



Häufigkeit von Schilddrüsenkrebs in der Samtgemeinde Elm-Asse und der Gemeinde Dettum in den Diagnosejahren 2015 - 2019

Oldenburg, November 2022

Registerstelle des EKN
Joachim Kieschke MPH - Ärztlicher Leiter der Registerstelle
Iris Urbschat
Industriestr. 9
26121 Oldenburg
Tel. 0441 361056-12

Inhaltsverzeichnis

1	HINTERGRUND	3
2	METHODIK	3
2.1	HYPOTHESE UND VORGEHEN	3
2.2	UNTERSUCHUNGSGEBIET UND VERGLEICHREGION	4
2.3	GESCHÄTZTE VOLLZÄHLIGKEIT DER ERFASSTEN KREBSNEUERKRANKUNGEN IM BEZIRK BRAUNSCHWEIG UND IM LANDKREIS WOLFENBÜTTEL	5
2.4	FALLVALIDIERUNG ZUR QUALITÄTSSICHERUNG	5
3	ERGEBNISSE	6
4	DESKRIPTIVE AUSWERTUNG FÜR DIE FRÜHERE SAMTGEMEINDE ASSE	6
5	ABSCHLIEßENDE EINSCHÄTZUNG	7
6	ZUSAMMENFASSUNG	7
7	LITERATUR	8
8	ANHANG	9
8.1	MELDEVERFAHREN UND STRUKTUREN DES EKN	9
8.2	EPIDEMIOLOGISCHE MAßZAHLEN	10
8.3	SCHLIEßENDE STATISTIK	11

Häufigkeit von Schilddrüsenkrebs in der Samtgemeinde Elm-Asse und der Gemeinde Dettum in den Diagnosejahren 2015 - 2019

1 Hintergrund

Der Landkreis Wolfenbüttel hat das Epidemiologische Krebsregister Niedersachsen (EKN) mit Anfrage vom 30.05.2018 um eine Auswertung zur Häufigkeit von Schilddrüsenkrebs für die Diagnosejahre 2015 - 2019 gebeten. Das Untersuchungsgebiet soll dabei die Samtgemeinde (SG) Elm-Asse und die Gemeinde Dettum umfassen. Die Auswertung soll für alle Altersgruppen und alle Geschlechter zusammengefasst erfolgen. Im Falle einer statistisch signifikanten Erhöhung in dem genannten Untersuchungsgebiet sollen die Fallzahlen für die einzelnen (Mitglieds-)Gemeinden deskriptiv dargestellt werden.

Hintergrund ist eine Sonderauswertung des EKN aus dem Jahr 2010 [1], in der eine statistisch auffällige Häufung von Schilddrüsenkrebs und hämatologischen Erkrankungen für die frühere SG Asse in den Diagnosejahren 2002 - 2009 festgestellt wurde. In einer Folgeauswertung aus dem Jahr 2016 [2] wurde die Erhöhung von Schilddrüsenkrebs für die Diagnosejahre 2010 - 2014 bestätigt; hämatologische Krebserkrankungen waren dagegen unauffällig.

Die Ergebnisse der beantragten Auswertung werden in diesem Bericht vorgestellt.

2 Methodik

Zur Auswertung kommen die dem EKN gemeldeten Krebsneuerkrankungsfälle für Schilddrüsenkrebs (ICD-10 C73) der Diagnosejahre 2015 - 2019 einschließlich der DCO-Fälle ('death certificate only'). DCO-Fälle sind dem EKN ausschließlich über eine Todesbescheinigung bekannt geworden; als Diagnosejahr wird für sie das Sterbejahr angenommen. Datenstand ist September 2022.

2.1 Hypothese und Vorgehen

Die durchgeführten Analysen¹ sind auf die formulierte Fragestellung der Anfrage ausgerichtet.

Als statistische Nullhypothese wurde formuliert:

„Die Anzahl an Schilddrüsenkrebsneuerkrankungen ist für beide Geschlechter zusammengefasst in dem Untersuchungsgebiet für die Diagnosejahre 2015 - 2019 kleiner oder gleich der erwarteten Fallzahl“.

Die Alternativhypothese lautet:

¹ Erläuterungen zu den epidemiologischen Maßzahlen - siehe Anhang

„Die Anzahl an Schilddrüsenkrebsneuerkrankungen ist für beide Geschlechter zusammengefasst in dem Untersuchungsgebiet für die Diagnosejahre 2015 - 2019 größer als die erwartete Fallzahl“.

Gemäß der Fragestellung wird einseitig getestet. Als Signifikanzniveau wird 5 % festgelegt.

Die Berechnungen wurden überwiegend mit der für Krebsregister spezifischen Auswertungssoftware CARESS durchgeführt. Die Berechnung der p-Werte erfolgte mit der R-Routine Exact Poisson test².

2.2 Untersuchungsgebiet und Vergleichsregion

Untersuchungsgebiet für diesen Bericht ist die SG Elm-Asse und die Gemeinde Dettum. Die SG Elm-Asse liegt südöstlich der kreisfreien Stadt Wolfenbüttel im Landkreis Wolfenbüttel (s. **Abbildung 1**), der zum Bezirk Braunschweig gehört. Die in früheren Auswertungen untersuchte SG Asse fusionierte im Rahmen einer Gebietsreform im Jahr 2015 mit der SG Schöppenstedt zu der neuen SG Elm-Asse. Sie umfasst die Mitgliedsgemeinden Denkte, Wittmar, Remlingen-Semmenstedt, Roklum, Hedeper und Kissenbrück (frühere SG Asse) sowie Winnigstedt, Uehrde, Vahlberg, Kneitlingen, Schöppenstedt, Dahlum (frühere SG Schöppenstedt). Die Gemeinde Dettum grenzt im Norden an die SG Elm-Asse an; sie wird auf Wunsch des Landkreises aufgrund der räumlichen Nähe in das Untersuchungsgebiet mit einbezogen.

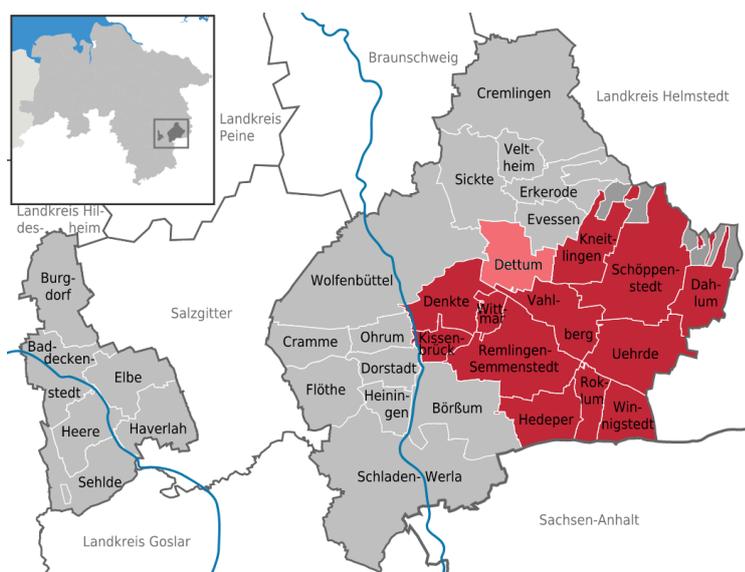


Abbildung 1: Das Untersuchungsgebiet innerhalb des Landkreises Wolfenbüttel: Samtgemeinde Elm-Asse (dunkelrot) und Gemeinde Dettum (hellrot).

Karte: Wikipedia, modifiziert. [NordNordWest](https://de.wikipedia.org/wiki/NordNordWest), Lizenz: [Creative Commons by-sa-3.0 de](https://creativecommons.org/licenses/by-sa/3.0/de/)

² R version 2.11.1 © 2010 The R Foundation for Statistical computing
'poisson.test(x, T = 1, r = 1, alternative = c("greater"))'

Im Untersuchungszeitraum 2015 - 2019 wohnten durchschnittlich 19.664 Einwohner im Gebiet der SG Elm-Asse und der Gemeinde Dettum.

Vergleichsregion für diese Untersuchung ist – wie schon im Bericht von 2016 – der Bezirk Braunschweig, dessen alters- und geschlechtsspezifische Raten zur Berechnung der zu erwartenden Fallzahlen genutzt werden. Die durchschnittliche jährliche Bevölkerung der Vergleichsregion Bezirk Braunschweig beläuft sich auf knapp 1,6 Millionen Einwohner (**s. Tabelle 1**). Für das Untersuchungsgebiet liegt der Anteil der Bevölkerung bei 1,2 % der Vergleichsregion.

Tabelle 1: Bevölkerungszahlen für das Untersuchungsgebiet Samtgemeinde (SG) Elm-Asse und Gemeinde Dettum sowie für die Vergleichsregion Bezirk Braunschweig in den Jahren 2015 - 2019 (gemittelt je Jahr)

	Bevölkerung (gemittelt je Jahr)
Untersuchungsgebiet gesamt (SG Elm-Asse + Gemeinde Dettum)	19.664
Vergleichsregion Bezirk Braunschweig	1.594.649

2.3 Geschätzte Vollzähligkeit der erfassten Krebsneuerkrankungen im Bezirk Braunschweig und im Landkreis Wolfenbüttel

Gemäß der Inzidenz des Robert Koch-Instituts (RKI Nds. 2019) liegt der geschätzte Erfassungsgrad für Schilddrüsenkrebs für den Bezirk Braunschweig sowie den Landkreis Wolfenbüttel für den Gesamtzeitraum 2015 - 2019 bei über 95 %. Eine Vollzähligkeit von mindestens 90 % gilt bundesweit als wesentliche Voraussetzung für wissenschaftlich fundierte Aussagen zur Häufigkeit von Krebsneuerkrankungen.

2.4 Fallvalidierung zur Qualitätssicherung

Routinemäßig werden die Angaben der Meldenden zu Wohnadressen der Patientinnen und Patienten auf Plausibilität geprüft – ein Routineabgleich der Wohnadressen mit Meldeamtsdaten erfolgt nicht. Im Rahmen von Sonderauswertungen werden für die Fälle aus dem Untersuchungsgebiet die dem EKN vorliegenden Informationen noch einmal bezüglich der Qualität der Diagnosesicherung, der Angabe der Diagnosekodierung, der Zusammenführung verschiedener Meldungen zu einem Fall und der Plausibilität der gemeldeten Wohnortangaben geprüft. Dies kann gegebenenfalls zu gezielten Nachrecherchen bei Meldenden oder Einwohnermeldeämtern führen. Die im Rahmen dieser Auswertung durchgeführten Validierungsschritte ergaben keine Hinweise auf Qualitätsdefizite.

3 Ergebnisse

Die Ergebnisse zur Häufigkeit von Schilddrüsenkrebs in der SG Elm-Asse und der Gemeinde Dettum werden für den Diagnosezeitraum 2015 - 2019 in **Tabelle 2** dargestellt. Die beobachtete Anzahl an Neuerkrankungen wird mit der erwarteten Anzahl – für alle Altersgruppen und alle Geschlechter zusammengefasst – verglichen.

Tabelle 2: Häufigkeit von Schilddrüsenkrebs in der Samtgemeinde (SG) Elm-Asse und der Gemeinde Dettum in den Diagnosejahren 2015 - 2019 (inklusive DCO-Fällen, für alle Geschlechter zusammengefasst)

SG Elm-Asse und Gemeinde Dettum	Neuerkrankungen beobachtet	Neuerkrankungen erwartet*	SIR (beobachtet/erwartet)	95% KI des SIR (zweiseitig)	p-Wert (einseitig)
Schilddrüsenkrebs (ICD-10 C73)	8	8,4	0,95	0,41 - 1,87	0,6059

**Vergleichsregion Bezirk Braunschweig, gerundet*

In den Diagnosejahren 2015 - 2019 werden in der SG Elm-Asse und der Gemeinde Dettum insgesamt 8 Schilddrüsenkrebsneuerkrankungen beobachtet bei 8,4 erwarteten Fällen (SIR 0,95). Der p-Wert liegt mit 0,6059 über dem Signifikanzniveau von 0,05; somit ist das SIR statistisch unauffällig. Die Nullhypothese kann nicht verworfen werden, d. h., die beobachtete Anzahl an Schilddrüsenkrebs ist für beide Geschlechter zusammengefasst im Zeitraum 2015 - 2019 in der SG Elm-Asse und der Gemeinde Dettum bei einer angenommenen Irrtumswahrscheinlichkeit von 5 % statistisch nicht erhöht.

Die beobachteten Fälle verteilen sich auf verschiedene Mitgliedsgemeinden, es liegen keine kleinräumigen Häufungen innerhalb des Untersuchungsgebiets vor.

4 Deskriptive Auswertung für die frühere Samtgemeinde Asse

Die EKN-Berichte aus den Jahren 2010 [1] und 2016 [2] bezogen sich jeweils auf die frühere SG Asse. Mit dem Ziel der Vergleichbarkeit der aktuellen Ergebnisse mit denen der früheren Berichte wird nachfolgend zusätzlich eine deskriptive Auswertung durchgeführt. In dieser wird für den Untersuchungszeitraum 2015 - 2019 geprüft, ob die Häufigkeit von Schilddrüsenkrebs auch für die frühere SG Asse inzwischen unauffällig ist. Die Bevölkerung beläuft sich für die frühere SG Asse auf 9.138 Einwohner (gemittelt je Jahr). Für alle Geschlechter zusammengefasst werden hier 4 Fälle beobachtet bei 4,0 erwarteten Fällen. Die Auswertung zeigt, dass in der früheren SG Asse die Häufigkeit von Schilddrüsenkrebs ebenfalls nicht mehr erhöht ist.

5 Abschließende Einschätzung

Die Auswertung zeigt keine Erhöhung für Schilddrüsenkrebsneuerkrankungen in dem Gebiet der SG Elm-Asse und der Gemeinde Dettum im Beobachtungszeitraum 2015 - 2019. Auch in der früheren SG Asse ist die Häufigkeit von Schilddrüsenkrebs im Beobachtungszeitraum unauffällig.

Das EKN sieht daher keinen Anlass für eine Weiterbeobachtung.

6 Zusammenfassung

In einer Sonderauswertung des EKN von 2010 [1] war für die frühere SG Asse für die Diagnosejahre 2002 - 2009 eine auffällige Häufung für die Diagnosen Leukämien (ICD-10 C91-95), Leukämien und Lymphome insgesamt (ICD-10 C81-96) sowie Schilddrüsenkrebs (ICD-10 C73) festgestellt worden.

Eine Folgeauswertung des EKN von 2016 [2] zeigte auf, dass für die frühere SG Asse für die Diagnosejahre 2010 - 2014 keine Erhöhung mehr für Leukämien und für die Gesamtgruppe der Leukämien und Lymphome bestand. Die Weiterbeobachtung wurde daher für Leukämien und Lymphome beendet. Die Häufigkeit von Schilddrüsenkrebs war dagegen weiterhin signifikant erhöht (9 Fälle beobachtet bei 3,8 erwarteten Fällen).

Der Landkreis Wolfenbüttel beantragt daher in einer Anfrage aus dem Jahr 2018 eine Auswertung zur Häufigkeit von Schilddrüsenkrebs für die Diagnosejahre 2015 - 2019 mit der Bitte, das Untersuchungsgebiet auszudehnen auf die nach einer Gebietsreform im Jahr 2015 fusionierte SG Elm-Asse und die an dieses Gebiet angrenzende Gemeinde Dettum. Alle Geschlechter sollen zusammengefasst betrachtet werden. Als Vergleichsregion soll der Bezirk Braunschweig herangezogen werden.

Der vorliegende Bericht zeigt, dass die Schilddrüsenkrebshäufigkeit in dem Gebiet der SG Elm-Asse und der Gemeinde Dettum nicht erhöht ist (8 Fälle beobachtet bei 8,4 erwarteten Fällen, SIR 0,95, $p = 0,6059$).

Da das aktuelle Untersuchungsgebiet (SG Elm-Asse und Gemeinde Dettum) von dem früheren Untersuchungsgebiet (frühere SG Asse) abweicht, wird zusätzlich eine deskriptive Auswertung für die frühere SG Asse durchgeführt. In der früheren SG Asse werden in den Diagnosejahren 2015 - 2019 für alle Geschlechter zusammengefasst 4 Fälle beobachtet bei 4,0 erwarteten Fällen. Auch für die frühere SG Asse zeigt sich somit für Schilddrüsenkrebs keine Erhöhung mehr.

Eine Weiterbeobachtung wird daher vom EKN nicht empfohlen

7 Literatur

- (1) EKN-Registerstelle: Auswertung des EKN zur Krebshäufigkeit in der Samtgemeinde Asse. Oldenburg, 16. Dezember 2010
- (2) EKN-Registerstelle: Folgeauswertung des EKN zur Häufigkeit von Krebsneuerkrankungen in der Samtgemeinde Asse. Oldenburg, Oktober 2016
- (3) Epidemiologisches Krebsregister Niedersachsen, Registerstelle (Hrsg.): Krebs in Niedersachsen 2020, Jahresbericht, Oldenburg, 2021
- (4) Horn M, Vollandt R: Multiple Tests und Auswahlverfahren, Stuttgart, Jena, New-York, 1995
- (5) Victor A, Elsässer A, Hommel G, Blettner M: Wie bewertet man die p-Wert-Flut? Dtsch Arztebl Int 2010; 107(4): 50–6 DOI: 10.3238/arztebl.2010.0050

8 Anhang

8.1 Meldeverfahren und Strukturen des EKN

Seit 2003 erfasst das EKN flächendeckend Krebserkrankungen in Niedersachsen.

Die Meldeverfahren des EKN sind im Gesetz über das Epidemiologische Krebsregister Niedersachsen (GEKN) geregelt. Bis 2013 gab es eine Meldepflicht nur für Ärztinnen und Ärzte ohne direkten Patientenkontakt (Pathologien). Mit dem 1. Januar 2013 wurde in Niedersachsen für alle Ärztinnen und Ärzte, die eine Tumorerkrankung feststellen oder behandeln, eine Meldepflicht anstelle der bis dahin bestehenden Meldeberechtigung eingeführt. Seit der Betriebsaufnahme des neu gegründeten Klinischen Krebsregisters Niedersachsen (KKN) am 1. Juli 2018 erfassen KKN und EKN gemeinsam über ein Melderportal beim KKN die Meldungen zu Krebserkrankungen. Die Meldungen werden an das EKN weitergeleitet. Eine kleinräumige Zuordnung ist für alle Meldungen vorgesehen. Darüber hinaus erhält das EKN von den Meldeämtern die Angaben zu Verstorbenen und von den Gesundheitsämtern die Todesbescheinigungen.

Das EKN besteht aus zwei Arbeitseinheiten: Die Vertrauensstelle (VST), die organisatorisch am NLGA in Hannover angesiedelt ist, erfasst und prüft die eingehenden Meldungen zu Krebserkrankungen; seit 2018 findet dies in Kooperation mit dem KKN statt. Die VST leitet die Meldungen an die Registerstelle (RST) in Oldenburg weiter.

Vor der Weiterleitung werden die Personendaten aus Datenschutzgründen in der VST verschlüsselt (pseudonymisiert). Dafür werden nicht dechiffrierbare Kontrollnummern gebildet. Zusätzlich wird ein rückverschlüsselbares Chiffprat aus den Personendaten gebildet, um Betroffene kontaktieren zu können und ggf. weitere Auskünfte erhalten oder Untersuchungen durchführen zu können. Die VST entscheidet darüber hinaus über Anträge auf Herausgabe und Nutzung von Daten, koordiniert Anfragen aus der Bevölkerung und berät die kommunalen Gesundheitsbehörden bei Krebsclusteranfragen.

Die Aufgaben der Registerstelle bestehen in der weiteren Bearbeitung und Zusammenführung von Meldungen, der langfristigen Speicherung der Registerdaten sowie der Erstellung epidemiologischer Routine- und Sonderauswertungen [3]. Über die Kontrollnummern können in der Registerstelle die verschiedenen verschlüsselten Meldungen zu einer Person zusammengeführt werden.

Der Arbeitsbereich Umweltepidemiologie des NLGA ist bei Anfragen mit möglichen Umweltfaktoren als Auslöser einer vermuteten Häufung von Krebserkrankungen maßgeblich beteiligt.

8.2 Epidemiologische Maßzahlen

Fallzahlen

Über die Geschlechter zusammengefasste Fallzahlen umfassen die Kategorien Männer, Frauen, Divers, Unbekannt. Die Berechnung der alters- und geschlechtsstandardisierten erwarteten Fallzahlen (s.u.) bezieht sich aus methodischen Gründen auf die Kategorien Männer und Frauen, da nur für diese Nennerangaben vorliegen.

Altersspezifische Raten

Die altersspezifischen Raten werden gebildet aus der Anzahl von Krebsfällen in einer Altersklasse, dividiert durch die durchschnittliche Bevölkerung der jeweiligen Altersklasse, wobei die Ergebnisse pro 100.000 der Bezugsbevölkerung angegeben werden.

Altersstandardisierte Rate

Altersstandardisierten Raten lassen Vergleiche von verschiedenen Regionen zu, die eine unterschiedliche Altersstruktur aufweisen. Die altersstandardisierte Rate beschreibt, welche Krebshäufigkeit vorliegen würde (auf 100.000 Personen bezogen), wenn die Altersstruktur der beobachteten Bevölkerung derjenigen der Standardbevölkerung (Standardbevölkerung Europa) entsprochen hätte. Die Standardisierung wird anhand von 18 Fünfjahres-Altersklassen vorgenommen (0-4, ..., 85+).

Erwartete Fallzahl

Die erwartete Fallzahl gibt an, wie viele Krebsneuerkrankungsfälle in einer Region zu erwarten sind, wenn für die einzelnen Geschlechter und Altersgruppen der Wohnbevölkerung dieser Region die altersspezifischen Krebsneuerkrankungsraten der Vergleichsregion herangezogen werden. Die erwartete Fallzahl für die jeweilige Region wird anhand der Vergleichsregion – hier Bezirk Braunschweig – berechnet.

SIR

Das standardisierte Inzidenz-Verhältnis (SIR für 'standardized incidence ratio') gibt den Quotienten zwischen der Anzahl beobachteter und erwarteter Krebsneuerkrankungsfälle an. Das SIR ist daher genau dann 1, wenn beobachtete und erwartete Neuerkrankungsfallzahlen übereinstimmen; ein erhöhtes SIR, z.B. von 1,10 besagt, dass in der Untersuchungsregion eine um 10 % höhere Krebshäufigkeit vorliegt als erwartet. Ein SIR unter 1 weist auf eine vergleichsweise niedrigere Krebshäufigkeit für die jeweilige Diagnose hin als erwartet.

Der Zeitpunkt des Auftretens einer Erkrankung unterliegt einem Zufallsprozess. Die Aussage allein, dass das SIR ober- oder unterhalb des Wertes 1 liegt, hat keine Aussagekraft, da die beobachteten SIR im Normalfall um die 1 schwanken. Um daher beurteilen zu können, ob es zu statistisch auffälligen Abweichungen der beobachteten Anzahl von der erwarteten gekommen ist, muss eine Annahme zur statistischen Verteilung der Neu-

erkrankungen getroffen werden. Hierzu wird für die beobachtete Fallzahl in der Untersuchungsregion, die den Zähler des SIR bildet, eine Poissonverteilung angenommen. Mit dieser Annahme können Konfidenzintervalle für das SIR sowie Hypothesen- bzw. Signifikanztests für das SIR abgeleitet werden.

Das beobachtete SIR kann auch als Schätzung für das „wahre SIR“ betrachtet werden, das sich auf einen längeren Zeitraum oder eine größere Population bezöge.

Ergänzend zum Punktschätzer des SIR werden Vertrauens- bzw. Konfidenzintervalle [KI] angegeben, die einen Wertebereich schätzen, der das ‚wahre‘ SIR mit vorgegebener Wahrscheinlichkeit $(1 - \alpha)$ überdeckt. Ein 95%-KI überdeckt somit das wahre SIR mit 95%iger Wahrscheinlichkeit. Je schmaler ein KI ausfällt, desto präziser ist die Schätzung. Wenn das KI den Wert 1,00 nicht überdeckt, so kann die Hypothese, dass das wahre SIR 1 betrage, mit einer Irrtumswahrscheinlichkeit von α verworfen werden. In diesem Bericht werden ausschließlich zweiseitige 95%-KI angegeben.

Der bei einem SIR angegebene p-Wert gibt die Wahrscheinlichkeit an, dass der geschätzte oder ein noch größerer Wert für das SIR beobachtet wird, obwohl das „wahre“ SIR 1,00 beträgt. Ist der p-Wert kleiner als das für den Einzelvergleich vorgegebene Signifikanzniveau, ist die Hypothese, dass die Inzidenzrate nicht erhöht bzw. das wahre SIR nicht erhöht ist, zu verwerfen.

8.3 Schließende Statistik

Bei einem statistischen Test oder Hypothesentest wird eine im Vorfeld formulierte Hypothese anhand der vorliegenden Daten überprüft, um eine Aussage über ihre Gültigkeit treffen zu können. Da die Daten Zufallsprozessen unterliegen, wird die Aussage zur Gültigkeit einer Hypothese nie hundertprozentig sicher sein, sondern gewissen Irrtumswahrscheinlichkeiten unterliegen. Ein statistischer Test kann insofern als Entscheidungsregel unter Kontrolle von Unsicherheiten verstanden werden.

Aufstellung einer statistischen Hypothese und Prinzip des indirekten Beweises

Grundsätzlich ist der klassische Hypothesentest eine Form eines „indirekten Beweises“: Eine aufgestellte Hypothese wird so lange beibehalten, wie es die beobachteten Daten zulassen. Wenn hingegen die Beobachtungen unter der Hypothese zu unwahrscheinlich sind, erscheint die Annahme gerechtfertigt, dass die aufgestellte Hypothese nicht stimmt und sie daher verworfen werden kann. Dieses indirekte Vorgehen führt dazu, dass man die (statistische) Hypothese bzw. sogenannte Nullhypothese entsprechend scheinbar umgekehrt formulieren muss:

- Möchte man statistisch nachweisen, dass ein neues Medikament besser ist als das alte, formuliert man die Nullhypothese, das neue Medikament habe höchstens so

gute Erfolgsaussichten wie das alte. Die Alternativhypothese besagt, dass das neue Medikament besser sei. Zeigt sich nach der Prüfung anhand der Daten, dass das neue Medikament deutlich bessere Wirkungen erzielt hat, „muss“ die Nullhypothese zugunsten der Alternativhypothese verworfen werden. Dies führt dann zu dem (beabsichtigten) Nachweis, dass das neue Medikament besser ist.

- Um zu prüfen, ob in einer Untersuchungsregion die Wahrscheinlichkeit einer bestimmten Krebserkrankung höher ist als normalerweise, stellt man die Hypothese auf, dass diese Wahrscheinlichkeit höchstens genauso hoch ist. Liegen dann derartig viele Fallmeldungen vor, dass die Hypothese zu verwerfen ist, wird der Schluss gezogen, dass die Wahrscheinlichkeit für diese Krebserkrankung erhöht ist.

Die nachzuweisende Aussage tritt somit in der Alternativhypothese auf.

Irrtumswahrscheinlichkeiten

Die Irrtumswahrscheinlichkeit, die Nullhypothese auf Grund der vorliegenden Daten zu verwerfen, obwohl sie an sich richtig ist, wird als „Fehler 1. Art“ oder auch „ α -Fehler“ bezeichnet. Dieser Fehler wird beim klassischen Hypothesen- bzw. Signifikanztest kontrolliert. Dazu wird vor der Datenauswertung ein sogenanntes Signifikanzniveau des Tests vorgegeben, häufig 1 % oder 5 %. Das Signifikanzniveau ist der maximal zulässige Wert für den Fehler 1. Art. D.h. die Wahrscheinlichkeit, die Hypothese zu verwerfen, obwohl sie richtig ist, entspricht für den Test höchstens dem Signifikanzniveau.

Daneben gibt es noch den „Fehler 2. Art“ bzw. den „ β -Fehler“. Er gibt entsprechend an, mit welcher Wahrscheinlichkeit die Nullhypothese beibehalten wird, obwohl die Alternativhypothese korrekt ist. Die Wahrscheinlichkeit, die Nullhypothese abzulehnen, wenn sie falsch ist, wird demgegenüber als Güte oder „Power“ des Tests beschrieben.

Über den β -Fehler bzw. die Wahrscheinlichkeit, die Nullhypothese korrekterweise abzulehnen ($1 - \beta$), können Aussagen zur Güte des Tests getroffen werden:

- So kann für eine Untersuchungsregion abgeschätzt werden, mit welcher Wahrscheinlichkeit die Nullhypothese, es läge keine Erhöhung vor, abgelehnt wird, wenn es in Wirklichkeit eine relevante Erhöhung gäbe. Diese Wahrscheinlichkeit fällt umso höher aus, je größer der tatsächliche prozentuale Effekt ist. Aber auch die Größe der Untersuchungsregion spielt eine Rolle: In einer größeren Gemeinde ist bei derselben Effektstärke das Verwerfen der Nullhypothese wahrscheinlicher als bei einer kleineren Gemeinde.

Teststatistik

Eine Teststatistik muss in der Lage sein, für die aufgeworfene Fragestellung möglichst gut zwischen Null- und Alternativhypothesen zu unterscheiden und dabei möglichst alle verfügbaren Informationen zu nutzen:

- Man möchte anhand von Stichproben prüfen, ob die Bewohner einer Stadt A größer sind als die der Stadt B. Hier könnte man jeweils nur die ersten Personen der beiden Stichproben anhand ihrer Körpergrößen vergleichen. Dabei würden aber die übrigen Stichprobeninformationen verschenkt, d.h. die Körpergrößen der übrigen ebenfalls in den Stichproben erhobenen Bewohner blieben bei einem derartigen Test unberücksichtigt. Sinnvoller ist es, die durchschnittliche Körpergröße beider Stichproben miteinander zu vergleichen und darüber ein statistisches Testverfahren abzuleiten.

Für ein Testproblem können somit konkurrierende Teststatistiken bestehen, die anhand verschiedener Gütekriterien verglichen werden können. So sind viele der standardmäßig eingesetzten Testverfahren in dem Sinne optimal, dass es unter den getroffenen statistischen Annahmen keine Tests mit einer höheren Power gibt.

Die Entscheidungsregel und der Begriff der Signifikanz:

Für die Teststatistik wird ein kritischer Wert³ berechnet, so dass die Wahrscheinlichkeit dafür, dass die Teststatistik größer als dieser kritische Wert ist, maximal der Irrtumswahrscheinlichkeit entspricht. Überschreitet die Teststatistik diesen kritischen Wert, so kann die Nullhypothese verworfen werden. Das Ergebnis ist signifikant.

Werte unterhalb dieses kritischen Wertes werden auch als Annahmebereich des Tests bezeichnet, da mit ihnen die Nullhypothese nicht verworfen wird. Das Ergebnis des Tests ist nicht signifikant.

Ein signifikanter Test gibt nur an, dass der aus den konkreten Daten resultierende Wert des Tests unter der Nullhypothese unwahrscheinlich ist und zwar höchstens so wahrscheinlich wie die vorgegebene Irrtumswahrscheinlichkeit. Eine Aussage zur Wahrscheinlichkeit von Null- oder Alternativhypothese wird damit nicht getroffen. Entweder wird anhand der Daten die Nullhypothese verworfen oder aber sie muss beibehalten werden. Die Nullhypothese wird damit aber keineswegs „bestätigt“ oder gar „bewiesen“.

Ein „signifikantes Ergebnis“ kann auch zustande kommen, wenn unzutreffende Annahmen über den den Daten zu Grunde liegenden Zufallsprozess gemacht wurden. Auch

³ Dies bezieht sich auf einseitig formulierte Testprobleme, bei denen allein hohe Werte der Teststatistik gegen die Nullhypothese sprechen. Es gibt aber auch zweiseitig formulierte Testprobleme, bei denen sowohl sehr kleine wie große Werte gegen die Nullhypothese sprechen. So werden Abweichungen von Normwerten, die bspw. in der Produktion eingehalten werden müssen, meist zweiseitig überprüft. D. h. Abweichungen in beide Richtungen führen zur Ablehnung der Hypothese, 'die Normwerte werden eingehalten', und somit zur Neu-Adjustierung der Maschine.

können systematische Erhebungsfehler, die die Daten beeinflussen könnten, nicht kontrolliert werden. Eine Testentscheidung hängt somit ab von den Hypothesen, der gewählten Teststatistik, dem vorgegebenen Signifikanzniveau, dem beobachteten Effekt sowie dem gewählten Stichprobenumfang.

p-Wert – „Signifikanztest“ als Alternative zum „Hypothesentest“

Das Konzept des p-Wertes bzw. des „Signifikanztests“ stellt ein induktives Verfahren dar, von der Stichprobe auf die Grundgesamtheit zu schließen. Es wurde parallel zum Konzept des Hypothesentests entwickelt. Diese beiden Konzepte werden allerdings in den allermeisten Statistikeinführungen nicht klar voneinander abgegrenzt, sondern vermischt und als einheitliches Konzept dargestellt.

Dabei benötigt jedoch der Signifikanztest keine ausformulierte Alternativhypothese, sondern beschränkt sich auf Aussagen zur Hypothese. Der p-Wert gibt dabei die Wahrscheinlichkeit der beobachteten oder extremeren Daten unter der Hypothese an. Je kleiner der p-Wert ist, desto mehr Evidenz spricht gegen die aufgestellte Hypothese. Ein p-Wert von 0,001 scheint mehr Sicherheit in die Entscheidung, eine Hypothese zum Signifikanzniveau von 5 % zu verwerfen, zu vermitteln, als ein deutlich größerer p-Wert von 0,049, der aber zur selben Entscheidung führt.

Der enge Zusammenhang zum Hypothesentest besteht darin, dass ein p-Wert unter der vorgegebenen Irrtumswahrscheinlichkeit (1. Art) zur Ablehnung der Hypothese führt. Somit kann anhand des p-Wertes bewertet werden, ob ein entsprechend formuliertes Testproblem zu einem signifikanten oder nicht-signifikanten Ergebnis führt.

Der p-Wert darf aber keinesfalls, wie häufig fälschlich angegeben, als Irrtumswahrscheinlichkeit für die getroffene Entscheidung oder gar als die Wahrscheinlichkeit der Hypothese selbst interpretiert werden: Es wird schließlich die Wahrscheinlichkeit von Beobachtungen unter einer Hypothese, nicht jedoch die Wahrscheinlichkeit einer Hypothese unter den Beobachtungen betrachtet.