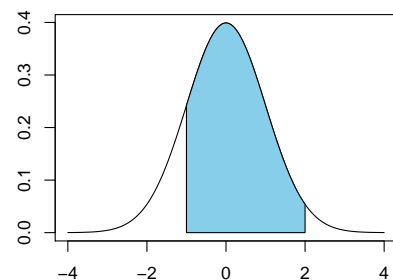


# Grundbegriffe der Wahrscheinlichkeitsrechnung

## Statistik I

Sommersemester 2009

- Wiederholung
- Non-parametrische Regression
- Wahrscheinlichkeitsrechnung
  - Wahrscheinlichkeitsbegriff
  - Eigenschaften von Wahrscheinlichkeiten
    - Notation/Begrifflichkeit
    - Eigenschaften
    - Vereinigungs- und Schnittmenge
    - Konditionale Wahrscheinlichkeiten
  - Zufallsvariablen & -verteilungen
    - Diskrete Verteilungen
    - Kontinuierliche Verteilungen
- Zusammenfassung



## Zum Nachlesen

- ▶ Agresti/Finlay: Kapitel 4.1-4.3
- ▶ Rudas, Tamas: Probability Theory. A Primer. Thousand Oaks u. a.: 2004

## Interaktion

- ▶ Standardmäßig additives Zusammenwirken
  - ▶ Arbeitslosigkeit hat Effekt auf Wahl des FN
  - ▶ Anteil Zuwanderer hat Effekt auf Wahl des FN
  - ▶ Beides *unabhängig* voneinander
- ▶ Interaktion: Effekte verstärken sich gegenseitig/schwächen sich gegenseitig
  - ▶ Wirkung von Arbeitslosigkeit +
  - ▶ Wirkung von Zuwanderung +
  - ▶ Wirkung von Arbeitslosigkeit  $\times$  Zuwanderung (positiv oder negativ)
- ▶ Bildung einer Produktvariablen; Effekt schätzen wir für normale Variable

## Daten

Département	Zuwanderer	Arbeitslosigkeit	Produktvariable
Ain	7.7	4.7	36.1
Aisne	2.5	8.5	21.5
Allier	2.9	6.8	19.4
...	...	...	...

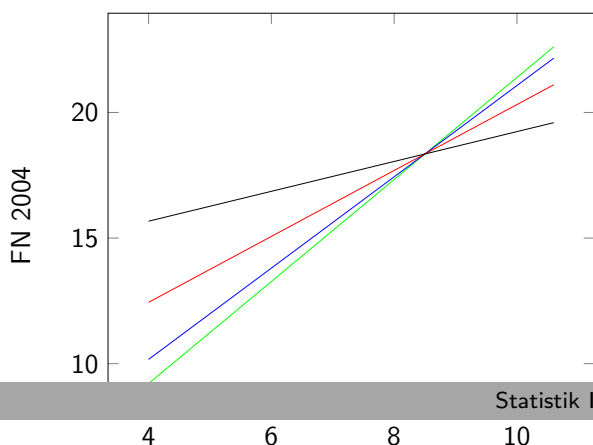
- ▶ Wenn Arbeitslosigkeit und Zuwanderung additiv wirken
- ▶ Hat Produktvariable keinen Effekt (Koeffizient von 0)
- ▶ Darstellung am besten graphisch

## Arbeitslosigkeit, Zuwanderung, FN-Wahl

### Modell mit Interaktion

$$\text{FN 2004} = -3.2 + 2.5 \times \text{Arbeitslosenquote} + 1.7 \times \text{Zuwandererquote} - 0.2 \times \text{Arbeitslosenquote} \times \text{Zuwandererquote}$$

- ▶ 2.5: Effekt der Arbeitslosenquote wenn Zuwandererquote = 0
- ▶ 1.7: Effekt der Zuwandererquote wenn Arbeitslosenquote = 0
- ▶ Für alle anderen Konstellationen: einsetzen und ausrechnen

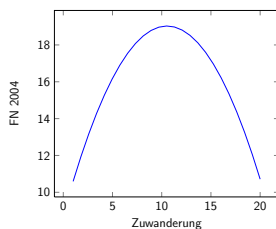


Zuwandererquoten: 2.6% (grün), 3.7%, 6.3%, 10% (schwarz);  $R^2 = 0.25$

- ▶ Interpretation?
- ▶ Effekt der ALQ bei niedriger Zuwanderung am stärksten
- ▶ Effekt Zuwanderung stark und positiv bei niedriger ALQ
- ▶ Schwächt sich ab mit höherer ALQ
- ▶ *Negativ* bei  $ALQ > 8.5$  ( $1.7/0.2$ )

## Non-lineare Regression

- ▶ Linearer Zusammenhang nicht immer plausibel
  - ▶ Höhere Zuwandererquote verunsichert ethnische Franzosen → Anstieg FN
  - ▶ Höhere Zuwandererquote → mehr Zuwanderer, die nicht FN wählen
  - ▶ Effekt der Zuwanderung schwächer/negativ für höhere Niveaus von Zuwanderungsquote
- ▶ Modelliert durch zusätzlichen Effekt für quadrierte Zuwanderungsquote
- ▶ Negativer Koeffizient → nach unten offene Parabel



## Was ist an unseren Modellen „parametrisch“?

- ▶ Alle bisher betrachteten Modelle haben einen oder mehrere Parameter
- ▶ D.h. mehr oder minder einfache Funktionen mit einigen wenigen Zahlen (Koeffizienten)
- ▶ Sparsam und relativ flexibel
- ▶ Aber: Annahmen
  - ▶ Funktionale Form korrekt
  - ▶ Verteilungsannahmen

## Was sind nicht-parametrische Verfahren?

- ▶ Keine Annahmen über Verteilung
- ▶ Keine vorgegebene funktionale Form
- ▶ Daten „sprechen“ (Setzen relativ viele Datenpunkte voraus)
- ▶ Vielzahl von unterschiedlichen Verfahren (nicht nur Regressionmodelle)
- ▶ Zwei Hauptgebiete
  - ▶ Verteilungsfreie Hypothesentests
  - ▶ **Non-parametrische Regression** → Vielzahl von Verfahren
- ▶ Ein wichtiges Verfahren: Lokale polynomiale Regression (= loess = scatter plot smoother)

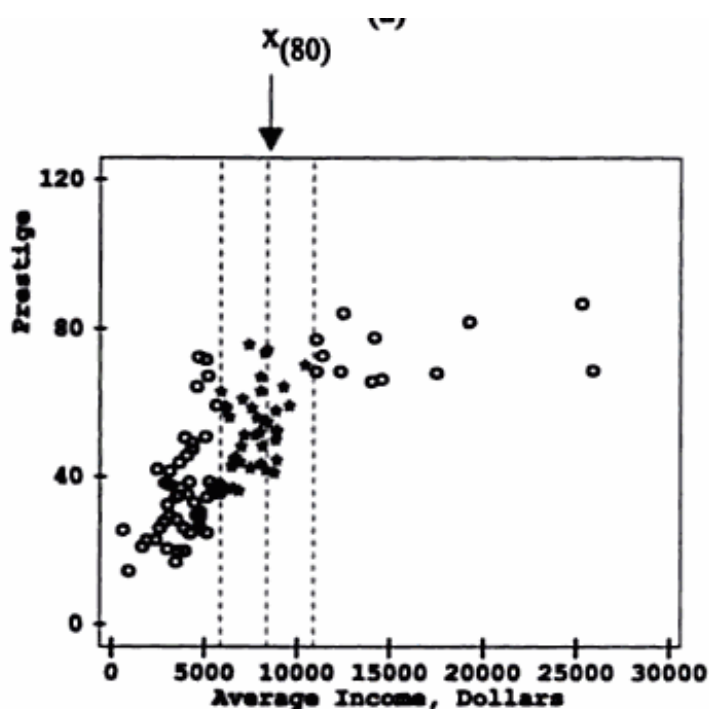
## Wie funktioniert die „polynomiale lokale Regression“

- ▶ Zwei Bestandteile
  1. „Polynomial“:  $y = \beta_0 + \beta_1 x_1 + (\beta_2 x^2) \dots$
  2. „Lokal“: eigene Regressionen für lokale Umgebungen → insgesamt sehr flexible Form
- ▶ (Im wesentlichen) graphisches Verfahren, keine Parameter
- ▶ Erfordert sehr viel Rechenleistung

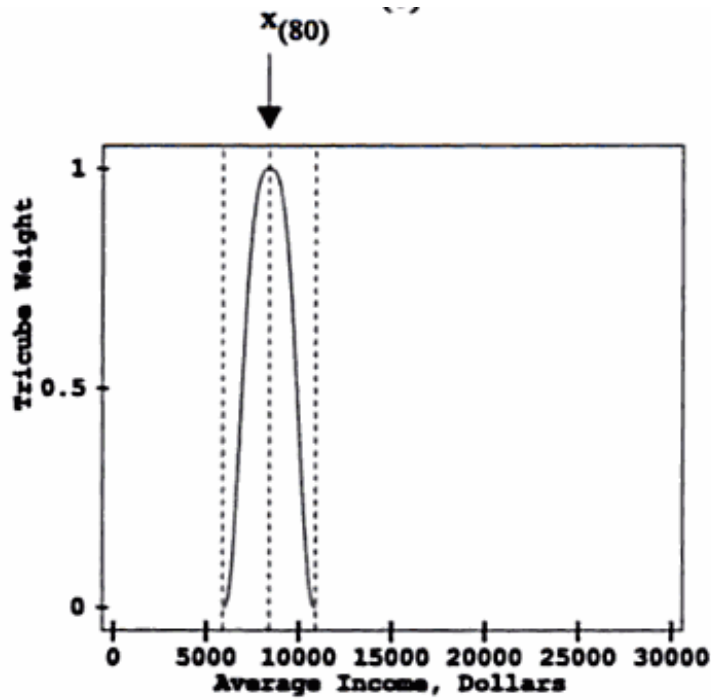
## Vorgehensweise: FN 2004 und Zuwandererquote

- ▶ Datensatz nach  $x$ -Variable (Zuwandererquote) sortieren
- ▶ Anteil der Meßwerte („Fenster“) festlegen, die in lokalen fit mit einbezogen werden (z. B. 40%)
- ▶ Festlegen von „fokalen“ Werten für  $x$ , an denen lokale Regression berechnet wird (z. B. reale Werte von  $x$ )
- ▶ Festlegen einer Gewichtungsfunktion
  - Fokaler  $x$ -Wert mit maximalem Gewicht
  - Abfallendes Gewicht für andere Werte nach Distanz vom fokalen Wert
- ▶ Gewichtete Regression von  $y$  auf  $x$ , ggf.  $x^2, x^3$  an jedem fokalen  $x$ -Wert
- ▶ Schätzwert für  $y$
- ▶ Glatte Kurve verbindet (non-parametrische) erwartete Werte für  $y$  über den gesamten Wertebereich von  $x$

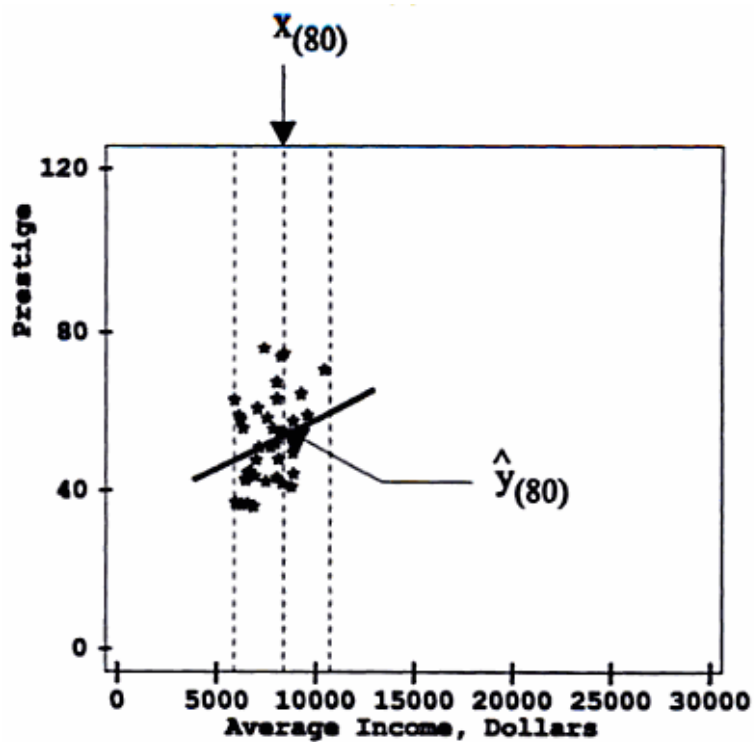
## „Fenster“



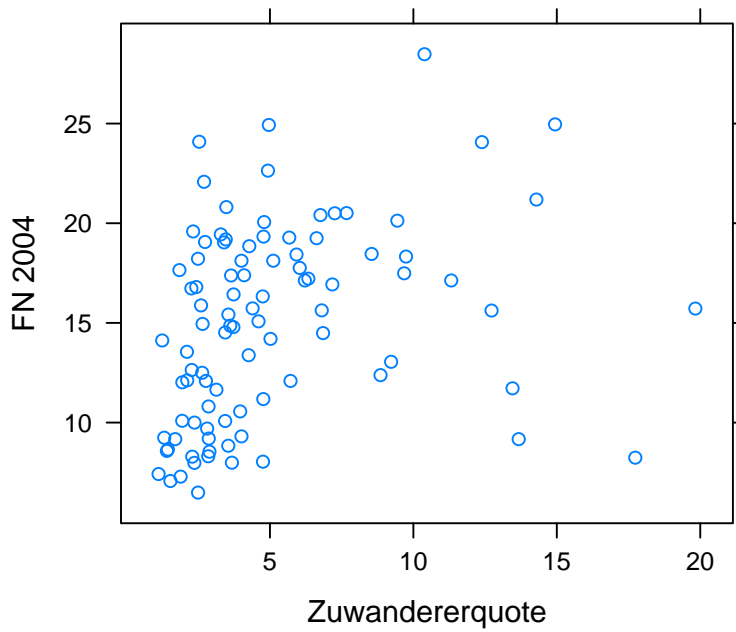
# Gewichtungsfunktion



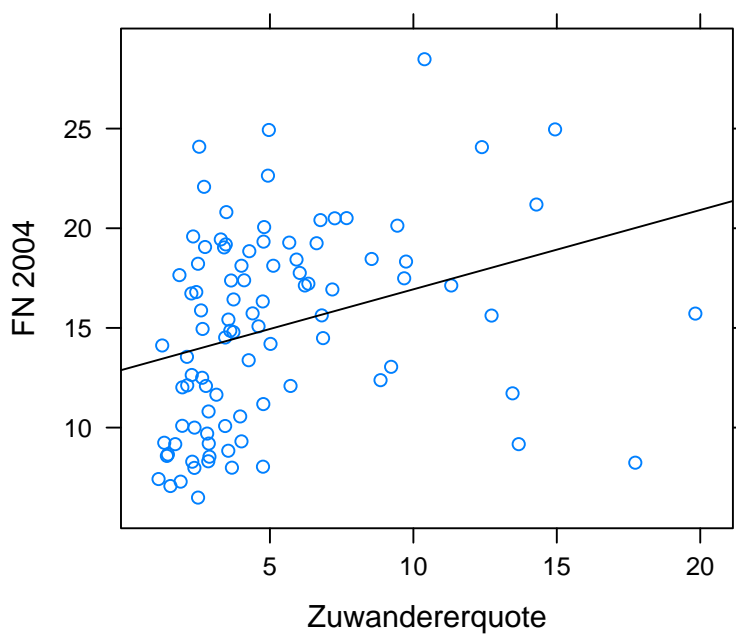
# Lokale Regression



## FN 2004 und Zuwanderer

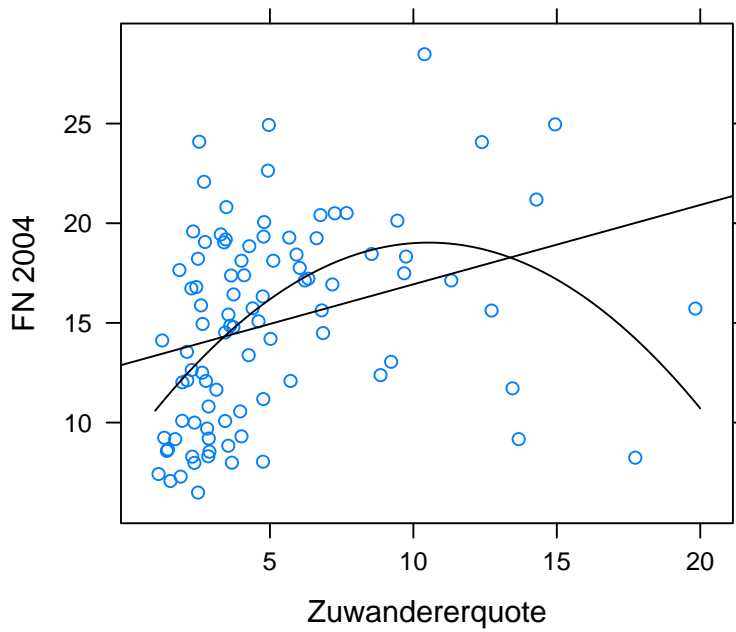


## FN 2004 und Zuwanderer: lineare Regression

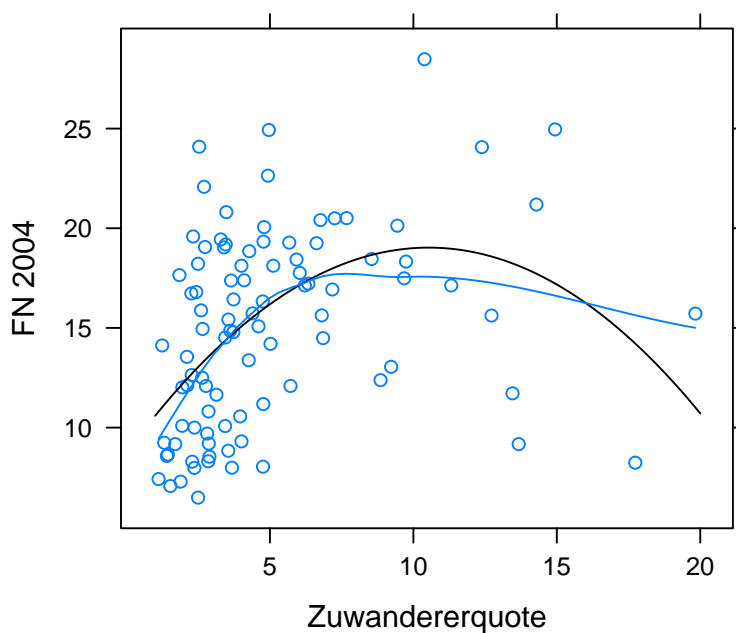




## FN 2004 und Zuwanderer: nicht-lineare Regression



## FN 2004 und Zuwanderer: non-parametrische Regression



## Woher kommen Wahrscheinlichkeiten?

- ▶ Politikwissenschaftliche Theorien/Modelle nicht deterministisch
  - ▶ Inglehart: Früher Wohlstand → Postmaterialismus ...
  - ▶ Almond/Verba: Übereinstimmung Struktur/Kultur → Stabilität des Systems
  - ▶ Lipset/Rokkan: Staat-Kirche-Konflikt → Christliche/konservative Parteien
- ▶ Theoretische Ansätze beschreiben eher eine *Tendenz*
  - ▶ Häufiger postmaterialistisch
  - ▶ Stabiler als andere
  - ▶ *Wahrscheinlicher* als in anderen Systemen
- ▶ Regressionsmodelle: stochastische Komponente
  - ▶ Nichtgemessene Einflüsse
  - ▶ Als *zufällig* betrachtet
- ▶ Zufallsstichproben

## Was ist der „frequentistische“ Wahrscheinlichkeitsbegriff?

- ▶ Zufallsexperiment
  - ▶ Ergebnis hängt vom Zufall ab (z. B. Befragung)
  - ▶ Beliebig oft wiederholbar; jeweils neues zufälliges Ergebnis
- ▶ Häufigkeit Ergebnis (z. B. CDU-Wähler) / Zahl der Wiederholung = relative Häufigkeit
- ▶ Wahrscheinlichkeit: Grenzwert der relativen Häufigkeit wenn Zahl der Wiederholungen  $\rightarrow \infty$
- ▶ Grundlage der inferenzstatistischen Standardverfahren
- ▶ Scheinbar einfach und anschaulich
- ▶ *Modell* für wissenschaftliche Untersuchungen/soziale Prozesse
- ▶ Probleme: Wiederholbarkeit und Stabilität der Wahrscheinlichkeit

## Welche Alternativen gibt es?



Thomas Bayes,  
1702-1761

- ▶ Frequentistische Wahrscheinlichkeit: objektiv
- ▶ Keine Wiederholbarkeit (z. B. Klausur): subjektive Wahrscheinlichkeit
- ▶ Alternative statistische Rahmentheorie → Bayesianische Statistik
- ▶ Seit ca. 25 Jahren praktisch anwendbar
- ▶ Steigende Bedeutung für die Politikwissenschaft seit ca. 10 Jahren

## Ereignisse

- ▶ Ergebnis eines Zufallsexperimentes: *Ereignis*
  - ▶ Elementareignis (outcome)
  - ▶ Ereignis als Kombination von Elementareignissen (event)
- ▶ Ereignisse werden oft von Großbuchstaben symbolisiert
- ▶ Alle möglichen Ergebnisse (Elementareignisse) eines Zufallsexperimentes bilden die Ergebnismenge  $\Omega$
- ▶ Z. B. Münzwurf:  $\Omega = \{K, Z\}$
- ▶ Oder Würfel-Wurf:  $\Omega = \{1, 2, 3, 4, 5, 6\}$
- ▶ Abzählbare vs. über-abzählbare Ergebnismengen: Würfel vs. exaktes Lebensalter
- ▶ Komplexe Experimente/Ereignisse, z. B. Zahl/Anteil Unionswähler
- ▶ Abzählbaren Elementarereignissen (und ihren Kombinationen) sind Wahrscheinlichkeiten zugeordnet (relative Häufigkeit)

## Eigenschaften von Wahrscheinlichkeiten

- ▶ Wahrscheinlichkeit von Ereignis  $K = P(K)$  z. B.  $P(K) = 0.5$  (faire Münze)
- ▶ Ein *unmögliches* Ereignis hat eine Wahrscheinlichkeit von 0
- ▶ Ein *sicheres* Ereignis hat eine Wahrscheinlichkeit von 1
- ▶ Die Summe der Wahrscheinlichkeiten aller Elementarereignisse für ein Zufallsexperiment ist 1
- ▶ Allgemeiner: Die Wahrscheinlichkeit für die Vereinigung einer Reihe von wechselseitig ausgeschlossenen (disjunkten) Elementarereignissen ist gleich der Summe ihrer Einzelwahrscheinlichkeiten
- ▶ Das Komplement eines Ereignisses  $A^C$  ist die Menge aller Ereignisse, in denen  $A$  *nicht* eintritt
- ▶ Die Wahrscheinlichkeit von  $A^C = P(A^C) = 1 - P(A)$

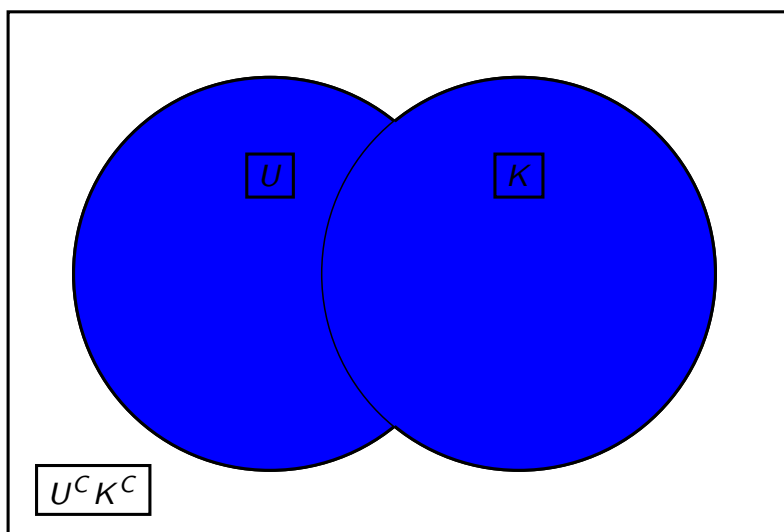
## Beispiel: Europawahl

- ▶ Befragung einer zufällig ausgewählten Wählerin als Zufallsexperiment
  - ▶ Drei Elementarereignisse: Union (U), SPD (S), Andere (A)
  - ▶ Wahrscheinlichkeiten:  $P(U) = 0.3$ ;  $P(S) = 0.2$ ;  $P(A) = 0.5$
- ▶ Elementarereignisse schließen sich gegenseitig aus (disjunkt)
- ▶ „Schnittmenge“ zwischen  $U$  und  $S$  (gemeinsames Auftreten) leer:  $U \cap S = \{\}$
- ▶  $P(US) = 0$  (gemeinsames Auftreten)
- ▶ „Vereinigungsmenge“:  $U$  oder  $S$  oder beides (hier nicht möglich)  $U \cup S = \{U, S\}$
- ▶ Wahrscheinlichkeit für  $U$  oder  $S$ :  
 $P(U \cup S) = P(U + S) = P(U) + P(S) = 0.3 + 0.2$

## Allgemeiner: „Summe“ und „Produkt“ von Ereignissen

- ▶ Nicht-disjunkte Elementarereignisse: Wahlverhalten wie vorher und Konfession katholisch/nicht katholisch ( $K, K^C$ )
- ▶  $P(K) = 0.4; P(K^C) = 0.6$
- ▶ Kombination von Elementarereignissen
- ▶ „Summe“ ( $U + K$ ):
  - ▶ *Mindestens eines* von zwei oder mehr Ereignissen tritt ein
  - ▶ Entspricht Vereinigungsmenge
- ▶ „Produkt“ ( $UK$ )
  - ▶ *Zwei (bzw. alle)* Elementarereignisse treten ein
  - ▶ Entspricht Schnittmenge

## Vereinigungsmenge



Alle Unionswähler Alle Katholiken  $UK$ : Unionswähler **und** katholisch  $U + K$ : Alle Unionswähler + alle Katholiken - Schnittmenge (würde sonst zweimal gezählt) =  $UK^C + U^C K + UK$

## Wahrscheinlichkeit: Auftreten von zwei Ereignissen

- ▶ Disjunkt
  - ▶ Keine Schnittmenge
  - ▶ Wahrscheinlichkeit 1 + Wahrscheinlichkeit 2
  - ▶ Z. B.  $P(U + S) = P(U) + P(S) = 0.3 + 0.2$
- ▶ Nicht disjunkt
  - ▶ Schnittmenge = Wahrscheinlichkeit für gemeinsames Auftreten
  - ▶ Wahrscheinlichkeit 1 + Wahrscheinlichkeit 2  
- Wahrscheinlichkeit für *gemeinsames* Auftreten  
(muß bekannt sein)

## Implikation

- ▶ Ein Elementarereignis ist *sicher*, wenn ein anderes beobachtet wurde
- ▶ Z. B. Elementarereignis Unionsbürger  
( $P(B) = 0.9; P(B^C) = 0.1$ )
- ▶ Wenn *irgendeine* Wahlentscheidung beobachtet wurde, muß  $B$  vorliegen
- ▶ Wahlentscheidung „SPD“ impliziert Unionsbürgerschaft:  
 $S \subseteq B$
- ▶ Gegenteil der ausschließenden Disjunktion
- ▶ Implikation und Disjunktion Extrembeispiele für *konditionale* Wahrscheinlichkeit

## Drei Typen von Wahrscheinlichkeiten

1. Marginale Wahrscheinlichkeit
    - ▶ Wahrscheinlichkeit insgesamt
    - ▶ Z. B. Wahrscheinlichkeit, einen Katholiken auszuwählen: 0.4
  2. Gemeinsame Wahrscheinlichkeit
    - ▶ Schnittmenge
    - ▶ Z. B. Wahrscheinlichkeit, einen katholischen Unions-Wähler auszuwählen: 0.24 (fiktiv)
  3. Konditionale Wahrscheinlichkeit
    - ▶ Wahrscheinlichkeit für ein Ereignis (z. B.  $K$ )
    - ▶ Wenn ein anderes Ereignis bereits eingetreten ist (z. B.  $U$ )
    - ▶ Symbolisch:  $P(K|U)$
- ▶ Alle drei Wahrscheinlichkeiten hängen eng zusammen

## Gemeinsame Wahrscheinlichkeit → konditionale Wahrscheinlichkeit

- ▶ Wie groß ist die Wahrscheinlichkeit, daß ein Unions-Wähler katholisch ist?
- ▶ Wie groß ist der Anteil katholischen Unions-Wähler (Schnittmenge) an allen Unionswählern?
- ▶ D. h.:  $\frac{\text{gemeinsame Wahrscheinlichkeit } (UK)}{\text{Wahrscheinlichkeit Union } (U)}$
- ▶ Bzw.  $P(K|U) = \frac{P(UK)}{U} = \frac{0.24}{0.3} = 0.8$
- ▶ Umgekehrt: marginale und konditionale Wahrscheinlichkeit → gemeinsame Wahrscheinlichkeit (Multiplikation)

## Unabhängigkeit und Indifferenztabelle

- ▶ Wenn  $A$  von  $B$  unabhängig ist und umgekehrt
- ▶ Ist  $P(A|B) = P(A)$  (konditionale = marginale Wahrscheinlichkeit)
- ▶ Und  $P(B|A) = P(B)$
- ▶ Wahrscheinlichkeit für  $= P(AB)$ ?

$$P(A|B) = \frac{P(AB)}{P(B)} = P(A)$$

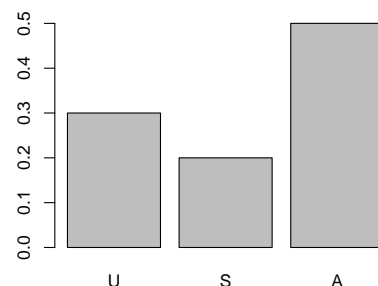
$$P(A|B) \times P(B) = P(AB)$$

$$P(A) \times P(B) = P(AB)$$

- ▶ Konstruktion der Indifferenztabelle:  
marginale Wahrscheinlichkeiten multiplizieren

## Was ist eine Zufallsverteilung?

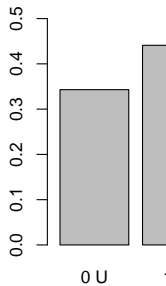
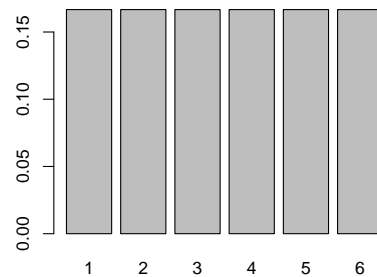
- ▶ Zufallsvariable  
(z. B. Wahlentscheidung)  
hat eine *Verteilung*
- ▶ Verteilung:  
gesamte Wahrscheinlichkeit  
(=1) verteilt sich auf Ereignisse
- ▶ Analog zu Verteilungen  
mit relativen Häufigkeiten
- ▶ Faßt Ereignisse/Wahrscheinlichkeiten des Zufallsexperimentes  
kompakt zusammen
- ▶ Kann häufig durch Formel beschrieben/approximiert werden





## Wichtige diskrete Verteilungen

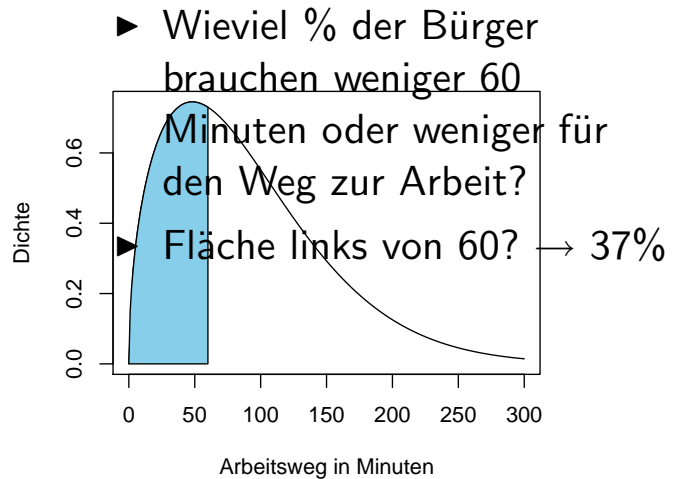
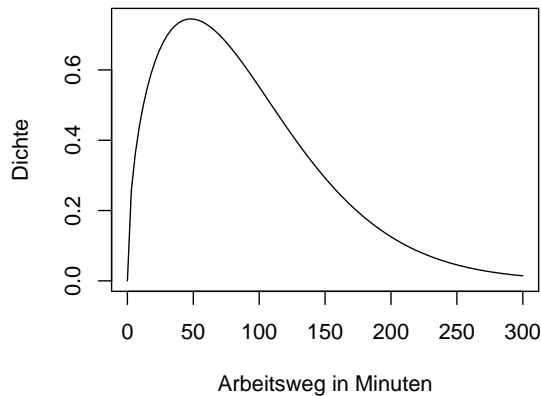
- ▶ **Gleichverteilung:**  
Alle Ereignisse gleich  
wahrscheinlich (z. B. Würfel)
- ▶ **Binomialverteilung:**  
komplexe Experimente
- ▶ Wie wahrscheinlich  
ist es, hintereinander  
drei Unionswähler zu befragen?
- ▶ Multinomialverteilung: Generalisierung der Binomialverteilung
- ▶ Balken (Wahrscheinlichkeiten) addieren sich stets zu eins



## Warum kontinuierliche Verteilungen?

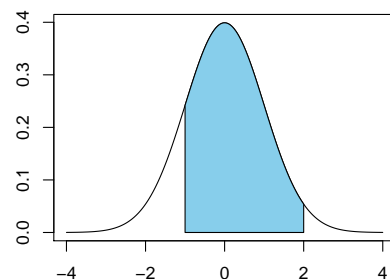
- ▶ Wahrscheinlichkeit  $\approx$  relative Häufigkeit eines Ereignisses
- ▶ Ok für abzählbare Ereignisse (diskrete Zufallsvariablen)
- ▶ Was tun mit kontinuierlichen Zufallsvariablen?
- ▶ Histogramm  $\rightarrow$  Dichteschätzung
- ▶ Fläche unter Dichtefunktion = 1 (Gesamtwahrscheinlichkeit)
- ▶ Wahrscheinlichkeit für Wert aus einem bestimmten Bereich =  
Fläche über diesem Intervall

## Dauer: Fahrt zur Arbeit



## Normalverteilung

- Bekannteste und wichtigste aller kontinuierlichen Zufallsverteilungen
- Wahrscheinlichkeitsdichte für Werte zwischen  $-\infty$  und  $+\infty$
- Fläche unter der Kurve = 1
- Über dem Intervall  $[-1;2]$  liegen rund 82% der Fläche
- D. h. in  $\approx 4$  von 5 Fällen zieht man einen Wert aus diesem Intervall



# Zusammenfassung

- ▶ Wahrscheinlichkeit  $\approx$  relative Häufigkeit bei häufiger Wiederholung
- ▶ Einfache Regeln/Eigenschaften
- ▶ Wahrscheinlichkeiten wichtig für Sozialwissenschaftler: Verhalten, Stichprobenziehung