VO Statistik

Sitzung 5: Da bin ich mir unsicher: Wahrscheinlichkeit

Dominik Duell

Universität Innsbruck

Etwas Admin

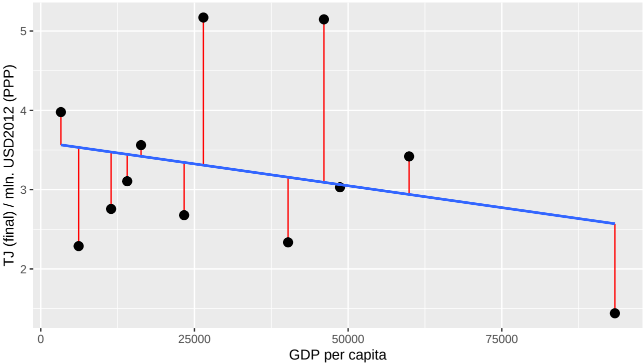
- Hausaufgabe 1 ist online
- etwas später, Pop Quiz: Folge dem Code, der an die Wand projeziert ist

Plan für heute

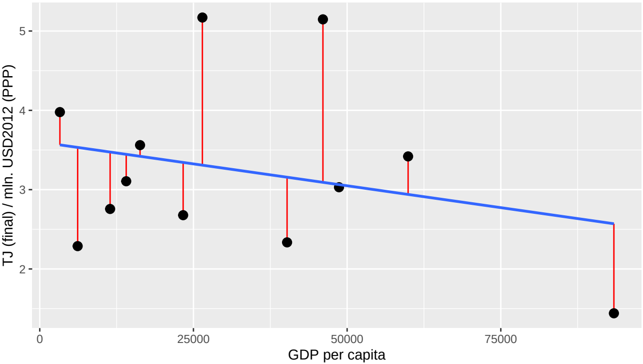
1. Modelle statistischer Zusammenhänge
2. Wahrscheinlichkeit

Modelle statistischer Zusammenhänge

Was ist unsere beste Schätzung?

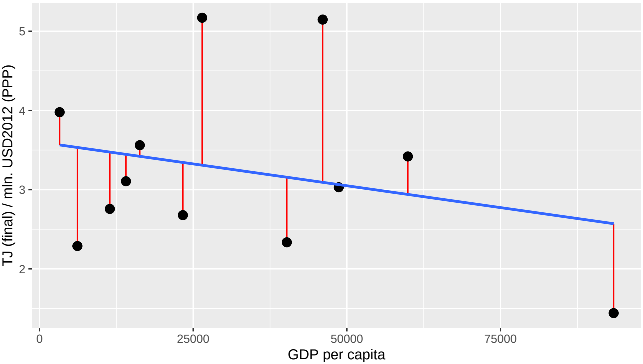


Was ist unsere beste Schätzung?



Die Linie, welche die Summe der quadrierten Fehler minimiert.

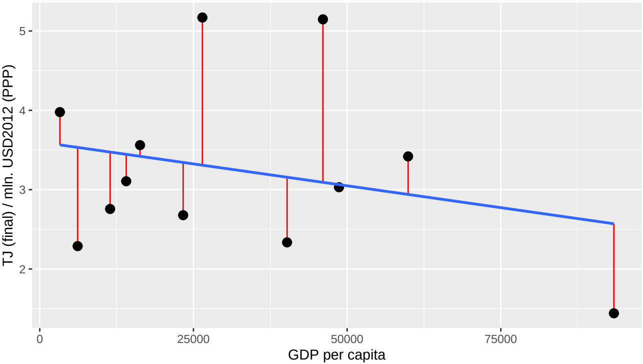
Was ist unsere beste Schätzung?



Die Linie, welche die Summe der quadrierten Fehler minimiert.

Methode der kleinsten Quadrate

Was ist unsere beste Schätzung?



Die Linie, welche die Summe der quadrierten Fehler minimiert.

Methode der kleinsten Quadrate

Wo kommt diese beste Schätzung her?

Wir wenden die **Methode der kleinsten Quadrate** auf unsere Regressionsmodel an:

$$emissions_i = b_0 + b_1 GDP_i + \epsilon_i$$

- wir wollen, dass die Summe der quadrierten Fehler in unserer Schätzung so klein wie möglich ist

$$\sum_{i=1}^{N} \epsilon_i^2$$

- Unseren Fehler ϵ können wir auch anders schreiben

$$\sum_{i=1}^{N} (emissions_i - b_0 - b_1 GDP_i)^2$$

Wo kommt diese beste Schätzung her?

- Die Daten für *emissions* und *GPP* sind gegeben durch unsere Stichprobe. - Wie kriegen wir jetzt die Regressionskoeffizienten, welche die Gleichung unten so klein wie möglich macht?

$$\sum_{i=1}^{N}(emissions_i - b_0 - b_1 GDP_i)^2$$

Wo kommt diese beste Schätzung her?

- Die Daten für *emissions* und *GPP* sind gegeben durch unsere Stichprobe. - Wie kriegen wir jetzt die Regressionskoeffizienten, welche die Gleichung unten so klein wie möglich macht?

$$\sum_{i=1}^{N}(emissions_i - b_0 - b_1 GDP_i)^2$$

Ableitung nach b_1 und b_0, um Extremwerte der Funktion zu finden!

Wo kommt diese beste Schätzung her?

Machen wir mal b_1

- $\frac{\partial \sum_{i=1}^{N} (y_i - b_0 - b_1 x_i)^2}{\partial b_1} = 0$

Wo kommt diese beste Schätzung her?

Machen wir mal b_1

- $\frac{\partial \sum_{i=1}^{N}(y_i - b_0 - b_1 x_i)^2}{\partial b_1} = 0$
- $\sum_{i=1}^{n} y_i^2 - 2b_0 y_i - \ldots = 0$

Wo kommt diese beste Schätzung her?

Machen wir mal b_1

- $\frac{\partial \sum_{i=1}^{N}(y_i - b_0 - b_1 x_i)^2}{\partial b_1} = 0$
- $\sum_{i=1}^{n} y_i^2 - 2b_0 y_i - \ldots = 0$
- $\ldots$

Wo kommt diese beste Schätzung her?

Machen wir mal b_1

- $\frac{\partial \sum_{i=1}^{N}(y_i - b_0 - b_1 x_i)^2}{\partial b_1} = 0$
- $\sum_{i=1}^{n} y_i^2 - 2b_0 y_i - \ldots = 0$
- $\ldots$
- $\sum_{i=1}^{n} -2(y_i - b_0 - b_1 x_i)x_i = 0$

Wo kommt diese beste Schätzung her?

Machen wir mal b_1

- $\frac{\partial \sum_{i=1}^{N} (y_i - b_0 - b_1 x_i)^2}{\partial b_1} = 0$
- $\sum_{i=1}^{n} y_i^2 - 2b_0 y_i - \ldots = 0$
- $\ldots$
- $\sum_{i=1}^{n} -2(y_i - b_0 - b_1 x_i) x_i = 0$
- $b_1 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n} (x_i - \overline{x})^2}$

Wahrscheinlichkeit

Warum, warum nur?

Warum, warum nur?

- wir wollen eine grundlegend unsichere und von Zufall getriebene Welt verstehen

Warum, warum nur?

- wir wollen eine grundlegend unsichere und von Zufall getriebene Welt verstehen
- jegliche deskriptive und kausale Inferenz über diese Welt basiert auf Schlussfolgerungen bei Unsicherheit

Warum, warum nur?

- wir wollen eine grundlegend unsichere und von Zufall getriebene Welt verstehen
- jegliche deskriptive und kausale Inferenz über diese Welt basiert auf Schlussfolgerungen bei Unsicherheit
- (Inferenz-)Statistik beschäfigit sich genau mit solchen Schlussfolgerungen bei Unsicherheit

Warum, warum nur?

- wir wollen eine grundlegend unsichere und von Zufall getriebene Welt verstehen
- jegliche deskriptive und kausale Inferenz über diese Welt basiert auf Schlussfolgerungen bei Unsicherheit
- (Inferenz-)Statistik beschäfigit sich genau mit solchen Schlussfolgerungen bei Unsicherheit

$\Rightarrow$ Wahrscheinlichkeit ist die formale Sprache der Unsicherheit und des Zufalls!

Wo sehen wir Wahrscheinlichkeit in der quantitativen Analyse durch Statistik?

```
read.csv('data/indicators.csv') %>%
  lm(GDPPerCapita~euJoin2004,data=.) %>%
  summary()
```

```
##
## Call:
## lm(formula = GDPPerCapita ~ euJoin2004, data = .)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8960.1 -2741.9  -474.1  2538.8 16372.9
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)               11128.9      335.3   33.19   <2e-16 ***
## euJoin2004No EU Member   -7607.7      443.6  -17.15   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4694 on 455 degrees of freedom
##   (56 observations deleted due to missingness)
## Multiple R-squared:  0.3926, Adjusted R-squared:  0.3912
## F-statistic: 294.1 on 1 and 455 DF,  p-value: < 2.2e-16
```

Begrifflichkeiten

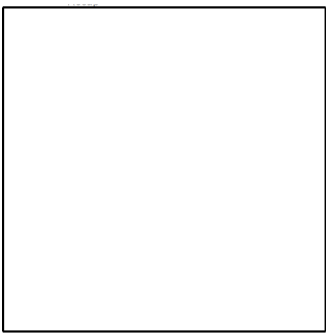
Definitionen I

- **Menge**: abzählbare Sammlung von Dingen: Menge an Champions League teams {Manchester City, AC Mailand, Real Madrid, ...}
- **Elemente**: enthalten in der Menge
- **Zufallsexperiment**: Momentaufnahme der Welt, die aber wiederholt werden kann.
- **Ergebnis**: alles das was in einem bestimmten Zufallsexperiment passieren kann
- **Ergebnisraum**: Menge alle möglichen Ergebnisse eines Zufallsexperiments.

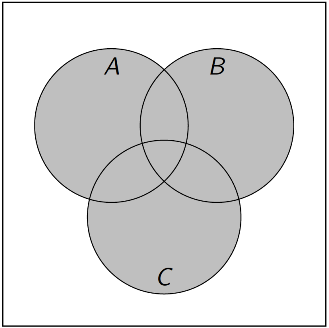
Definitionen II

- **Ereignis**: jede Sammlung möglicher Ergebnisse eines Experiments
- **Ereignisraum**: jede sich gegenseitig ausschließende, kollektiv erschöpfende Sammlung von Ereignissen eines Experiments.
- **Zusammengesetzte Ereignisse**: zusammengesetzt aus zwei oder mehreren einfachen Ereignissen – sind entweder **unabhängig** oder **bedingt**

Ergebnisraum



Ereignisse



Beispiel

- Werfe einen 6-seitigen Würfel einmal:
  - Ergebnisraum:

Beispiel

- Werfe einen 6-seitigen Würfel einmal:
  - Ergebnisraum: 1,2,3,4,5,6
    - Ereignisse: Werfe 1, Werfe 3, Werfe # größer als 4, ...

Beispiel

- Werfe einen 6-seitigen Würfel einmal:
  - Ergebnisraum: 1,2,3,4,5,6
    - Ereignisse: Werfe 1, Werfe 3, Werfe # größer als 4, ...
- Werfe eine faire Münze drei mal:
  - Ergebnisraum:

Beispiel

- Werfe einen 6-seitigen Würfel einmal:
  - Ergebnisraum: 1,2,3,4,5,6
    - Ereignisse: Werfe 1, Werfe 3, Werfe # größer als 4, ...
- Werfe eine faire Münze drei mal:
  - Ergebnisraum: TTT, TTH, THT, THH, HTT, HTH, HHT, HHH
  - Ereignisse:

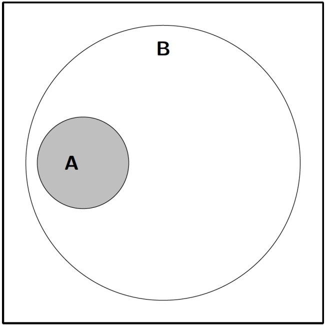
Beispiel

- Werfe einen 6-seitigen Würfel einmal:
  - Ergebnisraum: 1,2,3,4,5,6
    - Ereignisse: Werfe 1, Werfe 3, Werfe # größer als 4, ...
- Werfe eine faire Münze drei mal:
  - Ergebnisraum: TTT, TTH, THT, THH, HTT, HTH, HHT, HHH
  - Ereignisse: Werfe 3 T, Werfe 1 T on 1st and 1 H on 3rd, ...

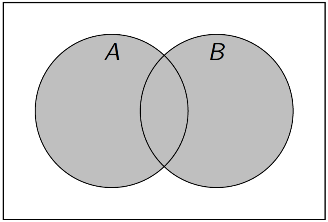
Arbeiten mit Mengen

- **Gegenereignis**: $A' = \{X : X \notin A\}$
- **Teilmenge**: A ist eine Teilmenge von B wenn jedes Element von A auch in B ist: $A \subset B \Leftrightarrow \forall X\, X \in A, X \in B$
- **Gleichheit**: $A = B \Leftrightarrow A \subset B, B \subset A$
- **Vereinigung von A, B**: $A \cup B = \{X : X \in A \,\text{or}\, X \in B\}$
  $A_1 \cup A_2 \cup \ldots A_n = \bigcup_{i \leq n} A_i$
- **Durchschnitt von A,B**: $A \cap B = \{X : X \in A \,\text{and}\, X \in B\}$
  $A_1 \cap A_2 \cap \ldots A_n = \bigcap_{i \leq n} A_i$
- **Leere Menge**: enthält keine Elemente, $\emptyset$

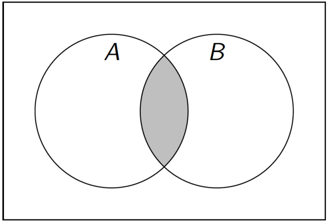
A als Teilmenge von B



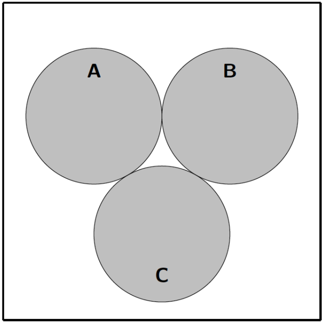
Vereinigung von A und B, $A \cup B$ oder $A + B$



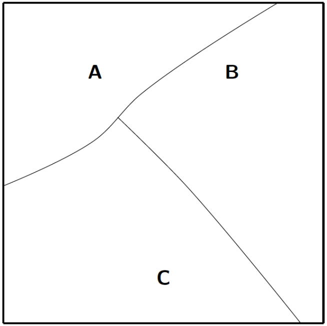
Durchschnitt von A und B, $A \cap B$ oder AB



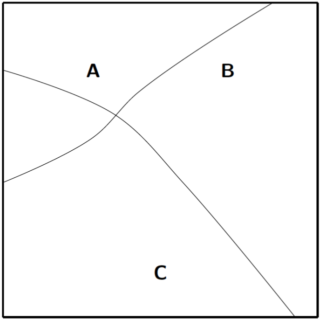
Gegenseitig ausschließend



Kollektiv erschöpfend



Kollektiv erschöpfend, aber nicht gegenseitig ausschließend

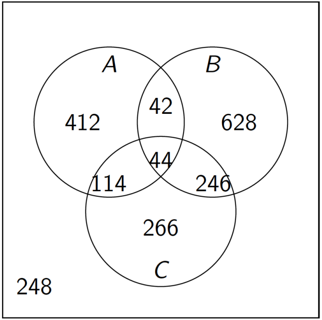


Beispiel

Nehmen wir das folgende Zufallsexperiment an: Wir fragen 2000 Personen nach deren Rauch und Trinkgewohnheiten. Nehmen wir an, es ergeben sich folgende Ereignisse aus diesem Zufallsexperiment:

- 612 rauchen,
- 960 tirnken,
- 670 sind älter als 25,
- 86 trinken und rauchen,
- 290 trinken und sind älter als 25,
- 158 rauchen und sind älter als 25,
- 44 trinken, rauchen und sind älter als 25,
- 248 trinken nicht, rauchen nicht und sind 25 und jünger.

Beispiel



Wahrscheinlichkeitsmaß

Wahrscheinlichkeitsmaß

- Nimm an, wir machen ein Zufallsexperiment mit Ereignisraum S,
- eine Funktion mit reellen Werten $\mathbb{R}$ für S nennt man **Wahrscheinlichkeitsmaß** wenn es die folgenden Merkmale hat:

Wahrscheinlichkeitsmaß

- Nimm an, wir machen ein Zufallsexperiment mit Ereignisraum S,
- eine Funktion mit reellen Werten $\mathbb{R}$ für S nennt man **Wahrscheinlichkeitsmaß** wenn es die folgenden Merkmale hat:

1. nicht negative $\#$: für jedes Ereignis A, $p(A) \geq 0$

2. Wahrscheinlichkeit von S ist 1: $p(S) = 1$

3. Für zwei Ergebnisse, die nicht zur gleichen Zeit stattfinden können, gilt: $AB = \emptyset$, $P(A + B) = p(A) + p(B)$

Wahrscheinlichkeitsmaß

Informell ausgedrückt:

$$\text{Prob(Ereignis)} = \frac{\#\text{ an Möglichkeiten, wie das Ereignis eintreten kann}}{\text{Gesamt }\#\text{ der möglichen Ergebnisse}}$$

Wahrscheinlichkeitsmaß

Informell ausgedrückt:

$$\text{Prob(Ereignis)} = \frac{\#\text{ an Möglichkeiten, wie das Ereignis eintreten kann}}{\text{Gesamt }\#\text{ der möglichen Ergebnisse}}$$

Beispiele

- Werfe Münze einmal, was ist die Wahrscheinlichkeit von "heads?"

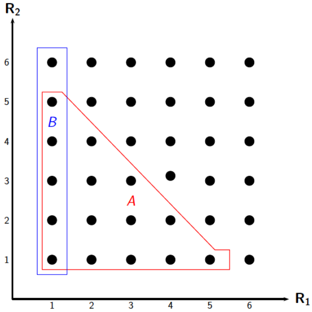
Beispiele

- Werfe Münze einmal, was ist die Wahrscheinlichkeit von "heads?"
  - # möglicher Ergebnisse ist 2
  - # der Ereignisse in den "heads" als Ergebnis möglich ist ist 1
  - daher $p(E) = \frac{1}{2}$
- Werfe Münze drei mal, was ist die Wahrscheinlichkeit von "heads at 2nd toss?"
  - # möglicher Ergebnisse ist $2^3 = 8$: TTT, TTH, THT, THH, HTT, HTH, HHT, HHH
  - # der Ereignisse in den "heads at 2nd toss" möglich ist ist 4: THT,THH,HHT,HHH
  - daher $p(E) = 4/8 = 1/2$

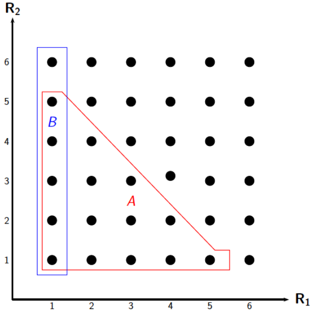
Bedingte Wahrscheinlichkeit

Bedingte Wahrscheinlichkeit

- *Wahrscheinlichkeit eines Ereignis gegeben ein anderes Ereignis*

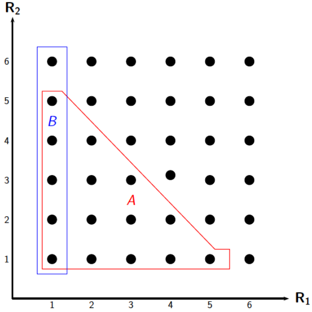


Bedingte Wahrscheinlichkeit



- Ereignis A: $R_1 + R_2 < 7$, Ereignis B: $R_1 = 1$
- $P(A) = \frac{15}{36} = .42$
- $P(B) = \frac{6}{36} = \frac{1}{6}$

Bedingte Wahrscheinlichkeit



- $P(B|A) = 5/15 = 1/3.$

Unabhängigkeit

Unabhängigkeit

- zwei Ereignisse sind unabhängig, wenn Wissen über das eine Ereignis uns nichts über das andee Ereignis sagt*
- formelle Definition:
  - A and B sind unabhängig iff $P(A|B) = P(A)$
  - Für N Ereignisse: N Ereignisse $A_1, \ldots, A_N$ sind gegenseitig unabhängig iff $P(A_i|A_j, A_k, \ldots, A_p) = P(A_i) \, \forall \, i \neq j, k, \ldots, p$ mit $1 \geq i, j, k, \ldots, p \geq N$

Bedingte Unabhängigkeit

Ereignis A und B sind bedingt unabhängig iff
$P(AB|C) = P(A|C)P(B|C)$

Bayes' Regel

Satz der totalen Wahrscheinlichkeit

- wir beginnen mit dem **Satz der totalen Wahrscheinlichkeit**

$$p(A) = p(A \cap B) + p(A \cap B')$$

- dies impliziert, dass das Ereignis A aus der Summe der bedingten Wahrscheinlichkeiten hergeleitet werden kann:

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

  - weiter oben hatten wir schon dies: $P(A \cap B) = P(A|B)P(B)$

Beispiel

- Event A: es regnet am Wahltag
- Event B: ein zufällig ausgewählter Wähler i geht wählen
- Wie berechnen wir $P(B|A)$?

Beispiel

- Event A: es regnet am Wahltag
- Event B: ein zufällig ausgewählter Wähler *i* geht wählen
- Wie berechnen wir $P(B|A)$?

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

Satz von Bayes

Satz von Bayes

- Glaube über die Welt bevor wir neue Belege gesichtet haben: **prior**
- sobald wir neue Belege gesichtet haben: **posterior**, $P(B|A)$
- **Satz von Bayes**:
  Für zwei *binäre* Ereignisse A und B gilt:

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

Satz von Bayes

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

- Verhältnis von
  - Wahrscheinlichkeit das A und B zur gleichen Zeit geschieht und
  - Wahrscheinlichkeit das A und B zur gleichen Zeit geschieht **und** Wahrscheinlichkeit das A und **B'** zur gleichen Zeit geschieht.

Beispiel

Say, we want to know whether observing a recession, R, tells us something about whether the government is of high competence HQ, average competence AQ, or incompetent I. Say, we know that whether a recession occurs is negatively related to the level of competence of the government (from theory or empirics):

- $P(R|HQ) = .1$
- $P(R|AQ) = .2$
- $P(R|I) = .5$

These numbers seem reasonable. When a government is highly competent, the probability of a recession is much lower than when it is incompetent.

Beispiel

Further assume that we have a *flat prior* on what level of competence describes the government absent any other information: $P(HQ) = P(AQ) = P(I) = 1/3$, that is we think that each level of government competence is equally likely before we have seen new evidence.

Now, what's $P(HQ|R)$? Let's apply Bayes:

$$P(HQ|R) = \frac{P(R|HQ)P(HQ)}{P(R|HQ)P(HQ) + P(R|AQ)P(AQ) + P(R|I)P(I)}$$

Let's plug in numbers:

$$P(HQ|R) = \frac{1/10 \times 1/3}{1/10 \times 1/3 + 1/5 \times 1/3 + 1/2 \times 1/3} = 1/8$$

Zufallsvariablen

Zufallsvariablen

- eine **Zufallsvariable** ist eine Funktion
  - die jedem Ereignis im Ereignisraum einen Wert gibt.
  - verbindet die Wahrscheinlichkeit das ein bestimmtes Ergebnis eintritt, das bestimmte Ereignisse formen und die Variablen, die wir benutzen, um die Welt zu beschreiben.

Beispiel

- werfe eine Münze zehn mal, Ereignis: "number of heads"
- "number of heads" ist eine Zufallsvariable, genannt X, die jedem möglichen Ereignis im Ereignisraum einen Wert gibt.
- Ereignisraum: 0 heads, 1 head, 2 heads, . . ., 10 heads über 10 Würfe verteilt

$$X = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

Wahrscheinlichkeitsverteilung einer Zufallsvariable

Wahrscheinlichkeitsverteilung einer Zufallsvariable

- die Wahrscheinlichkeitsverteilung
  - ist definiert über alle möglichen Werte der Zufallsvariable

Beispiel

- werfe eine Münze zehn mal, Ereignis: “number of heads”
- erinnert euch an das Wahrscheinlichkeitsmaß von oben
- beginnen wir mit $X = 0$: 10 Würfe, no “head”

Beispiel

- werfe eine Münze zehn mal, Ereignis: "number of heads"
- erinnert euch an das Wahrscheinlichkeitsmaß von oben
- beginnen wir mit $X = 0$: 10 Würfe, no "head"
  - gibt nur eine Möglichkeit, wie das passieren könnte
- Dann brauchen wir noch die Gesamtzahl aller möglichen Ergebnisse:

Beispiel

- werfe eine Münze zehn mal, Ereignis: “number of heads”
- erinnert euch an das Wahrscheinlichkeitsmaß von oben
- beginnen wir mit $X = 0$: 10 Würfe, no “head”
  - gibt nur eine Möglichkeit, wie das passieren könnte
- Dann brauchen wir noch die Gesamtzahl aller möglichen Ergebnisse: $2^{10} = 1024$

$$p(X = 0) = \frac{1}{1024} = 0.0009765625$$

Beispiel – Fortsetzung

- werfe eine Münze zehn mal, Ereignis: “number of heads”
- erinnert euch an das Wahrscheinlichkeitsmaß von oben
- machen wir weiter mit $X = 1$: 10 Würfe, einmal “head”

Beispiel – Fortsetzung

- werfe eine Münze zehn mal, Ereignis: “number of heads”
- erinnert euch an das Wahrscheinlichkeitsmaß von oben
- machen wir weiter mit $X = 1$: 10 Würfe, einmal “head”
  - kann auf 10 verschiedene Weisen passieren

Beispiel – Fortsetzung

- werfe eine Münze zehn mal, Ereignis: “number of heads”
- erinnert euch an das Wahrscheinlichkeitsmaß von oben
- machen wir weiter mit $X = 1$: 10 Würfe, einmal “head”
  - kann auf 10 verschiedene Weisen passieren
- Gesamtzahl aller möglichen Ergebnisse:

Beispiel – Fortsetzung

- werfe eine Münze zehn mal, Ereignis: “number of heads”
- erinnert euch an das Wahrscheinlichkeitsmaß von oben
- machen wir weiter mit $X = 1$: 10 Würfe, einmal “head”
  - kann auf 10 verschiedene Weisen passieren
- Gesamtzahl aller möglichen Ergebnisse: $2^{10} = 1024$

$$p(X = 1) = \frac{10}{1024} = 0.009765625$$

etc. ...

Beispiel – Fortsetzung

- hier ist die komplette **Wahrscheinlichkeitsverteilung** der Zufallsvariable X, "Anzahl an heads":

X	0	1	2	3	4	5	6
P(X)	.0010	.0098	.0439	.1172	.2051	.2461	.2051

X	7	8	9	10
P(X)	.1172	.0439	.0098	.0010

$\rightarrow$ Sitzungsnotizen

Wahrscheinlichkeitsverteilung einer Zufallsvariable

- die Wahrscheinlichkeitsverteilung einer Zufallsvariable nutzt die Gesetze der Wahrscheinlichkeitstheorie, um uns zu sagen, welche Werte eine Zufallsvariable können wir erwarten. Wie bilden wir solch eine Verteilung
  - aus empirischen Häufigkeitsverteilungen
  - aus theoretischen Wahrscheinlichkeitsverteilungen

Die Verteilungen gibt es in verschiedenen Formen:
+ für diskrete Variablen: Wahrscheinlichkeitsfunktion, probability mass function (PMF) + für stetige Variablen: Dichtefunktion, probability density function (PDF)

Wahrscheinlichkeitsfunktion (PMF)

- gibt eine Wahrscheinlichkeit für jedes Ereignis im Ereignisraum
- gibt eine Wahrscheinlichkeit für jeden Wert (Ausprägung) eine Zufallsvariable

$$P(X = x)$$

- Wahrscheinlichkeitsaxiome finden Anwendung

Gemeinsame Wahrscheinlichkeit von Zufallsvariablen

- $P_{X,Y}(x,y)$
- Wahrscheinlichkeit, dass X und Y die Werte x und y annehmen
- Wahrscheinlichkeitsaxiome finden Anwendung:

1. $\sum_x \sum_y P_{X,Y}(x,y) = 1$
2. $\sum_x P_{X,Y}(x,y) = p_Y(y)$
3. $\sum_y P_{X,Y}(x,y) = p_X(x)$
4. $0 \leq P_{X,Y}(x,y) \leq 1$

Bedingte Wahrscheinlichkeitsfunktion

- $P_{X,Y}(x|y)$
- bedingte Wahrscheinlichkeit, dass X den Wert x annimmt, gegeben dass Y den Wert y annimmt

$$P_{X,Y}(x|y) = \frac{P_{X,Y}(x,y)}{P_Y(y)}$$

Dichtefunktion (PDF)

- gibt eine Wahrscheinlichkeit für jedes Ereignis im Ereignisraum
- gibt eine Wahrscheinlichkeit für jeden Wert (Ausprägung) eine Zufallsvariable

$$f_X(x)$$

* Wahrscheinlichkeitsaxiome finden Anwendung

Unabhängigkeit und bedingte Unabhängigkeit

- Zufallsvariableen X und Y sind unabhängig iff

$$P_{X,Y}(y|x) = P_Y(y)\,\forall x, y$$

- Zufallsvariableen X und Y sind unabhängig gegeben Ereignis A iff

$$P_{X,Y}(x, y|A) = P_{X|A}(x|A)P_{Y|A}(y|A)\,\forall x, y$$

Erwartungswert einer Zufallsvariable

Erwartungswert einer Zufallsvariable

$$E[X] = \sum_{i=1}^{N} x_i p_X(x_i) = \overline{X}$$

Verteilung der Zufallsvariable

- Erster Zentralmoment

$$E[(x-\overline{x})^1] = 0.$$

Verteilung der Zufallsvariable

- Erster Zentralmoment

$$E[(x - \overline{x})^1] = 0.$$

- Zweiter Zentralmoment: Varianz

$$E[(x - \overline{x})^2] = E[x^2] - E[x]^2$$

Verteilung der Zufallsvariable

- Erster Zentralmoment

$$E[(x-\overline{x})^1] = 0.$$

- Zweiter Zentralmoment: Varianz

$$E[(x-\overline{x})^2] = E[x^2] - E[x]^2$$

- Dritter Zentralmoment: Schiefegrad
- Vierter Zentralmoment: Wölbung

Beispiel: Binomische Wahrscheinlichkeitsverteilung

Binomische Wahrscheinlichkeitsverteilung

- ist die Verteilung für jede Zufallsvariable, die auf binären Ergebnissen beruht.

Binomische Wahrscheinlichkeitsverteilung

- ist die Verteilung für jede Zufallsvariable, die auf binären Ergebnissen beruht.
- wie kommts?

Binomische Wahrscheinlichkeitsverteilung

- ist die Verteilung für jede Zufallsvariable, die auf binären Ergebnissen beruht.
- wie kommts?
- binäre Ergebnisse folgen der sogenannten Bernoulli-Wahrscheinlichkeitsverteilung:

Binomische Wahrscheinlichkeitsverteilung

- ist die Verteilung für jede Zufallsvariable, die auf binären Ergebnissen beruht.
- wie kommts?
- binäre Ergebnisse folgen der sogenannten Bernoulli-Wahrscheinlichkeitsverteilung:

$$P(x) = p \quad \text{and} \quad P(x') = 1 - p.$$

- die Zufallsvariable ist dann die Summe der beobachteten Bernoulli-Ereignisse:

$$Y = \sum_{i=1}^{n} X_i.$$

Binomische Wahrscheinlichkeitsverteilung

Lass uns das mal herleiten:

- wie viele Möglichkeiten gibt es um y Ereignisse aus, sagen wir, 3 Versuchen zu bekommen?

Binomische Wahrscheinlichkeitsverteilung

Lass uns das mal herleiten:

- wie viele Möglichkeiten gibt es um y Ereignisse aus, sagen wir, 3 Versuchen zu bekommen? Da hilft uns der **binomische Koeffizient**:

$$\binom{3}{y} = \frac{3!}{y!(3-y)!}$$

- was kriegen wir dann für $Y = 2$ aus 3 Versuchen:

$$\binom{3}{2} = \frac{3!}{2!(3-2)!} = \frac{3 \times 2 \times 1}{2 \times 1 \times 1} = 3$$

- erinnert euch an das Beispiel der drei Münzwürfe, die

Binomische Wahrscheinlichkeitsverteilung

P(Ergebnis)	Ergebnis
$p \times p \times (1-p)$	HHT
$p \times (1-p) \times p$	HTH
$(1-p) \times p \times p$	THH

Binomische Wahrscheinlichkeitsverteilung

P(Ergebnis)	Ergebnis
$p \times p \times (1-p)$	HHT
$p \times (1-p) \times p$	HTH
$(1-p) \times p \times p$	THH

$$P(Y=2)\binom{3}{2}p^2(1-p)^{3-2} = 3 \times .25 \times .5 = .375 = 3/8.$$

Binomische Wahrscheinlichkeitsverteilung

Hier ist die allgemeine **binomische Wahrscheinlichkeitsverteilung**:

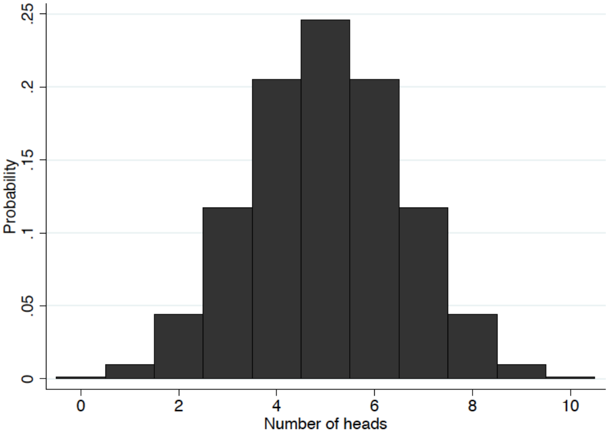
$$p(Y = y|n, p) = \binom{n}{y} p^y (1-p)^{n-y}.$$

Binomische Wahrscheinlichkeitsverteilung

X	0	1	2	3	4	5	6
P(X)	.0010	.0098	.0439	.1172	.2051	.2461	.2051

X	7	8	9	10
P(X)	.1172	.0439	.0098	.0010

Binomische Wahrscheinlichkeitsverteilung



Was solltet ihr von heute mitnehmen?

1. Verstehe, das ein Regressionskoeffizient von jeglichen quantitativen Daten berechnet werden kann. Ob dieser inhaltliche Bedeutung hat, ist eine Frage unserer Theorie.
2. Kenne die Grundbegriffe der Wahrscheinlichkeitsrechnung, sei fähig eine informelle Definition des Wahrscheinlichkeitsmaß zu geben
3. Verstehe wie unsere Beobachtung der Welt, unser Zufallsexperiment, sich in Zufallsvariablen mit Wahrscheinlichkeitsverteilungen übersetzt.

Was gibt's nächste Woche

1. Inferenzstatistik: wie ziehen wir Rückschlüsse aus unseren Daten mit Hilfe von empirisschen oder theoretischen Wahrscheinlichkeitsverteilungen.
2. Welche Annahmen (mehrheitlich über die Beschaffenheit unserer Daten) müssen zutreffen, damit wir valide Statistiken erhalten und Inferenzstatistik anwenden können (Beispiele: t-test, Regression)