Medikamententest

Medikamente sollen Krankheiten heilen oder Symptome lindern. Wie aber kann man nachweisen ob sie das tun? Denn wenn man es eingenommen hat weiß man ja nicht wie es einem gegangen wäre, wenn man es nicht eingenommen hätte. Man kann also nicht an sich selber testen, denn es fehlen Vergleichsmöglichkeiten!

In der medizinischen Forschung werden daher Vergleichsstudien durchgeführt. Dies findet z.B. zwischen Personen, die ein neu entwickeltes Medikament einnehmen und Personen, die ein Placebo (ein Mittel, das keinen Wirkstoff enthält) einnehmen statt. Ziel der Vergleichsstudien ist es zu zeigen, dass ein neues Medikament mehr Personen heilt als das Placebo.

Beispiel: 40 Personen einer Testgruppe T werden mit einem neu entwickelten Grippeimpfstoff geimpft. 30 Personen einer Placebogruppe P werden mit einem Placebo geimpft (man hat leider nicht mehr Freiwillige gefunden). Über die Winterzeit blieben in Gruppe T 34 Personen von der Grippe verschont, in Gruppe P waren dies 21 Personen.
Aus Sicht des Medikamentenherstellers sind nur zwei Hypothesen relevant:

- Der neue Impfstoff ist besser als das Placebo (diese Hypothese nennt man Arbeitshypothese und benutzt das Symbol H_1).
- Der neue Impfstoff ist nicht besser als das Placebo (diese Hypothese nennt man Nullhypothese mit dem Symbol H_0).

a) Begründe, ob du im Beispiel H_1 oder H_0 zustimmen würdest.
b) Mit welcher Wahrscheinlichkeit heilt das Medikament eine Person, wenn man annimmt, dass H_0 stimmt?
c) Angenommen, man hätte eine Gruppe mit 40 Personen mit dem Placebo geimpft:
  i. Berechne jeweils die Wahrscheinlichkeit, dass mindestens 30, 31,..., 39 Personen grippefrei bleiben. Und umgekehrt:
  ii. Berechne, wie viele Personen mindestens grippefrei bleiben müssen, damit die Wahrscheinlichkeit hierfür höchstens 20%, 15%, ..., 5% beträgt.
d) Welche Aussagen passen zu $\sum_{k=33}^{40} \binom{40}{k} \cdot 0.7^k \cdot 0.3^{40-k} \approx 6\%$?
  i. Die Wahrscheinlichkeit, dass in einer Placebogruppe mindestens 33 Personen gesund bleiben beträgt etwa 6%.
  ii. In etwa 6 von 100 Placebogruppen bleiben mindestens 33 Personen gesund.
  iii. In etwa 6% aller Gruppen bleiben mindestens 33 Personen gesund, wenn das Medikament wirkungslos ist.
  iv. Bleiben mindestens 33 Personen durch das Medikament gesund, dann nimmt man in etwa 6 von 100 Gruppen an, das Medikament ist wirksamer als das Placebo, obwohl es das tatsächlich aber nicht ist.
  v. Wenn mindestens 33 Personen durch das Medikament gesund bleiben, dann beträgt die Wahrscheinlichkeit das es irrtümlich als wirksamer als ein Placebo erachtet wird etwa 6%.

Die Wahrscheinlichkeit, dass man ein Medikament irrtümlich für wirksam hält nennt man **Irrtumswahrscheinlichkeit, Fehler 1. Art** oder **α-Fehler**.
Wie in c) ii. wird dieser Wert für Untersuchungen vorgegeben, damit berechnet und entschieden werden kann, ob ein Impfstoff als wirksam akzeptiert wird.

e) Wie müsstest du a) zu einem α-Fehler von höchstens 5% beantworten?

Lösungen

a) Offensichtlich ist, dass mit dem Impfstoff 85% und mit dem Placebo 70% der Personen grippeverschont bleiben. Deswegen ist der Impfstoff wahrscheinlich besser aber nicht zwangsläufig (wie wahrscheinlich ist aber offen). Denn die 34 Personen könnten auch eine zufällige Abweichung nach oben sein. Weitere Argumente dafür: Der Stichprobenumfang ist gering und es gibt keine Informationen über die Repräsentativität der Stichprobe (z.B. Geschlecht, Alter, Gesundheit, Beruf). Man kann also beide Hypothesen rechtfertigen.

b) Für H_0 darf der Impfstoff nicht besser als das Placebo sein. Deswegen entspricht die Wahrscheinlichkeit p dem Anteil 21 von 30 Personen, also gilt $p \leq 0{,}7$. Wenn mit p gerechnet werden soll, bietet sich die „Randwahrscheinlichkeit" p=0,7 an.

c) i. Mind. 30 Personen: $\sum_{k=30}^{40} \binom{40}{k} \cdot 0.7^k \cdot 0.3^{40-k}$ *=binomcdf(40,0.7,30,40)*≈30,9%

Mind. 31 Personen: $\sum_{k=31}^{40} \binom{40}{k} \cdot 0.7^k \cdot 0.3^{40-k}$ *=binomcdf(40,0.7,31,40)*≈19,6%

Mind.	32	33	34	35	36	37	38	39
Wkt.	11,1%	5,5%	2,4%	0,9%	0,3%	0,06%	0,01%	0,001%

ii. Aus c) i. folgt

Höchstens 20%: Mindestens 31 Personen
Höchstens 15%: Mindestens 32 Personen
Höchstens 10%: Mindestens 33 Personen
Höchstens 5%: Mindestens 34 Personen

d) Alle Aussagen sind richtig.
i. Folgt aus c)
ii. 6 von 100 ist eine Umformulierung von 6%.
iii. Wenn das Medikament wirkungslos ist, heilt es so gut wie das Placebo.
iv. Eine andere Formulierung von iii.
v. Eine andere Formulierung von iv.

e) Mit 34 von 40 grippefrei gebliebenen Personen beträgt der α-Fehler 2,4%. Damit wird die Vorgabe „höchstens 5%" für den α-Fehler eingehalten, der Impfstoff ist demnach wirksamer als das Placebo, wobei man mit dieser Annahme bei etwa 2 von 100 Untersuchungen irrt.

Medikamententest - Theorie

Ein wenig Testtheorie: Forschungslogisch ist es nicht seriös mit Zahlenbeispielen zu begründen, da hiermit nur widerlegt werden kann[1].
Deswegen kann man mit Zahlen auch nicht begründen, dass H_1 gilt, dass also ein Medikament besser als ein anderes ist.
Medikamentenhersteller haben also nur die Möglichkeit mit den vorliegenden Zahlen H_0 zu widerlegen, um empirische Unterstützung für das logische Gegenteil der Nullhypothese (also H_1) zu erhalten. Mit Hilfe der Wahrscheinlichkeitsrechnung geht man folgendermaßen vor[2] (diese Vorgehensweise nennt sich **Hypothesentest**):

- Bestimme die zu H_0 gehörende Wahrscheinlichkeit p, dass das Medikament genau so gut/schlecht, wie das Placebo heilt (für den Grippeimpfstoff gilt $p \leq 70\%$, da 21 von 30 = 0,7 Personen der Gruppe P gesund blieben).
- Lege die Irrtumswahrscheinlichkeit α selber fest oder recherchiere, was bei Studien jeweils üblich oder gefordert ist (z.B. $\alpha = 5\%$).
- Mit Hilfe von $\sum_{k=33}^{40} \binom{40}{k} \cdot 0{,}7^k \cdot 0{,}3^{40-k}$ = *„binomcdf(40,0.7,33,40)“* ≈ 5,5% > α und $\sum_{k=34}^{40} \binom{40}{k} \cdot 0{,}7^k \cdot 0{,}3^{40-k}$ = *„binomcdf(40,0.7,34,40)“* ≈ 2,4% < α kann man folgern, dass die Wahrscheinlichkeit das Medikament bei 34 von 40 gesund gebliebenen Personen irrtümlich wirksamer als das Placebo zu erachten etwa 2,4% beträgt und somit kleiner als 5% ist (bei 33 von 40 gesund gebliebenen Personen gilt dies nicht).
- Mit 34 von 40 gesund gebliebenen Personen folgt also, dass H_0 unter der vorgegebenen Irrtumswahrscheinlichkeit 5% widerlegt und die Studie erfolgreich war.
- Mit 33 von 40 gesund gebliebenen Personen kann man (streng genommen) keine Folgerung treffen, denn mit Zahlenbeispielen kann man H_0 nicht begründen.

Aufgaben:

a) 80 Personen einer Testgruppe T werden mit einem neu entwickelten Bellsuchtimpfstoff geimpft. 50 Personen einer Placebogruppe P werden mit einem Placebo geimpft. Über die Urlaubszeit blieben in T 70 Personen von der Bellsucht verschont, in P waren dies 40 Personen. Begründe, ob der neue Wirkstoff mit einer Irrtumswahrscheinlichkeit von 5% erfolgreich ist.

b) 200 Männer erhalten täglich eine Email mit Witzen, um sie glücklicher zu machen. 185 von ihnen geben nach einem Monat an hierdurch glücklich zu sein. 150 Männer kann man überreden an dieser Studie ohne Email teilzunehmen. 15 von ihnen geben nach einem Monat an nicht glücklich zu sein. Begründe, ob die Email mit einer Irrtumswahrscheinlichkeit von 10% glücklich macht.

c) 320 Schülern am Rivius Gymnasium gibt man Vitamin C vorbeugend gegen Erkältung im Winter. 200 von diesen Schülern halten das für Quatsch und nehmen das Vitamin nicht ein. Im Verlauf des Winters erkrankten 25 der mit Vitamin behandelten und 50 der nicht mit Vitamin behandelten Schüler. Begründe, ob die Vitamin C Behandlung mit einer Irrtumswahrscheinlichkeit von 20% erfolgreich ist.

---

[1] Dieses Prinzip der **Falsifizierbarkeit** ist eine der Grundaussagen des **kritischen Rationalismus**, den Sir Karl Popper mit seinem Werk „Logik der Forschung“ (1935) begründet hat.

[2] Die Grundlagen dieser Idee hat Ronald Aylmer Fischer in seinem Buch „Statistical Methods for Research Workers“ (1925) zusammengefasst.

Lösungen

a) 40 von 50 Personen der Placebogruppe bleiben von Bellsucht verschont.
Für H_0 gilt, dass der Impfstoff genau so gut wie das Placebo heilt. Die Wahrscheinlichkeit, dass der Impfstoff vor Bellsucht schützt beträgt also $p \le 40/50 = 80\%$.
70 von 80 Personen der Testgruppe bleiben von Bellsucht verschont. Hierzu gilt

$$\sum_{k=70}^{80} \binom{80}{k} \cdot 0.8^k \cdot 0.2^{80-k} = \textit{binomcdf(80,0.8,70,80)} \approx 5{,}6\% > 5\% = \alpha$$

Also wird der vorgegebene α-Fehler nicht unterschritten. Demzufolge kann H_0 nicht widerlegt werden. Über die Wirksamkeit des neuen Bellsuchtimpfstoffs kann man also keine Folgerung treffen.

b) 135 von 150 Personen der Email-freien Gruppe (=Placebogruppe) geben an glücklich zu sein.
Für H_0 gilt, dass eine Email genau so glücklich macht wie keine Email mit Witzen. Die Wahrscheinlichkeit, dass eine Email mit Witzen glücklich macht beträgt also $p \le 135/150 = 90\%$.
185 von 200 Personen macht die Witze-Email glücklich. Hierzu gilt

$$\sum_{k=185}^{200} \binom{200}{k} \cdot 0.9^k \cdot 0.1^{200-k} = \textit{binomcdf(200,0.9,185,200)} \approx 14{,}3\% > 10\% = \alpha$$

Also wird der vorgegebene α-Fehler nicht unterschritten. Demzufolge kann H_0 nicht widerlegt werden. Über die Wirksamkeit der Witze-Email kann man also keine Folgerung treffen (das wäre auch sehr seltsam).

c) 150 von 200 Schüler, die kein Vitamin C einnehmen (=Placebogruppe) bleiben von einer Erkältung verschont.
Für H_0 gilt, dass Vitamin C genau so gut vor Erkältung schützt wie kein Vitamin C. Die Wahrscheinlichkeit, dass Vitamin C vor Erkältung schützt beträgt also $p \le 150/200 = 75\%$.
95 von 120 Schülern, die Vitamin C einnehmen bekommen keine Erkältung. Hierzu gilt

$$\sum_{k=95}^{120} \binom{120}{k} \cdot 0.75^k \cdot 0.25^{120-k} = \textit{binomcdf(120,0.75,95,120)} \approx 17{,}2\% < 20\% = \alpha$$

Also wird der vorgegebene α-Fehler unterschritten. Demzufolge ist H_0 widerlegt, wodurch H1 (Vitamin C hilft gegen Erkältung) mit einer Irrtumswahrscheinlichkeit von 20% (genauer: 17,2%) gilt.

Medikamententest - Alternativen

Einen Medikamententest durchzuführen ist nicht zwingend auf nur eine Art möglich. Dies soll durch folgendes Beispiel erläutert werden:

Grippeimpfung

In einer medizinischen Studie wurden im Winter von 250 Teilnehmern 200 gegen Grippe geimpft. 60 der geimpften und 40 der nicht geimpften Teilnehmer erkrankten im Frühjahr an Grippe. Ist hiermit die Wirksamkeit der Grippeimpfung mit einer Irrtumswahrscheinlichkeit von 5% gesichert?

Sinnvoll wäre es zunächst anzunehmen, dass eine Impfung die Lage der Patienten verbessern sollte (denn sonst könnte man sich den ganzen medizinischen Forschungsaufwand sparen). Das kann bedeuten, dass mit Impfung weniger Patienten krank werden als ohne Impfung. Die Impfung ist aber auch erfolgreich, wenn mehr Patienten gesund bleiben als ohne Impfung. Hieraus ergeben sich zwei Arten die Wirksamkeit des Impfstoffs zu testen:

1. Möglichkeit:

H_1: Mit Grippeimpfung erkranken weniger Personen an Grippe (als ohne Impfung), d.h. $p<4/5$. Daraus folgt

H_0: Die Grippeimpfung hat keine vor Grippe schützende Wirkung, d.h. $p=4/5$

2. Möglichkeit:

H_1: Mit Grippeimpfung bleiben mehr Personen gesund (als ohne Impfung).

Unüblich (aber rechnerisch ebenso legitim) ist die Hypothese, dass ohne Impfung mehr Patienten krank werden als mit Impfung. Denn es lässt sich ethisch kaum vertreten, dass man eine Gruppe nicht impft um zu zeigen, dass mehr Personen krank werden (denkbar wäre höchstens, dass Medikament auf dem Markt ist und seine Wirksamkeit angezweifelt wird). Alternativ können ohne Impfung aber auch weniger Patienten gesund bleiben als mit Impfung. Demzufolge sind folgende Hypothesentest denkbar:

3. Möglichkeit:

H_1: Ohne Grippeimpfung erkranken mehr Personen an Grippe (als mit Impfung).

4. Möglichkeit:

H_1: Ohne Grippeimpfung bleiben weniger Personen gesund (als mit Impfung).

Aufgaben:

a) Wie sind für die einzelnen Möglichkeiten die Bedingungen für p zu H_1 und entsprechend H_0 zu wählen? Wie lautet jeweils H_0?
b) Ist die Wirksamkeit der Grippeimpfung auf dem 5%-Niveau für alle Möglichkeiten gesichert?
c) Früher wurde Kindern häufig von den Eltern gesagt, dass sie nach dem Essen von Kernobst (z.B. Äpfeln) nichts trinken sollten, da sie sonst Bauchweh bekämen. Angenommen, man würde einen Test mit 300 Kindern durchführen, die alle 3 Äpfel essen. 200 Kinder sind vorsichtig und möchten nach dem Verzehr nichts trinken, 100 Kinder trinken danach. Von den Kindern die nichts getrunken haben klagten 70 Kinder nach dem Verzehr über Bauchweh. Von den Kindern, die nach den Äpfeln tranken, meinten 40 Bauchweh zu verspüren.
   Bilde und prüfe alle 4 Möglichkeiten, ob die Elternweisheit auf Grundlage der Daten auf dem 5%-Niveau haltbar ist.

Mögliche Lösungen mit Hilfe der BV und des GTR

a) **Zur 2. Möglichkeit**:

H_1: Mit Grippeimpfung bleiben mehr Personen gesund (als ohne Impfung), d.h. p>1/5. Daraus folgt

H_0: Die Grippeimpfung hat keine vor Grippe schützende Wirkung, d.h. p=1/5

Zur 3. Möglichkeit:

H_1: Ohne Grippeimpfung erkranken mehr Personen an Grippe (als mit Impfung), d.h. p>3/10. Daraus folgt

H_0: Nicht zu impfen verschlechtert die Erkrankungschancen nicht, d.h. p=3/10

Zur 4. Möglichkeit:

H_1: Ohne Grippeimpfung bleiben weniger Personen gesund (als mit Impfung). d.h. p<7/10. Daraus folgt

H_0: Nicht zu impfen verschlechtert die Erkrankungschancen nicht, d.h. p=7/10

b)

60 von 200 geimpften erkranken, 140 von 200 geimpften bleiben gesund

40 von 50 nicht geimpften erkranken, 10 von 50 nicht geimpften bleiben gesund

Zur 1. Möglichkeit: Mit Grippeimpfung erkranken weniger (60 von 200) als ohne Impfung (40 von 50=4/5). Also gilt H_0: p=4/5

Es gilt $\sum_{k=0}^{60}\binom{200}{k}\cdot\left(\frac{4}{5}\right)^k\cdot\left(\frac{1}{5}\right)^{200-k}$ = binomcdf(200,4/5,0,60) ist fast 0, also deutlich unter 5%.

Also muss H_0 mit einer Irrtumswkt. von 5% verworfen werden.

Zur 2. Möglichkeit: Mit Grippeimpfung sind mehr Personen gesund (140 von 200) als ohne Impfung (10 von 50=1/5). Also gilt H_0: p=1/5

Also: $\sum_{k=140}^{200}\binom{200}{k}\cdot\left(\frac{1}{5}\right)^k\cdot\left(\frac{4}{5}\right)^{200-k}$ = binomcdf(200,1/5,140,200) ist fast 0, also deutlich unter 5%.

Also muss H_0 wieder mit einer Irrtumswkt. von 5% verworfen werden.

Zur 3. Möglichkeit: Ohne Grippeimpfung erkranken mehr Personen (40 von 50) als mit Impfung (60 von 200). Also gilt H_0: p=3/10

Also: $\sum_{k=40}^{50}\binom{50}{k}\cdot\left(\frac{3}{10}\right)^k\cdot\left(\frac{7}{10}\right)^{50-k}$ = binomcdf(50,3/10,40,50) ist fast 0, also deutlich unter 5%.

Also muss H_0 wieder mit einer Irrtumswkt. von 5% verworfen werden.

Zur 4. Möglichkeit: Ohne Grippeimpfung sind weniger Personen gesund (10 von 50) als mit Impfung (140 von 200). Also gilt H_0: p=7/10

Also: $\sum_{k=0}^{10}\binom{50}{k}\cdot\left(\frac{7}{10}\right)^k\cdot\left(\frac{3}{10}\right)^{50-k}$ = binomcdf(50,7/10,0,10) ist fast 0, also deutlich unter 5%.

Also muss H_0 wieder mit einer Irrtumswkt. von 5% verworfen werden.

c) Grundsätzlich: H_1 zu bilden macht in diesem Fall natürlich nur dann Sinn, wenn man z.B. auf Grundlage von Erfahrung annimmt, dass man nach dem Verzehr von Kernobst besser nichts trinkt.

70 von 200 Kinder die nicht trinken haben Bauchweh, 130 von 200 nicht

40 von 100 Kindern die trinken haben Bauchweh, 60 von 100 nicht

1. Möglichkeit:

H_1: Ohne Trinken haben weniger Kinder Bauchweh (70 von 200) als mit trinken (40 von 100=0,4), also p<0,4. Daraus folgt

H_0: Nicht zu trinken hat keinen Einfluss auf Bauchweh, d.h. p=0,4.

Es gilt $\sum_{k=0}^{70}\binom{200}{k}\cdot 0{,}4^k \cdot 0{,}6^{200-k}$ = binomcdf(200,0.4,0,70) ≈ 8,4%, also deutlich über 5%. Also kann H_0 nicht verworfen werden.

2. Möglichkeit:

H_1: Ohne Trinken sind mehr Kinder beschwerdefrei (130 von 200) als wenn sie trinken (60 von 100=0,6), d.h. p>0,6. Daraus folgt

H_0: Nicht zu trinken hat keinen Einfluss auf Schmerzfreiheit, d.h. p=0,6.

Es gilt $\sum_{k=130}^{200}\binom{200}{k}\cdot 0{,}6^k \cdot 0{,}4^{200-k}$ = binomcdf(200,0.6,130,200) ≈ 8,4%, also deutlich über 5%. Also kann H_0 wieder nicht verworfen werden.

3. Möglichkeit:

H_1: Mit Trinken haben mehr Kinder Bauchweh (40 von 100) als wenn sie nicht trinken (70 von 200), d.h. p>7/20. Daraus folgt

H_0: Trinken hat keinen Einfluss auf Bauchweh, d.h. p=7/20.

Es gilt $\sum_{k=40}^{100}\binom{100}{k}\cdot\left(\frac{7}{20}\right)^k\cdot\left(\frac{13}{20}\right)^{100-k}$ = binomcdf(100,7/20,40,100) ≈ 17,2%, also deutlich über 5%. Also kann H_0 wieder nicht verworfen werden.

4. Möglichkeit:

H_1: Mit Trinken sind weniger Kinder beschwerdefrei (60 von 100) als wenn sie nicht trinken (130 von 200), d.h. p<13/20. Daraus folgt

H_0: Trinken hat keinen Einfluss auf Bauchweh, d.h. p=13/20.

Es gilt $\sum_{k=0}^{60}\binom{100}{k}\cdot\left(\frac{13}{20}\right)^k\cdot\left(\frac{7}{20}\right)^{100-k}$ = binomcdf(100,13/20,0,60) ≈ 17,2%, also deutlich über 5%. Also kann H_0 wieder nicht verworfen werden.