

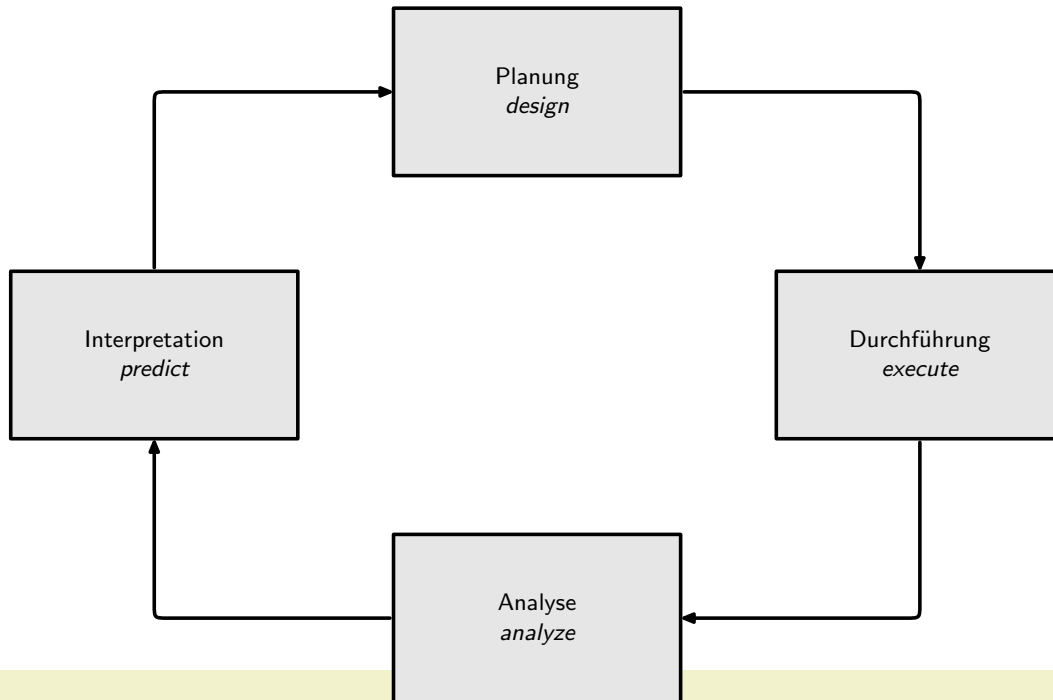
Bayes-Verfahren und frequentistischer Ansatz in der medizinischen Biometrie

Johannes Hüsing, Institut für Medizinische Informatik, Biometrie und Epidemiologie, Universität Essen

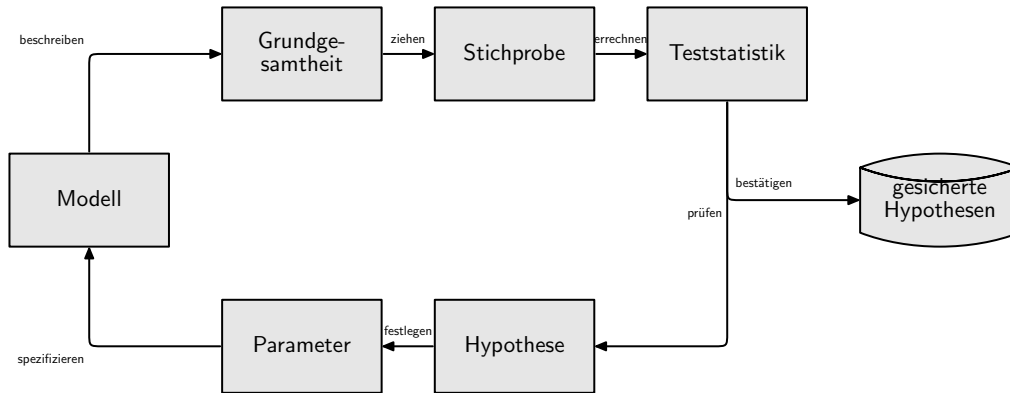
Homburg (Saar), den 19. März 2001

Prinzip

Der experimentelle Zyklus (Nelder 1999)

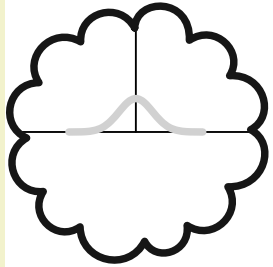


Neyman-Pearson-Ansatz

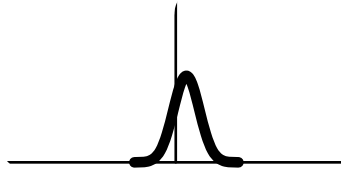


Diese Sichtweise geht von objektiven Wahrheiten aus, deren durch abgesicherte Hypothesen auf die Spur gekommen wird.

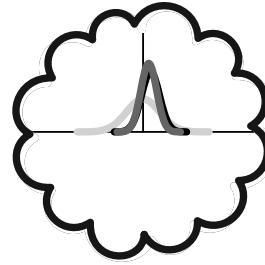
Schema des Bayes-Ansatzes



A-priori-
Annahme



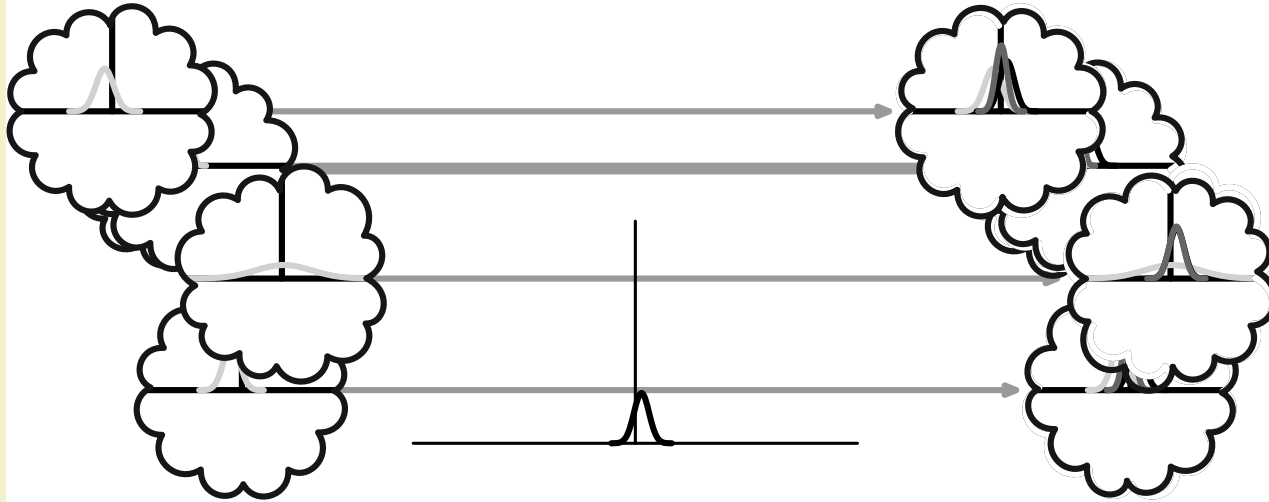
Ergebnis des
Experiments



A-posteriori-
Annahme

- Vorstellung über den Lageparameter wird als a-priori-Verteilung formuliert.
- Durch das Ergebnis des Experiments wird diese Vorstellung aufgefrischt und in eine a-posteriori-Verteilung überführt.
- Bei Normalverteilung: Je unspezifischer die a-priori-Verteilung, umso ähnlicher ist die a-posteriori-Verteilung dem Ergebnis des Experiments.

Idee des Bayes-Ansatzes



Unter verschiedenen Vorannahmen führt ein und derselbe Versuch zu verschiedenen Schlussfolgerungen über den Parameter.

Gegenüberstellung

Der Likelihood-Ansatz liefert ein Messergebnis und eine Aussage über seine Genauigkeit.

Der Likelihood-Ansatz fragt nach den Modellvoraussetzungen.

Der Bayes-Ansatz liefert ein Messergebnis und eine Aussage über den Kenntnisstand des Experimentators.

Der Bayes-Ansatz fragt nach den a-priori-Verteilungen.

Berechnung

Satz von Bayes – zwei Ereignisse

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

Satz von Bayes – stetige Zufallsvariablen

Likelihood im frequentistischen Ansatz:

$$f(y|\vartheta) = f_{Y,\Theta}(y, \theta|\theta = \vartheta)$$

A-posteriori-Verteilung in Bayes-Verfahren:

$$f(\vartheta|y) = \frac{f_Y(y|\vartheta)f_\Theta(\vartheta)}{\int_{\text{Tr}\Theta} f_Y(y|s)f_\Theta(s)ds}$$

Wenn sowohl die Likelihood als auch die a-priori-Verteilung normalverteilt sind, ergibt sich die a-posteriori-Verteilung als Normalverteilung, deren Lageparameter das mit den reziproken Varianzen gewogene Mittel aus den Lageparametern ist. Da dies häufig vorkommt, wird als Parameter die „Präzision“ als Kehrwert der Varianz an ihre Stelle gesetzt.

Interpretation

Irrtümer in diagnostischen Tests

Sensitivität/Spezifität vs. positiver/negativer prädiktiver Wert

Sensitivität und Spezifität sind bedeutend, um die Trennschärfe eines Tests zu beurteilen.

Positiver und negativer prädiktiver Wert sind im Einzelfall interessanter und haben auch Public-Health-Bedeutung (*number needed to diagnose*).

Irrtümer bei statistischen Tests

Die Wahrscheinlichkeiten für den Fehler erster und zweiter Art treffen Aussagen über die Güte des Testinstruments und sind als Qualitätskriterien gut geeignet.

Im Nachhinein ist aber möglicherweise die Wahrscheinlichkeit, im konkreten Fall zu irren, interessanter (zum Beispiel für Pharmaunternehmen: Entwicklung eines Therapeutikums fortsetzen?)

Klinische Prüfungen

Ein positiver prädiktiver Wert hängt von der Prävalenz der Krankheit ab.

Die a-posteriori-Wahrscheinlichkeit, dass eine studienbelegte Arbeitshypothese auch zutrifft, hängt von der Prävalenz aller zutreffenden Arbeitshypothesen ab.

Beide Aussagen sind grob vereinfachend.

Diagnostische Tests vs. klinische Prüfungen (zweiter Versuch)

Ein positiver prädiktiver Wert hängt von

- der Prävalenz der Krankheit
- dem Gespür des überweisenden Hausarztes
- dem Zusammentreffen der Begleitscheinungen, der Anamnese

ab. Insgesamt von der a-priori-Wahrscheinlichkeit.

Die a-posteriori-Wahrscheinlichkeit, dass eine studienbelegte Arbeitshypothese auch zutrifft, hängt von

- der Prävalenz der zutreffenden Arbeitshypothesen
- dem Gespür des Forschers
- der Einbettung der Arbeitshypothese in Vorversuche

ab. Insgesamt von der a-priori-Wahrscheinlichkeit.

Planung

Grundlage der Planung

Gehe von a-posteriori-Wahrscheinlichkeiten aus:

V „Verum besser als Placebo“

S „Studie liefert positives Ergebnis für Verum“

Bei klassischem Ansatz: Kontrolliere $\alpha = P(S|\bar{V})$ und wähle n so, dass $\beta = P(\bar{S}|V)$ einen bestimmten Wert nicht überschreitet.

Hier: Die a-posteriori-Wahrscheinlichkeiten $1 - \alpha^* = P(\bar{V}|\bar{S})$ und $1 - \beta^* = P(V|S)$ sollten einen festgelegten Gewissheitsgrad nicht unterschreiten.

Auswahl der Randwahrscheinlichkeit

Um die Relation zwischen α , β , α^* und β^* zu beschreiben, wird die Randwahrscheinlichkeit $P(S)$ benötigt. Vorschlag: Mit relativer Häufigkeit positiver Studienergebnisse aus einer relevanten Gesamtheit von vergangenen Studien schätzen.

Dann können aus $P(S)$ sowie den vorgegebenen α^* , β^* die frequentistischen Fehlerwahrscheinlichkeiten berechnet werden. Darauf werden die bekannten Methoden zur Fallzahlplanung angewandt.

Stopregel ist vorhanden

Vorteile

Stopregel ist vorhanden

In Likelihood-Ansätzen muss bei Zwischenauswertungen darauf geachtet werden, dass das globale Niveau über alle in der Zwischenzeit gewonnenen Tests insgesamt das vorgegebene Niveau α nicht überschreitet.

Im Bayesschen Ansatz kann man nach jeder hinzugekommenen Beobachtung seine a-posteriori-Verteilung dem neuesten Stand anpassen und aufhören, wenn die Wahrscheinlichkeit für keinen Unterschied unter ein vorgegebenes α^* sinkt.

Probleme

Nullmengenypothesen

Bei einer stetigen Verteilung des Parameters ist in der Regel $P(\vartheta = 0) = 0$.

Das ist kein wirkliches Problem: Ein Frequentist glaubt auch nicht an die Wahrheit der Nullhypothese.

Stattdessen: Klinische Relevanz formulieren ($P(|\vartheta| \leq \Delta) > 0$).

Berechnung

Ein Integral über mehrere Nuisance-Parameter ist numerisch kaum beherrschbar.

$$\int_{\mathbf{R}} \int_{\mathbf{R}_+} f_{\Theta}(\vartheta|\beta, \tau) dF_B(\beta) dF_T(\tau)$$

Aber durch stochastische Integration wird es beherrschbar (Folgevorträge)

ethische Bedenken

Im frequentistischen Ansatz: Vor Beginn der Studie keine Überlegenheit nachgewiesen $\Rightarrow H_0$ und H_1 können stimmen.

Bayes-Ansatz:

- Patient versucht, persönliches Risiko zu minimieren (Philipson und DeSimone 1997).
- Risiko ist Verlust mal a-priori-Wahrscheinlichkeit, Nutzen ist Gewinn mal a-priori-Wahrscheinlichkeit.
- Abwägung von Heilung und Nebenwirkungen.
- Probleme sind die punktförmige Nullhypothese (sonst Präferenz) und die Wichtung der Heilung gegen die Nebenwirkungen (individuell verschieden).

möglicher Ausweg

„Community equipoise“ (Kadane 1996): Experten formulieren a-priori-Wahrscheinlichkeiten für Heilung unter Therapie und Kontrolle, abhängig von anderen prognostischen Faktoren. Während der Studie werden diese Wahrscheinlichkeiten durch die laufenden Ergebnisse verändert.

Ein Patient, bei dem die Experten verschiedener „Meinung“ sind (momentane a-posteriori-Verteilungen führen zu Bevorzugung des Einen oder Anderen), wird randomisiert.

a-priori-Wahrscheinlichkeiten formulieren

Wie bekommt man Anwender dazu, ihre Vorstellungen in Form einer Wahrscheinlichkeitsverteilung zu notieren?

Im Allgemeinen haben diejenigen, die bereit sind, eine klinische Prüfung durchzuführen, eine a-priori-Verteilung, die die Therapiegleichheit nicht kategorisch ausschließt.

Wie testet man Adäquatheit der a-priori-Verteilungen?

Es werden a-priori-Verteilungen auf Parameter der Lage und Streuung angenommen.

Innerhalb des Bayes-Ansatzes ist es schwer, die angenommene Verteilung anhand der Daten zu kontrollieren (Nelder 1999). Box (1980) schlägt vor, den Nenner der Bayesschen Formel zu untersuchen:

$$f(y) = \int_{\Theta} f_Y(y|\vartheta) dF_{\Theta}(\vartheta)$$

Literatur

Box GEP. (1980) Sampling and Bayes' inference in scientific modelling and robustness. *Journal Royal Statistical Society A* **143(4)**, 383-430.

Lee SJ, Zelen M (2000). Clinical trials and sample size considerations: another perspective. *Statistical Science* **15 (2)**, 95-119.

Kadane J (Hrsg.). (1996) *Bayesian Methods and Ethics in a Clinical Trial Design*. Wiley.

Nelder JA. (1999) Statistics for the millennium. From statistics to statistical science. *The Statistician* **48(2)**, 257-269.

Philipson T, DeSimone J. (1997) Experiments and subject sampling. *Biometrika* **84(3)**, 619-630.