
Soumis le : 12 Juillet 2013

Forme révisée acceptée le : 16 Septembre 2014

Email de l'auteur correspondant :

ahlem_stat@live.fr

Nature & Technology

Bayesian prediction using two-stage designs in experimental trials

Hayet Merabet^a, Ahlam Labdaoui^a

*^aLaboratory of Applied Mathematics and Modeling, Department of Mathematics, University Constantine1
25000 Constantine, Algeria*

Abstract

Prediction provides discipline and pragmatic importance to empirical research. Two stages design is commonly used in phase II experimental trials. This design possesses good frequentist properties and allows early termination of the trial when the interim data indicate that the experimental regimen is inefficacious. The design with the predictive probability approach provides an excellent alternative for conducting multi-stage phase II trials. It is efficient and flexible and possesses desirable statistical properties. Often, preliminary experimental information is already available as a “pilot”, where a first experience that we ask for confirmation of results. Formally, we consider the following situation: Given a first sample of data, we want to plan an experiment (or a new sample) to have good chances of getting the relief sought if the experiment is not abandoned. We propose the procedure based on the notion of satisfaction index which is a function of the p-value and we expect, given the available data to calculate an estimate of satisfaction for future data as Bayesian predictive index conditional on previous observations. To illustrate the proposed procedure, several models have been studied by choosing the prior distribution justifying the motivations of objectivity or neutrality that underlie the analysis of experimental data.

Key words: Predictive Bayesian approach, experimental trials, p-value, two-stage design.

1. Introduction

In many situations in experimental trials, the goal of statistical analysis is to predict values of a future sample. In clinical trials, sequential methods are used when there are formal interim analyses. An interim analysis is an analysis intended to assess treatment effect with respect to efficacy or safety at any time prior to the completion of a clinical trial. Because interim analysis results may introduce bias to subsequent clinical evaluation of the subjects who enter the trial, all interim analyses should be carefully planned in advance and described in the study protocol. The concept of sequential statistical methods was

originally motivated by the need to obtain clinical benefits under certain economic constraints. For a trial with a positive result, early stopping means that a new product can be exploited sooner. If a negative result is indicated, early stopping ensures that resources are not wasted. Sequential methods typically lead to saving in sample size, time, and cost when compared with the standard fixed sample procedures. Interim analyses enable management to make appropriate decisions regarding the allocation of limited resources for continued development of the promising treatment. Under special circumstances, there may be a need for an interim analysis that was not planned originally. In this case, a protocol amendment describing the rationale for such an interim analysis should be implemented prior to any clinical data being unblinded. The acceptance and application of Bayesian methods in virtually all branches of

science and engineering have significantly increased over the past few decades [1]. This increase is largely due to advances in simulation-based computational tools for implementing Bayesian methods. The Bayesian approach provides greater flexibility in the statistical methodology of experimental trials, it can be found in this respect [2, 3]. The aim of an exploratory experimental trial is to determine whether a new intervention is promising for further testing in confirmatory experimental trials. Most exploratory clinical trials for example are designed as single-arm trials using a binary outcome with or without interim monitoring for early stopping. Prediction models are important in various fields, including medicine, physics, meteorology, and reliability. Prediction models will become more relevant in the medical field with the increase in knowledge on potential predictors of outcome, e.g. from genetics. Also, the number of applications will increase, e.g. with targeted early detection of disease, and individualized approaches to diagnostic testing and treatment. The current era of evidence-based medicine asks for an individualized approach to medical decision-making. Evidence-based medicine has a central place for meta-analysis to summarize results from randomized controlled trials; similarly prediction models may summarize the effects of predictors to provide individualized predictions of a diagnostic or prognostic outcome. In this work, we define indices of satisfaction and prediction of satisfaction related to hypothesis testing, we calculate the predictive probabilities of achieving a successful result at the end of the trial using the analysis prior in order to stop the trial in case of low or high efficacy [4, 5, 6], in several models (Poisson, Binomial, Gaussian). Bayesian modeling will be used. We treated our applications given by software: Matlab and R. The outline of the paper is as follows. Section 2 introduces the basic idea of sequential prediction in experimental design. In section 3, we propose a study predictive sue the binomial model, we present an

illustrative example and conclude with a discussion in Section 4.

2. Model selection in the experimental planning

2.1. Bayesian predictive design

We use the Bayesian framework as a tool to design clinical trials with desirable frequentist properties. Taking the Bayesian approach, we derive an efficient and flexible design.

Statistical methodology has already been used [7, 8]. Remember that, the Bayesian model was introduced in the context of J. Mr. Grouin [9] and after improved in [10, 11, 12].

We specify the experimental context by choosing $(P_\theta)_{\theta \in \Theta}$ a family of probability observations on a space Ω and where Θ is the space of unknown parameters and is to test the null hypothesis Θ_0 against the alternative hypothesis Θ_1 . In classical asymmetric design test problems, such a situation is generally in the experimenter, a desire to put in evidence a significant result, that is to say, to conclude the rejection of the null hypothesis.

2.2. Satisfaction index

If adopted a procedure deterministic test, relative to a level α , leading to partition Ω in a region of not-rejection $\Omega_0^{(\alpha)}$ and a rejection region $\Omega_1^{(\alpha)}$, an index particularly simple satisfaction is the indicator function of $\Omega_0^{(\alpha)}$. It is satisfied if the result is significant at α , if dissatisfied. But very often users want rather face an outcome that seems likely to lead to their rejection of Θ_0 , know what its degree of significance; that is to say, know how far the results appear significant.

Using the fact that any reasonable test technique leads to a family of not-rejection regions $\Omega_0^{(\alpha)}$ in the sense of

decreasing inclusion when α increases, that is to say, when our precautions s' mitigate, then use a new index of satisfaction, a little less rough than the previous one, denoted $\Phi^{(\alpha)}$, and defined as a function of p-value, the simplest as zero on the region of not-rejection $\Omega_0^{(\alpha)}$ and equal to $(1-p)$, or more generally $(\alpha-p)^l$ with $l > 0$ otherwise. In other words it offers a satisfaction index which is zero if it is not detected significant and otherwise a decreasing function of the p-value, and therefore, the more p is small and the experimenter believes that the result is significant [13] remind that the p-value is considered as a measure of credibility to attach to the null hypothesis that practitioners often use to meet several critical and disadvantages of approach Neymann of Pearson. The value of this index satisfaction and an extended family of indices in the concept of predicting satisfaction of a sample future as a first sample.

2.3. Prediction of satisfaction

Experimental contexts that we have mentioned in the introduction often lend themselves to analysis in several phases [14], and we limit ourselves to two phases and the situation, which corresponds to the requirements in the experimental trials, where the first phase is that indicative and is intended only to consider whether to resume testing for a second phase, conducted independently of the first and of which only the test result based on the conclusion of which is the ultimate purpose of the study. We note here Ω , Ω' and Ω'' sets complete results, the results of the first phase and the results of the second phase ($\Omega = \Omega' + \Omega''$).

It is in this context that has proposed to introduce a Bayesian model with a prior distribution on Θ and the family of probabilities $(P_\theta)_{\theta \in \Theta}$ on Ω . He sees in this model the probability of Ω'' , influenced by the outcome of the first phase ω' , which we denote by $P_{\Omega''}^{\omega'}$, recall that, according to the usual Bayesian terminology, the term

predictive probability, the probability P_Ω on the space of complete results, which is used here is the probability on Ω'' which is deduced by conditioning by ω' . We find as a prediction on the view of the first phase of a significant result in the second phase, the value $P_{\Omega''}^{\omega'}(\Omega_1^{(\alpha)})$, where $\Omega_1^{(\alpha)}$ is the rejection region of the classical test made on the basis of the results of the second phase. It is in this sense that here is practice both classical statistics and Bayesian statistics.

We propose here, more typically associate with any satisfaction index on the second phase index forecasting is the mathematical expectation with respect to $P_{\Omega''}^{\omega'}$, satisfaction provided by Consider the second phase of the experiment and the predicted using the first.

It is shown elementarily that the value ω' an index of prediction can also be obtained as the expectation with respect to the posterior distribution based on ω' , the average value of the index of satisfaction related to the law sampling the second phase. The problem that arises is that of the calculation of this hope in situations of tests, for a choice of prior distribution. Several models are considered to illustrate the Bayesian predictive procedure proposed.

2.4. Statistical inference for the design

We define the indices of satisfaction and anticipation of satisfaction related to a decreasing hypothesis test as a function of the p-value, satisfaction is higher than the null hypothesis is rejected more broadly, that is to say that p-value is small. We consider the case of a two-step procedure, which is often done in the case of clinical trials where these satisfaction indices are interesting protocols and when the inference concerns an effect evaluated from the future sample only.

Consider the satisfaction provided by the second phase of the experiment and the predicted using the first. We saplings under which the statistician "wants" to observe a

significant result, i.e. reject the null hypothesis Θ_0 . His "satisfaction" will be greater in the case of rejection, and even generally much larger than the observation that led to the rejection is more significant.

Being fixed α , a level α test defined by the critical first satisfaction index region $\Omega_1^{(\alpha)}$:

$$\phi(\omega'') = 1_{\Omega_1^{(\alpha)}}(\omega'')$$

The default of the above rudimentary index is that it expresses a satisfaction in "all or nothing". It is interesting to take into account to what level will the result always appears significant. One thus uses a new index of satisfaction defined by:

$$\begin{aligned} \phi(\omega'') &= 0 \text{ if } p(\omega'') \geq 1 - \alpha \\ &= L(p(\omega'')) \text{ else.} \end{aligned}$$

Where L is a decreasing function. We can generalize this procedure to a family of limited indices defined by:

$$L(p) = (1 - p)^l \text{ where } l \geq 0.$$

It is preferable to choose limited indexes because of their easier interpretation. In the case where $l=1$, $1 - \phi(\omega'')$ is the p -value and in the case where $l=0$, one finds the indicator function of the critical region.

For the sequel, we choose $l=1$, $L(p) = (1 - p)$ therefore

$$\begin{aligned} \phi(\omega'') &= 0 \text{ if } p(\omega'') \geq 1 - \alpha \\ &= (1 - p) \text{ else.} \end{aligned}$$

Based on the fact that most clinical trials meeting "legal" requirements (imposed by the control authorities for the authorization of placing drugs on the drug market) use as primary criterion of evaluation the significance level of a frequentist test, which is no else than the p -value. May we recall for this purpose that the p -value is always regarded as a measure of credibility to be attached to the null hypothesis that practitioners often use to answer several

criticisms and disadvantages of the Neymann Pearson approach

Recall that $p = \inf \{ \beta, \omega'' \in \Omega_1^{(\beta)} \}$ is what practitioners note the associated ω'' and is called the p -value, it is considered a measure of credibility to be attached to the null hypothesis and practitioners often use to meet several critical and disadvantages of the approach Neymann - Pearson, you can see why. Therefore, the more that p is, the more the practitioner Considers that the result is significant.

An indicator of prediction is given by:

$$\begin{aligned} \pi(\omega') &= \int_{\Omega_1^{(\alpha)}} \phi(\omega'') P_{\Omega''}^{\omega'}(d\omega'') \\ &= \int_{\Theta} \left(\int_{\Omega_1^{(\alpha)}} \phi(\omega'') P_{\Omega''}^{\theta}(d\omega'') \right) P_{\Theta}^{\omega'}(d\theta'). \end{aligned}$$

It is noticed that $\int_{\Omega_1^{(\alpha)}} \phi(\omega'') P_{\Omega''}^{\theta}(d\omega'')$ generalizes the power of the test in the logic of the index of satisfaction proposed. Therefore, this index of prediction can be used to determine whether the trial should be stopped early due to efficacy/futility or continued because the current data are not yet conclusive; it is the experimenter to take the final decision.

3. Applications of prediction models

It is proposed to calculate the prediction of satisfaction in several models where the law of the unknown parameter θ is a conjugate prior or non informative. The sequential aspect can be a particularly innovative element relative to existing technology, it helps alleviate studies multiphase, also authorized the statistical analysis, one observation at a time, sometimes is ethical or economic as it allows to stop quickly and least late the experience

3.1. Poisson model

For the Poisson model, the mean number of failures per unit time is the unknown model parameter.

Independent identically distributed observations are carried out, the first result is a sequence of n observations Poisson

parameter θ , Y_t ($t=1, 2, \dots, T$) and is chosen as prior distribution for θ the gamma distribution $G(\underline{\alpha}, \underline{\beta})$, where $\underline{\alpha} > 0$ and $\underline{\beta} > 0$. For reasons of completeness we set

$\omega' = \sum_{t=1}^T Y_t$ we know that θ / ω' follows the

Law $G(\bar{\alpha}, \bar{\beta})$, where $\bar{\alpha} = \underline{\alpha} + \omega'$ and $\bar{\beta} = (\underline{\beta}^{-1} + T^{-1})$. In

the case of a frequentist test, the threshold α , where the null hypothesis is of the type: $\theta \leq \theta_0$, the satisfaction index is given for a future observation of ω'' Poisson parameter θ by:

$$\phi^{(\alpha)}(\omega'') = 0 \text{ if } \omega'' < q_0$$

$$= \sum_{s=0}^{\omega''-1} e^{-n\theta_0} \frac{(n\theta_0)^s}{s!} \text{ if } \omega'' \geq q_0,$$

$$\text{where } q_0 = \inf \left\{ u; \sum_{s=0}^{u-1} e^{-n\theta_0} \frac{(n\theta_0)^s}{s!} \geq 1 - \alpha \right\},$$

secondly the density f of the predictive distribution of conditional ω'' knowing ω' is:

$$\begin{aligned} f(\omega''/\omega') &= \int_0^1 f(\omega''/\theta) f(\theta/\omega') d\theta \\ &= \int_0^1 \left[\frac{\theta^{\omega''} \exp(-\theta)}{\omega''!} \right] \left[\left(\bar{\beta} \bar{\alpha} \Gamma(\bar{\alpha}) \right)^{-1} \theta^{\bar{\alpha}-1} \exp(-\theta/\bar{\beta}) \right] d\theta \\ &= \left(\frac{(\bar{\beta}^{-1} + 1)^{-(\bar{\alpha} + \omega'')}}{\omega''! (\bar{\beta} \bar{\alpha} \Gamma(\bar{\alpha}))} \right) \end{aligned}$$

Therefore, the prediction becomes

$$\pi(\omega) = \sum_{\omega'=q_0}^{\infty} \sum_{s=0}^{\omega'-1} e^{-\theta_0} \frac{(\theta_0)^s}{s!} \left(\frac{(\bar{\beta}^{-1} + 1)^{-(\bar{\alpha} + \omega')}}{\omega''! \bar{\beta} \bar{\alpha} \Gamma(\bar{\alpha})} \right)$$

3.2. The Bernoulli model

The concept of sequential statistical methods was originally motivated by the need to obtain clinical benefits under certain economic constraints. For a trial with a positive result, early stopping means that a new product can be exploited sooner. If a negative result is indicated, early stopping ensures that resources are not wasted.

Several phase II clinical trial designs are proposed in the statistical literature almost of them are conducted in two stages. Let x_i be the dichotomous response variable [15], which assumes value 1 if the clinicians classify the patient i as responder to the treatment t and 0 otherwise. In a typical two-stage design, T patients are accrued and treated at the first stage. If the observed number of treatment successes

$\omega' = \sum_{i=1}^T X_i$ is less than or equal to r_1 , the experiment stops for lack of treatment efficacy. Otherwise, the trial continues to the second stage, which involves additional patients.

We denote θ the probability that an individual suffering from a disease is cured with the treatment t . It is also considered that the medication (treatment related t) may be marketed only if $\theta \geq \theta_0$. From a statistical point of view, we can formulate the problem using the following test:

$$H_0 : \theta \leq \theta_0.$$

We work in the framework of the sampling model where we assume that are realizations of independent random variables X_i and even Bernoulli parameter θ , again for the

sake of completeness we take $\omega' = \sum_{i=1}^T x_i$. If we choose as

prior distribution for θ a beta law $B(\underline{\alpha}, \underline{\delta})$ is then known

that the posterior distribution of θ/ω' is still a beta law $B(\bar{\alpha}, \bar{\delta})$ with $\bar{\alpha} = \underline{\alpha} + \omega'$ and $\bar{\delta} = \underline{\delta} + T + \omega'$

The satisfaction index for observation $\omega'' = 0$ or 1 is:

$$\phi(\omega'') = 0 \text{ if } \omega'' < q_0$$

$$= \sum_{s=0}^{\omega''-1} \theta_0^s (1-\theta_0)^{1-s} \text{ if } \omega'' \geq q_0,$$

with

$$q_0 = \inf \left\{ u; \sum_{s=u}^1 \theta_0^s (1-\theta_0)^{1-s} \leq \alpha \right\}.$$

On the other hand, the predictive density ω''/ω' is no other than the beta-binomial:

$$f(\omega''/\omega') = \int_0^1 f(\omega''/\theta) f(\theta/\omega') d\theta$$

$$= \int_0^1 \theta^{\omega''} (1-\theta)^{(1-\omega'')} [\beta(\bar{\alpha}, \bar{\delta})]^{-1} \theta^{\bar{\alpha}-1} (1-\theta)^{\bar{\delta}-1} d\theta$$

$$= \left[\frac{\Gamma(\bar{\alpha} + \omega'') \Gamma(\bar{\delta} + 1 - \omega'')}{\Gamma(\bar{\alpha} + \bar{\delta} + 1)} \right] \left[\frac{\Gamma(\bar{\alpha} + \bar{\delta})}{\Gamma(\bar{\alpha}) \Gamma(\bar{\delta})} \right]$$

Hence the prediction of satisfaction is

$$\pi(\omega) = \sum_{\omega'=q_0}^1 \sum_{s=0}^{\omega'-1} \theta_0^s (1-\theta_0)^{1-s} \frac{B(\bar{\alpha} + \omega', \bar{\delta} + 1 - \omega')}{B(\bar{\alpha} + \bar{\delta})}$$

Moreover, due to ethical considerations, the phase II clinical studies are planned as a multi-stage design to ensure that the trials do not last too long if the treatment shows a clear inadequateness.

3.3. The Binomial model

Note that we can generalize the (3.2) Bernoulli model and take T' additional patients, then $\omega'' = \sum_{i=1}^n X_i''$, The satisfaction index is:

$$\phi(\omega'') = 0 \text{ if } \omega'' < q_0$$

$$= \sum_{t=0}^{\omega''-1} C_N^t \theta_0^t (1-\theta_0)^{N-t} \text{ if } \omega'' \geq q_0,$$

where

$$q_0 = \inf \left\{ u; \sum_{t=u}^{T'} C_{T'}^t \theta_0^t (1-\theta_0)^{T'-t} \leq \alpha \right\}.$$

Then the Bayesian prediction distribution of ω''/ω'

$$f(\omega''/\omega') = \int_0^1 f(\omega''/\theta) f(\theta/\omega') d\theta$$

$$= \int_0^1 C_{T'}^{\omega''} \theta^{\omega''} (1-\theta)^{(T'-\omega'')} [\beta(\bar{\alpha}, \bar{\delta})]^{-1}$$

$$\times \theta^{\bar{\alpha}-1} (1-\theta)^{\bar{\delta}-1}$$

$$= C_{T'}^{\omega''} \frac{[\beta(\bar{\alpha} + \omega'', \bar{\delta} + T' - \omega'')]}{[\beta(\bar{\alpha}, \bar{\delta})]}$$

From here on, observed the response of the first step ω' , the prediction $\pi(\omega')$ is

$$\sum_{\omega'=q_0}^{T'} \sum_{t=0}^{\omega'-1} C_{T'}^t \theta_0^t (1-\theta_0)^{T'-t} \frac{C_{T'}^{\omega'} \beta(a + \omega' + \omega'', b + T + T' - \omega' - \omega'')}{\beta(a + \omega', b + T + \omega')}$$

We illustrate this Bayesian design with two kinds of prior distributions.

3.3.1 An illustrative example for a uniform prior

Prediction models are important in various fields, including physics, meteorology, and reliability. The predictive probability design under the Bayesian framework provides an ideal environment for learning. For example, in reliability, in the case of failure probabilities of launch vehicles, we need to estimate the number of new launch vehicles that will succeed in, say, T future launches scheduled during the next calendar year, we can see [16].

Calculating the predictive distribution for the number of successes of a new launch vehicle. Assume that a uniform distribution $B(1, 1)$ is used to model the prior distribution on the launch vehicle success probability. With this prior distribution, we previously found the prediction when

$$T = T' = 20 \text{ and } \theta_0 = 0.5$$

Simulation results

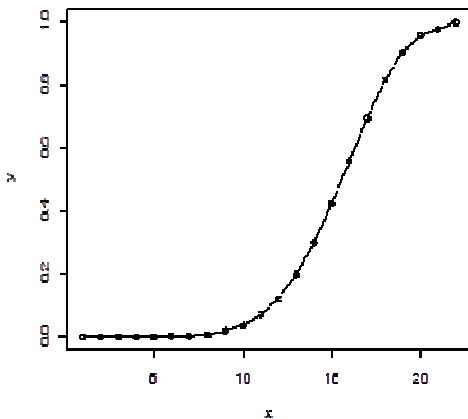


Fig. 1. the graphical presentation of prediction with uniform prior

Table I: Index of prediction with uniform prior $B(1, 1)$

Vehicle	Observations ω'	$\pi(\omega')$
1	0	0.00000006
2	1	0.000005
3	2	0.000015
4	3	0.000076
5	4	0.00031
6	5	0.00108
7	6	0.00314
8	7	0.00797
9	8	0.01807
10	9	0.0371
11	10	0.0701
12	11	0.1221
13	12	0.1977
14	13	0.2987
15	14	0.4220
16	15	0.5591
17	16	0.6957
18	17	0.8153
19	18	0.9040
20	19	0.9559
21	20	0.9765

We note in Table I, where ω' is included in $[0, 17]$, the result of $\theta_0 = 0.5$ varies from 0.00000006 to 0.8153.

Therefore we conclude H_0 for $\omega' < 17$. On the other hand, when ω' is included in $[18, 20]$ the result $\pi(\omega')$ varies from 0.9040 to 0.9765. In this case, instead of deciding for $H_1: \theta > 0.5$ considering the predictive probability is the weighted average of the indicator of a test which runs until the end of the study.

3.3.2 An illustrative example for non informative prior

The aim of exploratory clinical trials, such as phase II trials and proof-of-concept studies, is to determine whether a new intervention is promising for further testing in confirmatory clinical trials, such as phase III randomised controlled trials.

The clinical trial, a prospective study to evaluate the effect of interventions in humans under prespecified conditions, is a standard and integral part of modern medicine. Many adaptive and sequential approaches have been proposed for use in clinical trials to allow adaptations or modifications to aspects of a trial after its initiation without undermining the validity and integrity of the trial.

In all rigor, the Jeffreys rule gives different priors for the different designs, since it is based on the Fisher information, showd that Jeffreys prior offers new perspectives for the development of Bayesian procedures with good frequentist properties in hypothesis testing procedures.

Suppose two imaging modalities (e.g., CT vs.MRI) for diagnosing lung cancer are to be compared on the basis of test accuracy (sensitivity, specificity, and the area under the ROC curve) [17].

Suppose $T = T' = 20$ are the sample sizes of the two groups and the prior probability of the null hypotheses is $\theta_0 = 0.5$,

The predictive probability at each point $\pi(\omega')$ is calculated via simulation in table II.

Simulation results

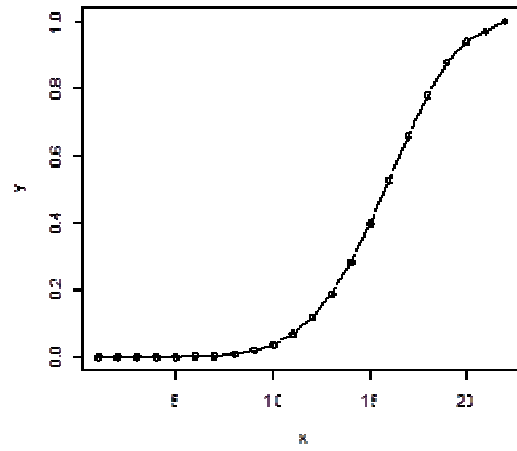


Fig. II. The graphical presentation of prediction with non informative prior

Table. II: Representation of prediction with non-informative prior B (0.5, 0.5).

Patients	Observations ω'	$\pi(\omega')$
1	0	0.000007
2	1	0.000006
3	2	0.00002
4	3	0.00010
5	4	0.00039
6	5	0.00125
7	6	0.0034
8	7	0.0083
9	8	0.0183
10	9	0.0367
11	10	0.0679
12	11	0.1166
13	12	0.1869
14	13	0.2809
15	14	0.3965
16	15	0.5269
17	16	0.6602
18	17	0.7818
19	18	0.8779
20	19	0.9404
21	20	0.9708

We note in Table II where ω' is included in $[0, 18]$, the result of $\theta_0 = 0.5$ varies from 0.0000007 to 0.8153. Therefore we conclude H_0 for $\omega' < 18$. On the other hand, when ω is included in $[18, 20]$ the result p varies from 0.9040 to 0.9765. In this case, instead of deciding for H_1 considering the predictive probability is the weighted average of the indicator of a test which runs until the end of the study.

As can be concluded from two tables when the parameters of the beta law are equal and equal to 0.5 and the value of $\theta_0 = 0.5$ we get the best convergence predictive probability.

The predictive probability approach still offers a consistent way to evaluate the strength of the treatment efficacy based on the observed data.

3.4 Gaussian model

Several cases are considered according to the choice of the a priori and the study of the test.

We perform independent observations and of same normal random variable $N(\theta, \sigma^2)$. In all that follows, Φ (resp. ϕ) indicates the cumulative distribution function (resp. the density) of the distribution $N(0, 1)$.

The first result is $\omega' = \frac{\sum_{t=1}^T Y_t}{T}$ of distribution $N(\theta, \sigma_1^2)$

with $\sigma_1^2 = \frac{\sigma^2}{T}$, and the second result $\omega'' = \frac{\sum_{t=1}^{T'} Y_t}{T'}$, ω'' is

still a normal distribution $N(\theta, \sigma_2^2)$ with $\sigma_2^2 = \frac{\sigma^2}{T'}$.

We wish to test a null assumption of type $\theta \leq \theta_0$. The distribution of the result ω'' is obviously stochastically increasing in θ ,

We use here a usual test ranging on ω'' , whose critical region is $]q_0, +\infty[$, where $q_0 = \theta_0 + \sigma_2 u_{\alpha}^+$, u_{α}^+ indicating the upper α quantile of the standard normal distribution

$N(0, 1)$: $\Phi(u_{\alpha}^+) = 1 - \alpha$. The satisfaction index is naturally defined as:

$$\begin{aligned} \phi(\omega'') &= 0 \quad \text{if } \omega'' < q_0 \\ &= \Phi\left(\frac{\omega'' - \theta_0}{\sigma_2}\right) \quad \text{if } \omega'' \geq q_0 \end{aligned}$$

To study the prediction we distinguish two cases:

3.4.1. Known variance: variance σ^2 is known but θ unknown

1-If we choose a conjugate prior, we can see [18] if θ has a normal prior distribution $N(\delta, \tau^2)$, then the posterior distribution of θ knowing ω' is a normal distribution:

$$N\left(\frac{\tau^2 \omega' + \sigma_1^2 \delta}{\sigma_1^2 + \tau^2}, \frac{\sigma_1^2 \tau^2}{\sigma_1^2 + \tau^2}\right)$$

Finally the forecasting satisfaction is given by:

$$\pi(\omega') = \int_{q_0}^{+\infty} \left(\int_{-\infty}^{\omega'' - \theta_0} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \right) v(\omega''/\omega') d\omega''$$

Where $v(\omega''/\omega')$ is the conditional distribution density of ω'' knowing ω' .

One quotes that, being σ_1^2 , σ_2^2 , δ , α and ω' given, one wants to calculate

$$\pi(\omega') = \frac{1}{s} \int_{q_0}^{+\infty} \left(\Phi\left(\frac{\omega'' - \theta_0}{\sigma_2}\right) \right) \phi\left(\frac{\omega'' - m}{s}\right) d\omega''$$

$$\text{where } m = \frac{\tau^2 \omega' + \sigma_1^2 \delta}{\sigma_1^2 + \tau^2}, \quad s^2 = \sigma_2^2 + \frac{\tau^2 \sigma_1^2}{\sigma_1^2 + \tau^2}.$$

by change of variable, one obtains

$$\pi(\omega') = \int_{-bx' + d\theta'_0 + tu_\alpha^+}^{\infty} \left(\Phi \left(\frac{z - (-bx' + d\theta'_0)}{t} \right) \right) \varphi(z) dz$$

$$\text{Where, } x' = \frac{\omega' - \delta}{\tau}, \quad \theta'_0 = \frac{\theta_0 - \delta}{\tau}, \quad \sigma'_1 = \frac{\sigma_1}{\tau}, \quad \sigma'_2 = \frac{\sigma_2}{\tau},$$

$$b = \left[(1 + \sigma_1^2) (\sigma_1^2 + \sigma_2^2 + \sigma_1^2 \sigma_2^2) \right]^{\frac{1}{2}}, \quad d = (1 + \sigma_1^2) b, \quad t = \sigma_2^2 d.$$

It is noted that $\pi(\omega')$ only depends of the three real numbers which we will call essential parameters: two parameters of scale, d and t, in the expressions of which intervene only the ratios of variances $\frac{\sigma_1^2}{\tau^2}$ and $\frac{\sigma_2^2}{\tau^2}$ and a

location parameter $a = -bx' + d\theta'_0$.

At threshold α and at fixed scale parameters, a modification of θ_0 and δ has a translation effect on $\pi(\omega')$ if θ'_0 increases by $\Delta\theta'_0$, the representing curve of π undergoes a horizontal adjustment of amplitude $\frac{d}{b} \Delta\theta'_0$ where $\frac{d}{b} = 1 + \sigma_1^2$.

In order to carryout the calculations of $\pi(\omega')$ using a Monte-Carlo method. We rewrite $\pi(\omega')$ in the form

$$\pi(\omega') = [1 - \Phi(a + tu_\alpha^+)] \int_{\mathbb{R}} \left(\Phi \left(\frac{z - a}{t} \right) \right) \frac{\varphi(z)}{1 - \Phi(a + tu_\alpha^+)} 1_{[a + tu_\alpha^+, \infty[}(z) dz$$

where $\frac{\varphi}{1 - \Phi(a + tu_\alpha^+)} 1_{[a + tu_\alpha^+, \infty[}$ is the probability density Q,

deduced from the cumulative distribution function of the standard normal distribution by conditioning by the event $[a + tu_\alpha^+, \infty[$.

The Monte-Carlo method then consists in approaching by [19, 20].

$$[1 - \Phi(a + tu_\alpha^+)] \left[\frac{1}{N} \sum_{i=1}^N \left(\Phi \left(\frac{Z_i - a}{t} \right) \right) \right],$$

(2)

where the Z_i are N realizations of the probability Q. The pulling of the Z_i proceeds in the following way:

- U_i : is drawn according to the uniform distribution on $[0,1]$,

- $V_i = \Phi(a + tu_\alpha^+) + (1 - \Phi(a + tu_\alpha^+))U_i$, i.e., that V_i follows the uniform distribution on $[\Phi(a + tu_\alpha^+), 1]$,

- $Z_i = \Phi^{-1}(V_i)$, i.e., that Z_i follows the distribution Q.

Simulation setting

$T=10, T'=20, \sigma^2=1, N=50, \theta_0=0, \delta=0, \tau=1, Pas=0,05$.

We present in Figure (III), the comparative curves of the prediction in the case of the index of the literature and the proposed index defined as a function of the p -value. We can see that these curves are very close, but they break away when x is rather large, i.e., superior to θ_0 , which conveys well the interest of the consideration of the p -value in the index of satisfaction that the only rejection of the assumption is all the more informative since ω' is larger,

which proves well that the index that we propose is better.

(29)

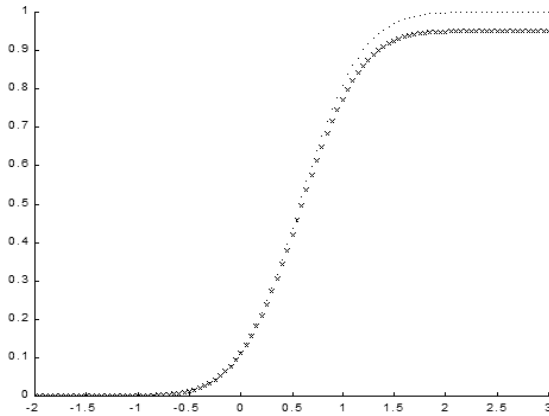


Fig. III: Representation of Forecasting

2 - If we choose as a priori θ a prior non informative in the sense of Jeffreys, we know that θ/ω' follows a

$$N(\omega', \sigma_1^2).$$

On the other hand, the predictive density ω''/ω' is

$$N(\omega', \sigma_1^2 + \sigma_2^2)$$

we immediately deduce the prediction $\pi(\omega')$ given by:

$$\pi(\omega') = \int_{q_0}^{+\infty} \Phi\left(\frac{\omega'' - \theta_0}{\sigma_2}\right) f(\omega''/\omega') d\omega''.$$

Make a change of variable:

$$z_0 = \frac{q_0 - \omega'}{\sqrt{\sigma_1^2 + \sigma_2^2}} \quad \text{and} \quad z = \frac{\omega'' - \omega'}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

It has just so:

$$\pi(\omega') = \int_{z_0}^{+\infty} \Phi\left(\frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma} z + \frac{\omega' - \theta_0}{\sigma}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

This integral is approximated by Monte-Carlo methods, the figure (IV), represents the prediction.

Simulation setting

$T=10, T'=10, \sigma^2=1, N=50, \theta_0=0, \delta=0, \tau=1, Pas=0,05.$

$Pas =0.05, \text{sigma1}=0.31623, \text{sigma2}=0.31623, \text{delta}=0,$

$\text{tau}=1, \text{alpha}=0.05,$

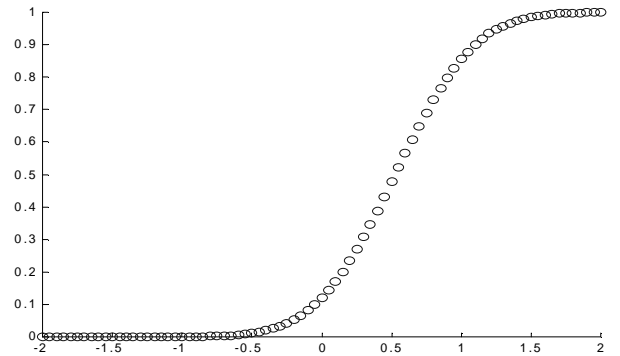


Fig.IV. Prediction in the case of non informative prior.

3.4.2 Unknown variance

If the mean θ and variance σ_1^2 are both unknown,

assume that the prior distribution for θ and σ_1^{-2} is the

normal gamma $NG(\underline{\mu}, \underline{T}^{-1}, \underline{s}^{-2}, \underline{\nu})$.

We know that, the posterior distribution of θ and σ_1^{-2} is

still $NG(\bar{\mu}, \bar{T}^{-1}, \bar{s}^{-2}, \bar{\nu})$ with:

$$\bar{\mu} = \bar{q}(\bar{T}^{-1} \mu + T \omega'), \quad \bar{q} = (\bar{T}^{-1} + T)^{-1} \quad \text{and} \quad \bar{\nu} = \nu + T$$

On the other hand, the predictive density ω''/ω' is given

by:

$$f(\omega''/\omega') = \int_0^{\infty} \left(\omega''/\bar{\mu}, \sigma_1^2 [1 + \bar{T}^{-1}]\right) f_{\gamma}(\sigma_1^{-2}/\bar{s}^{-2}, \bar{\nu}) d\sigma_1^{-2}$$

$$\begin{aligned}
&= \int_0^{\infty} \left(2\pi \sigma_1^2 \left[1 + \bar{T}^{-1} \right] \right)^{-1/2} \exp \left[-\frac{(\omega'' - \bar{\mu})^2}{2 \sigma_1^2 \left[1 + \bar{T}^{-1} \right]} \right] \times \\
&\left(\left[\frac{2}{\bar{v}s} \right]^{-1} \Gamma(\bar{v}/2) \right)^{-1} \sigma_1^{-2(\bar{v}-2)/2} \exp \left(-\frac{1}{2} \sigma_1^{-2\bar{v}s} \right) d \sigma_1^{-2} \\
&= f_t \left(\omega'' / \bar{\mu}, s^{-2} \left[1 + \bar{T}^{-1} \right], \bar{v} \right)
\end{aligned}$$

Where f_t is the density of the Student \bar{v} degrees of freedom, mean $\bar{\mu}$ and scale parameter: $s^{-2} \left[1 + \bar{T}^{-1} \right]$.

We explicitly derived prediction:

$$\pi(\omega') = \int_{z_0}^{+\infty} \Phi \left(\frac{\omega' - \theta_0}{\sigma} \right) f_t \left(\omega'' / \bar{\mu}, s^{-2} \left[1 + \bar{T}^{-1} \right], \bar{v} \right) d\omega''$$

We limit ourselves to present the prediction only explicitly because its approximation is mean by the Monte Carlo method exactly as 3.4.1.

4. Conclusion

In this paper, we propose Bayesian two-stage designs for experimental trials. The main contribution of our work was to closely related to two-stage and more in general sequential procedures in experimental data for early termination due to futility. This is a well debated issue in the recent literature, in particular Bayesian stochastic curtailment, i.e. early stopping for futility based on predictive distributions. The Bayesian predictive approach enables stopping the trial early or conversely extending it to an adequate size, in a sequential perspective, as illustrated in our examples in clinical protocols or in reliability. This fits particularly well with the methodology of adaptive designs. An extra advantage of our design is that easy to

compute the predictive probability for different models in experimental trials.

ACKNOWLEDGEMENTS

The authors are grateful to the anonymous reviewers for their useful suggestions and constructive comments that led to improvements in the paper.

References:

- [1] Y-Guosheng, Phase II trial design with Bayesian adaptive randomization and predictive probability, *Appl. Statist.* (2012) 61, Part 2, pp. 219–235
- [2] S.BERNARDO., *Bayesian theory*, Wiley, 1994.
- [3] B. Lecoutre., *Bayesian methods for experimental data analysis*, In C.R. Rao, J. Miller, D.C. Rao, (eds.), *Handbook of statistics: Epidemiology an medical statistics*, (vol. 27), Amsterdam: Elsevier, (2008), 775–812.
- [4] J.J. Lee, DD. Liu , A predictive probability design for phase II cancer clinical trials, *Clin Trials.* (2008); 5(2):93-106.
- [5] V .Sambucini, A Bayesian predictive two-stage for phase II clinical trials, *Statists. Med.* (2008) ; 27 :1199 -1224
- [6] S. Teramukai, T. Daimon and S. Zohar, A Bayesian predictive sample size selection design for single-arm exploratory clinical trials, *statistics in medicine*, (2012), 10.1002/sim.5505.
- [7] BW .Brown, J. Herson, Atkinson, N and ME Rozelle. Projection, from previous studies: a Bayesian and frequentist compromise. *Controlled clinical trials*, 1987, 8, 29-44.
- [8] SC. Choi and PA. Pepple, Monitoring clinical trials based on predictive probability of signicance, *Biometrics*, (1989), 45, 317-323.

- [9] JM. Grouin., Procédures bayésiennes prédictives pour les essais expérimentaux, thèse de Doctorat de l'Université René Descartes, 1994, Paris.
- [10] Merabet, H. and Raoult J. P., Les indices de satisfaction, outils classiques et bayésiens pour la prédiction statistique, ASU, XXVII^{èmes} Journées de Statistique, Jouy-en-Josas. (1995).
- [11] H. Merabet, Méthodologie des essais préliminaires. Calcul de prévision de satisfaction dans le cas Gaussien, Rev Stat Appl, LII (3), (2004), 93-106.
- [12] H. Merabet, Index and prevision of satisfaction in exponential models for clinical trials. STATISTICA, anno LXIV, n.3, (2004), 441-453.
- [13] G. Koop, D.J. Poirier, and J. L. Tobias, Bayesian Econometric Methods, 2007.
- [14] C. M. C. Shein Chung, Adaptive design methods in clinical trials, Chapman & Hall/CRC, 2011.
- [15] P. Bunouf, B. Lecoutre, Bayesian priors in sequential binomial design, Comptes Rendus Mathematique; (2006) 343, 5; 339-344.
- [16] S. Michael, G. Hamada Alyson, C. Wilson. Shane Reese Harry F. Martz, Bayesian Reliability, Springer Series in Statistics, 2008.
- [17] H. Merabet, A. Labdaoui, Bayesian Prediction of Performance in Clinical Trials, J. of Math and System Science 3 (2013) 342-348.
- [18] C.P. Robert, Bayesian choice, Springer-Verlag, Paris. 2006.
- [19] CP. Robert, The Bayesian Choice, 2nd ed, springer, 2007.
- [20] CP. Robert and G. Casella, Monte Carlo Statistical Methods, Springer, 2004.