Bayesian Statistics

Planning and Analyzing a Flexible Bayesian Clinical Trial

Telba Irony, Ph.D.
Chief, General and Surgical Devices Branch
Division of Biostatistics
Office of Surveillance and Biometrics
Center for Devices and Radiological Health
Food and Drug Administration

Public Meeting: Draft Bayesian Guidance **July 27, 2006**

Outline

1. Section 4: Bayesian Statistics
2. Section 5: Planning a Bayesian Trial (5.1 to 5.4; 5.7, 5.8)
3. Appendix 9.4: Simulations to Obtain Operating Characteristics
4. Section 6: Analyzing a Bayesian Trial (6.4, 6.5)

4. Bayesian Statistics

4.1 Introduction

Statistics: Learning from evidence in presence of uncertainty

Unknown state of nature: parameter

Probability measures *one's* uncertainty about a state of nature ("parameter")

Probability changes as information accrues

Bayes Theorem

Example

Clinical trial to evaluate a new treatment

delta: Treatment effect (parameter) - unknown

p(delta) : prior probability

Data: Obtained from the clinical trial and described by the Likelihood Function

Objective: Posterior probability on delta

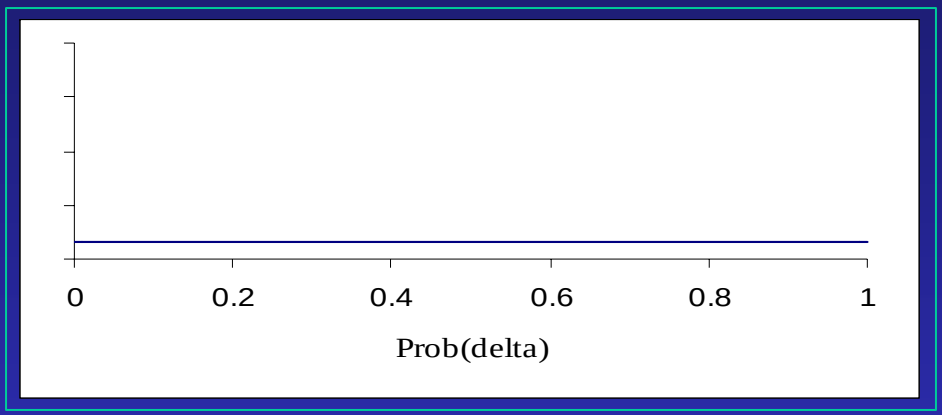
Question

What is the probability that the treatment effect is large enough to be clinically significant?

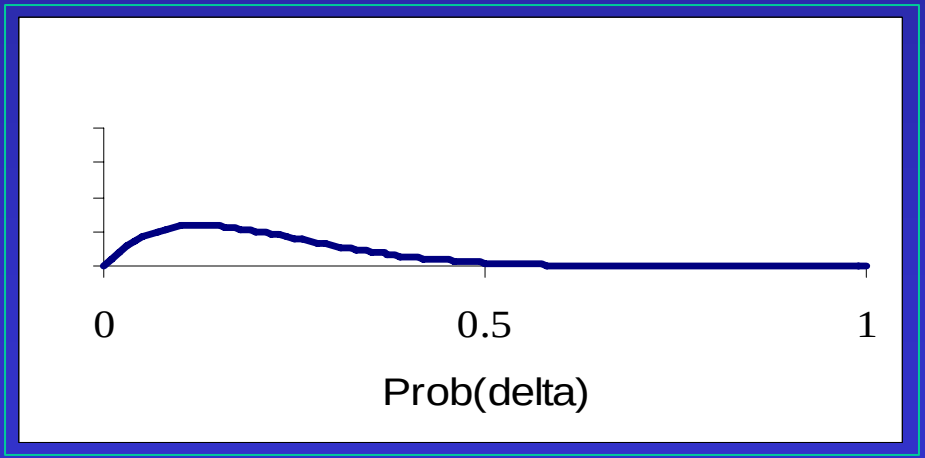
4.2 Prior distributions

Probability expresses one's uncertainty about delta

My uncertainty
Uniform
Less informative



Expert's uncertainty
More informative



4.3 Likelihood Function

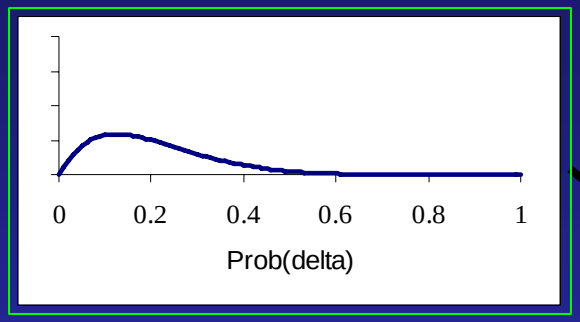
It is the mathematical model that relates the observed data to the parameter of interest

4.4 Posterior Distribution

It is the probability of each possible value of the endpoint, conditional on the prior information and on the observed data

Today

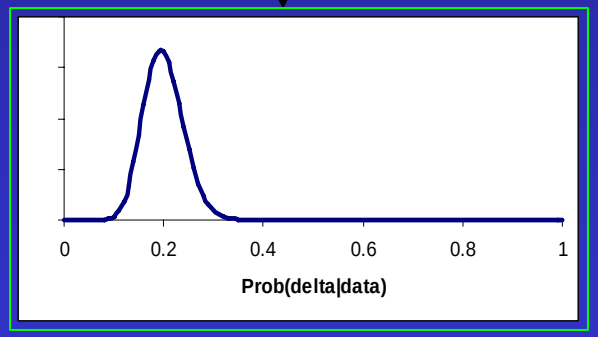
Prior



Trial: Likelihood

Data

Bayes



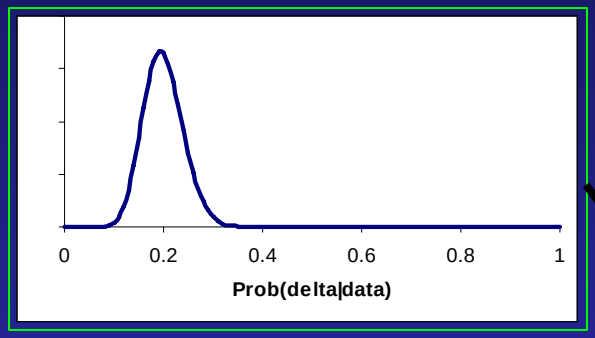
Posterior

Want to learn more?

Today's posterior is tomorrow's prior!

Tomorrow

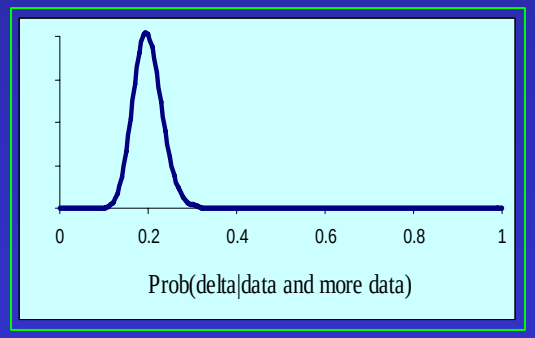
Prior



Trial: Likelihood

More Data

Bayes

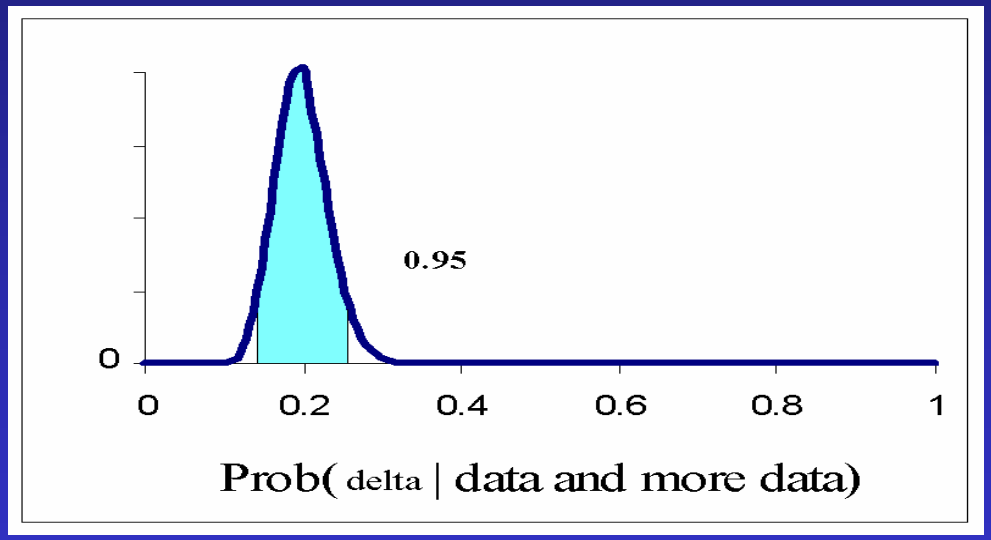


Posterior

The more information gathered, the sharper the posterior distribution may become !

If enough data are collected, the prior information is washed away

All relevant information about the unknown state of nature is contained in the posterior distribution



Credible Interval for the parameter of interest

[0.15;0.25]

4.5 Predictive Distributions

- Special type of posterior distribution
- Give the probability of future or unobserved events given the data that have been already observed
- Give the probability of the outcomes for future patient (s) in midcourse of a trial
- Give the probability that a new patient will be successful when implanted with a device once the device is on the market

4.5 Uses of the Predictive Distribution

- Helps to decide when to stop a trial
- Helps physicians and patients make decisions about the use of a device intervention (labeling)
- Predicts a clinical outcome from a valid surrogate
- Adjusts trial results for missing data

4.6 Exchangeability

- Patients in a trial are exchangeable if they belong to the same population of patients
- If patients in a trial are exchangeable with the patients in the population for which a device is intended, then the trial can be used to estimate safety and effectiveness of the device for the intended population
- Two trials (or samples) are exchangeable if they are elements of the same super-population of trials

4.7 The Likelihood Principle

- Central to Bayesian Approach
- All information about the parameter of interest contained in the data from a clinical trial is in the likelihood function
- If changes to the experimental design do not change the likelihood function (except for a constant factor), then the posterior distribution remains the same and the conclusions should not be modified

4.7 The Likelihood Principle

A trial may be altered in several ways without changing the likelihood function:

- **Modification in sample size**
- **Interim looks for stopping early**
- **Changing the randomization rate in midcourse**
- **Adaptive designs in general**

4. 8 Difference in approaches Bayesian vs. Frequentist

Bayesian

- Inferences based on the posterior
- Follows to the Likelihood Principle
- Inferences are based on the “parameter space”

Frequentists

- Inferences based on p-values
- Do not follow the Likelihood Principle
- Inferences are based on the “sample space”

5. Planning a Bayesian Trial

5.1 Sound Clinical Trial Design

- Objectives of the trial
- Endpoints to be evaluated
- Conditions under which the trial will be conducted
- Population to be investigated
- Planned statistical analysis
- Randomization, bias minimization, masking, type of analysis
- Valid for both Bayesian and frequentist trials

5.2 Selecting the relevant endpoints or parameters of interest

- Clinically relevant
- Directly observable
- Related to the device claims

5.3 Covariates

One should plan to collect information on covariates that may affect patient outcome for future adjustment

5.4 Choosing a Comparison: Control

- Gold standard: Concurrent controls
- Less rigorous:
  - Self controls
  - Historical controls
  - Problems: covariate adjustment, placebo effect, regression to the mean

Active Controls vs. Inactive Controls

Bayesian trials are especially suited to demonstrate not only that the new device is not inferior to an active control but also superior to an inactive control

5.5 Prior Distributions

5.6 Hierarchical Models

Dr. Gene Pennello

5.7 Sample Size Determination

One sizes a trial to gather enough information to make a decision without wasting resources or putting too many patients at risk

The sample size in a trial depends on:

- Prior information
- Likelihood model used
- Variability of the sample
- Decision criteria

5.7 Sample Size Determination

Traditional frequentists need to pre-determine the sample size because they work on the sample space

Bayesian sample sizes may be continuously updated as information is accrued

Instead of sample size, Bayesians tend to pre-specify a stopping criterion based on the desired amount of information about the parameter of interest:

- Length of the credible interval
- Posterior distribution

At any point in a Bayesian trial, the expected additional sample size needed to stop the trial can be computed

5.6 Sample Size Determination Recommendations

- Decide in advance the minimum sample size considering both safety and effectiveness endpoints to verify:
  - Model assumptions
  - Prior distribution
  - Controlling type I error
- Decide in advance the maximum sample size:
  - Ethical
  - Regulatory
  - Economical

5.8 Assessing the Operating Characteristics of a Bayesian Design

- The Likelihood Principle allows for flexibility in Bayesian Trial designs
- Often Bayesian trials are designed with a minimum sample size, several interim looks, and a maximum sample size
- At each interim look, the trial may stop either for success or for futility

5.8 Assessing the Operating Characteristics of a Bayesian Design

For regulatory purposes, a thorough evaluation of the operating characteristics of the trial design should be provided at the planning stage:

- Probability of type I error (approving a device that should not be approved)
- Probability of type II error (not approving a device that should be approved)
- Sample size distribution or expected sample size

Appendix 9.4 (page 34)
Simulations to Obtain Operating Characteristics

- Simulations results should be presented at the IDE stage
- Simulated trials should mimic the proposed trial by considering the same:
  - Prior information
  - Sample size
  - Interim analyses
  - Possible modifications of the trial in midcourse
- How large could the type I error rate be?
  - It depends on the submission
  - Confidence on the prior information

Simulations to Obtain Operating Characteristics

- **Assessing type I error rate:**
  - **parameters are fixed at the borderline values for which the device should NOT be approved**
  - **Proportion of successful trials = type I error rate**
- **Assessing type II error rate (and Power)**
  - **Parameters are fixed at plausible values for which the device should be approved**
  - **Proportion of unsuccessful trials = type II error rate = (1 – Power)**
- **Sample size distribution and expected sample size may be easily obtained for each scenario**
- **CDRH recommends several likely scenarios be simulated**

Example : Evaluating the experimental design

For each rate, simulated 1000 trials

Simulated Rates (θ)	% Succes	Expected Sample Size	Error type	Error Rate
0.2	0.6%	426	I	0.6%
0.15	2.7%	529	I	2.7%
0.11	4.9%	761	I	4.9%
Target	9.2%	786	?	?
0.099	22.7%	778	II	77.3%
0.095	45.7%	612	II	54.3%
0.09	80%	594	II	20%
0.08	89.3%	510	II	10.7%
0.07	95.6%	428	II	4.4%
0.06	99.7%	327	II	0.3%

Controlling the Type I error rate

Remedies to decrease type I error rate

- **Increase the posterior or predictive probability value to define a successful trial (96%, 98%, etc)**
- **Increase the number of patients before the first interim analysis**
- **Discount the prior information***
- **Reduce the number of interim analyses**
- **Increase the maximum sample size**
- **Any combination of the above**

*** Increase the number of patients before the first interim look; Iteratively increase the variance of the prior information until the type I error rate reduces to an acceptable level**

6. Analyzing a Bayesian Trial
6. 4 Uses of Predictive Probabilities

- **Predictive probability to decide to stop a trial**
  - **Success or futility**
  - **Exchangeability: key issue**
- **Predicting outcomes for future patients**
  - **Exchangeability**
  - **Labeling**
- **Predicting (imputing) missing data**
  - **Exchangeability**
  - **Assumptions: are the patients missing at random?**
  - **If not, sensitivity analysis should be performed**

6.4 Uses of Predictive Probabilities

- Predicting a clinical outcome from a surrogate
- Example:
  - Patients have measurements at 2 follow-up visits
  - Predictions for the second follow-up visits may be made if
    - Some patients have results from both follow-up visits
    - There is correlation between the early and later results

6. Analyzing a Bayesian Trial
6.5 Interim Analyses

- Interim analyses should be pre-planned
- Simulations should be performed in advance
- Stopping may occur when the posterior probability of the hypothesis of interest is large enough
- Stopping may occur by the use of predictive probability of a successful trial
  - Large: success
  - Small: futility
- Decision Analysis: considers the cost of decision errors