



03.06.2020

Statistische Tests und fachliche Kritik an Testverfahren einer vorläufigen Auswertung von Christian Drosten & Co.

Anlass

Das Team um Christian Drosten hat eine zweite Version der Preprint-Publikation zur Viruslast von SARS-CoV-2 bei Kindern und Erwachsenen auf den Server der Charité hochgeladen (siehe Primärquelle). In den vergangenen Tagen löste die erste Version dieser vorläufigen Studie eine Welle medialer Berichterstattung aus. Ein Teil der Berichte griff auch die fachliche Kritik einiger Statistiker auf, die Bedenken hinsichtlich der verwendeten Methoden des statistischen Testens äußerten. Die Arbeitsgruppe um Christian Drosten überarbeitete daraufhin die Analyse.

Dieses Fact Sheet soll einen Einblick in die Hintergründe des statistischen Testens generell und in Bezug auf die Preprint-Publikation geben.

Übersicht

Grundlagen statistischer Tests	1
Anwendungsfall: Preprint-Studie der Gruppe um Christian Drosten	2
Statistische Tests in der Praxis	2
Interpretation von statistischen Tests	2
Fachliche Kritik an der Preprint-Studie der Gruppe um Christian Drosten in Bezug auf die Testgüte	3
Weitere Recherchequellen	4

Grundlagen statistischer Tests

- ▶ Statistische Tests können dabei helfen, Gesetzmäßigkeiten in Daten zu entdecken.
- ▶ Reale Prozesse folgen jedoch keinen strikten mathematischen Formeln.
- ▶ Jede Beobachtung, also jeder Datenpunkt, weicht allerdings um einen zufälligen Faktor von dieser Gesetzmäßigkeit ab.



- ▶ Für eine bestimmte Datenmenge kann man daher nur entscheiden, ob es plausibel ist, dass sie der untersuchten Gesetzmäßigkeit folgen, nicht ob sie ihr tatsächlich folgen.
- ▶ Statistische Tests helfen dabei, für gegebene Datenmengen zu entscheiden, welche von zwei möglichen Gesetzmäßigkeiten den Daten wahrscheinlicher zugrunde liegt.

Anwendungsfall: Preprint-Studie der Gruppe um Christian Drosten

- ▶ Die Studie der Arbeitsgruppe um Christian Drosten untersucht die Frage, ob die Viruslast bei Kindern anders ist als bei Erwachsenen.
- ▶ Es gibt also zwei mögliche Gesetzmäßigkeiten (Hypothesen):
 - Nullhypothese (H_0): Es liegt kein Unterschied in der Viruslast zwischen Kindern und Erwachsenen vor.
 - Alternativhypothese (H_1): Es gibt einen Unterschied zwischen der Viruslast von Kindern und Erwachsenen.
- ▶ Aufgrund der oben genannten zufälligen Schwankungen, die in der realen Welt vorkommen, kann nur die Aussage getroffen werden, ob die Nullhypothese in Bezug auf die vorliegenden Daten unwahrscheinlich ist; diese Aussage können statistische Tests treffen.

Statistische Tests in der Praxis

- ▶ Grundvorgehen ist immer dasselbe:
 - Aus Daten wird eine Maßzahl errechnet, die sogenannte Teststatistik
 - Als Vergleichsgröße dient die Wahrscheinlichkeitsverteilung dieser Maßzahl
 - Die errechnete Teststatistik wird mit der Wahrscheinlichkeitsverteilung dieser Maßzahl verglichen. Bezogen auf die Viruslast: Der errechnete Mittelwert für die Viruslast bei Kindern wird verglichen mit der Verteilung aller Viruslast-Werte
 - ein Beispiel für eine Wahrscheinlichkeitsverteilung ist die Normalverteilung, oder auch Glockenkurve, bei der die Maßzahlen in einem mittleren Segment häufiger vorkommen und damit wahrscheinlicher sind als Maßzahlen an den äußeren Rändern
- ▶ Ist der beobachtete Wert unter der Verteilungsannahme hinreichend unwahrscheinlich – zum Beispiel viel größer als erwartet –, verwirft man die Nullhypothese und nimmt die Alternativhypothese an; im aktuellen Beispiel: Es existiert ein signifikanter Unterschied zwischen der Viruslast von Erwachsenen und Kindern.
- ▶ Ist die Teststatistik nicht hinreichend (signifikant) unwahrscheinlich, kann das Zutreffen der Nullhypothese nicht ausgeschlossen werden; das führt aber nicht gleichzeitig dazu, dass die Nullhypothese angenommen werden kann.

Interpretation von statistischen Tests

- ▶ wie angedeutet, sind die beiden Ausgänge – Nullhypothese verwerfen oder nicht verwerfen – in ihrer Interpretation nicht gleichberechtigt; das liegt am Prinzip der Tests selbst
- ▶ Grundsätzlich können beim Testen zwei Fehler entstehen:
 - Man kann die Nullhypothese verwerfen, obwohl sie eigentlich wahr ist (Fehler erster Art).



- Man kann die Nullhypothese nicht verwerfen, obwohl die Alternative eigentlich wahr ist (Fehler zweiter Art).
- ▶ beide Fehler liegen im Zufallsfaktor der realen Welt begründet: Die Gesetzmäßigkeiten sind nicht direkt ersichtlich, also weiß man auch nicht, wenn man von der Wahrheit abweicht, da sie nicht bekannt ist.
- ▶ Statistische Tests sind so konstruiert, dass die Irrtumswahrscheinlichkeit für den Fehler erster Art festgelegt werden kann im sogenannten Signifikanzniveau.
 - ein üblicher Wert ist dabei fünf Prozent
 - unter der Annahme, dass die Nullhypothese wahr ist, wird also in fünf Prozent der Fälle die Nullhypothese fälschlicherweise verworfen.
- ▶ Es ist nicht möglich, für den Fehler zweiter Art ebenfalls eine solche Grenze festzulegen, da ein Test nicht für zwei gewählte Werte optimiert werden kann.
- ▶ Aus diesem Grund müssen die beiden Ausgänge bei einem Test unterschiedlich bewertet werden:
 - wird die Nullhypothese verworfen, kann die Alternativhypothese mit einer Fehlerwahrscheinlichkeit von den festgelegten fünf Prozent angenommen werden.
 - kann die Nullhypothese nicht verworfen werden, kann umgekehrt keine Aussage über die Wahrscheinlichkeit getroffen werden, dass die Nullhypothese wirklich wahr ist, da der Fehler zweiter Art unbekannt ist.
- ▶ Insbesondere zwei Effekte können dabei eine Rolle spielen: Die Stichprobe ist zu klein oder der gewählte Test hat nicht ausreichend Güte.
- ▶ Eine bestimmte Fragestellung (ein Hypothesenpaar) kann in der Regel mit verschiedenen Tests untersucht werden. Der Fehler erster Art kann für alle Tests gleich eingestellt werden – zum Beispiel fünf Prozent. Die Unterschiede liegen in der Größe des Fehlers zweiter Art.
- ▶ Statt von dem Fehler zweiter Art wird auch von der Güte (Power) eines Tests gesprochen. Die Güte ist definiert als 100 Prozent minus des Fehlers zweiter Art. Kann die Nullhypothese nicht verworfen werden, kann dies daran liegen, dass der Test nicht ausreichend Güte hat und ein geeigneterer Test einen signifikanten Unterschied gefunden hätte.

Fachliche Kritik an der Preprint-Studie der Gruppe um Christian Drost in Bezug auf die Testgüte

- ▶ Einer der Kritikpunkte an der Studie von Drost und Kollegen war, dass die gewählten Tests zu wenig Power haben
- ▶ Sie wählten Tests, die die Eigenschaft der Daten nicht nutzen, dass die verwendeten Altersgruppen einer natürlichen Ordnung unterliegen, 10- bis 20-Jährige sind jünger als 50- bis 60-Jährige. Dieses Niveau der Daten nennt man Ordinalskala.
- ▶ Da die Tests der verschiedenen Altersgruppen nicht signifikant waren, konnte die Nullhypothese nicht verworfen werden.
- ▶ Aufgrund der oben beschriebenen Eigenschaften kann bei so einem Ausgang keine Aussage über die Hypothesen getroffen werden.



fact sheet

Weitere Recherchequellen

Bortz J (1999): *Statistik*. DOI: 10.1007/978-3-662-10031-8.

Genschel U et al. (2005): *Schließende Statistik*. ISBN: 978-3-540-21838-8.



fact sheet

Ansprechpartner in SMC Lab

Lars Koppers

Gastwissenschaftler im SMC Lab

Ansprechpartnerin in der Redaktion

Marleen Halbach

stellvertretende Redaktionsleiterin und Redakteurin für
Medizin und Lebenswissenschaften

Telefon +49 221 8888 25-0

E-Mail redaktion@sciencemediacenter.de

Disclaimer

Dieses Fact Sheet wird herausgegeben vom Science Media Center Germany. Es bietet Hintergrundinformationen zu wissenschaftlichen Themen, die in den Schlagzeilen deutschsprachiger Medien sind, und soll Journalisten als Recherchehilfe dienen.

SMC-Fact Sheets verstehen sich nicht als letztes Wort zu einem Thema, sondern als eine Zusammenfassung des aktuell verfügbaren Wissens und als ein Hinweis auf Quellen und weiterführende Informationen.

Sie haben Fragen zu diesem Fact Sheet (z. B. nach Primärquellen für einzelne Informationen) oder wünschen Informationen zu anderen Angeboten des Science Media Center Germany? Dann schicken Sie uns gerne eine E-Mail an redaktion@sciencemediacenter.de oder rufen Sie uns an unter +49 221 8888 25-0.

Impressum

Die Science Media Center Germany gGmbH (SMC) liefert Journalisten schnellen Zugang zu Stellungnahmen und Bewertungen von Experten aus der Wissenschaft – vor allem dann, wenn neuartige, ambivalente oder umstrittene Erkenntnisse aus der Wissenschaft Schlagzeilen machen oder wissenschaftliches Wissen helfen kann, aktuelle Ereignisse einzuordnen. Die Gründung geht auf eine Initiative der Wissenschafts-Pressekonferenz e.V. zurück und wurde möglich durch eine Förderzusage der Klaus Tschira Stiftung gGmbH.

Nähere Informationen: www.sciencemediacenter.de

Diensteanbieter im Sinne RStV/TMG

Science Media Center Germany gGmbH
Schloss-Wolfsbrunnenweg 33
69118 Heidelberg
Amtsgericht Mannheim
HRB 335493

Redaktionssitz

Science Media Center Germany gGmbH
Rosenstr. 42–44
50678 Köln

Vertretungsberechtigte Geschäftsführer

Beate Spiegel, Volker Stollorz

Verantwortlich für das redaktionelle Angebot (Webmaster) im Sinne des §55 Abs.2 RStV

Volker Stollorz

