U.M.

Sensitivität und Spezifität

S sei ein Symptom oder allgemein ein Indikator für ein zu diagnostizierendes Merkmal M. Dann sind die bedingten Wahrscheinlichkeiten $P(S|M)$ und $P(\neg S|\neg M)$ die *Sensitivität* bzw. die *Spezifität* von S. Diese Begriffe werden häufig anhand einer 4-Felder-Tafel illustriert:

Tabelle 1: Vierfelderdarstellung von Sensitivität und Spezifität ($N = a + b + c + d$)

	Merkmal		
Symptom	+	−	Σ
+	a	b	a + b
−	c	d	c + d
Σ	a + c	b + d	N

Die Anwendung der Formel für die bedingte Wahrscheinlichkeit

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

liefert dann die Formeln

$$P(S|M) = \frac{a}{a+c}, \quad P(\neg S|\neg M) = \frac{d}{b+d} \tag{1}$$

Offenbar ist $P(S|M) = 0$ für $a = 0$, $P(\neg S|\neg M) = 0$ für $d = 0$, und $P(S|M) = 1$ für $a > 0$, $c = 0$, und $P(\neg S \neg M) = 1$ für $d > 0$ und $b = 0$. Also können sowohl die Sensitivität wie auch die Spezifität beide "sehr groß" werden, d.h. Werte nahe 1 oder gleich 1 annehmen, wie auch "sehr klein" werden, d.h. Werte nahe 0 oder gleich 0 annehmen.

Die Illustration der Sensitivität und der Spezifität anhand der Tabelle 1 suggeriert eine Alles-oder-nichts-Art des Vorhandenseins von S bzw. M: entweder S ist vorhanden oder nicht, und M it entweder vorhanden oder nicht. Diese Vorstellung von S und M ist aber oft nicht gerechtfertigt; sowohl S wie auch M können mehr oder weniger ausgeprägt sein. Gegeben sei ein Test, dessen Durchführung zu einem Score X führt. Technisch gesehen ist X eine zufällige Veränderliche, von deren Wert auf das Vorhandensein von M geschlossen werden soll, – die Situation ist also vergleichbar mit der, in der für eine "Messung" entschieden werden soll, ob sie nur "Rauschen" oder "Signal plus Rauschen" repräsentiert. Die diagnostizierende Person muß sich für eine der beiden Möglichkeiten entscheiden. Sie kann bzw. muß dazu einen kritischen Wert x_c wählen und gemäß

$$X \leq x_c \Rightarrow \text{Rauschen}, \quad X > x_c \Rightarrow \text{Signal + Rauschen} \tag{2}$$

entscheiden. Dann ist[1]

$$P(\neg S|\neg M) = \int_{-\infty}^{x_c} f_n(x)dx \tag{3}$$

$$P(S|M) = \int_{x_c}^{\infty} f_{sn}(x)dx = 1 - \int_{-\infty}^{x_c} f_{sn}(x)dx \tag{4}$$

f_n ist die Dichte von X, wenn kein Signal vorhanden ist, f_{sn} ist die Dichte von X, wenn ein Signal vorhanden ist (signal + noise).

Für f_n bzw. f_{sn} werden häufig Gauß-Verteilungen (Normalverteilungen) angenommen:

$$f_n = \frac{1}{\sigma_n\sqrt{2\pi}} \exp\left(-\frac{(x-\mu_n)^2}{2\sigma_n^2}\right), \quad f_{sn}(x) = \frac{1}{\sigma_{sn}\sqrt{2\pi}} \exp\left(-\frac{(x-\mu_{sn})^2}{2\sigma_{sn}^2}\right) \tag{5}$$

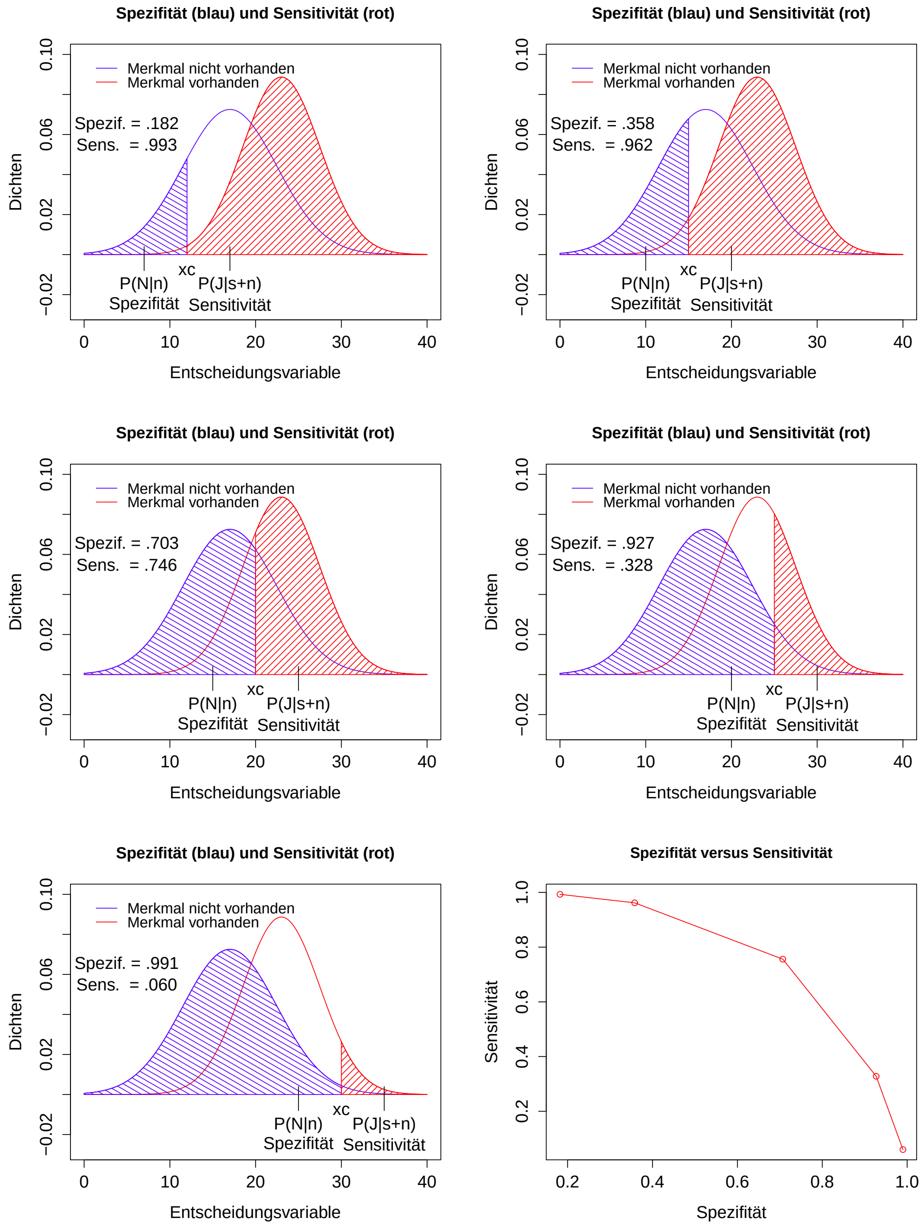
mit $\mu_n = \mathbb{E}(X|n)$, $\sigma_n^2 = Var(X|n)$, $\mu_{sn} = \mathbb{E}(X|sn)$, $\sigma_{sn}2 = Var(X|sn)$. Diagnostiker müssen nun entscheiden, welchen Wert für x_c sie wählen wollen. Alle enstprechenden Betrachtungen der SDT (Signal-Entdeckungs-Theorie) übertragen sich auf diese Situation. Im Prinzip handelt es sich dabei um ein Abwägen der Kosten für der verschiedenen möglichen Entscheidungen, insbesondere die eines falschen Alarms und die eines "Miss", d.h. des Nichtendeckens des "Signals", also des Merkmals.

Hier stehen die Werte der Sensitivität und der Spezifität in Abhängigkeit von einem gegebenen x_c im Vordergrund. Diese Werte hängen von den Werten von σ_n^2 und σ_{sn}^2 relativ zur Differenz $\mu_{sn} - \mu_n$ ab, d.h. vom Ausmaß der Überlappung der Dichten f_n und f_{sn}. Ist die Überlappung (praktisch) gleich Null, so erhält man den Fall $b = c = 0$ in Tabelle 1. Im Übrigen hängen die Werte von b und c in der Tabelle von σ_n^2 und σ_{sn}^2 ab.

Abbildung 3 zeigt, wie sich die Sensitivität und die Spezifität zueinander verhalten, wenn x_c von kleineren zu größeren Werten variiert wird. f_n und f_{sn} überlappen sich, und $\sigma_n^2 \neq \sigma_{sn}^2$. In der Abbildung rechts unten sind die Sensitivitätswerte gegen die Spezifitätswerte aufgetragen worden – eine kleine Spezifität impliziert eine große Sensitivität und umgekehrt. Für $\sigma_n^2 \neq \sigma_{sn}^2$ ist der Graph asymmetrisch, und für $\sigma_n^2 = \sigma_{sn}^2$ ist der Graph symmetrisch (vergl. Abbildung 3). Allgemein gilt: je weniger sich die Dichten überlappen, desto mehr ähnelt die Kurve der rechten Kurve in Abbildung 3.

[1] Für diejenigen, die nicht mit der Differential- und Integralrechnung vertraut sind: $\int$ ist das Integralzeichen; es geht auf Leibniz zurück und ist ein lang gezogenes S – für 'Summe', denn ein Integral ist im Prinzip eine Summe von Flächeninhalten von Rechtecken der Höhe $f(x)$ und der Breite dx, wobei allerdings dx ein Symbol ist. Für bestimmte Funktionen lassen sich die Integrale explizit als Formal angeben, für viele Funktionen ist dies allerdings nicht der Fall. In diesem Fall muß man das Integral "numerisch" approximieren; dazu werden dann statt der dx tatsächlich Werte $\Delta x > 0$ für die Berechnung der Flächeninhalte der Rechtecke und damit der Integrale gewählt.

Abbildung 1: Trade-off von Spezifität und Sensitivität



Nichtüberlappende Verteilungen Verteilungen ergeben sich natürllch auch, wenn die Varianzen der Verteilungen verschieden sind; im Folgenden sind der Einfachheit halber identische Varianzen gewählt worden. In diesem Fall ergeben sich (wie bei ROC-Kurven) symmetrische Graphen, wenn die Sensitivität gegen die Spezifität aufgetragen wird. Für $s_1 = s_2 = 2.5$ überlappen sich die Verteilungen, für $s_1 =$

Abbildung 2: Spezifität und Sensitivität, keine Überlappung ($\mu_n = 12, \mu_{sn} = 25$, $\sigma_n = \sigma_{ns} = .95$) für $x_c = (\mu_n + \mu_{ns})/2$)

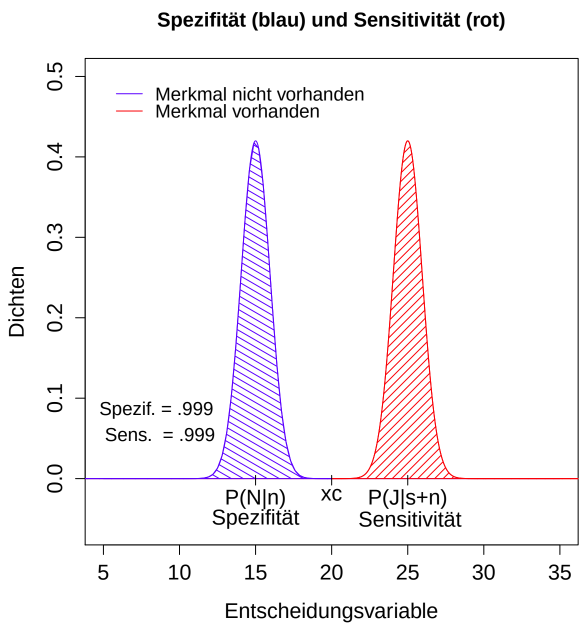
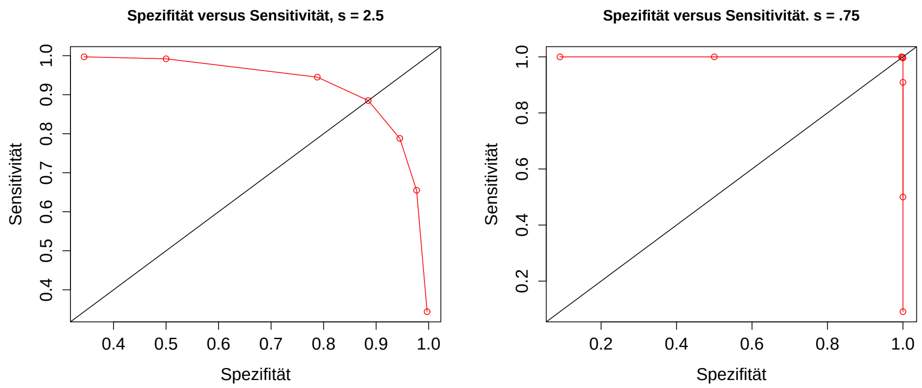


Abbildung 3: Trade-off von Spezifität und Sensitivität, II – links; Verteilungen mit Überlappung, rechts: ohne Überlappung ($\mu_n = 17, \mu_{sn} = 23$)



$s_2 = .75$ nicht; in diesem Fall ist $b = c = 0$. Bei identischen Varianzen ist die Sensitivität gleich der Spezifität, wenn $x_c = (\mu_n + \mu_{s+n})/2$, und die Sensitivität und die Spezifität nehmen einen maximalen Wert an; dieser Wert liegt auf der Diagonalen des jeweiligen Graphen.