Bayesian Monitoring of Clinical Trials: Examples Using Conjugate Priors

by Olga Korosteleva, CSU Long Beach

My background

- Ph.D. in Statistics, Purdue University, West Lafayette, IN, May 2002
- Statistical Analyst, 3F Therapeutics, Inc., Lake Forest, CA, Dec. 2002 – Dec. 2004
- Author, *Clinical Statistics: Introducing Clinical Trials, Survival Analysis, and Longitudinal Data Analysis*, Jones and Bartlett Publishers, 2009.

What is a Clinical Trial ?

- Clinical Trial is a study of risk and benefits of a new therapy (for example, drug, biological device, surgical procedure) proposed for use in humans.
- Endpoint of a trial is the measure of a target outcome, for example, at least 20% decrease of excess body fat, or not more than 5% death rate.

Data Monitoring in Clinical Trials

- Standard Non-Bayesian Approach
  - Required sample size of a trial is statistically predetermined by conducting *power analysis*.
  - Trial continues until at least that many subjects have been accrued.
  - One statistical test of efficacy of the new therapy is carried out at the end of the trial.

Data Monitoring in Clinical Trials

- Bayesian Sequential Procedure
  - Knowledge of product efficacy is updated via the Bayesian formula as new data become available.
  - Applied when investigators are confident in product's efficacy.
  - Allows to stop the trial earlier.
  - If a trial is not stopped earlier due to non-satisfactory performance of the tested product, the trial is stopped when non-Bayesian approach dictates it to stop and a standard test is carried out.

Who was Bayes?

Rev. Thomas Bayes (c. 1702 – 17 April 1761) was a British mathematician and Presbyterian minister.

Has formulated a specific case of the theorem that bears his name: Bayes' theorem, which was published posthumously.

Bayesian Sequential Procedure

- Consider a clinical endpoint. Call it Θ.
- Researchers are interested in testing the null hypothesis

$$H_0 : \Theta \text{ belongs to a set } \Omega_0$$

(product is not effective)

against the alternative hypothesis

$$H_1 : \Theta \text{ belongs to a set } \Omega_1$$

(product is effective)

where Ω_0 and Ω_1 complement each other.

- In Bayesian monitoring, the endpoint Θ is modeled as a random variable.
- The primary knowledge about its possible values is summarized by a *prior density*

$$\pi(\theta) = f_{\Theta}(\theta)$$

- Investigators who have a *s*trong belief in efficacy of the tested product chose an *enthusiastic prior* (also called *optimistic prior*),

which assumes that the alternative hypothesis $H_1 : \Theta \in \Omega_1$ is more likely to hold than the null $H_0 : \Theta \in \Omega_0$, that is, $P(H_1) = \int_{\Omega_1} \pi(\theta) d\theta > 0.5$.

- The other choice of prior distribution is called *skeptical prior* (or *pessimistic prior*). It is used by investigators who are cautious about the tested product, and let data prove or disprove efficacy. It is assumed that $P(H_1) < 0.5$.

- Bayesian hypotheses testing is based on the *posterior distribution* of Θ, given the data from trial.
- The *posterior density* is computed according to the Bayes' formula

$$f_{\Theta}(\theta \mid data) = \frac{f(data \mid \Theta = \theta)\pi(\theta)}{\int f(data \mid \Theta = \theta)\pi(\theta)d\theta}$$

where $f(data \mid \Theta = \theta)$ is the likelihood function.

- The posterior probability of the alternative hypothesis is computed according to the formula

$$P(H_1 \mid data) = \int_{\Omega_1} f_\Theta(\theta \mid data) d\theta$$

- The decision of accepting or rejecting the alternative hypothesis is based on the following rule:

- If the posterior probability of H_1 is
  - smaller than 0.05, trial is stopped and H_1 is rejected;
  - larger than 0.95, trial is stopped and H_1 is accepted;
  - between 0.05 and 0.95, trial continues.

- If the trial is not stopped earlier and reaches its predetermined size, then the trial should be stopped, and a non-Bayesian statistical test should be performed on the data.

Conjugate Priors

Computation of the posterior density $f_{\Theta}(\theta \mid data)$ is simplified if the prior density $\pi(\theta)$ is *conjugate* to the likelihood function $f(data \mid \Theta = \theta)$.

By definition, a prior density is *conjugate* to a likelihood function if posterior density has the same algebraic form as the prior density.

We will consider three examples:

1. A gamma prior is conjugate to a Poisson likelihood function.
2. A normal prior is conjugate to a normal likelihood function (called *self-conjugate*).
3. A beta prior is conjugate to a binomial likelihood function.

Example 1: (Poisson-gamma) Bayesian Heart Valve Testing

- A clinical trial is conducted to test the performance of a new heart valve.
- The endpoint of the trial is the rate of certain valve-related complication.
- A *rate* of a complication is defined as the total number of cases divided by the total number of years accrued by all patients in the trial (called *number of patient-years*).

Ex. 9 cases in 500 patient-years = 0.018 or 1.8% rate.

- From a list of all possible valve-related complications, *endocarditis* (inflamation of heart lining) is chosen as the *primary endpoint*, because it is the most rare event and it takes the longest time to be detected.
- The *primary endpoint* is the one that is used in power analysis to pre-determine the required sample size of a trial.
- The FDA requires that a trial should continue for a minimum of 800 patient-years.

- The complication rate R for the new heart valve is compared to a historical value for endocarditis $R_h = 0.012$.
- The null $H_0: R \geq 2\,R_h = 0.024$ is tested against $H_1 : R < 2\,R_h = 0.024$. Note that the null hypothesis indicates that the new valve performs much worse than the historical one. If the null is accepted, the valve should not be marketed.

- We make an assumption that the number of endocarditis events N during a time period T has a Poisson distribution with mean RT and probability function

$$P(N = n) = \frac{(RT)^n \exp(-RT)}{n!}, \quad n = 0, 1, 2, \ldots$$

- From the Bayesian viewpoint, R is modeled as a random variable.

- To simplify calculations, the prior distribution of *R* is chosen from a family of distributions conjugate to Poisson, that is, the prior is taken as a *Gamma (a, b)* for some parameters *a* and *b* that are yet to be determined. The prior density of *R* is

$$\pi(r) = \frac{r^{a-1}\exp(-r/b)}{\Gamma(a)b^a}, \qquad r, a, b > 0.$$

- The posterior distribution of *R*, given that *n* endocarditis events have been observed during time *t*, is $Gamma\left(n+a,\left(t+1/b\right)^{-1}\right)$ with density

$$f_R(r \mid n,t) = C(rt)^n \exp(-rt)\ r^{a-1}\exp(-r/b)$$

$$= C_1\, r^{n+a-1}\exp\left(-r\left(t+1/b\right)\right)$$

where $C = \left(n!\Gamma(a)b^a\right)^{-1}$ and $C_1 = \dfrac{\left(t+1/b\right)^{n+a}}{\Gamma(n+a)}$.

- The posterior probability that the alternative hypothesis is correct is

$$P(H_1 \mid data) = P(R < 0.024 \mid n,t) = \int_0^{0.024} f_R(r \mid n,t)dr$$

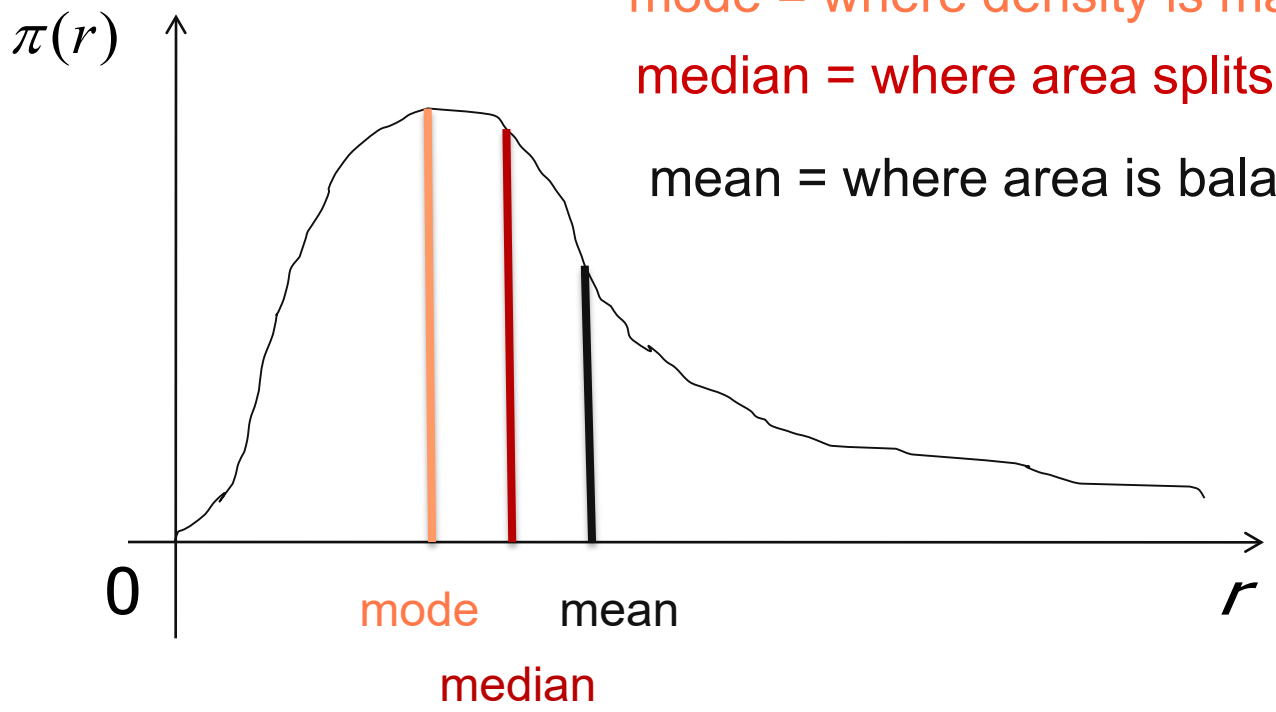
$$= \frac{(t+1/b)^{n+a}}{\Gamma(n+a)} \int_0^{0.024} r^{n+a-1} \exp\left(-r(t+1/b)\right) dr$$

- If $P(H_1 \mid \text{data}) < 0.05$, reject H_1
- If $P(H_1 \mid \text{data}) > 0.95$, accept H_1
- Otherwise, the trial continues

- It remains to specify the parameters *a* and *b*.
- Here is the trick:

Gamma density is *unimodal* (that is, has one peak) and *right-skewed* (that is, has a long right tail).

It looks like this:



Note that mode < median < mean

Therefore, $P(R < \text{mode}) < 0.5 < P(R < \text{mean})$.

- If investigators are inclined to use enthusiastic prior, then they should take the mean of the prior distribution to be equal to 0.024. This gives the opportunity to specify any desired prior probability of the true H_1 larger than 0.5. Indeed,

$$0.5 < P(R < \text{mean}) = P(R < 0.024) = P(H_1)$$

- For a skeptical prior, the mode should be chosen equal to 0.024. In this case,

  $P(H_1) = P(R < 0.024) = P(R < mode) < 0.5$

  and $P(H_1)$ can take on any value below 0.5.

- For a *Gamma* (*a, b*) distribution, the mode equals (*a-1*)*b* and the mean is *ab*.

- Thus, *a* and *b* can be computed numerically from the equations:

$$ab = 0.024$$ (for an enthusiastic prior)

$$(a-1)b = 0.024$$ (for a skeptical prior)

$$P(H_1) = P(R < 0.024) = \int_0^{0.024} \pi(r)dr$$

$$= \int_0^{0.024} \frac{r^{a-1}\exp(-r/b)}{\Gamma(a)b^a}dr$$

where $P(H_1)$ is specified by investigators.

Numerical Example

- Suppose investigators would like to use a skeptical prior with the probability of the true alternative equal to $P(H_1) = 0.4$.

- Therefore, the parameters of the prior density a and b satisfy

$$(a-1)b = 0.024 \quad \text{and} \quad 0.4 = \int_0^{0.024} \frac{r^{a-1}\exp(-r/b)}{\Gamma(a)b^a}dr \ .$$

- The minimum required length of the trial without the Bayesian monitoring is 800 patient-years.
- Suppose that investigators decide a priori to conduct interim Bayesian analyses at $t = 400$ and $t = 600$ patient-years.
- We will be looking for n such that the posterior probability of true H_1, $P(H_1 \mid n, t)$ is less 0.05, or larger than 0.95. The stopping rules are summarized in the following table:

t	n	$P(H_1 \mid n, t)$	t	n	$P(H_1 \mid n, t)$
400	**2**	**0.9688**	400	16	0.0505
	3	0.9421		**17**	**0.0317**

If <=2 cases, valve marketed; >=17 not marketed; 3-16 trial continues

600	**6**	**0.9643**	600	21	0.0668
	7	0.9399		**22**	**0.0450**

If <=6 cases, valve marketed; >=22 not marketed; 7-21 trial continues until 800 patient-years, when it is stopped and the standard maximum likelihood test is carried out.

Example 2: (normal-normal) Lowering Blood Pressure

- A new drug is tested for efficacy in lowering blood pressure. Let μ_{tr} and μ_c be the true mean percentage reduction in blood pressure in the treatment and control group, respectively. Investigators test

$$H_0 : \mu_{tr} \leq \mu_c \text{ against } H_1 : \mu_{tr} > \mu_c .$$

- Let $X_{tr} \sim N(\mu_{tr}, \sigma^2)$ and $X_c \sim N(\mu_c, \sigma^2)$ be the random variables representing blood pressure reduction in the treatment and control groups, respectively. Thus, $X_{tr} - X_c \sim N(\mu_{tr} - \mu_c,\ 2\sigma^2)$ for some known σ^2.

- We put a conjugate normal prior on the difference in means, $\mu_{tr} - \mu_c \sim N(\delta,\ \sigma_0^2)$.

- The posterior after observing $n,\ \bar{x}_{tr}, \bar{x}_c$ is also normal with mean $\left(\frac{\delta}{\sigma_0^2} + \frac{(\bar{x}_{tr} - \bar{x}_c)n}{2\sigma^2}\right) \Big/ \left(\frac{1}{\sigma_0^2} + \frac{n}{2\sigma^2}\right)$ and variance $\left(\frac{1}{\sigma_0^2} + \frac{n}{2\sigma^2}\right)^{-1}$.

- It remains to determine δ and σ_0^2. The mean δ can be elicited by asking investigators explicitly what they think the most likely value of $\mu_{tr} - \mu_c$ is. The investigators should also specify $P(H_1)$.

- Then we will have an equation for σ_0

$$P(H_1) = P(\mu_{tr} - \mu_c > 0) = \int_0^{\infty} \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{(x-\delta)^2}{2\sigma_0^2}\right) dx$$

$$= \int_{-\delta/\sigma_0}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-z^2/2) dz$$ after substitution $z = (x-\delta)/\sigma_0$

$$= 1 - \Phi(-\delta/\sigma_0)$$ where Φ is the *N*(0,1) cdf.

Thus, $\sigma_0 = \dfrac{-\delta}{\Phi^{-1}(1 - P(H_1))}$.

Numerical Example

- Suppose from previous similar studies it is known that the standard deviation is $\sigma = 15$.
- Investigators would like to use an optimistic prior. They are confident with $P(H_1) = 0.7$ that, on average, the difference in reduction in blood pressure between the groups is $\delta = 5$.
- Hence,

$$\sigma_0 = \frac{-\delta}{\Phi^{-1}\left(1 - P(H_1)\right)} = \frac{-5}{\Phi^{-1}(0.3)} = \frac{-5}{-0.5244} = 9.5347$$

- The required sample size for a non-Bayesian monitoring (with a single z-test for comparison of two means at the end of the trial) is computed via power analysis and is 97 patients per group.
- Suppose investigators decide a priori to conduct an interim Bayesian analyses when n=50 patients per group are accrued.

- We will be looking for $\bar{x}_{tr} - \bar{x}_c$ such that the posterior probability

$$P(H_1 \mid n, \bar{x}_{tr} - \bar{x}_c) = 1 - \Phi\left(- \frac{\left(\frac{\delta}{\sigma_0^2} + \frac{(\bar{x}_{tr} - \bar{x}_c)n}{2\sigma^2} \right)}{\sqrt{\frac{1}{\sigma_0^2} + \frac{n}{2\sigma^2}}} \right)$$

is less 0.05, or larger than 0.95. The stopping rules are summarized in the following table:

n	$\bar{x}_{tr} - \bar{x}_c$	$P(H_1 \mid n, \bar{x}_{tr} - \bar{x}_c)$	n	$\bar{x}_{tr} - \bar{x}_c$	$P(H_1 \mid n, \bar{x}_{tr} - \bar{x}_c)$
50	**-5.7**	**0.0490**	50	4.6	0.9474
	-5.6	0.0523		**4.7**	**0.9507**

- If <=-5.7, the trial is stopped and the drug not marketed;
- If >=4.7, the trial is stopped and drug marketed;
- If -5.6 to 4.6, the trial continues until 97 patients per group are accrued, at which point it is stopped and the standard z-test is carried out.

Example 3: (binomial-beta) False Positive Alarms by Defibrillators

- A study is conducted in which *N* patients with heart arrhythmia are implanted defibrillators (a device that delivers a therapeutic dose of electrical energy to ailing heart).
- Investigators are interested in testing whether the chance of false positive alarms by these defibrillators within the first year of use is low.

- Let *X* be the number of false positives (assuming one incident per patient). Then $X \sim Binomial(N, p)$ where *p* is the probability of a false positive alarm.
- The hypotheses are $H_0 : p \geq p_0$ and $H_1 : p < p_0$.
- We put a *Beta(a,b)* prior on *p*, which is conjugate to binomial.

- The posterior distribution of *p* after observing the number of false positive alarms *x* is

$Beta(x+a, N-x+b)$ with density

$$f_p(p \mid x) = Cp^x(1-p)^{N-x}\, p^{a-1}(1-p)^{b-1}$$

$$= \frac{p^{x+a-1}(1-p)^{N-x+b-1}}{B(x+a, N-x+b)}$$

where $a, b > 1$ and $0<p<1$.

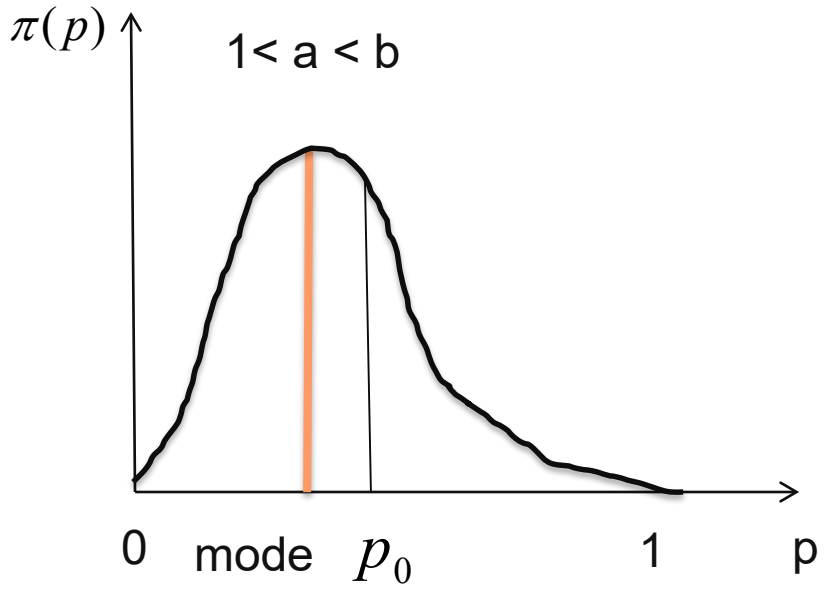
- To elicit the values of *a* and *b*, we ask investigators for the prior probability that the alternative hypothesis is true

$$P(H_1) = P(p < p_0) = \int_0^{p_0} \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)} dp$$

and also for the most likely value of *p*

$$Mode = \frac{a-1}{a+b-2}$$

To ensure a unique solution for *a* and *b*, we must require that $Mode < p_0$.


$\pi(p)$
1< a < b
0
mode
p_0
1
p

Numerical Example

- Suppose *N=100* and $H_1 : p < 0.3$ is being tested.
- Suppose investigators would like to use a skeptical prior with the probability of the true alternative equal to *P*(*H1*) =0.45.
- The elicited value of the mode is 0.25.
- Therefore, the values of parameters *a* and *b* are *a=1.7755* and *b=3.3265.*

- The posterior probability of the alternative is computed as

$$P(H_1 | x) = \int_0^{p_0} \frac{p^{x+a}(1-p)^{N-x+b-1}}{B(x+a, N-x+b)} dp$$

with $p_0 = 0.3, N = 100, a = 1.7755, b = 3.3265$

- The Bayesian stopping rule is summarized as follows:

x	$P(H_1 \mid x)$	x	$P(H_1 \mid x)$
22	**0.9585**	37	0.0679
23	0.9342	**38**	**0.0448**

- If <=22, the trial is stopped and the defibrillator marketed;
- If >=38, the trial is stopped and defibrillator not marketed;
- If 23 to 37, the trial continues.
- After one year, the trial is stopped and a standard maximum likelihood test is carried out.