

Rückblick auf das letzte Semester und Plan für das kommende Semester

Dr. Mariana Nold

Institut für Soziologie, Fakultät für Sozial- und Verhaltenswissenschaften, Lehrstuhl für empirische Sozialforschung und Sozialstrukturanalyse

19. Juni 2017



Organisatorisches

Organisatorisches

Organisatorisches

Die beschreibende Statistik und ihre Intention

Die deskriptive Statistik hat die Intention einen gegebenen Datensatz zu beschreiben, dazu nutzt sie

- Lagemaße wie Mittelwert, Modus, Median und andere empirische Quantile
- Streuungsmaße wie empirische Varianz oder Interquartilsabstand
- Kreuztabellen und bedingte relative Häufigkeiten, insbesondere die Spaltenprozentuierung
- Grafiken wie die empirische Verteilungsfunktion, Boxplots, Histogramme oder (gruppierte) Balkendiagramme

Die schließende Statistik und ihre Bedeutung

Was macht die induktive Statistik aus?

Antwort von Yudi Pawitan

Uncertainty is pervasive in problems that deal with the real world, but statistics is the only branch of science that puts systematic effort into dealing with uncertainty. Statistics is suited to problems with inherent uncertainty due to limited information, it does not aim to remove uncertainty, but in many cases it merely quantifies it, uncertainty can remain even after analysis is finished.

Ich möchte Ihnen in diesem Semester einen Einblick in methodischen Ansätze (systematic efforts) geben, die die klassische Inferenz bietet.

Beispiel PISA

Warum sind viele wissenschaftliche Fragestellungen von Unsicherheit durchdrungen?

- Wir wollen basierend auf den PISA-Daten eine Antwort auf die Frage finden, ob die Leseleistung von 15-jährigen Mädchen in Deutschland besser ist als die von 15-jährigen Jungen.
- Unsere Stichprobe enthält nicht die Information über die Leseleistung aller 15-Jährigen in Deutschland.
- Wenn wir eine Aussage über die Grundgesamtheit machen, kann diese Aussage falsch sein. Die Wissenschaftlichkeit der Aussage entsteht dadurch, dass wir ihre Unsicherheit quantifizieren.

Beispiel Unfalltod

Warum sind viele wissenschaftliche Fragestellungen von Unsicherheit durchdrungen?

- Wir wollen eine Aussage darüber machen, ob die Wahrscheinlichkeit für den Unfalltod auf deutschen Autobahnen gestiegen ist.
- Die Zahl der Unfalltoten ist von 358 im Jahr 2012 auf 387 im Jahr 2013 gestiegen. Ist das dann eine “wirkliche” Veränderung?
- Die Statistik fragt: Hat sich etwas an dem stochastischen Mechanismus geändert der diese Unfalltoden hervorbringt.
- Um wie viel muss die Zahl steigen bzw. sinken, damit man davon ausgeht, dass sich dieser Mechanismus verändert hat?
- Diese Frage kann nicht willkürlich beantwortet werden. Es muss klar sein, wann man von einer signifikanten Veränderung spricht.

Inferenzkonzepte

- Durch welche Legitimation ist es erlaubt davon zu sprechen, dass Mädchen besser lesen als Jungen oder dass die Anzahl der Unfalltoden gestiegen ist?
- Wenn ich die entsprechenden statistischen Tests rechne, dann ist der Unterschied in der Leseleistung signifikant, der Unterschied in der Zahl der Unfalltoden nicht.
- Was bedeutet das eigentlich?
- Unter einem Inferenzkonzept versteht man ein statistisches Konzept, das die Schlussfolgerung von Beobachtungen auf Hypothesen rechtfertigt.

Klassische Inferenz

- Wie werden innerhalb der Statistik Schlüsse von Beobachtungen auf Hypothesen gerechtfertigt?
- Angesichts der Vielfalt wissenschaftstheoretischer und philosophischer Ansätze zur Erklärung des empirischen Forschungsansatzes ist es nicht verwunderlich, dass es auf diese Frage innerhalb der Statistik mehrere Antworten gibt und nicht eine einzige, in sich geschlossene statistische Inferenztheorie existiert.
- Die drei bekanntesten Inferenztheorien sind die Likelihood-Inferenz, die Bayes-Inferenz und die klassische Inferenz.
- Wir beschäftigen uns in diesem Semester mit Ansätzen aus der klassischen Inferenz.

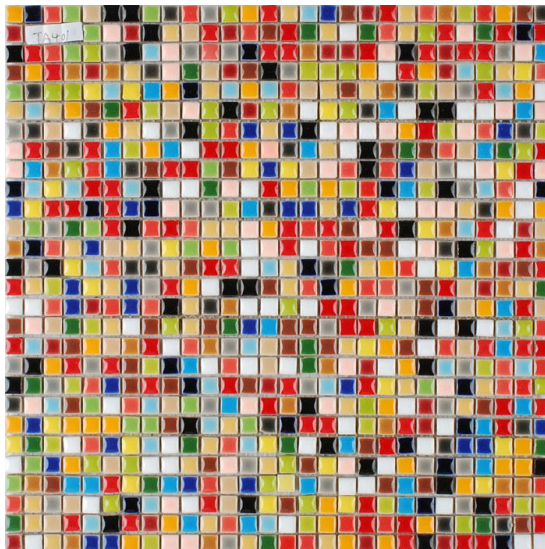
Empirische Forschung ist vergleichbar damit ein Mosaik zu restaurieren



Restauration: Was es zu beachten gilt

- 1 Transparenz ist wichtig. Man muss sagen, welche Teile man zu Beginn hat und durch welche Annahmen und Methoden man zu dem Gesamtbild kommt.
- 2 Das Bild das, nach der Restauration entsteht, kann (in Teilen) fehlerhaft sein. Es ist mit Unsicherheit behaftet.
- 3 Je weniger Teile man zu Beginn hat bzw. je komplizierter das Bild ist, desto größer ist die Unsicherheit über das Ergebnis.
- 4 Es gibt viele Möglichkeiten zu dem Ergebnis zu kommen, es gibt nicht den einen richtigen Weg.
- 5 Konstruktiv-kritisches Nachfragen daher ist sinnvoll. Man braucht den Blick von unterschiedlichen Personen (aus verschiedenen Disziplinen) um die Güte des Bildes beurteilen zu können.

Gibt es überhaupt ein Muster?



Können wir aus den gegebenen Informationen etwas schließen?



Wie viel darf hier fehlen, damit man eine Chance hat das Muster zu finden?



Welche Teile dürfen fehlen, damit man eine Chance hat das Muster zu finden?



Werkzeugkiste für die Restauration

Was gehört rein:

- 1 Kreativität und Fingerspitzengefühl
- 2 Sachwissen über den Hintergrund der Daten (der Teile, die man hat)
- 3 Grundlagen der Methoden der empirischen Sozialforschung
- 4 Grundlagen der deskriptiven Statistik, eine gute Beschreibung der Daten ist sehr viel wert
- 5 Grundlagen der Inferenzstatistik, insbesondere der Test- und Schätztheorie
- 6 Die Fähigkeit Statistik-Software zu nutzen

Zufallsvorgang

Definition: Zufallsvorgang

Ein Zufallsvorgang führt zu einem von mehreren, sich gegenseitig ausschließenden Ergebnissen. Es ist vor der Durchführung ungewiss, welches Ergebnis tatsächlich eintreten wird.

- Die Ziehung einer Stichprobe ist ein Beispiel für einen Zufallsvorgang.
 - ▶ Befragung von Studierenden nach gezahlter monatlicher Miete
- Viele Prozesse die wir beobachten können sind Zufallsvorgänge.
 - ▶ Beispiel: Wie viel Personen nehmen heute an der Vorlesung teil?

Zufallsexperiment

Definition: Zufallsexperiment

Bei Beobachtungsstudien, Befragungen oder allgemeinen Stichprobenerhebungen sind im Gegensatz zu Experimenten die Rahmenbedingungen i. a. nicht kontrollierbar bzw. bekannt. Man spricht von einem Zufallsexperiment, wenn ein Zufallsvorgang unter kontrollierten Bedingungen abläuft und somit unter gleichen Bedingungen wiederholbar ist.

Unter Rahmenbedingungen bzw. Bedingungen sind alle Umstände zu verstehen, die einen Einfluss auf den Vorgang haben. Das Würfeln ist ein Beispiel für ein Zufallsexperiment.

Chaotisches System

Definition: Chaotisches System

Oft beschreiben wir unerklärte Phänomene mit Hilfe von probabilistischen Gesetzen, obwohl man diese Phänomene rein deterministisch beschreiben könnte. Es ist eine Möglichkeit mit Prozessen umzugehen, die zu komplex sind, um sie einfach vorherzusehen. Es ist dann für uns ungewiss, welches Ereignis tatsächlich eintreten wird, auch wenn man es theoretisch vorhersagen könnte.

- Wir unterscheiden nicht, ob ein Prozess eigentlich deterministisch ist und einfach zu komplex um den Ausgang absehen zu können oder ob es ein wirklicher stochastischer Prozess ist.
- Pendelbewegungen als auch das Würfeln können als chaotische Systeme verstanden werden.

Zufallsvorgang

Der Begriff Zufallsvorgang ist daher ein Überbegriff für Zufallsexperimente und chaotische Systeme. Charakteristisch für den Zufallsvorgang sind zwei Eigenschaften

- 1 Man kennt im Vorfeld bereits die möglichen Ausgänge, wobei unbekannt ist, welcher Eintritt.
- 2 Es hängt "vom Zufall" ab, welchen Ausgang man beobachtet.

Befragung nach Taschengeld als Zufallsvorgang

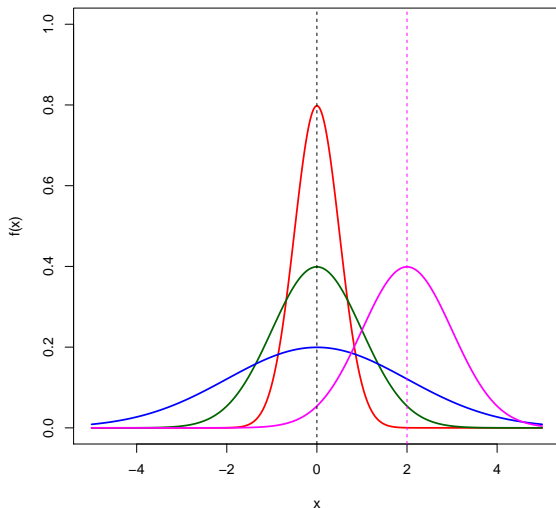
Im Folgenden wollen wir die Normalverteilung nutzen als probabilistisches Gesetz um die fiktive Befragung von Kindern in einer Schule nach der Höhe ihres Taschengeldes als Zufallsvorgang zu beschreiben.

Rückblick: Die Normalverteilung

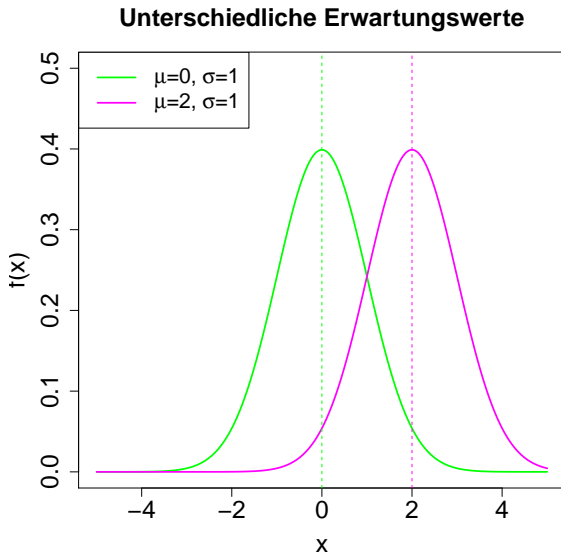
Die Normalverteilung ist ein erste Beispiel für ein probabilistisches Gesetz.
(Erklären was eine Verteilung ist)

Die unterschiedlichen Formen der Normalverteilung

Dichten unterschiedlicher Normalverteilungen

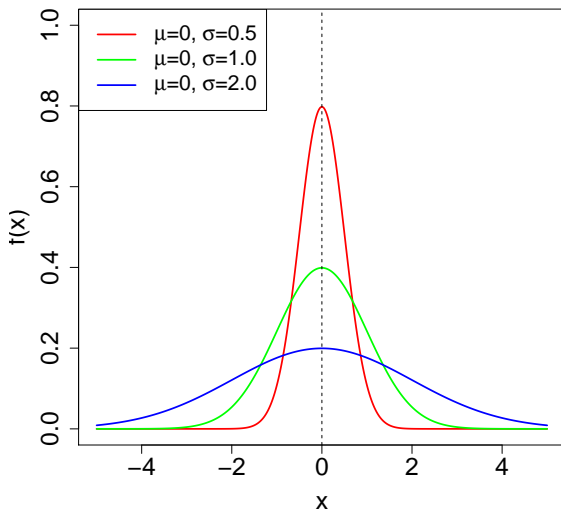


Was bewirkt eine Veränderung des Mittelwerts?



Was bewirkt eine Veränderung der Standardabweichung?

Unterschiedliche Standardabweichungen



Bekommen Jungen mehr Taschengeld als Mädchen?

Bitte lesen sie die beiden Artikel

① Statistik und Wahrheit - Hauptsache spektakulär

<http://www.spiegel.de/wissenschaft/mensch/taschengeld-gibt-es-wirklich-einen-gender-pay-gap-a-1100475.html>
Erscheinungsdatum: 27.7.2016

② Umfrage: Jungen bekommen mehr Taschengeld als Mädchen

<http://www.spiegel.de/lebenundlernen/schule/jungen-bekommen-mehr-taschengeld-als-maedchen-a-1161915.html>
Erscheinungsdatum: 8.8.2017

Machen Sie sich ein paar Stichpunkte zu dem Ergebnis,
dass der zweite Artikel Ihnen Nahe legt und kommentieren Sie es kritisch.