ABSTRACT

BAYESIAN MONITORING OF CLINICAL TRIALS: EXAMPLES USING CONJUGATE PRIORS

By

Mary Tromp

January 2015

A clinical trial can save time and resources if it incorporates Bayesian monitoring. Generally speaking, conducting Bayesian analysis is a computationally intensive task. However, in the special case of hypotheses testing for clinical trials, and, moreover, when conjugate prior distributions of parameters are used, computational complexity is reduced remarkably. This thesis presents three examples where the Bayesian monitoring is achieved with a prior density of a parameter and the likelihood function of the data belonging to conjugate families of distributions. The first example studies a heart valve trial with a Poisson rate of adverse events and a gamma prior distribution of the rate. The second example focuses on testing certain drug efficacy for lowering high blood pressure, with self-conjugate normal family of distributions. In the third example, the probability of a false positive alarm produced by a heart defibrillator is modeled with beta prior distribution conjugate to binomial likelihood function.

BAYESIAN MONITORING OF CLINICAL TRIALS: EXAMPLES USING CONJUGATE PRIORS

A THESIS

Presented to the Department of Mathematics and Statistics

California State University, Long Beach

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Applied Statistics

Committee Members:

Olga Korosteleva, Ph.D. (Chair)
Alan Safer, Ph.D.
Morteza Ebneshahrashoob, Ph.D.

College Designee:

Tangan Gao, Ph.D.

By Mary Tromp

B.S., 2010, California State University, Long Beach

January 2015

TABLE OF CONTENTS

LIST OF TABLES

TABLE Page

CHAPTER 1

INTRODUCTION

Interim monitoring of clinical trials using Bayesian methods will be examined in this paper. As stated by Brown, Schriger, and Barrett (2010) frequentist statistics have little over Bayesian statistics, other than an aura of objectivity that does not withstand critical scrutiny for interim analyses. Under the Bayesian setting, Lee and Liu (2008) developed a sequential design on the basis of the predictive probability approach. They constructed stopping rules on the basis of the predictive probability, which corresponds to the probability of rejecting the null hypothesis at the end of the study, providing the same trend as that observed continues. The trial is stopped if the predictive probability exceeds a chosen efficacy cutoff value or is lower than a futility cutoff value. By choosing the proper cutoff value, the design preserves the types I and II error rates and improves the trial efficiency. Additionally, the design is flexible and easy to implement. It is also robust when the study conduct deviates from the original design.

Compared with the frequentist methods, Bayesian methods are better alternatives for constructing combined designs involving both sequential stopping and adaptive randomization. This is because Bayesian inference relies on the posterior distributions of the parameters, which automatically consider the dependence among the observations.

If the posterior distributions are in the same family as the prior probability in Bayesian probability theory, then the prior is called a conjugate prior for the likelihood function. There are three families of conjugate priors that this paper will examine. The

conjugate priors are Poisson-gamma conjugate families, normal self-conjugate family and binomial-beta conjugate families. Korosteleva (2009) discussed Poisson-gamma conjugate families but there has not been much literature discussing interim monitoring of clinical trials using Bayesian methods for the other families.

Okogbaa et al. (2010) say that careful choice of the prior likelihood probability structures could ensure computationally efficient posteriors. Conjugate families are efficient from a computational point of view since we can often compute the posterior distribution through a simple formula involving the family without resorting to Bayes's theorem directly. In any Bayesian analysis the aim is to obtain posterior estimates for some parameters or functions of parameters. In a limited number of cases.(e.g., in the case of conjugate priors) such estimates may be directly obtained.

The Food and Drug Administration (FDA) until recently required frequentist analysis in trials presented in support if a new drug application or new medical devices. The FDA (2010b, 30-35) now welcomes trials with Bayesian analysis and appears to be ready to consider trials that use adaptive methodologies. If it is part of the clinical trial plan, you may use a predictive probability at an interim point as the rule for stopping your trial. If the predictive probability that the trial will be successful is sufficiently high (based on results at the interim point), you may be able to stop the trial and declare success. If the predictive probability that the trial will be successful is small, you may stop the trial for futility and cut losses.

CHAPTER 2

DESCRIPTION OF BAYESIAN MONITORING OF CLINICAL TRIALS WITH CONJUGATE PRIORS

In a standard non-Bayesian approach the required sample size is statistically predetermined by conducting a power analysis. The trial must continue until at least that many subjects have been accrued. At the end of the trial one statistical test of efficacy of the new therapy is carried out.

When using a Bayesian sequential procedure, knowledge of product efficacy is updated using the Bayes formula as new data become available. The Bayesian sequential approach is applied when the investigators are confident in the product's efficacy because this approach allows an earlier termination of the trial. If a trial is not stopped earlier due to a 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.

Consider a clinical endpoint Θ which is modeled as a random variable. The null hypothesis H_0: Θ *belongs to a set* Ω_0, is tested against the alternative hypothesis H_1: Θ *belongs to a set* Ω_1 where the sets Ω_0 and Ω_1 complement each other. The prior knowledge about the endpoints possible values is summarized by a *prior density* $\pi(\theta) = f_{\Theta}(\theta)$.

Investigators who have strong belief in efficacy of the tested product chose an *enthusiastic 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 $P(H_1) = \int_{\Omega_1} \pi(\theta) d\theta > 0.5$.

If the investigators do not have such a strong belief, then their choice of prior distribution will be a *skeptical prior*. In this case, it is assumed $P(H_1) < 0.5$.

Bayesian hypothesis testing is based on the *posterior distribution* of Θ, given the data from the trial. The posterior density is computed using 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 using the formula

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

If the posterior probability of H_1 is smaller than 0.05, the trial is stopped, and H_1 is rejected. If the probability is greater than 0.95, the trial is stopped, and H_1 is accepted, otherwise the 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.

Computation of the posterior density $f_{\Theta}(\theta \mid data)$ is simplified if the prior density $\pi(\theta)$ is conjugate to the likelihood function $f_{\Theta}(\theta \mid data)$. A prior density is conjugate to a likelihood function if posterior density has the same algebraic form as the prior density.

In this thesis, three examples of conjugate priors will be considered a) a gamma prior is a conjugate to a Poisson likelihood function, b) a normal prior is a conjugate to a normal likelihood function, and c) a beta prior is conjugate to a binomial likelihood function.

CHAPTER 3

EXAMPLE OF TRIAL MONITORING WITH POISSON-GAMMA CONJUGATE FAMILIES OF DISTRIBUTIONS

A clinical trial is conducted to test the performance of a new heart valve. The endpoint of the trial is the rate of valve-related complications. The rate of a complication is defined as the total number of cases divided by the total number of years accrued by all patients (patient-years) in the trial (e.g., 9 cases in 500 patient-years gives the complication rate = 9/500=0.018 or 1.8%).

From a list of all possible valve-related complications, endocarditis, inflammation of heart lining, is chosen as the primary endpoint because it is the least frequently observed complication and takes the longest time to be observed. The primary endpoint is the one that is used in power analysis to pre-determine the required sample size of a trial. The FDA (2010a) requires that a trial should continue for a minimum of 800 patient-years according to the FDA guidance.

The complication rate R for the new heart valve is compared to the historical endocarditis rate $R_h = 0.012$. The null hypothesis, $H_0: R \geq 2R_h = 0.024$, is tested against $H_1: R < 2R_h = 0.024$. 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.

An assumption is made that the number of endocarditis events N during a time period T has a Poisson distribution with mean RT and probability function

$$P(N=n)=\frac{(RT)^n e^{-RT}}{n!},\ \ n=0,1,2,...$$

From the Bayesian viewpoint, *R* is modelled as a random variable. To simplify calculations, the prior distribution of *R* is chosen from a family of distributions conjugate to Poisson. 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

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

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

The posterior probability that the alternative hypothesis is correct is

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

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

If $P(H_1|data)<0.5$, the alternative hypothesis H_1 is rejected, if $P(H_1|data)>0.95$, then the H_1 is accepted, otherwise, the trial continues.

The parameters, *a* and *b,* must be specified. Note that the gamma density is unimodal, which means it has one peak, and right-skewed, which means it has a long right tail. For unimodal right-skewed distributions, the mean-median-mode inequality hold true. It states that the mean is always larger than the median, which, in turn, is

always larger than the mode. Therefore, $P(R<mode) < P(R<median) < P(R<mean)$. 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, $0.5=P(R<median) < P(R<mean) = P(R<0.024) = P(H_1)$. For a skeptical prior, the mode should be chosen equal to *0.024*. In this case, $P(H_1) = P(R<0.024) = P(R<mode) < 0.5=P(R<median)$, and $P(H_1)$ can take on any value below *0.5*. For a *Gamma(a, b)* distribution, mode equals *(a-1)b* and mean is *ab*.

Thus, *a* and *b* are the solution of the following two equations:

Equation 1. $ab = 0.024$ for an enthusiastic prior, and $(a-1)b = 0.024$ for a skeptical prior

and

Equation 2. $$P(H_1) = P(R < 0.024) = \int_0^{0.024} \pi(r)dr = \int_0^{0.024} \frac{r^{a-1}\exp(-r/b)}{\Gamma(a)b^a}dr$$

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

Now a specific numerical example will be presented. Suppose investigators would like to use a skeptical prior with the probability of the true alternative equal to $P(H_1)=0.4$. The parameter of the prior density *a* and *b* satisfy $(a-1)b = 0.024$ and

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

The minimum required length of the trial is *800* patient-years per FDA (2010b). Suppose investigators decide to conduct interim Bayesian analysis at *t=400* and *t=600* patient-years. We will be looking for *n* such that the posterior probability of true H_1,

$P(H_1|n, t)$, is less than 0.05 or larger than 0.95. A Mathematica code for calculations is presented in Appendix. The results are summarized in Table 1 in the same appendix.

From the table, at 400 patient-years, if $n \leq 2$, the trial is stopped, and the heart valve is marketed; if $n \geq 17$, then the trial is stopped, and the valve is not marketed, otherwise the trial continues until 600 patient-years are accrued. At 600 patient-years, if $n \leq 6$, then the trial is stopped and the valve is marketed; if $n \geq 22$, then the trial is stopped, and the valve is not marketed; otherwise, the trial continues to 800 patient-years, and a standard non-Bayesian test for valve efficacy is performed.

CHAPTER 4

EXAMPLE OF TRIAL MONITORING WITH NORMAL SELF-CONJUGATE FAMILY OF DISTRIBUTIONS

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: $\mu_{tr} \leq \mu_c$ 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. And let's assume that σ^2 is known.

We have that $X_{tr} - X_c \sim N(\mu_{tr} - \mu_c, 2\sigma^2)$. Using a Bayesian approach, we put a conjugate normal prior on the difference in means, $\mu_{tr} - \mu_c \sim N(\delta, \sigma_0^2)$.

The posterior distribution of $\mu_{tr} - \mu_c$, after observing $n, \overline{x_{tr}}, \overline{x_c}$, is also normal with mean $\left(\frac{\delta}{\sigma_0^2} + \frac{(\bar{x}_{tr} - \bar{x}_c)n}{2\sigma^2}\right) / \left(\frac{1}{\sigma_0^2} + \frac{n}{2\sigma^2}\right)$ and variance $\left(\frac{1}{\sigma_0^2} + \frac{n}{2\sigma^2}\right)^{-1}$. This can be seen by noticing that the posterior density is proportional up to a multiplicative constant.

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 $z = (x-\delta)/\sigma_0$]

$= 1 - \Phi\left(-\delta/\sigma_0\right)$ where Φ is the *N(0,1)* cumulative distribution function. Thus,

$$\sigma_0 = \frac{-\delta}{\Phi^{-1}\left(1 - P(H_1)\right)}.$$

Next, we will present a numeric 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 per Korosteleva (2009). Suppose investigators decide to conduct an interim Bayesian analysis when $n=50$ patients per group are accrued. We will be looking for $X_{tr} - X_c$ 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 than 0.05, or larger than 0.95.

The calculations are simple enough to be done in Excel. The results of calculations are summarized in Table 2 in Appendix. The Bayesian monitoring

procedure is as follows: when *n* has reached *50* in both groups and $\bar{x}_{tr} - \bar{x}_c \leq -5.7$, then the trial is should be stopped, and the drug should not be marketed. If, on the other hand, $\bar{x}_{tr} - \bar{x}_c \geq 4.7$, then the trial should be stopped, and the drug should be marketed. In case $-5.7 < \bar{x}_{tr} - \bar{x}_c < 4.7$, the trial should continue until *97* patients per group are accrued, at which point the trial should be stopped, and the standard z-test should be carried out.

CHAPTER 5

EXAMPLE OF TRIAL MONITORING WITH BINOMIAL-BETA CONJUGATE FAMILIES OF DISTRIBUTIONS

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

Let X be the number of false positives, 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 family of distributions

The posterior distribution of p, after observing the number of false alarms x, is $Beta(x+a, N-x+b)$ with density

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

$$= \frac{\binom{N}{x} p^{x+a-1}(1-p)^{N-x+b-1} / B(a,b)}{\int_0^1 \left(\binom{N}{x} y^{x+a-1}(1-y)^{N-x+b-1} / B(a,b)\right) dy}$$

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

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 $Mode < p_0$. A numeric example follows. 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 hypothesis equal to $P(H_1)=0.45$, and suppose also that the elicited value of the mode is 0.25. The posterior probability of the alternative is computed as

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

Table 3 in Appendix shows the results of computations. The Mathematica code is given in the same appendix. The Bayesian stopping rule is summarized as follows:

If the number of false alarms $x \leq 22$, then the trial should be stopped, and the defibrillator should be marketed; if $x \geq 38$, then the trial should be stopped, and the defibrillator should not be marketed. If x is between 23 and 38, then the trial should continue. After one year, the trial should be stopped and a standard maximum likelihood test should be carried out.

APPENDIX

MATHEMATICA CODE

Poisson-Gamma Example

Mathematica Code

In[1]:= a/. FindRoot[Gamma[a,a-1]/Gamma[a]==0.6, {a,5}]

Out[1]= 7.81438

a=7.81438, b=0.024/(a-1), P(H1|t,n)=1-Gamma[n+a,0.024*t+a-1]/Gamma[n+a]

In[2]:= 1-Gamma[9.81438, 16.41438]/Gamma[9.81438] t=400, n=2

Out[2]= 0.96882

In[3]:= 1-Gamma[10.81438, 16.41438]/Gamma[10.81438] t=400, n=3

Out[3]= 0.942084

In[4]:= 1-Gamma[23.81438, 16.41438]/Gamma[23.81438] t=400, n=16

Out[4]= 0.0505347

In[5]:= 1-Gamma[24.81438, 16.41438]/Gamma[24.81438] t=400, n=17

Out[5]= 0.031655

In[6]:= 1-Gamma[13.81438, 21.21438]/Gamma[13.81438] t=600, n=6

Out[6]= 0.964337

In[7]:= 1-Gamma[14.81438, 21.21438]/Gamma[14.81438] t=600, n=7

Out[7]= 0.939906

In[8]:= 1-Gamma[28.81438, 21.21438]/Gamma[28.81438] t=600, n=21

Out[8]= 0.066778

In[9]:= 1-Gamma[29.81438, 21.21438]/Gamma[29.81438] t=600, n=22

Out[9]= 0.0449617

TABLE 1. Results of Bayesian Monitoring in Poisson-Gamma Example

t	n	$P(H1 \mid n, t)$	t	n	$P(H1 \mid n, t)$
400	**2**	**0.9688**	400	16	0.0505
	3	0.9421		**17**	**0.0317**
600	**6**	**0.9643**	600	21	0.0668
	7	0.9399		**22**	**0.0450**

Normal-Normal Example

TABLE 2. Results of Bayesian Monitoring in Normal-Normal Example

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

Binomial-Beta Example

Mathematica Code

In[1]:= a/. FindRoot[Beta[0.3, a, 3*a-2]/Beta[a, 3*a-2]==0.45,{a,1}]

Out[1]= 1.77546

a=1.77546, b=3a-2, N=100, P(H1|x)= Beta[0.3,a+x,N-x+b]/Beta[a+x,N-x+b]

In[2]:=Beta[0.3,23.77546,81.326]/Beta[23.77546,81.326] x=22

Out[2]= 0.958525

In[3]:= Beta[0.3,24.77546,80.326]/Beta[24.77546,80.326] x=23

Out[3]= 0.934213

In[4]:= Beta[0.3,38.77546,66.326]/Beta[38.77546,66.326] x=37

Out[4]= .0679384

In[5]:= Beta[0.3,39.77546,65.326]/Beta[39.77546,65.326] x=38

Out[5]= 0.0448417

TABLE 3. Results of Bayesian Monitoring in Binomial-Beta Example

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

REFERENCES

REFERENCES

Brown, A.M., Schriger, DL., and Barrett, TW. Outcome Measures, Interim Analysis, and Bayesian Approaches to Randomized Trials. Annals of Emergency Medicine, 2010: 216-224

Food and Drug Administration. 2010a. *Draft Guidance for Industry and FDA Staff: Heart Valves: Investigational Device Exemption (IDE) and Premarket Approval (PMA) Applications*. Accessed 12/10/2014. http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm193096.htm.

Food and Drug Administration. 2010b. *Guidance for Industry and FDA Staff: Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials.* Accessed 12/10/2014. http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071121.pdf

Korosteleva, O. 2009. *Clinical Statistics: Introducing Clinical Trials, Survival Analysis, and Longitudinal Data Analysis*. Sudbury, MA: Jones and Bartlett Publishers.

Lee, J.J., and D. D. Liu. 2008. "A predictive probability Design for Phase II Cancer Clinical Trials." *Clinical Trials* 5(2): 93-106.

Okogbaa, O.G., L. Devarakonda, W. Otieno, and V. Albino. 2010. "Analysis of Sequential Monitoring Schemes Using Natural Conjugate Priors." *International Journal of Reliability, Quality and Safety Engineering* 17(1): 57-87.