Tests: Sensitivität, Spezifität und der Prävalenzfehler

Ralf Blendowske

25. Mai 2021

Zusammenfassung

Wie groß ist die Wahrscheinlichkeit, dass ich an COVID19 erkrankt bin, wenn ich ein positives Testergebnis habe? Bei seltenen Erkrankungen, wie Covid19, wird die Wahrscheinlichkeit nach einem positivem Test tatsächlich erkrankt zu sein, oft überschätzt; denn die Antwort hängt stark von der Verbreitung der Krankheit ab. Der sogenannte Prävalenzfehler vernachlässigt diesen Einfluss. Die Größen Sensitivität und Spezifität werden als entscheidende Parameter von Tests an einigen Beispielen diskutiert und ihre Vorhersagekraft mit Bezug auf die Prävalenz erläutert.

1 Modell-Beispiel

Eine Gruppe von $N = 20$ Menschen wird auf eine Krankheit getestet. Das Ergebnis ist binär, und der Test ist daher entweder positiv (p) oder negativ (n). Ein Viertel der Menschen ist tatsächlich krank ($N_K = 4$) und der Rest ($N_G = 16$) gesund. Die Prävalenz beträgt $\mathrm{Pr} = 4/16 = 0{,}25$. Der Test erkennt $r_p = 3$ der Erkrankten richtig als positiv und einen fälschlicherweise nicht, $f_n = 1$. Bei den Gesunden werden $f_p = 3$ fälschlicherweise als positiv getestet , und $r_n = 13$ Personen haben richtigerweise ein negatives Ergebnis, siehe Abb. 1. Die Tabelle fasst diese Ergebnisse zusammen

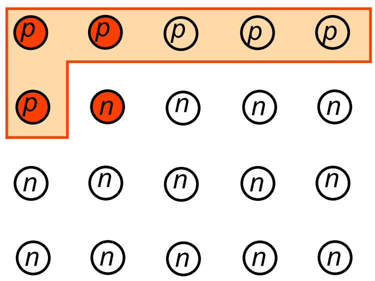


Abbildung 1: Beispiel für die Verbreitung einer Krankheit (rot) in einer Gruppe von 20 Menschen mit den Testergebnissen positiv (p) und negativ (n). Markiert sind alle 6 positiven Testergebnisse, wovon 3 falsch positiv sind.

	K	G	$\sum$
p	3	3	6
n	1	13	14
$\sum$	4	16	20

	K	G	$\sum$
p	r_p	f_p	N_p
n	f_n	r_n	N_n
$\sum$	N_K	N_G	N

Tabelle 1: Ergebnisse des Tests: der Status ist (K)rank oder (G)esund. Das Testergebnis ist (p)ositiv oder (n)egativ. Rechts wird die Tabelle mit den Variablen für die (r)ichtigen und (f)alschen Testergebnisse und den Anzahlen der Kranken N_K bzw. Gesunden N_G gezeigt. Die Summe der positiven N_p und der negativen N_n Testergebnisse stehen in der letzten Spalte.

2 Kenngrößen

Die **Prävalenz** bezeichnet den Anteil der Kranken N_K an der Gesamtanzahl N der Population

$$\mathrm{Pr} = \frac{N_K}{N}$$

Im Hinblick auf die Testergebnisse ergibt sich die Anzahl der Kranken sich aus der Summe der richtig positiv und der falsch negativ getesteten Personen

$$N_K = r_p + f_n$$

Analog folgt für die Anzahl der Gesunden

$$N_G = r_n + f_p$$

Die Güte eines Tests wird durch zwei Kenngrößen beschrieben. Wieviele Kranke tatsächlich als krank erkannt werden, das wird durch die **Sensitivität** beschrieben. Sie ist die Verhältniszahl der richtig als krank getesteten zur Anzahl der Kranken insgesamt

$$\mathrm{Se} = \frac{r_p}{N_K} = \frac{r_p}{r_p + f_n}$$

Diese Größe wird auch als **Richtig-Positiv-Rate** bezeichnet.

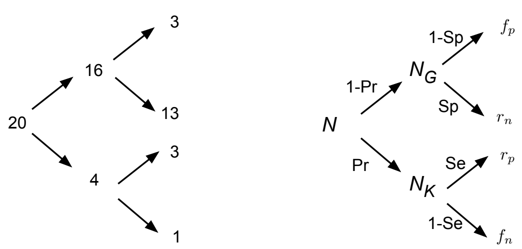
Der Anteil, der als richtig erkannten Gesunden, wird durch die **Spezifität** beschrieben

$$\mathrm{Sp} = \frac{r_n}{N_G} = \frac{r_n}{r_n + f_p}$$

Eine andere Bezeichnung ist **Falsch-Negativ-Rate**.

In unserem Beispiel ergeben sich diese Kennzahlen zu $P = 4/20 = 0{,}2$, $\mathrm{Se} = 3/4 = 0{,}75$ und $\mathrm{Sp} = 13/16 = 0{,}8125$.

Die Ergebnisse des Beispiels können allgemein durch einen Verzweigungsbaum beschrieben werden



Die Wahrscheinlichkeit bei einem positiven Test tatsächlich erkrankt zu sein, berechnet sich aus dem Verhältnis der richtig positiven Fälle zur Gesamtzahl der positiven Fälle

$$W_K = \frac{3}{3+3} = 0{,}50$$

Das Ergebnis liegt deutlich unter der Sensitivität von 0,75.

Aus dem Verzweigungsbaum kann man die allgemeine Gleichung für W_K ableiten:

$$W_K = \frac{r_p}{r_p + f_p} = \frac{\mathrm{Pr} \cdot \mathrm{Se}}{\mathrm{Pr} \cdot \mathrm{Se} + (1 - \mathrm{Pr})(1 - \mathrm{Sp})}$$

Für sehr kleine Prävalenzen kann die folgende Näherung verwendet werden

$$W_K \underset{Pr<<1}{=\longrightarrow} \frac{Pr \cdot \mathrm{Se}}{(1 - \mathrm{Sp})}$$

Der (Vor)-Faktor Pr zeigt, dass die Wahrscheinlichkeit tatsächlich erkrankt zu sein, wenn ein positiver Test vorliegt, entscheidend von der Prävalenz abhängt. Die fälschliche Vernachlässigung dieses Faktors heißt **Prävalenzfehler**.

Analog gilt für die Wahrscheinlichkeit tatsächlich gesund zu sein, wenn ein negativer Test vorliegt

$$W_G = \frac{r_n}{r_n + f_n} = \frac{(1 - \mathrm{Pr})\mathrm{Sp}}{\mathrm{Pr} \cdot (1 - \mathrm{Se}) + (1 - \mathrm{Pr})\mathrm{Sp}}$$

Wiederum kann für sehr kleine Prävalenzen die folgende Näherung verwendet werden

$$W_G \underset{Pr<<1}{=\longrightarrow} 1 - \frac{\mathrm{Pr} \cdot (1 - \mathrm{Se})}{\mathrm{Sp}}$$

Je kleiner die Prävalenz ist, desto größer ist die Wahrscheinlichkeit bei einem negativen Test auch tatsächlich gesund zu sein.

Die Größen W_K und W_G heißen auch **positive** und **negative Vorhersagewerte**.

3 Reales Beispiel

Für einen Laien-Selbsttest gibt die Firma ViroMed die Gütezahlen ihres Tests an:

$$\mathrm{Se} = 0{,}9538 \quad \mathrm{Sp} = 0{,}9960$$

Für die Berechnungen wird die Prävalenz aus der 7-Tages-Inzidenz, zunächst unter Vernachlässigung der Dunkelziffer, geschätzt. Als Beispiel kann etwa der Wert $\mathrm{Pr} = 60/100000$ gelten.

Mit diesen numerischen Werten beträgt die Wahrscheinlichkeit bei einem positiven Test tatsächlich krank zu sein

$$W_K = 0{,}13$$

Die Wahrscheinlichkeit bei einem negativen Test dennoch krank zu sein ergibt sich zu

$$1 - W_G = 2{,}8 \cdot 10^{-5}$$

Bei einer Verdreifachung der Prävalenz $\mathrm{Pr} = 180/100000$ steigen diese Werte wie folgt an

$$W_K = 0{,}36 \quad 1 - W_G = 1{,}1 \cdot 10^{-4}$$

Näherungsweise folgt die Erhöhung der Werte der Prävalenz um einen Faktor 3. Als Funktion der Prävalenz zeigt die Abb. 2 den Verlauf der Wahrscheinlichkeit W_K.

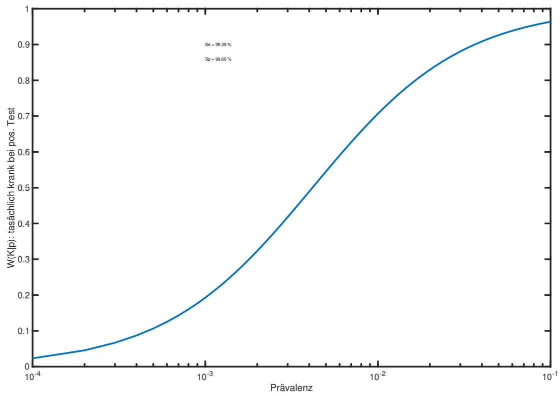


Abbildung 2: Die Wahrscheinlichkeit nach einem positiven Test tatsächlich krank zu sein als Funktion der Prävalenz

Der größte Anstieg ist beim Übergang vom Promille- zum Prozentbereich der Prävalenz zu beobachten. Die große Wert der Sensitivität legt einen hohen Wert W_K nahe. Dies ist aber nur im Fall hoher Prävalenzen der Fall. Mit fallendem Anteil der Erkrankten an der Gesamtbevölkerung verliert das Testergebnis seine Bedeutung. In diesen Fällen werden oft zusätzliche Informationen wie etwa das Vorliegen von Symptomen benötigt, um die A-Priori-Wahrscheinlichkeit zu erhöhen, indem die Gesamtbevölkerung auf einen relevanten Anteil eingeschränkt wird. Wenn bei Vorliegen von Symptomen oder Kontakten zu Erkrankten, die Wahrscheinlichkeit beispielsweise 5% beträgt, dass man erkrankt ist, dann steigt die „Prävalenz" auf diesen Wert und der Test hat eine ganz andere Aussagekraft.

4 Satz von Bayes

Der Satz von Bayes beschreibt die Beziehung zwischen bedingten Wahrscheinlichkeiten

$$P(A|B)P(B) = P(B|A)P(A)$$

für die beiden Ereignisse A und B. Die Größe $P(A|B)$ liest sich so: Wahrscheinlichkeit für das Eintreten von A unter der Bedingung von B, d.h. dass B eingetreten ist.

Uns interessiert die Fragestellung: wenn ein positiver Test (p) gegeben ist, wie groß ist dann die Wahrscheinlichkeit, dass die Person auch krank (K) ist, als Formelzeichen: $W(K|p)$?

Der Satz von Bayes lautet für diese Größen

$$W(K|p) = \frac{W(p|K)}{W(p)} W(K)$$

Dabei bedeutet $W(p|K)$ die Wahrscheinlichkeit, dass bei gegebener Krankheit der Test positiv ist, das ist aber die Sensitivität. Die Wahrscheinlichkeit einen Kranken zu finden $W(K)$ ist die Prävalenz. Sie beschreibt die A-Priori-Wahrscheinlichkeit einer Erkrankung, bevor („a priori“) ein Testergebnis vorliegt. Nach einem Test („a posteriori“) liegt eine zusätzliche Information vor, die zu einer Neuberechnung der Wahrscheinlichkeit führt. Die so erhalten Größe heißt A-posteriori-Wahrscheinlichkeit. Das kann $W(K|p)$ oder $W(K|n)$ sein. Wenn wir das Testergebnis als „Evidenz“ bezeichnen, dann beschreibt $W(K|p)$ die Wahrscheinlichkeit den Zustand „Krank“ bei gegebener Evidenz „p“. Im Unterschied dazu wird die Wahrscheinlichkeit $W(p|K)$ für das Vorliegen der Evidenz p unter der Bedingung des Zustandes K als „Likelihood“ bezeichnet.

Um die A-posteriori-Wahrscheinlichkeit nach einem positiven Test berechnen zu können benötigen wir die Wahrscheinlichkeit $W(p)$ für einen positiven Test. Sie ergibt sich durch die vollständige Zerlegung

$$W(p) = W(p|K)W(K) + W(p|G)W(G)$$

Diese Größe kann auch als Normierungsfaktor bezüglich aller Zustände, hier K und G, aufgefasst werden.

Damit erhalten wir durch Einsetzen der entsprechenden Bezeichnungen

$$W(K|p) = \frac{\mathrm{Se} \cdot \mathrm{Pr}}{\mathrm{Se} \cdot \mathrm{Pr} + (1 - \mathrm{Sp}) \cdot (1 - \mathrm{Pr})}$$

in Übereinstimmung mit dem Ergebnis aus dem Verzweigungsbaum.

5 Beispiel: Gesichtserkennung

Mittels eines Bildverarbeitungssystems sollen Personen überwacht werden, um gesuchte Straftäter zu finden. Bei der Gesichtserkennung liegt die Falsch-Positiv-Rate $(1 - \mathrm{Se})$ bei 1% und die Falsch-Negativ-Rate $(1 - \mathrm{Sp})$ ebenfalls bei 1%. Wenn sich unter 10000 überwachten Personen im Mittel 1 gesuchte Person befindet (Pr), wie groß ist dann die Wahrscheinlichkeit, dass bei einem Treffer (T) der Anlage tatsächlich ein Straftäter (S) gefunden wurde?

Die Wahrscheinlichkeit berechnet sich aus

$$\begin{aligned} W(S|T) &= \frac{\mathrm{Se} \cdot \mathrm{Pr}}{\mathrm{Se} \cdot \mathrm{Pr} + (1 - \mathrm{Sp}) \cdot (1 - \mathrm{Pr})} \\ &= \frac{0{,}99 \cdot 10^{-4}}{0{,}99 \cdot 10^{-4} + 0{,}01 \cdot (1 - 0{,}99 \cdot 10^{-4})} = 9{,}9 \cdot 10^{-3} \approx 1\% \end{aligned}$$

Selbst wenn sich unter 10000 Personen keine Straftäter befinden, wird es dennoch 100 „Treffer“ geben.