalt / Anforderungen (s.o.):

- a. Hypoth.-geflecht: logisch widerspruchsfrei, deduzierbar, falsifizierbar
- b. akzeptierte Verfahren: system., intersubjektiv überprüfbar
- c. Th.-bildung steht in Abhängigkeit von Empirie (=> nur Falsifikation, nicht Verifikation)
- d. kein Praxisbezug der Wissens., nur für Erkenntnis

für P&K-Wissenschaft [! Teildisziplin mit unterschiedl. / eigene Methodologien, hier: Emp. K & K-Forschung]:

Erkenntnisgegenstand: Strukturen / Fkt / Prozesse öffentl., durch techn. Medien vermittelter K

Erkenntnisinteresse: Bedeutung / Auswirkungen von MK auf soziale Sachverhalte / soziales Verhalten

Erkenntnisverfahren: sozialwissens. Verfahren (emp.- anal. Wissenschaftstheorie): Th.-bildung & Empirie

Ablaufmodell empirischer Sozialforschung:

Allgemein: Problem -> Einordnung (allg. Theorien) -> konkrete Theoriebildung (Hypothesengeflecht)

-> Hypth.-prüfung (Validierung) -> Messung -> Datenauswertung -> Hyp.-belastung -> Theoriediskus.

Detailliertes Modell:

- (1) Identifikation des Problemkontextes => Fragestellung
 - (2) Dimensionale Analyse (rekonstruktiv (Einordnung) / konstruktiv (Neubildung Zielsystem)
== wissenschaftliche Transformation => Hypothesenbildung (mit Hinblick auf Falsifikation)
 - (3) Operationalisierung:
 - (a) Untersuchungsobjekte definieren (genaue begriff. Def.)
 - (b) Meßkonzept festlegen:
Erhebungsverfahren (s. Verfahren), Merkm.dimens. (Variablen), M-ausprägungen (Kategorien), Meßniveau (Klassifiz., Intervallbestim.), Datenerhebungstechnik (EDV-Einsatz?, Hilfsmittel?)
 - (c) Analysekonzept:
Auswertungsmethode (qualitativ / quantitativ), in Zsh. mit Erhebungsverfahren
 - (4) Zwischenbilanz (Effektivitätsprüfung: Validität, Reliabilität)
 - (5) Realisierung I (Testphase, dann Datenerhebung (Messung))
 - (6) Realisierung II (Datenübertragung, -auswertung, -analyse; => aufbereitete Datensätze)
 - (7) Dateninterpretation: Hypothesenbelastung, Relativierung Daten <-> Hypothesen
- => wissens. Abhandl.: Problem.-Zsh., Th.-bildung., Methodendisk., Datenerheb., Hypth.-belast., Verwend.-Zsh.;

Begriffsdefinition in den Sozialwissenschaften:

Begriff: in Wort-Zsh. gestellter Vorstellungsinhalt (Definiens: schon da ; Definiendum: noch nicht definiert)

wissens. Anforderungen: präzise, eindeutig, klarer emp. Bezug (da Basis jeder wissens. Arbeit)

operationeller Begriff: Begriffbestimm. eines benötigten Parameters in Operationalisierungskonzept

Grundbegriffe in der Sozialforschung:

- 1) Hypothese: Annahmen über reale soz. Sachverhalte bzgl. Verhalten von Variablen
 - Zsh. / Unterschiede / Kausalitäten zw. Variablen mittels Verknüpfungsregeln (kausale, struktur)
 - Anforderungen: deduzierbar, falsifizierbar, im Geflecht logisch widerspruchsfrei
 - Fkt: Grundlage / Steuerungsinstrument / Ziel des gesamten F-Prozesses
 - Arten: ungerichtet – gerichtet ; einseitig – zweiseitig ; monoton, implikativ, äquivalent
- 2) Variable: eindeutig meßbare Eigenschaft einer Person / Gruppe / ...
 - Merkmalsdimension (Variable): 1 meßbare Eigens. mit mindest. 2 Kategorien
 - Merkmalsausprägung (Kategorie): alle konkreten Eigens., die die Variable annehmen kann
dabei: Anforderungen: disjunkt (nicht überlappend) , erschöpfend (alle Möglichkeiten)
Differenzierung: diskrete Kat. (kein Intervall) ; kontinuierl. Kat. (mit klaren Interv.)
 - Merkmalsträger: Person / Gruppe / ... , die Eigenschaft trägt
 - Merkmalsebene: Teilgebiet des realen soz. Sachverhalts, der durch Variable beschrieben wird
 - Typen:
 1. unabh. : unbeeinflußt von anderen (selbst Beeinflusser), direkt meßbar
 2. abh. : beeinflusst durch andere Var. , nur indirekt meßbar
 3. intervenierend: beeinflusst Wirkung zw. Anderen Variablen
 4. antezedierend: Variable, die mehrere Variablen beeinflusst
- 3) Kumulativität: Addition von Kategorien (o.ä.)
- 4) Objektivität: Ergebnis unabh. Vom Durchführenden (intersubjektive Überprüfbarkeit)
- 5) Reliabilität: Zuverlässigkeit (gleiches Ergebnis bei Wiederholung)
- 6) Validität: Gültigkeit (? : U'-design erfaßt wirkliches U'-interesse) [auf jedes Detail beziehbar]
- 7) Falsifikation: emp. Nachweisbarkeit / Operationalisierung möglich
- 8) Verifikation: Nachweis des allgemeingültigen Wahrheitsgehaltes

Datenerhebung: Messung & Meßniveaus:

system. Zuordnung von Werten zu Kategorien => Relationsdarstellung [(!) Isomorphie: ?: Mess. = = Realität]

WICHTIG: Check Operational. (Objekt., Reliab., Valid.) ; Meß- / Analyseverfahren abh. von Variablentypen;

Vorgehen: Festlegung des Meßniveaus (in Operat.) ; Meßdaten in Rohdatenmatrix (alle erhobenen) ->

Auswertung /Aufbereitung der Daten in festgelegtem Analysekonzept

Meßniveaus: Sortierung nach Aussagekraft

Skalentyp	Aussagekraft	Mittelwert	mögl. Transformat. (werterhaltend)	Beispiel
Nominalskala	Gleichheit	Modalwert (Mo)	eindeutige	
alternativ (2 Kat.)				Geschlecht
kategorial (> 2)				Familienstand
Ordinalskala	Rangfolge (Vergleichsoperator)	Median (Md)	Rangfolge bewahrend	Schulabschluß
Intervallskala	Rangfolge mit gleichen Intervallen	Arithm. Mittel (x)	positive lineare	Temperatur
Ratioskala	qualitative Rangfolge mit gleichen Interv. + absoluter Nullpkt.	geometr. Mittelwert	positiv proportionale	Einkommen
Absolutskala	Wie Ratioskala + festgelegte Einheit.	geometr. Mittelwert	positiv proportionale	Häufigkeiten

Nominal- & Ordinalskala: unmetrische / qualitative Niveaus ; Intervall- / Ratio- : metrische, quantitative

Repräsentativität von Stichproben:

in Praxis keine Vollerhebung möglich => Frage: Stichprobenergebnis = Vollerhebungsergebnis (imaginär) ?

=> Beurteilungskriterien: Stichprobe repräsentativ wenn $n > 30$, nur mögl. bei genaue defini. Grundgesamtheit

Kriterien: wofür, Auswahltech., Größe, Sonstige konkrete Besonderheiten ; unwichtig: Proportionalität

statistisch nicht mögl., aber stochastisch.: Ggsmtheit (X) ~ Stichpr. (x) + St-fehler (e) [mit Irrtumswahrsch.] ;

zwei Auswahltechniken für Stichproben-Probanden (mit Repräsentativitätsanspruch):

- 1) Zufallsauswahl: jedes Elem. von X muß gleiche reelle Chance haben, in Stichprobe zu kommen

Verfahren: einfache Zufallsauswahl: direkte, zufällige Auswahl von n Probanden aus X

geschichtete Auswahl: X wird gruppiert (auch mehrfach), dann wieder einfache Auswahl

(z.B. Random Route, Adress Route, ...)

Klumpenauswahl: x wird gruppiert (in Klumpen) -> Vollerhebung eines Klumpen

- 2) Quotenauswahl (bewußte Auswahl): aus Kriterienkatalog (Quotenplan) werden bewußt Probanden gewählt

Vorgehen: 1. Def. von X

2. Erstellen Quotenplan (Variablenbestimmung)

a) proportionale Auswahl: gleiche Verhältnis in X und x [Voraus.: Verteil. in X bekannt]

b) disproportionale Ausw.: Kategoriendifferenzierung => Vollerhebung einer Klasse

Vorteile:

Zufallsauswahl: theoret. stochastisches Argument

Quotenauswahl: Probanden erfüllen Voraussetz., Kosten, Ausfallquote, Weiternutzbarkeit der Daten

Verfahren: Auswahl abh. von Ziel, Inhalt, techn. Voraussetzungen, ...

Befragung:

Königsweg der Sozialforschung; WICHTIG: Antizipation (Kontrolle & system. Konstanzhaltung)

Problem: K-situation (Interviewer <-> Instrument <-> Befragter) -> Vergleichbarkeit, Sprache; Zuverlässigkeit;

Verfahren (klassifiziert nach Standardisierungsgrad):

1. unstrukturiert/ungelenkt (qualitativ): Exploration, Intensivinterview; für Feldexploration
(-): Dokumentation, Konstanz der Befragungssituation
2. strukturiert/gelenkt (quantitativ): halb-/vollstrukturiertes Interv.; für Hypothesenbelastung
(-): Sprach-/K-barrieren; (+): je größer theoret. Vorwissen, desto größer I-Gehalt

Arten:

postal. Umfrage (preisgünstig, gutes Instrument wichtig, geringer Rücklauf)

Telefoninterview (normal; CATI [computergestützt], Voice Processing [digit. Verarbeitung, verboten in Sozialf.])

Online-Befragung (Online-Panels [Überprüfung von Online-Aktiv.], klassis. Feldstudie)

Panel-Befragung (genau def. Probandengruppe über längere Zeit beobachtet [? Repräsentativität])

Häufigkeit / Wiederholung:

punktueller Befrag. (einmalig), mehrfach (gleiches Instrument, wechselnde Befragte / gleiche Befragte);

Instrument (Fragebogen):

20-30 min.; Reihenfolge, Aufbau (verdeckt / durchschaubar), Einbau von Trichter/Filter/Kontroll-Fragen;

Sprachniveau: allgem. Verständlichkeit, Validitätsprüfung durch Pretests (bzgl. Verständlichkeit)

Fragetypen: offen (Befragter muß aktiv antworten) -> qualitativ; geschlossen (Erinnerung) -> quantitativ;

Faktfragen, Meinungsfragen (schwer: Vergleichbarkeit, Hypoth.-belastung)

Skalierung der Antwortvorgaben: Intervallgrößen, gerade / ungerade Anzahl, eindeutige Gegensatzpaare

Befragungssituation:

Problem: Konstanzhaltung zwecks Vergleichbarkeit (K-situation nicht standardisierbar, menschl. Interaktion)

Interviewer: Ziel, Plan (mit Vorbereitung); Befragter: Erwartungsdruck, Antwortformulierung

Interviewsituation:

face-to-face (1 / 1), Gruppendiskus. (1 / n), Umfrage (0 / 1), Telefonbefrag. (1 / 1, nicht face-to-face)

Befragungsmedium:

Interv. protokolliert, beide protokollieren, Befragter füllt aus, elektron. Verarbeitung (Interv.)

Beobachtung:

zur Hypothesenprüfung / system. Feldexploration

Problem: selekt. Wahrnehm. / Fehlinterpret.: Verzerrungen (=> Schulung, detail. Instruktion, mehrere Beob.)

Verfahren: (generell: strukturiert <-> unstrukturiert (Sinn s.o.))

teilnehmend <-> nicht teilnehm.; offen <-> verdeckt; Feld <-> Labor; Fremdbeob. <-> Selbstbeob.

Inhaltsanalyse (Content Analysis, Bedeutungsanalyse):

system. hypoth.-geleitete Erhebung & Auswertung von Kommunikaten (verbale (Texte) & nonverbale) [MK]

[andere Def. (Berelson 1948): objektive, system., quantitative Beschreibung manifester K-inhalte; kritisiert]

Ziel / Ansätze:

1) deskriptiv: system. Inhaltsangabe (+ Typologie); keine Rückschlüsse auf Intention/Rezeption/Wirkung

2) ? Inferenz: Rückschluß auf externe Sachverhalte (Intention (diagnostisch), Rezept./Wirk (prognostisch)) ?

methodologische Problematik:

Bedeutungserfassung (valide & reliabel) => K-erfolg (quant. & qualit.) bzgl. des angenommenen K-modell:

Linearer Transfer (Laswell): abh. von Inhaltsübertragungsqualität

Reziprokes Modell: abh. von Interaktion der K-partner; (Wissens. muß Objektivität wahren)

Ablaufmodell:

1) Entscheidung für IA (nur offene Fragen, IA-Daten Hypoth.-belastungs-Fähigkeit ?)

2) Festlegung des Untersuchungsmaterials (im Hinblick auf Operationalisierung):

Gesamtheit (Zeitraum o.ä.); ggf. Stichprobe (Auswahltechn., s.o.); Untersuchungs- & Zählseinheiten

3) Definition des Kategoriensystems (erst testen, dann Codeplan + Codeerläuterungen)

Schwer: erschöpfend, objektive & valide Einheiten, abh. von U-variablen, Inhalt / Gewichtung

Codierung: trennscharfe Kategorienausprägungen, präzise Anweisung => Intercoder-Reliabilität

=> sehr komplexe Struktur

4) Pretests (bis Koeffizient (Intercoder-Reliabilität % Re-Test) > 85%)

5) Datenerhebung

6) Datenaufbereitung & -analyse

Teil B – Statistik

Statistik in P&K-Wissenschaft / Sozialwissenschaften:

P&K-Wissenschaft ist interdisziplinär => eine Teildisziplin mit sozialw. Erkenntnisverfahren
=> Statistik als Hilfsdisziplin der Datenverarbeitung (qualitative & quantitative Meßkonzepte)
speziell: Datenauswertung (s. Ablaufmodell emp. Sozialforschung) :
[Rohdatenmatrix (Urliste) -> primäre , sekundäre , kumulative Datenaufbereitung (jeweils höherwertig)]
Arten: deskriptiv (uni-, bi-, multivariate Verteil. [bei Vollerhebung]) ; Stichprobenanalyse (Schätzverfahren)

Datenformen:

Verteilung \ Methodologie	Beschreibung	Erklärung
univariat (1 Variable)	Verteilung, Maße zentr. Tendenz, Streumaße	%
bivariat (2 Vars, Zsh.)	Verteilungen, Zsh.- Kennwerte	unabh. Var. (Erklärungsfaktor) abh. Var. (Kriteriumsvariable)
multivariat (kompl.Var.-system)	kompl. Zusammenhangsanalyse (Faktoren-/Clusteranal.)	kompl. Kausalanalyse (Pfad / Regression (Beziehungen))

berechenbare Werte bzgl. Methodologie

Univariates:Tabellen:

Hauptinstrument der Datenpräsentation => wichtig: Aufmachung, Aussagefähigkeit, Angabe von Bezügen
Nutzungs- & Weiterverarbeitungsmöglichkeiten abh. vom Meßniveau !!
Arten: Häufigkeitstabelle, Histogramm (Stabdiagramm), Tortendiagramm

Änderung in Datenaufbereitung:

Rohdatenmatrix (erhobene Daten) -> Originalwerte (Häufigkeiten, prim. Aufbereit.)

→ Interpretationsklassen / Verteilungsdarstellung (Häufigkeiten in Wertgruppen, sek. Aufbereitung)

→ Interpret.-taugliche Daten / schrittweise gehäufte Fallzahlen (kumulierte Häufigkeiten, kumul. Aufber.)

Maße der zentralen Tendenz (TT für Mittelwerte): gegliedert nach Info-Gehalt

Sinn: Darstellungsvereinfachung (Informationsverdichtung)

1. Modalwert [Mo]: Merkmalsausprägung mit höchster Fallzahl

Berechnung: sofort einsichtig

Anmerkung: berechenbar ab Nominalskala-Niveau

2. Median [Md]: Wert, der Verteilung in 2 gleiche Hälften teilt

Berechnung: ungerade Fallanzahl: Wert des Falls, der Verteilung halbiert

gerade Fallanzahl : $(\text{Wert}[n/2] + \text{Wert}[(n+2)/2]) / 2$ (Grenzfälle addieren, durch 2)

Anmerkungen: berechenbar ab Ordinalskala-Niveau

unbeeinflusst von Extremwerten (Gegensatz zu 'arithmetisches Mittel')

\sum der absoluten Abweichungen vom Md ist ein Minimum

3. arithmetisches Mittel: \sum aller Beobachtungswerte dividiert durch ihre Anzahl

Berechnung: $x = 1/n \sum x_i * f_i$ [1 bis k] mit: n = Fälle, x_i = Meßwerte, f_i = Häufigkeiten, k = untersch. Meßw. ;

Anmerkungen: berechenbar ab Intervallskala – Niveau

beeinflusst von jedem Meßwert (=> mögl. Verfälschungen durch 'Ausreißer')

\sum aller Abweichungen von $x = 0$

\sum aller quadrierten Abweichungen von x ist ein Minimum

Maße der Dispersion (TT für Streumaße):

Sinn: Aussage für Homogenität der Verteilung (Ähnlichkeit)

für Skalenniveaus: Ordinal: Range, Quartile, Perzentile ; ab Intervall: \emptyset -Abweichung, Varianz, Standardabw.

Veranschaulichung Homogenität: homogen, wenn $1/n * \sum (x_i - x) \rightarrow 0$ [von 1 bis k]

=> Formel für Varianz: Mittelwert der Summe der quadrierten Abweichung von x

$s^2 = 1/n * \sum (x_i - x)^2 * f_i$; für Stichproben: $s^2 = 1/(n-1) * \sum (x_i - x)^2 * f_i$

=> Standardabweichung: Quadratwurzel der Varianz (zur besseren Anschaulichkeit)

$s = \sqrt{s^2}$

WICHTIG: Aussagekraft nur sinnvoll in Verbindung mit x !!

Bivariate: Beziehungen zw. 2 Variablen

Arten: bivar. Zsh.-annahme ($X \leftrightarrow Y$); bivar. Unterschiedsann. ($A \neq B \rightarrow Y$); bivar. Kausalann. ($X \rightarrow Y$);

bivariate Zusammenhänge von Verteilungen für polynome (> 2 Kategorien) Variablen:

Koeffizienten für Skalenniveaus: Nomial: Kontingenzkoeff.(C); Ordinal: Rangkoeff. (R); [? Unterschiede ?]
metrisch: Maßkorrelationskoeff. (r) [für symmetr. & eingipflig, sonst gruppieren oder Rangkoeff. (R)]

WICHTIG: Wenn zwei Niveaus verglichen werden, bestimmt das niedrige den Koeffizienten

Maßkorrelationskoeffizient: [andere nicht erläutert]

- Kennwert für linearen Zsh. zw. 2 Variablen (im Koordinatensystem: Annäherungsgerade)
- Voraussetzung: metrisches Skalenniveau, Nomialverteilung der beiden Variablen
- Aussage: +1.0 = höchster positiver Zsh. ; -1.0 = höchster negativer Zsh. ; 0.0 = kein Zsh.
dabei: $|r| > 0.3$: überzufällig [für $n > 100$], bessere Aussagekraft mit r^2]
- Kovarianz: Abweichungswert für beide Variablen: $\text{cov}_{xy} = 1/(n-1) * \sum (y_i - \bar{y}) (x_i - \bar{x}) * f_i$ [1-k]

Formel: Kovarianz / ((n-1) * s_x * s_y)

bivariate Unterschiede von Verteilungen für dichotome (2 Kategorien) Variablen:

Darstellungsmöglichkeiten:

Kreuztabelle: Var (1) in Zeilen, Var (2) in Spalten, jeweils am Md geteilt (\Rightarrow 4-Felder-Tabelle)

\Rightarrow Phi-Koeff.: 4 Felder= a,b,c,d $\Rightarrow \phi = \frac{a*d - b*c}{\sqrt{(a+c)*(c+d)*(a+b)*(b+d)}}$

Prozenttabelle: Darstellung der Verteilung einer Variable in mehreren Gruppen

dabei: Zeilen = Kategorien, Spalten = Gruppen, Angaben in % (\Rightarrow Vergleichsübersicht)

Wahrscheinlichkeitstheorie:

Anwendung: Signifikanzprüfung von Stichproben (s. Repräsentativität): Wiederholung / stochast. Beweis

Basisdefinitionen:

a) Begriffe:

Zufallsexperiment: beliebig wiederholbar ; keine Voraussagbarkeit des Ergebnisses

Zufallsereignis/-ergebnis (E): Wahrs. des Eintretens: $P(E) = f_E / n$ [n = Versuche, f_E = Anzahl/Ereig.]

Zufallsvariable X: Zuordnung von Werten für Ergebnisse eines Zufallsexper. \Rightarrow Wahrs.-verteilung

Elementarereignis: alle möglichen Ereignisse ; Menge aller Elementarereign.: Ereignisraum

\Rightarrow wenn alle Ereig. im Ereignisraum gleichwahrscheinlich: $P(E_i) = 1 / n$

\Rightarrow für Ereignisgruppe E_1 (k gleichwahrscheinliche Ereig.) : $P(E_1) = k / n$

Wahrscheinlichkeit (P) zw. $0 \leq P \leq 1$

b) Basisaxiome:

Additionssatz: $P(E_1 \vee E_2 \vee \dots \vee E_n) = P(E_1) + P(E_2) + \dots + P(E_n)$ [logisches oder] \Rightarrow Vereinigung

Multiplikation: $P(E_1 \wedge E_2) = P(E_1) * P(E_2)$ [logisches und] \Rightarrow Durchschnitt

Komplementarität: $P(p) + P(q) = 1$ mit: p= untersuchtes Ereig. ; q= alle anderen Ereig. aus E-raum

c) Permutation (Anzahl der mögl. Kombis einer Menge): $P_n = n!$ mit: P_n = Anzahl der mögl. Permutationen

Wahrscheinlichkeitsverteilungen:

abh. von Skalierung d. Zufallsvar. X : diskrete Kat.: Binomialverteilt ; stetige (kontinuierl.) Kat.: Normalverteilt.

Binomialverteilung ($B_n(p)$):

Voraussetz.: Ereignisraum ist begrenzt (diskrete Kategorisierung)

jedes Zufallsexperiment hat kleine Ausgangsbedingungen

\Rightarrow gesucht: $P(X = a)$ [Wahrscheinl., daß X den Wert a annimmt]

Berechnung: Binominalkoeffizient: $\binom{n}{a} = \frac{n!}{(n-a)! * a!}$ \Rightarrow Relation $p \leftrightarrow n$

Wahrs. für irgendein Ereign. aus E-raum: $P(X=x) = p^x * q^{(n-x)}$

Wahrs. für bestimmtes Ereign. aus E-Raum: $P_{n,x} = \binom{n}{x} * p^x * q^{(n-x)}$

Anmerkungen: je größer Differenz $p \leftrightarrow q$, desto asymmetrischer

Normalverteilung:

Allgemeines: beschreibt Fläche unter Zufallsvariable X, genauer: relative Lage der Werte bzgl. x

2/3 aller Fälle zw. $[x - s; x + s]$; 95% aller Fälle zw. $[x - 2s; x + 2s]$

theoretisch unendliche Anzahl von Elementarereignissen möglich

\Rightarrow gesucht: $P(X = x)$ [Wahrscheinl., daß Ereig. in bestimmtem Intervall landet]

Berechnung: Wahrs., daß Ereign. in Intervall a / b: $p(a \leq x \leq b) = \int_a^b f(x) dx$

\Rightarrow Normalverteilung: $f(x) = \frac{1}{\sigma \sqrt{2\pi}} * e^{-1/2 * ((x-\mu)/\sigma)^2}$

Anmerkungen: glockenförmig, symmetrisch um μ

Standardisierung (bessere Vergleichbarkeit): z-Transformation ($\Rightarrow x = 0, s=1$)

Stichprobentheorie:

bzgl. statistischem Signifikanztest, nur für schließende Studien (von Stichprobe auf Ggsmtheit)

Induktionsschluß: Schluß von Stichprobe auf Ggsmt

$x \sim \mu$, $s \sim \sigma$, $P \sim \Pi$ (x, s, P in Stichprobe; μ, σ, Π in Ggsmtheit); nur bei Normalverteilungen

Stichprobenfehler e (in jeder Stichprobe):

Berechnung (empirisch: Mittelwert der arithmetr. Mittel):

[e]: $\sigma_x = \sigma / \sqrt{n}$ aber: σ nicht bekannt \Rightarrow Schätzformel: $\sigma_x = s_x = s / \sqrt{n}$

also: $\mu \sim x \pm z_i * s / \sqrt{n}$ [Bestimmung des arithmetr. Mittels in der Ggsmtheit]

Schätzverfahren für P \rightarrow Π :

Standardfehler:

$\sigma_p = \sqrt{\Pi(100-\Pi) / n}$ aber: Π nicht bekannt \Rightarrow Schätzformel: $\sigma_p = s_p = \sqrt{p(100-P) / n} = \sqrt{p * q / n}$

$\Rightarrow \Pi \sim P \pm z_i * \sqrt{p * q / n}$

[praxisorientierte Faustregel (konservativ (Intervall größer als echtes): $\Pi = P \pm (1/\sqrt{n}) * 100$]

Beeinflussung des Intervalls:

Stichprobengröße, Sicherheit, statistische Werte

Hypothesenprüfung:

Vertrauensgrenzen: falls Überschneidungen in Ggsmtheit von μ_1, μ_2 : keine wissens. korrekten Aussagen mögl.

Voraussetzungen: deduzierbar, log. widerspruchsfrei, falsifizierbar (s. Definition oben)

Arten:

Alternativhypothesen (= inhaltl., wissens.): Unterschieds-, Zsh., Kausalhyp. ; gerichtet, ungerichtet

Nullhypothese (= statistische, formale): zur Prüfung stat. Sign. (besagt: Alternativ-Hyp. stat. nicht signifikant)

\Rightarrow wenn bestätigt, dann ist Alternativhypothese falsch

Testverfahren: Prüfung der relevanten Sachverhalte (Frage, Operational., Meßniveau, ..) \Rightarrow Richtigkeit

Bsp: χ^2 – Test: Formel zur Berechnung der Nullhypothese (s. Aufzeichnungen)