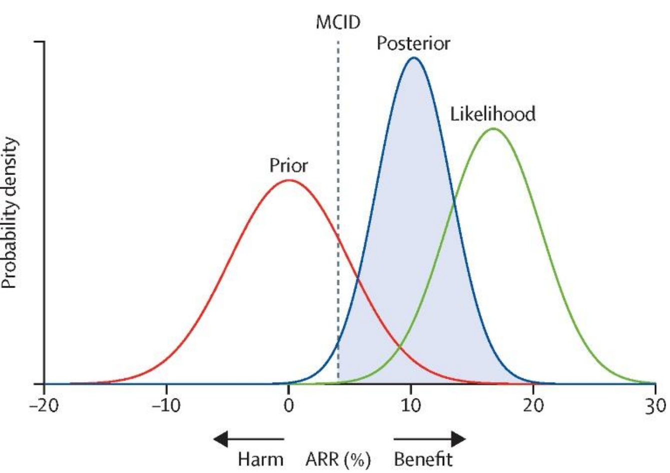
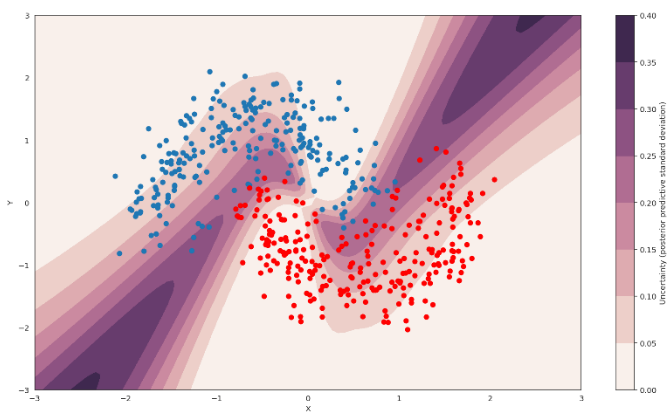
Bayes-Statistik

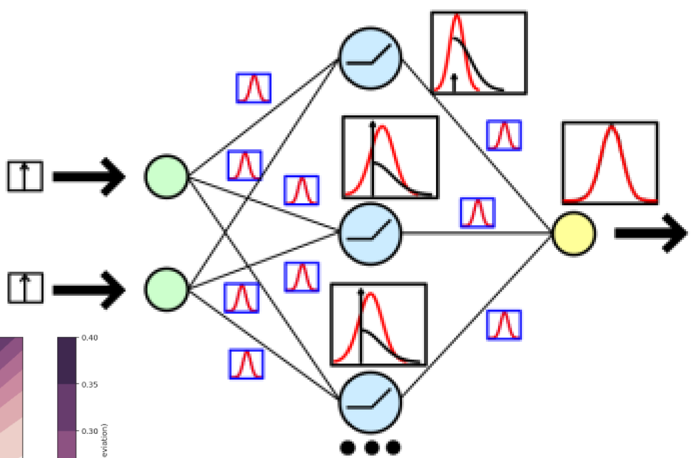
Ein Weg um Erwartungen, Glabuensgrade und Erfahrung in Acht zu nehmen

oder

Wie können wir aus ungenügende Daten entscheiden?







Gusztáv Schay
schay.gusztav@med.semmelweis-univ.hu

Der dichotomische Entscheidungsweg

Frage (von der med. Praxis)

umwandeln ins J/N Frage(n)

H0 auswahl

H0: unsere Daten gehören zu einer bekannten Referenz

Wie stark wir an der H0 glauben beeinflusst den Schwellenwert -> den nennen wir **Signifikanzniveau.**

α oder α_{max}

berechnen wir die bedingte Wahrsch. P(mindestens solche Abweichung | H0)

Entscheidung

P < sign.

P >= sign.

H0 ablehnen

H0 behalten

wird nicht abgelehnt

die Entscheidung ist mutmaßlich

Es gibt immer die Möglichkeit das wir einen Fehler machen, wegen der eingebaute Unsicherheit der Natur. Aber durchschnittlich, im langen lauf können wir optimieren.

die Entscheidung

		H_0 wird behalten	H_0 wird abgelehnt
Die Wahrheit (manchmal wird später erfahren)	H_0 ist wahr	**korrekte Entscheidung**	Type I. Fehler α
	H_0 ist falsch	Type II. Fehler β	**korrekte Entscheidung**

„Der bester Fachartzt ist der Pathologe. Er weiss alles. Leider ist es aber oft zu spät…"

wir können den maximal ertragbaren Wert von α mit dem Signifikanz einstellen.

Leistung des Tests: mit welcher Wahrscheinlichkeit wird H_0 abgelehnt wenn wir die Alternativhypothese auch genau kennen.

Alle unsere Aussagen und Beobachtungen sind eigentlich bedingt.
„hypo thesis” = „diejenige am unten, die am unten gestellte”

$P(A|H_0,C)$: nicht nur H0, aber auch die Versuchsbedingungen müssen ich Acht genommen werden

C ist schwer zu bestimmen!
-> wir brauchen *inter-subjective Vereinbarungsmöglichkeit.*

dann aber in **C** ist auch der Subjectum, Individuum beinhaltet!

P(Regen | **C**) oder P(die Operation war erfolgreich | **C**)

wir benutzen intuitive, nicht Frequenzbestimmte „Wahrscheinlichkeiten”

„Ich denke heute wird *wahrscheinlich* schönes Wetter sein.”
„keine Sorge, die Operation *sollte erfolgreich abgeschlossen sein*”

-> Können wir damit auch rechnen?

vorsicht, wir haben hier
$P(A|C)$ und $k_A|C$,
nur C lassen wir oft weg.

Häufigkeitsbestimmte Defintion: $P_A = \lim_{N\to\infty} k_A$

subjektive Wahrscheinlichkeit:
P_A = „***Glaubensgrad*** *darin, dass A passieren wird*"

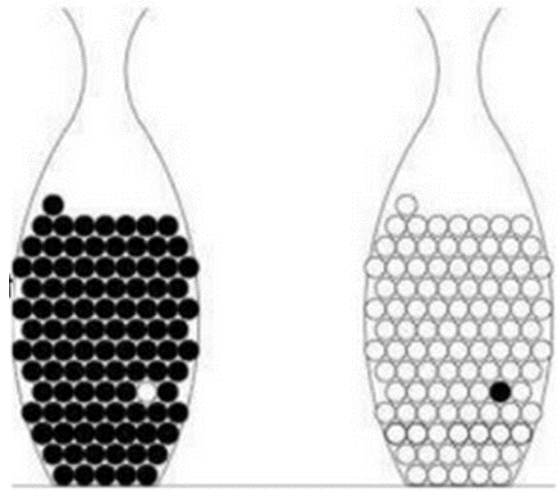
-> wir brauchen intersubjective Vereinbarung!
NUR dann ist ein *vernünftiger* P-Wert möglich
P(A|C) wird genauso bewertet von *jedem mit gleichen Kenntinssen (C)*.
(und in der gleichen Lage)

-> alle Kollegen - die den selben Laborergebnisse bekommen, und die selbe Facherfahrung und Wissen haben – werden die Wahrscheinlichkeiten über Ω (also die möglichen Diagnosewahrscheinlichkeiten) zu den selben Werten setzen.

alles ist dann relativ? wie können wir Werte vergleichen? → “Kalibration”

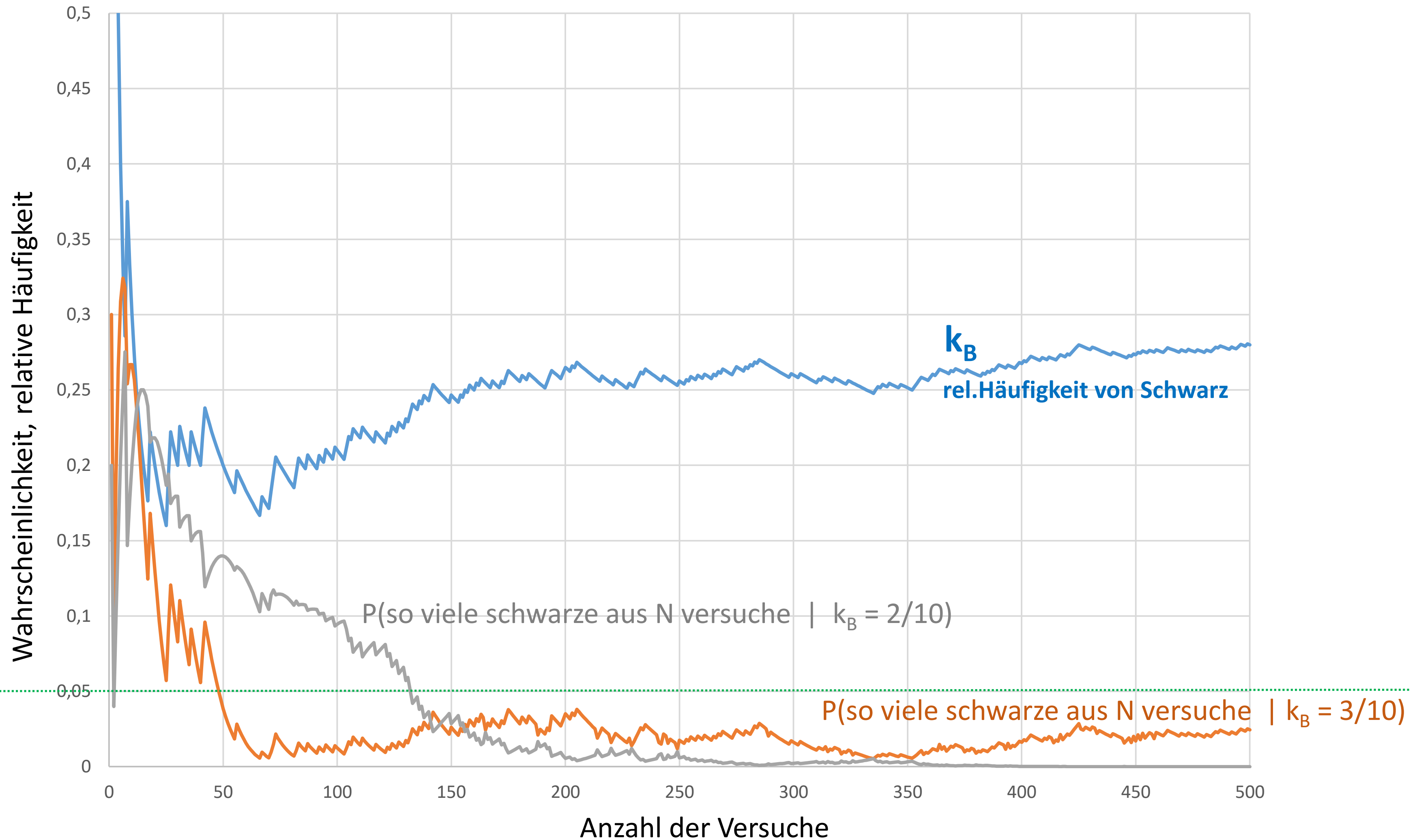
Das Urnenmodell.

alle Wahrscheinlichkeiten sind mit Urnen definierbar



P(A|C) = P(Schwarz gezogen| eine gegebene Urne)

Der dichotomische Entscheidungsweg ist manchmal problematisch...



Die Wahrheit: wir hatten eine 3/10 Urne.

algebra für Aussagen.

mit formalen Logik und vernunft können wir die Regeln aufbauen

$$\sim\sim a = a, \tag{1}$$

$$a \cdot b = b \cdot a, \quad (2) \qquad a \vee b = b \vee a, \quad (2')$$

$$a \cdot a = a, \quad (3) \qquad a \vee a = a, \quad (3')$$

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c = a \cdot b \cdot c, \tag{4}$$

$$a \vee (b \vee c) = (a \vee b) \vee c = a \vee b \vee c, \tag{4'}$$

$$\sim(a \cdot b) = \sim a \vee \sim b, \tag{5}$$

$$\sim(a \vee b) = \sim a \cdot \sim b, \tag{5'}$$

$$a \cdot (a \vee b) = a, \quad (6) \qquad a \vee (a \cdot b) = a. \quad (6')$$

~ = nicht
. = oder
v = und

(1) und noch 5 andere sind unabhängig, rest kann bewiesen werden

Nicht zu memorisieren!!

I. $(cb)|a = F(c|ba\ ,\ b|a)$

II. $\sim b|a = S(b|a)$

$F(x,y)$ und
$S(x)$ sind Funktionen zu bestimmen

-> wir suchen die einfachsten Möglichkeiten, dann:
$F(x,y) = F(x)*F(y)$, $F(x):=x$ and $S(x) := 1-x$.

-> das bringt uns zu die Kolmogorovsche Axiomen!

$b|a + \sim b|a = 1$
und
$a|a + \sim a|a = 1$
$\sim a|a = 0$
Sicherheit und Unmöglichkeit

nicht zu memorisieren!!

Die Bayes-Wahrscheinlichkeit ist breit:
keine unendliche Wiederholungsmöglichkeit ist nötig
ABER wenn die Häufigkeitsbestimmte Wahrscheinlichkeit existiert, dann gibt die Bayes-definition den selben Wert. (wegen die intersubjektive Vereinbarungsregel)

Die technische Rechnenregeln bleiben

Doch wir können mit Wahrscheinlichkeiten der Aussagen, Hypothesen auch rechnen!

1) für irgendeine **Aussage** *p haben wir* $0 \leq P(p) \leq 1$
   wenn p ist sicher dann $P(p) = 1$
   wenn p ist unmöglich $P(p)=0$
2) wenn *p* und *q* sich gegenseitig ausschließen dann $P(p \textit{ oder } q) = P(p) + P(q)$

$P(\textit{nicht-}p) = 1 – P(p)$

Wir können unsere Glaubensgrade standardisieren: → noch ein “Kalibrationsweg”
„Alles ist ein Glücksspiel oder Wette.”

Ein Glücksspiel (Lotto) wenn es *fair* gespielt wird, hat einen Preis gleich zu dem erwarteten Gewinn.

-> intersubjektive Vereinbarung:
in einem *fairen* Spiel sind die zwei Seiten im Balanz. (auf dem *langen Lauf* gewinnt und verliert niemand)

Wir setzen unsere Glaubensgrade in einem Spiel um:
Was für ein *fernünftliches* Glücksspiel würden wir auf Ereigniss A setzen?
(fernünftlich: erwarteter Balanz ist 0)

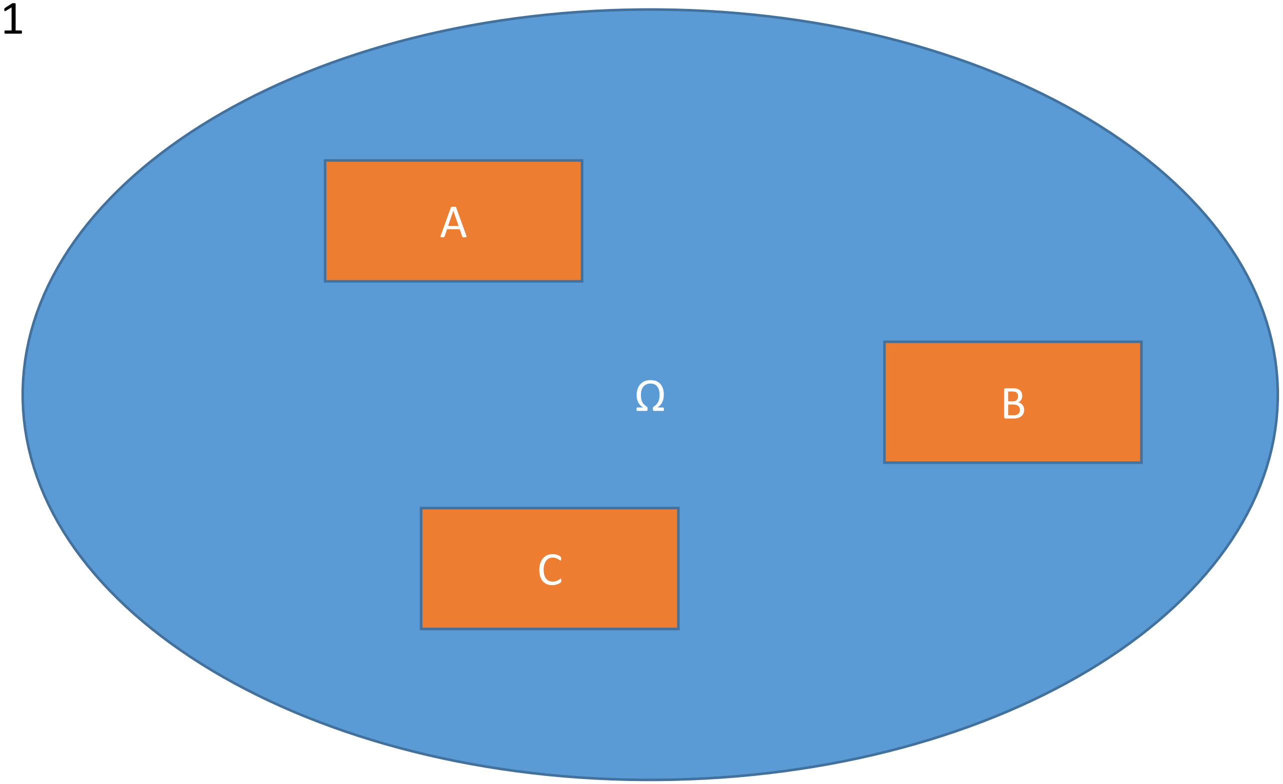
-> ähnilch zu der Urne, macht aber Vergleiche einfacher
z.B. 1:100000 ist die Wahrscheinlichkeit einer schweren Nebenwirkung.

1:x bedeutet P=1/(1+x) ist die Wahrscheinlichkeit von Ereigniss A worauf wir wetten.
n.B. :Odds war grade bei Wetten zuerst benutzt: Odds=P/(1-P) oder P=O/(1+O)

Cohärenz ist hier wichtig!

$P(\Omega)=1$

$P(A)+P(B)+P(C) <=1$



„Dutch book argument" -> nur kohärente Wahrscheinlichkeitsverteilungen führen zu fairen Glucksspiele! (Betrug beginnt mit Inkohärenz)
-> mehr siehe im Wahlfach…

Wozu ist sowas nützlich?

wir können Wahrscheinlichkeiten jetzt auch zu Hypothesen zuordnen

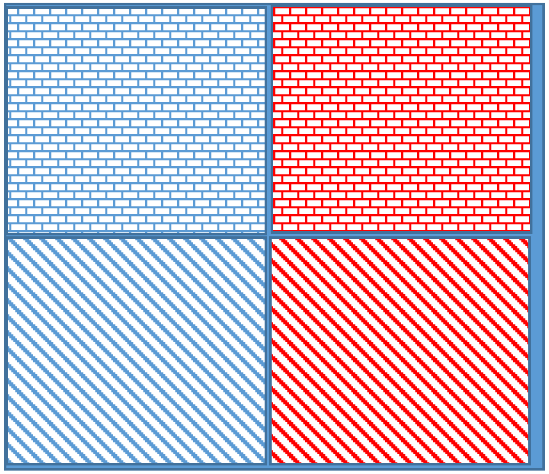
Diagnostik:
wir können die höchstwahrscheinlichste Diagnose wählen

wir können sogar berechnen wie viele bits ein Testergebniss uns bringt.
-> siehe Informationstheorie Vorlesung

Bedingte Wahrscheinlichkeit

P(A | B) = wahrscheinlichkeit von A ***wenn*** die Bedingung B ist wahr.

z.B. : unser Patient(inn) hat Fieber *wenn* Er/Sie COVID-19 angesteckt ist.
ich bekomme 5 in Statistik *wenn* ich alle Vorlesungen verstanden habe.



Wir sind nur an einen Unterraum von Ω interessiert.

P(blau | gitter) = wie viele blaue wir unter gitterte haben
= 1 blau UND gitter / 2 gitter = ½

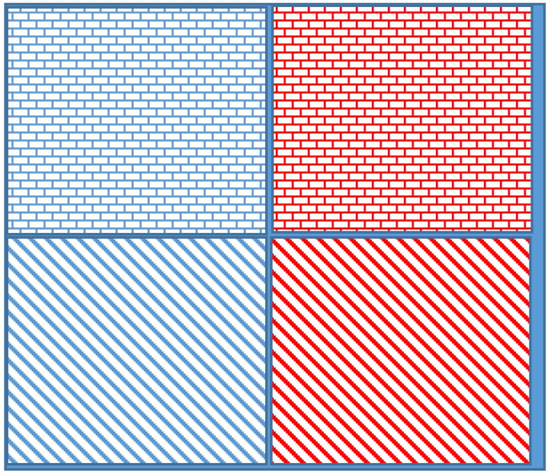
Bemerkungen:

1. für unabhängige: P(A|B=P(A)
2. für irgendwelche A,B: P(AB) = P(A|B)*P(B) Bayes-Regel oder Multiplikationsregel

Bayes-Regel

bedingte Wahrscheinlichkeit $P(A|H) = \frac{P(A \cdot H)}{P(H)}$

die „Ereignisse" sind hier A,H
H kann aber auch eine Hypothese sein!



P(AH) = P(blau UND gitter)
P(H)   = P(gitter)

P(A|H) = P(blau, wenn wir schon wissen das es gittern hat)

jetzt können wir für P(H) umsetzen: P(H) = P(AH)/P(A|H)
-> wir brauchen die gemeinsame und die bedingte Wahrscheinlichkeiten

Bayes-Regel

$$P(A|H) = \frac{P(A \cdot H)}{P(H)}$$

wenn P(A) und auch P(H) existieren, dann können wir P(AH) von zwei „richtungen” berechnen:

$$\mathrm{P}(\mathrm{A} \cdot H) = P(A|H) * P(H) = P(H|A) * P(A)$$

wovon aber: $$P(H|A) = \frac{P(A|H) * P(H)}{P(A)}$$

also wir können „H und A vertauschen”

Das bedeutet wir können die Wahrsch. von H **nach** der Beobechtung vom A berechen!

posterior Wahrscheinlichkeit

Vorsicht: für P(A) müssen wir ALLE wege kennen wodurch A passieren kann!

Epidemiologie: nur 2 Wege existieren da alles dichotom ist.

Wenn wir nur Zwei disjunkte Ausgänge haben, dann können wir relativ einfach rechnen.
also P(A) z.B. kann nur durch zwei Wege passieren.
Das nutzen wir in der Epdemiologie aus: Krank/Gesund und Test+/-

$$LR_+ = \frac{P(+|Krank)}{P(+|Gesund)} \qquad O_{Krankheit} = \frac{P(Krank)}{P(nicht-Krank)} = \frac{P(Krank)}{P(Gesund)}$$

$$O_{Krankheit\,|+} = \frac{P(Krank|+)}{P(Gesund|+)} = \frac{\frac{P(Krank\ und\ +)}{P(+)}}{\frac{P(Gesund\ und+)}{P(+)}} = \frac{P(Krank\ und+)}{P(Gesund\ und+)} = \frac{P(+|Krank) * P(Krank)}{P(+|Gesund) * P(Gesund)} = LR_+ * O_{Krankheit}$$

ähnlicherweise geht mit LR_

$$P(Krank|+) * P(+) = P(Krank\ und\ +) = P(+|Krank) * P(Krank)$$

prevalence = D/ALL

incidence = NEW cases over t time / number at risk

incidence RATE = incidence / t time

se	+|D	prevalence indep
sp	-|H	prevalence indep
false neg rate 1-se	-|D	prevalence indep
false pos rate 1-sp	+|H	prevalence indep
PPV	D|+	prevalence DEP
NPV	H|-	prevalence DEP
false alarm rate 1-PPV	H|+	prevalence DEP
false reassurance rate 1-NPV	D|-	prevalence DEP

RR_D	(D|R+)/(D|R-)	PPV/(1-NPV)
RR_H	(H|R+)/(H|R-)	(1-PPV)/NPV
LR+	+|D / +|H	se/(1-sp)
LR-	-|D / -|H	(1-se)/sp
OR_D	(O_D|R+) / (O_D|R-)	RR_D / RR_H
OR_H	O_H|R+ / O_H|R-	RR_H / RR_D
O_D = D/H	O_D_post = O*LR	

D: diseased	R+ risk factor present	O = p/1-p
H: healthy	R- risk factor NOT present	p = O/1+O
+/- Test result		

Beispiel: (Sie müssen es nicht nach-rechnen können!)

$$P(H|A) = \frac{P(A|H) * P(H)}{P(A)}$$

nehmen wir an, das die Wahrsch. von Magengeschwür sei P(H)=15% (unser Glaubensgrad)

Wir verlangen einen PCR-Test für *helicobacter pylori*, welches in 92% positive ist in Magengeschwür. Wir bekommen einen positiven Ergebniss.

Wir wissen auch das Sensitivität=0.95 und Spezifität=0.91, wobei Prävalenz von h.pylori Infektion in der Population ist 2%.

	+	-	
h.pylori angesteckt	1.90%	0.10%	2%
nicht angesteckt	8.82%	89.18%	98%

10.72% ist die Wahrsch. von Test+ unabhängig von andere Bedingungen.

P(A), kann durch zwei Wege passieren:

- H ist wahr
- H ist falsch

$P(A)=P(AH)+P(A\bar{H})$

also: $P(A) = P(H)*P(A|H) + P(\bar{H})*P(A|\bar{H})$

wenn unser Patien(inn) keinen Magengeschwür hat, können wir die generelle Daten des PCR-Tests benutzen

$$P(H|A) = \frac{0.92 \; * 0.15}{0.15 * 0.92 + (1 - 0.15) * 0.1072} = 0.602 = 60.2\%$$

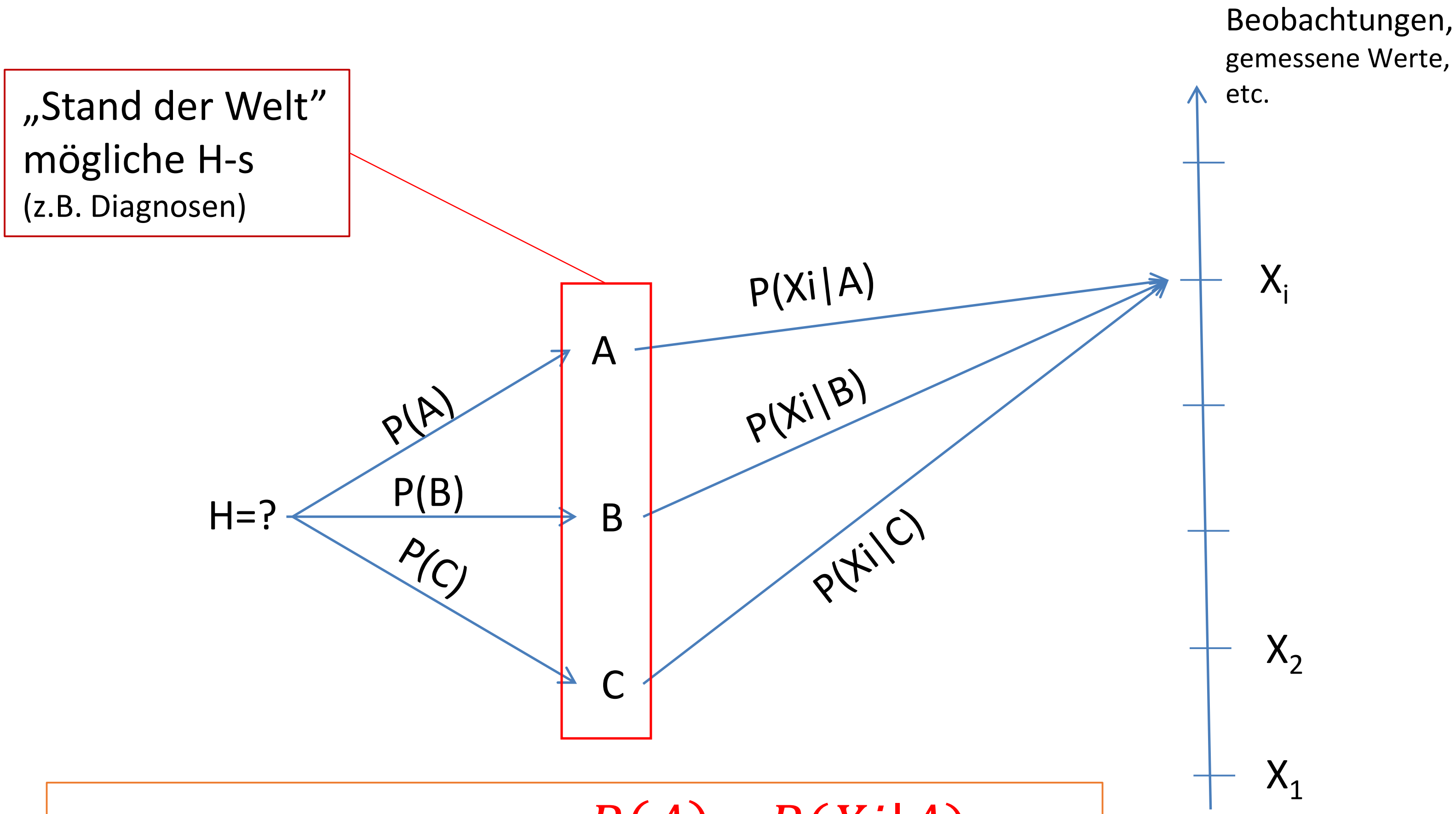
posterior Wahrscheinlichkeit

also **NACH** dem positiven Testergebniss sind wir deutlich sicherer in der Diagnose von Magengeschür.

-> ähnlich zu PPV, NPV.

-> Prävalenz hat einen großen Einfluss

Wir können Rückschlüsse ziehen



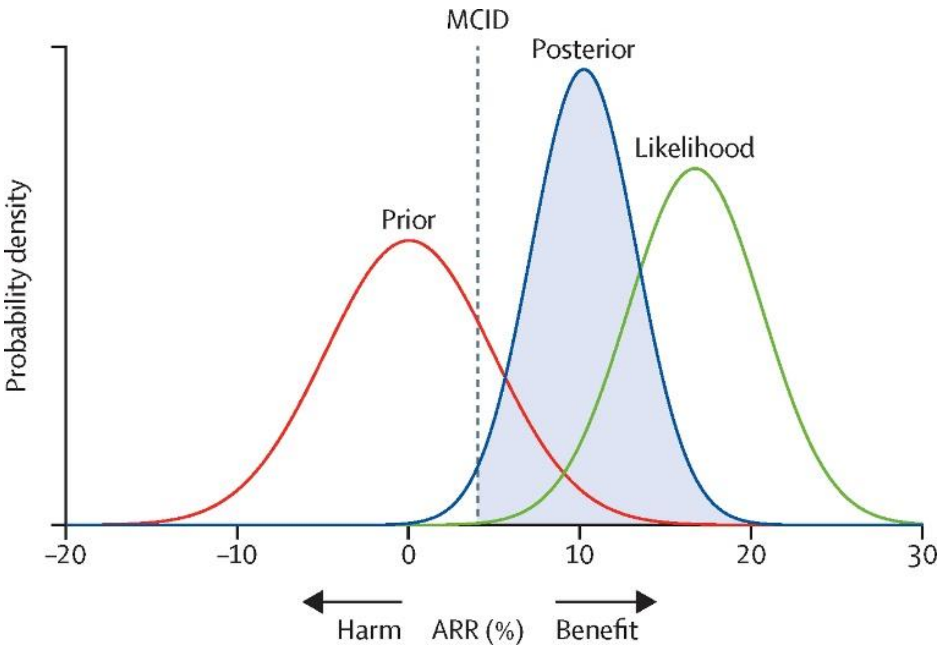
$$P(A|X = Xi) = \frac{P(A) * P(Xi|A)}{\sum_{k=A,B,C} P(k) * P(Xi|k)}$$

es geht auch wenn H kontinuierlich ist.

$$\boldsymbol{P(A|X=Xi)} = \frac{\boldsymbol{P(A)} * P(Xi|A)}{\sum_{k=A,B,C} \boldsymbol{P(k)} * P(Xi|k)}$$

$$\boldsymbol{f(h|x)} = \frac{\boldsymbol{f(h)} * P(x|h)}{\int_h \boldsymbol{f(h)} P(x|h)\,dh}$$

„Updating" das Aktualisieren der Glaubensgrade von dem priori Verteliung mit Hilfe der Daten, Beobachtungen, zu der neuen **posteriori Vertelung**.
wir müssen dazu die „vorwärz" Wahrscheinlichkeiten (Likelihood) kennen.



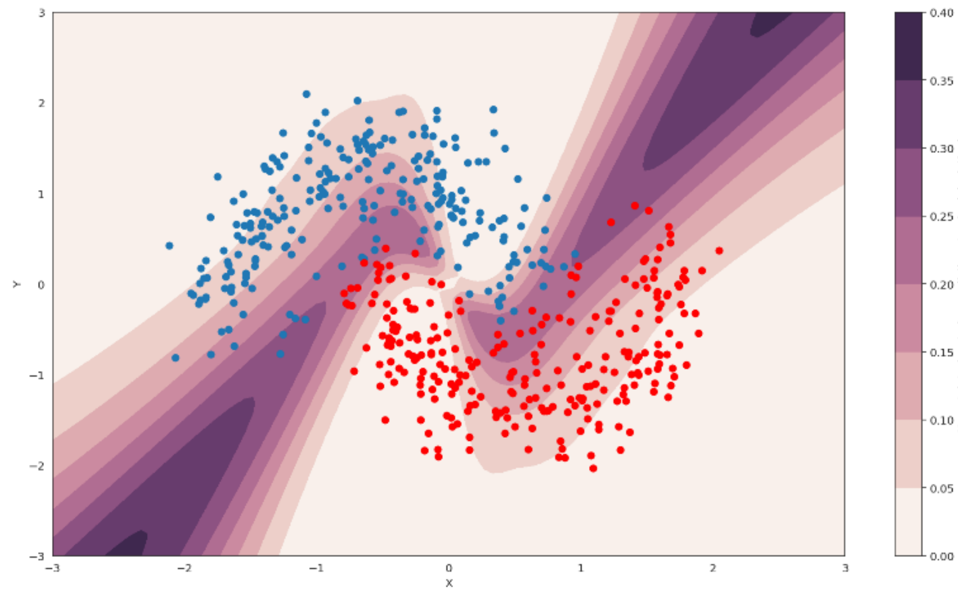
MCID=Minimal Clinically Important Difference
etwas was auch unser Patient(inn) fühlt.

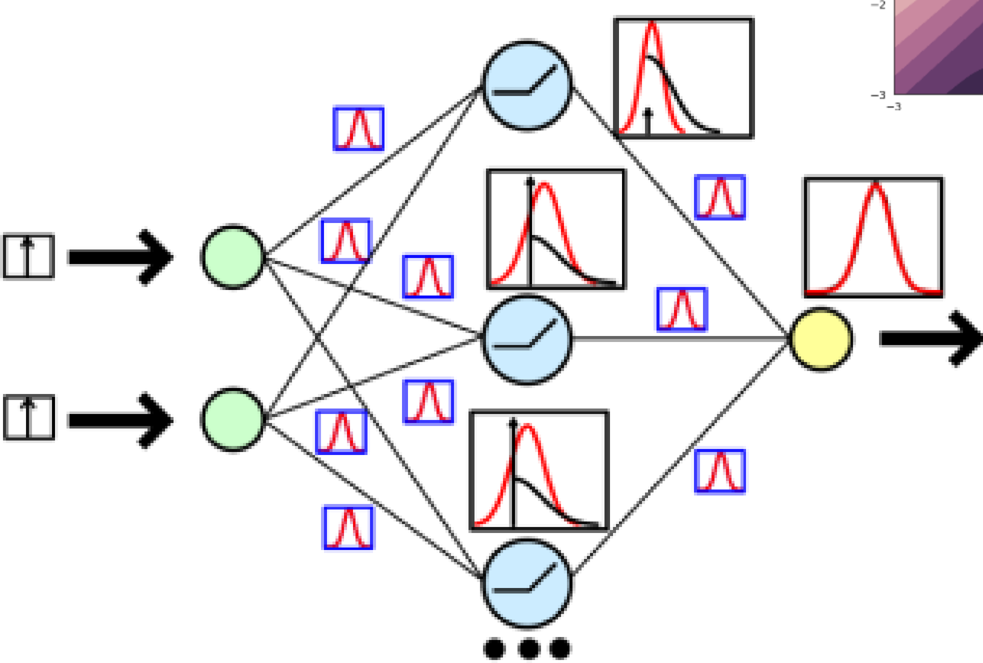
Bemerkung:

- das Prior muss existieren (mit dem Bayes-Definition ist es möglich)
- alle Likelihoods sind bekannt und wir können Ω voll abdecken.

Medizinische Verwendungen

Entscheidungs-unterstützung
Kategorisierung
Forschung
künstlicher Intelligenz
(neural netz)





ein bischen Entscheidungstheorie

Die Bayes-Methode hilft uns Einsicht zu bekommen, und durch Updating können wir unsere Glaubensgrade immer aktualisieren.

Wie können wir richtig Entscheinden?

Nutzen, Gewinn

Entscheidung -> Gewinn?

Wir entscheinden um den erwarteten Gewinn zu maximalisieren.

für **Möglichkeiten** haben wir Präferenzen, es gelten zwei Axiomen:

Vollständigkeit: $1 \preccurlyeq 2 \;\; oder \; 2 \preccurlyeq 1$

Transitivität: $1 \preccurlyeq 2 \;\; und \; 2 \preccurlyeq 3 \;\Longrightarrow 1 \preccurlyeq 3$

hier die $\preccurlyeq$ Relation bedeutet das wir auch bereit sind mehr zu „bezahlen” für den größeren erwarteten Gewinn.

Möglicherweise kennen wir nicht alle Möglichkeiten (z.B. Therapien), aber wenn neue auftauchen, dann können wir unsere Präferenzen vernünftig ergänzen.

unsere Präferenzen können mit einer **Gewinnfunktion** („utility function”) **(G)** beschreiben, welche Zahlen zu den Präferenzen zuordnet. (z.B. den erwarteten Gewinn)

vNM (von Neumann-Morgernstern) erwarteter Gewinntheorie:
G kann der “Preis” einer “Lotto” sein.

-> mehr im Wahlfach!

der Bayes-sche Entscheidungsweg