Statistik I Januar bis Februar in Schnelldurchlauf
(Grundlagen Wahrscheinlichkeit)

Um ökonomische Modelle, die Unsicherheit berücksichtigen zu verstehen, ist eine Kenntnis grundlegender wahrscheinlichkeitstheoretischer Konzepte notwendig

- Ökonomische Modelle unter Sicherheit sind oft realitätsfern und wenig hilfreich
- Individuen treffen Investitions- und Finanzierungsentscheidungen (auch Ihre Investition in Humankapital) unter Unsicherheit über zukünftige Kosten und Auszahlungen
- Die Bewertung von unsicheren zukünftigen Auszahlungen ist eines der großen Themen der wirtschaftswissenschaftlichen Forschung
- Die sowohl wissenschaftlich interessantesten als auch praxisrelevantesten ökonomischen Modelle sind unter Unsicherheit formuliert und nutzen grundlegene Konzepte der Wahrscheinlichkeitstheorie
- Ohne die Kenntnis des Instrumentarium der Wahrscheinlichkeitstheorie ist ein anspruchsvolles wirtschaftswissenschaftliches Studium kaum möglich

Klassische und Bayesianische Statistiker verbinden zwei unterschiedliche Sichtweisen mit dem Begriff „Wahrscheinlichkeit“

Wahrscheinlichkeit

Klassische/frequentistische Sicht
Konzeptioneller Rahmen:
Zufallsexperiment:
- mögliche Ergebnisse des Experiments bekannt
- Experiment kann wiederholt werden (unter gleichen Randbedingungen)
- Ergebnis eines konkreten Versuchs nicht mit Sicherheit vorhersehbar

relative Häufigkeit des Eintritts eines Ereignisses im wiederholten Experiment
→ Wahrscheinlichkeit für das Ereignis (wenn Anzahl der Wiederholungen groß wird)

Bayesianische Sicht
subjektive Wahrscheinlichkeit

Wahrscheinlichkeit als Quantifizierung (persönlicher) Einschätzungen (Meinungen, Hypothesen).

„System von Wahrscheinlichkeiten“

Eintritt von Ereignissen kann zu Revisionen der Wahrscheinlichkeiten führen

a-priori Wahrscheinlichkeiten für Hypothesen
⇒ Ereignis ⇒
a-posteriori Wahrscheinlichkeiten für Hypothesen

Auf der Kolmogorov‘schen Axiomatik baut sowohl die klassische als auch die Bayesianische Statistik auf

Für alle Ereignisse in der Ereignismenge ***E*** (Menge ***E*** enthält aus den Elementar-ereignissen konstruierte Ereignisse) muß ein Wahrscheinlichkeitsmaß *P()* die folgenden Eigenschaften aufweisen:

Ein beliebiges Ereignis aus E

1. Axiom: $0 \leq P(A)$

Das sichere Ereignis (also der Ereignisraum, die Menge, die alle möglichen Ergebnisse des Experiments enthält)

2. Axiom: $P(S) = 1$

3. Axiom: $P(A_1 \cup A_2 \cup A_3 \ldots) = P(A_1) + P(A_2) + P(A_3)$
falls sich $A_1, A_2, \ldots$ paarweise ausschließen, d.h. $A_i \cap A$
wenn $i \neq j$

Ereignisse aus E

Lies: Wenn sich Ereignisse A und B ausschließen, ist die Wahrscheinlichkeit, daß Ereignis A **oder** B eintreten die Summe der Wahrscheinlichkeiten der beiden Ereignisse

Der Additionssatz folgt aus der Kolmogorov'schen Axiomatik

Der Additionssatz folgt direkt aus dem dritten Axiom

Für beliebige Ereignisse A und B gilt

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Wahrscheinlichkeit, daß **sowohl** Ereignis A **als auch** Ereignis B eintreten (**Und-Verknüpfung**)

Wahrscheinlichkeit, daß Ereignis A oder Ereignis B eintritt (oder auch beide) **Oder-Verknüpfung**

Das Konzept der statistischen Wahrscheinlichkeit basiert auf der Konvergenz von relativen Häufigkeiten wenn die Zahl der Wiederholungen eines Zufallsexperiments groß wird

- Zufallsexperiment wird wiederholt durchgeführt
- Der Eintritt oder Nicht-Eintritt eines Ereignisses in jedem Versuch wird notiert
- Es wird die absolute und die relative Häufigkeit des Eintritts des Ereignisses in den wiederholten Versuchen berechnet
- Der Wert, zu dem die relative Häufigkeit konvergiert, ist die (statistische) Wahrscheinlichkeit für das Ereignis

Dies ist die klassische/frequentistische Definition von Wahrscheinlichkeit

Anwendbarkeit in den Wirtschaftswissenschaften auf den ersten Blick problematisch

Dennoch der konzeptionelle Rahmen in dem Tests von ökonomischen Modellen und die Schätzung von Modellparametern durchgeführt werden!

Laplace-Wahrscheinlichkeiten können berechnet werden, wenn der Ereignisraum endlich viele und gleich wahrscheinliche Elementarereignisse enthält

- Laplace-Wahrscheinlichkeiten sind Teil der klassischen Statistik

- Wenn der Ereignisraum endlich ist und jedes Elementarereignis gleich wahrscheinlich, sind Wahrscheinlichkeiten auch ohne die Wiederholung des Zufallsexperiments zu berechnen

- Vorgehen: Zähle Anzahl (*m*) Elementarereignisse (welche Ergebnisse des Experiments sind möglich).
  Prüfe, welche Elementarereignisse dazu führen, daß ein beliebiges Ereignis A eintritt und zähle die Anzahl dieser Elementarereignisse (*g*)

- Wahrscheinlichkeit für Ereignis *A* errechnet sich aus dem Verhältnis von *g* und *m*

Oft ist die Zahl der Elementarereignisse sehr groß (aber endlich). Dann hilft die Kombinatorik bei der Berechnung von Laplace-Wahrscheinlichkeiten

Bei der Berechnung der Zahl der Elementarereignisse helfen die folgenden Formeln

Ausgangspunkt ist ein Urnenmodell: Aus einer Menge von n Elementen sollen r zufällig ausgewählt werden. Das erweiterte Multiplikationsprinzip liefert:

$$n^r$$

Anzahl mögliche Stichproben mit Zurücklegen (mit Berücksichtigung der Anordnung)

$$n \cdot (n-1) \cdot (n-2) \cdot (n-3) \cdot \ldots \cdot (n-r+1)$$

Anzahl mögliche Stichproben ohne Zurücklegen mit Berücksichtigung der Anordnung

$$\frac{n!}{(n-r)!r!} \equiv \binom{n}{r}$$

Anzahl mögliche Stichproben ohne Zurücklegen (ohne Berücksichtigung der Anordnung)

$$n! = n \cdot (n-1) \cdot (n-2) \cdot (n-3) \ldots 2 \cdot 1$$

Anzahl der Möglichkeiten, eine Menge von n Elementen anzuordnen

Bedingte Wahrscheinlichkeiten sind auf einem durch die Bedingung eingeschränkten Ereignisraum definiert. Dies ist NICHT der Satz von Bayes!

Wahrscheinlichkeit für den Eintritt von Ereignis A **und** B

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

Lies: Wahrscheinlichkeit von Ereignis A, gegeben, daß Ereignis B eintritt

Die Wahrscheinlichkeit für das Ereignis, auf das bedingt wird, muß von Null verschieden sein

Bedingte Wahrscheinlichkeit in der klassisch/frequentistischen Sicht: Der Ereignisraum wird auf solche Elementarereignisse eingeschränkt, die den Eintritt von Ereignis B implizieren

Aus der Definition der bedingten Wahrscheinlichkeit folgt der Multiplikationssatz

Manchmal liegen Informationen über bedingte Wahrscheinlichkeiten vor, aber keine Informationen über gemeinsame Wahrscheinlichkeiten. Dann kann die gemeinsame Wahrscheinlichkeit wie folgt berechnet werden:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

Multiplikationssatz für beliebige Ereignisse A und B

Die Wahrscheinlichkeit für den Eintritt von Ereignis A **und** Ereignis B kann auf zwei Weisen ermittelt werden. Die Kenntnis der bedingten Wahrscheinlichkeiten ist hierfür notwendig

Sind zwei Ereignisse unabhängig, so ist die gemeinsame Wahrscheinlichkeit das Produkt der Wahrscheinlichkeiten der beiden Ereignisse

Bei Unabhängigkeit der Ereignisse A und B verändert der Eintritt oder Nichteintritt von Ereignis B **nicht** die Wahrscheinlichkeit für Ereignis A (und vice versa)

$$P(A|B) = P(A|\bar{B}) = P(A) \Leftrightarrow P(B|A) = P(B|\bar{A}) = P(B)$$

Die Berechnung der gemeinsamen Wahrscheinlichkeit mit der allgemeinen Formel ....

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

.... vereinfacht sich daher zu

$$P(A \cap B) = P(A) \cdot P(B)$$ **Multiplikationssatz für unabhängige Ereignisse**

Die totale Wahrscheinlichkeit drückt die Wahrscheinlichkeit für ein Ereignis als Funktion von bedingten Wahrscheinlichkeiten aus. Sieht umständlich aus...

Wir teilen den Ereignisraum in n sich ausschließende Ereignisse A_1, $A_{2.}$....,A_n auf.
Wir können (auf den ersten Blick umständliche) alternative Möglichkeit nutzen, die Wahrscheinlichkeit für ein Ereignis E berechnen
(E bezeichnet hier ein beliebiges Ereignis, nicht die Ergebnismenge)

Dazu nutzen wir das dritte Axiom (paarweise sich ausschließende Ereignisse)

Diese Ereignisse schließen sich paarweise aus: Entweder tritt E gemeinsam mit A_1 ein oder mit A_1 etc.

$$\begin{aligned} P(E) &= P(E \cap A_1) + P(E \cap A_2) + \ldots + P(E \cap A_n) \\ &= \sum_{i=1}^{n} P(E \cap A_i) \end{aligned}$$

Wir können auch die bedingten Wahrscheinlichkeiten verwenden....

$$\begin{aligned} P(E) &= P(E|A_1) \cdot P(A_1) + P(E|A_2) \cdot P(A_2) + \ldots P(E|A_n) \cdot P(A_n) \\ &= \sum_{i=1}^{n} P(E|A_i) \cdot P(A_i) \end{aligned}$$

Totale Wahrscheinlichkeit für Ereignis E

Für einen klassischen Statistiker ist es nur eine komplizierte Form, eine bedingte Wahrscheinlichkeit zu schreiben, für Bayesianer aber viel mehr: Satz von Bayes

Die übliche Definition einer bedingten Wahrscheinlichkeit

$$P(A_k|E) = \frac{P(A_k \cap E)}{P(E)}$$

Bayesianer unterscheiden: A_i sind Hypothesen, E ist ein Ereignis (Daten)

A-posteriori Wahrscheinlichkeit für die Hypothese A_i

A-priori Wahrscheinlichkeit für die Hypothese A_i

$$P(A_k|E) = \frac{P(E|A_k) \cdot P(A_k)}{\sum_{i=1}^{n} P(E|A_i) \cdot P(A_i)}$$

Im Nenner steht die totale Wahrscheinlichkeit für Ereignis E

Lies: Wie wahrscheinlich ist es, daß Ereignis E eintritt, gegeben, daß Hypothese A_i wahr ist

Für Bayesianer führt der Eintritt von Ereignissen zu einer Revision der Wahrscheinlichkeiten für das Hypothesen-System
Der Prozess ist dynamisch: **Bayesianisches Lernen**

Eine Zufallsvariable ist weder zufällig noch eine Variable

- Eine Zufallsvariable ist eine Abbildung aus dem Ereignisraum (Mengen von Ereignissen) in die Menge der reelen Zahlen (oft ist der Ereignisraum selbst schon die Menge der reelen Zahlen)

- Die Struktur des Wahrscheinlichkeitsraumes muß erhalten bleiben. Allerdings sind mit reelen Zahlen Rechenoperationen möglich, die mit Ereignissen (Mengen) nicht möglich waren

- Wir geben nicht mehr Wahrscheinlichkeiten für allgemeine Ereignisse an, sondern dafür, daß die Zufallsvariable Werte in einem vorgegebenen Intervall annimmt

- Ist der Wertevorrat einer Zufallsvariable endlich oder abzählbar unendlich, dann sprechen wir von einer diskreten Zufallsvariable

- Besteht der Wertevorrat nur aus 0 oder 1, so handelt es sich um eine binäre Zufallsvariable

- Ist der Wertevorrat der Zufallsvariable ein Intervall aus der rellen Achse (evtl. die Menge der reelen Zahlen selbst) sprechen wir von einer stetigen Zufallsvariablen

- Für stetige Zufallsvariable ist die Wahrscheinlichkeit dafür, daß die Zufallsvariable einen Wert x (beliebige reele Zahl) annimmt gleich Null! Nur die Wahrscheinlichkeit dafür, daß die Zufallsvariable Werte in einem Intervall annimmt kann ungleich Null sein

In ökonomischen Modellen mit Unsicherheit (über zukünftige Ereignisse) werden ökonomische Variable als Zufallsvariable aufgefaßt

- Wir werden Daten als Realisationen von Zufallsvariablen auffassen

- Der grundlegende Kontext des Zufallsexperiments wird aufrecht erhalten (wiederholte Durchführung unter kontrollierten Bedingungen möglich)

- Wir werden Wahrscheinlichkeitsverteilungen von Zufallsvariablen kennenlernen

- Dies ermöglicht es, Wahrscheinlichkeiten dafür anzugeben, daß ökonomische Variable bestimmte Werte annehmen

- Dies ermöglicht zum einen Prognosen, zum anderen Tests von Hypothesen, die aus ökonomischen Modellen abgeleitet werden

- Ökonomische Modelle unter Unsicherheit enthalten Zufallsvariable, für die Wahrscheinlichkeitsverteilungen angenommen werden

- Daten (d.h. Realisationen von Zufallsvariablen) werden zur Schätzung von Parametern verwendet, welche die zu Grunde liegenden Wahrscheinlichkeitsverteilungen beschreiben (d.h. wir schätzen Parameter von ökonomischen Modellen unter Verwendung von Daten)