Bedingte Wahrscheinlichkeit (2)

- Bei einem bestimmten Stamm weißer Mäuse tritt mit einer Wahrscheinlichkeit von p=0,1 ein Leberschaden ein. Außerdem weiß man, dass bei Mäusen mit Leberschaden in 2 von 3 Fällen Haarausfall auftritt, bei Mäusen ohne Leberschaden jedoch nur in einem von 9 Fällen.

- Wie groß ist die Wahrscheinlichkeit, dass bei einer Maus mit Haarausfall ein Leberschaden vorliegt?

$$P(LS) = 0{,}1 = \frac{1}{10}$$

$$P(HA \mid LS) = \frac{2}{3}$$

$$P(HA \mid \overline{L}\,\overline{S}) = \frac{1}{9}$$

Bedingte Wahrscheinlichkeit (3)

Satz von Bayes:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})}$$

Bedingte Wahrscheinlichkeit (4)

Satz von Bayes:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})}$$

$$P(LS \mid HA) = \frac{P(HA \mid LS) \cdot P(LS)}{P(HA \mid LS) \cdot P(LS) + P(HA \mid \bar{L}\bar{S}) \cdot P(\bar{L}\bar{S})}$$

$$= \frac{\frac{2}{3} \cdot \frac{1}{10}}{\frac{2}{3} \cdot \frac{1}{10} + \frac{1}{9} \cdot \frac{9}{10}}$$

$$= 0{,}400 = 40{,}0\%$$

Kenngrößen (Gütekriterien) der Tests (4)

	T-	T+	Zeilensumme
D-	RN $P(T-\cap D-)$	FP $P(T+\cap D-)$	RN+FP $P(D-)$
D+	FN $P(T-\cap D+)$	RP $P(T+\cap D+)$	FN+RP $P(D+)$

Sensitivität = $$P(T+|D+) = \frac{P(T+\cap D+)}{P(D+)} = \frac{RP}{(FN+RP)}$$

- Wahrscheinlichkeit für positiven Test bei tatsächlich Erkrankten
- Wahrscheinlichkeit für richtig positiven Test
- Bewertet die Empfindlichkeit des Tests für die Diagnose

Kenngrößen (Gütekriterien) der Tests (5)

	T-	T+	Zeilensumme
D-	RN $P(T-\cap D-)$	FP $P(T+\cap D-)$	RN+FP $P(D-)$
D+	FN $P(T-\cap D+)$	RP $P(T+\cap D+)$	FN+RP $P(D+)$

Spezifität = $$P(T-|D-) = \frac{P(T-\cap D-)}{P(D-)} = \frac{RN}{(RN+FP)}$$

- Wahrscheinlichkeit für negativen Test bei tatsächlich Nicht-Erkrankten
- Wahrscheinlichkeit für richtig negativen Test
- Bewertet die Eingrenzung des Tests auf die Diagnose, für die er charakteristisch sein soll

Prädiktive Werte (7)

	T-	**T+**
D-	RN $P(T-\cap D-)$	FP $P(T+\cap D-)$
D+	FN $P(T-\cap D+)$	RP $P(T+\cap D+)$
Spaltensumme	RN+FN $P(T-)$	FP+RP $P(T+)$

PPW =

$$P(D+|T+) = \frac{P(T+\cap \mathrm{D}+)}{P(T+)} = \frac{RP}{(FP+RP)}$$

- ⇨ Wahrscheinlichkeit, dass die Erkrankung (Diagnose) vorliegt, wenn der Test positiv ist
- ⇨ *Vorhersage*

Prädiktive Werte (8)

	T-	T+
D-	RN $P(T- \cap D-)$	FP $P(T+ \cap D-)$
D+	FN $P(T- \cap D+)$	RP $P(T+ \cap D+)$
Spaltensumme	RN+FN $P(T-)$	FP+RP $P(T+)$

NPW =

$$P(D-|T-) = \frac{P(T- \cap \mathrm{D}-)}{P(T-)} = \frac{RN}{(RN+FN)}$$

⇨ Wahrscheinlichkeit, dass die Erkrankung (Diagnose) ausgeschlossen ist, wenn der Test negativ ist.

⇨ *Vorhersage*

Ein erfundenes Beispiel (3)

Auf Ihre Anfrage erfahren Sie

zum Test:

- Sensitivität: Bei 99 von 100 Infizierten spricht der Test an.
- Falsch-positiv-Rate:
  Auch bei 2 von 100 Nichtinfizierten spricht der Test an.

zur Infektionshäufigkeit:

- Prävalenz: Unter Touristen infiziert sich von 1000 einer.

Ein erfundenes Beispiel (7)

	Test negativ (T-)	Test positiv (T+)	Zeilensumme
Nicht infiziert (D-)			**99900**
Infiziert (D+)			**100**
Spaltensumme			**100 000**

Eine "genügend große" Kohorte wählen:

Wegen Prävalenz = $0{,}001 \rightarrow N = 100\,000$,

also im Mittel 100 angesteckt (D+).

Ein erfundenes Beispiel (8)

	Test negativ (T-)	Test positiv (T+)	Zeilensumme
Nicht infiziert (D-)			99900
Infiziert (D+)	**1**	**99**	100
Spaltensumme			100 000

Sensitivität ist 99%, also
ist der Test bei 99 von 100 positiv.

Ein erfundenes Beispiel (9)

	Test negativ (T-)	Test positiv (T+)	Zeilensumme
Nicht infiziert (D-)	**97 902**	**1 998**	99900
Infiziert (D+)	1	99	100
Spaltensumme			100 000

Es gibt 2% falsch-Positive unter den nicht-Infizierten,
das sind 1 998.
Die verbleibenden 97 902 sind richtig-Negative.

Ein erfundenes Beispiel (10)

	Test negativ (T-)	Test positiv (T+)	Zeilensumme
Nicht infiziert (D-)	97 902	1 998	99900
Infiziert (D+)	1	99	100
Spaltensumme	**97 903**	**2097**	100 000

Im 4. Schritt werden die Spaltensummen gebildet, die sich selbstverständlich zu 100 000 summieren.

Ein erfundenes Beispiel (11)

	Test negativ (T-)	Test positiv (T+)	Zeilensumme
Nicht infiziert (D-)	97 902	1 998	99900
Infiziert (D+)	1	**99**	100
Spaltensumme	97 903	**2097**	100 000

Positiv-prädiktiver Wert (PPW):

99 von 2 097 positiv Getesteten haben wirklich die Krankheit D, das sind

$\mathrm{PP}W = P(D+|T+) = \frac{P(D+ \cap T+)}{P(T+)} = \frac{99}{2097} = 0{,}047$, also 4,7%.

Sie können sich also ziemlich sicher fühlen.

Ein erfundenes Beispiel (13)

Wie sicher ist ein negativer Testausgang in diesem Beispiel?

	Test negativ (T-)	Test positiv (T+)	Zeilensumme
Nicht infiziert (D-)	**97 902**	1 998	99900
Infiziert (D+)	1	99	100
Spaltensumme	**97 903**	2097	100 000

Unter 97 903 Testnegativen ist nur ein Infizierter zu erwarten. Damit ist die gesuchte Wahrscheinlichkeit, bei **negativem** Testausgang tatsächlich **nicht** infiziert zu sein (NPW), gegeben durch $P(D- \mid T-) = \frac{P(D- \cap T-)}{P(T-)} = \frac{97902}{97903} \approx 1$, also 100%. Bei negativem Testausgang kann man sich also sehr sicher fühlen.

Weitere Aufgabe (3)

	CLO-Test negativ (T-)	CLO-Test positiv (T+)	Summe
Nicht infiziert (D-)	**7760**	**240**	**8000**
Infiziert (D+)	**40**	**1960**	**2000**
Summe	**7800**	**2200**	**10000**

Sensitivität =
Spezifität =
PPW =
NPW =

Weitere Aufgabe (4)

	CLO-Test negativ (T-)	CLO-Test positiv (T+)	Summe
Nicht infiziert (D-)	**7760**	**240**	**8000**
Infiziert (D+)	**40**	**1960**	**2000**
Summe	**7800**	**2200**	**10000**

Sensitivität = 1960/2000 = 0,980 = 98,0%
Spezifität = 7760/8000 = 0,970 = 97,0%
PPW = 1960/2200 = 0,891 = 89,1%
NPW = 7760/7800 = 0,995 = 99,5%