The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://www.tandfonline.com/loi/utas20

# Relationship Between Bayesian and Frequentist Sample Size Determination

Lurdes Y.T Inoue, Donald A Berry & Giovanni Parmigiani

**To cite this article:** Lurdes Y.T Inoue, Donald A Berry & Giovanni Parmigiani (2005) Relationship Between Bayesian and Frequentist Sample Size Determination, The American Statistician, 59:1, 79-87, DOI: 10.1198/000313005X21069

**To link to this article:** http://dx.doi.org/10.1198/000313005X21069

Published online: 01 Jan 2012.

# Relationship Between Bayesian and Frequentist Sample Size Determination

Lurdes Y. T. INOUE, Donald A. BERRY, and Giovanni PARMIGIANI

Sample size determination is among the most commonly encountered tasks in statistical practice. A broad range of frequentist and Bayesian methods for sample size determination can be described as choosing the smallest sample that is sufficient to achieve some set of goals. An example for the frequentist is seeking the smallest sample size that is sufficient to achieve a desired power at a specified significance level. An example for the Bayesian is seeking the smallest sample size necessary to obtain, in expectation, a desired rate of correct classification of the hypothesis as true or false. This article explores parallels between Bayesian and frequentist methods for determining sample size. We provide a simple but general and pragmatic framework for investigating the relationship between the two approaches, based on identifying mappings to connect the Bayesian and frequentist inputs necessary to obtain the same sample size. We illustrate this mapping with examples, highlighting a somewhat surprising "approximate functional correspondence" between power-based and information-based optimal sample sizes.

KEY WORDS: Lindley information; Prior information; Probabilities of Type I and Type II errors.



## 1. INTRODUCTION

### 1.1 Overview

Sample size determination is among the most commonly encountered tasks in statistical practice. It is a routine component of the design of both clinical trials and epidemiologic studies, and it plays a critical role in the evaluation of most biomedical research proposals. A common approach is to use frequentist operating characteristics of hypothesis testing procedures. For example, one chooses the smallest sample size sufficient to achieve specified power and significance level. This approach is deeply engrained in statistical practice. Moreover, many Bayesian statisticians—including the authors—take this approach. Have we abandoned our statistical philosophy in so doing? In this article we give a negative answer.

The purpose of this article is to explore parallels between Bayesian and frequentist sample size calculations. We seek functional relationships. Obviously, a Bayesian and a frequentist statistician can come up with the same sample size for particular assumptions about the parameters of interest. The question we address is whether the two will continue to agree if those assumptions change. We provide a simple but general approach for investigating the relationships between the two approaches by identifying mappings that connect the Bayesian and frequentist inputs to sample size calculation. In the process we clarify the differences between Bayesian and frequentist approaches to sample size determination. Our message is ecumenical: Bayesians and frequentists are not very different!

### 1.2 Framework

In problems of sample size determination one seeks to achieve a balance between the cost of the experiment and the quality of the final analysis. As a result, a broad range of frequentist and Bayesian methods for sample size determination can be described as choosing the smallest sample size that is sufficient to achieve, in expectation, some set of goals. An example for the frequentist is seeking the smallest sample size that is sufficient to achieve a desired power for a specified significance level. An example for the Bayesian is seeking the smallest sample size necessary to obtain, in expectation, a desired rate for correctly identifying a hypothesis as true or false. When the goal of a study is estimating a parameter, goals can be, for example, the widths of confidence and probability intervals, respectively. We will furnish a common formal structure that will help identify differences between these two approaches, and we will establish correspondences between them.

We formalize the sample size selection goal by specifying a goal function $G$. Examples of this function include power, information, mean squared prediction error, size of confidence interval or probability interval, and classification error. We use subscripts B and F to denote Bayesian and Frequentist analyses, respectively. The function $G$ depends on the sample size and also other user-specified inputs. The vector $\mathbf{u}$ denotes a frequentist input (say significance level, target value within a composite alternative, desired size of confidence interval, etc.). The vector $\mathbf{v}$ denotes a Bayesian input (prior distribution of the hypotheses or parameter of interest, loss from incorrect classification, etc.). Sampling to achieve a prespecified criterion that increases with $n$ can then be represented as finding the smallest $n$ such that

$$G_{\mathrm{F}}(n, \mathbf{u}) \geq G^*_{\mathrm{F}} \quad \text{frequentist} \tag{1}$$

$$G_{\mathrm{B}}(n, \mathbf{v}) \geq G^*_{\mathrm{B}} \quad \text{Bayesian,} \tag{2}$$

Lurdes Y. T. Inoue is Assistant Professor, Department of Biostatistics, University of Washington, Box 357232, Seattle, WA 98195 (E-mail: linoue@u.washington.edu). Donald A. Berry is Professor and Chairman, Department of Biostatistics, M.D. Anderson Cancer Center, University of Texas, Box 447, Houston, TX 77030 (E-mail: dberry@odin.mdacc.tmc.edu). Giovanni Parmigiani is Associate Professor, Departments of Oncology and Biostatistics, Johns Hopkins University, 550 North Broadway, Baltimore, MD 21205 (E-mail: gp@jhu.edu). The work of Lurdes Y. T. Inoue was partially supported by a fellowship from CNPq, Ministry of Science and Technology of Brazil, process number 200123/96-3 and by the Career Development Funding from the Department of Biostatistics, University of Washington. The authors wish to thank the editor, associate editor, and referees for their helpful suggestions.

 DOI: 10.1198/000313005X21069

Table 1. Summary of Illustrative Examples Considered Here

| Likelihood | Hypothesis | Goal | Section |
|---|---|---|---|
| Normal | simple vs simple | power/classification error | 2.1 |
| Normal | composite vs composite | power/classification error | 2.2 |
| Normal | point null vs composite | power/classification error | 2.3 |
| Bernoulli | simple vs simple | power/classification error | 2.4 |
| Normal | composite vs composite | power/information | 3.2 |
| Bernoulli | simple vs simple | power/information | 3.3 |

where $G_{\mathrm{F}}^*$ and $G_{\mathrm{B}}^*$ denote, respectively, the desired levels of the frequentist and Bayesian goal functions.

Algorithmically, both Bayesian and frequentist approaches define a mapping from the set of user-specified inputs to an optimal sample size. Formally, solutions to the sample size minimization under constraints (1) and (2) exist when $G_{\mathrm{F}}$ and $G_{\mathrm{B}}$ are monotonic functions of $n$ and we can write the two mappings as

$$\begin{aligned} n_{\mathrm{F}}^* &= \min\{n \in \mathcal{N} : G_{\mathrm{F}}(n, \mathbf{u}) \geq G_{\mathrm{F}}^*\} \\ &= F(\mathbf{u}) \qquad \text{frequentist} \end{aligned} \tag{3}$$

$$\begin{aligned} n_{\mathrm{B}}^* &= \min\{n \in \mathcal{N} : G_{\mathrm{B}}(n, \mathbf{v}) \geq G_{\mathrm{B}}^*\} \\ &= B(\mathbf{v}) \qquad \text{Bayesian.} \end{aligned} \tag{4}$$

Here $\mathcal{N}$ denotes the set of integers. The functions F and B represent the algorithms for sample size determination, and depend on the specified objectives of experimentation.

For equal sample sizes, the functions F and B implicitly define two subsets of values of $\mathbf{u}$ and $\mathbf{v}$ such that $n_{\mathrm{F}}^* = n_{\mathrm{B}}^* = n^*$. If the prespecified goals $G_{\mathrm{F}}^*$ and $G_{\mathrm{B}}^*$ are kept constant, then varying $n^*$ and keeping $n_{\mathrm{F}}^* = n_{\mathrm{B}}^* = n^*$ defines two parallel partitions of the two sets of inputs. The family of such partitions reflects the relationship between the two algorithms.

As an example, consider testing a null hypothesis $\theta = \theta_0$ against a simple alternative $\theta = \theta_1$. Suppose a frequentist uses power to determine $n_{\mathrm{F}}^*(\theta_1)$ and a Bayesian uses expected information to determine $n_{\mathrm{B}}^*(\theta_1)$ and that the resulting sample sizes are equal. There is nothing particularly surprising about this circumstance because there are potentially many frequentist/Bayesian pairs in which both would choose the same sample size. However, if $\theta_1$ changes, then $n_{\mathrm{F}}^*(\theta_1)$ and $n_{\mathrm{B}}^*(\theta_1)$ change accordingly and one expects that they will be different. But suppose that $n_{\mathrm{F}}^*(\theta_1) = n_{\mathrm{B}}^*(\theta_1)$ for all possible $\theta_1$. That would demonstrate a remarkable stronger equivalence between using power in a frequentist approach and using information in a Bayesian approach.

### 1.3 Earlier Work and Outline of the Article

Choosing the sample size for an experiment lends itself naturally to decision-theoretic approaches. The first explicit decision-theoretic modeling of sample size determination is frequentist and can be traced back at least to the minimax approach of Wald (1950) and the fiducial approach of Grundy, Healy, and Rees (1956). The common frequentist approach of seeking the sample size to achieve a desired power at a specified significance level also has decision-theoretic roots (Neyman and Pearson 1933). Desu and Raghavarao (1990) provided an extensive discussion of frequentist methods to sample size determination. Fully Bayesian approaches to the sample size choice are in the context of utility-based optimization (Raiffa and Schlaifer 1961; Berger 1985; and Lindley 1997). Bayesians who take this perspective do not recognize themselves immediately in the "goal sampling" approach described earlier, although there are strong similarities between goal sampling and a constrained multiobjective decision problem in which one minimizes cost subject to a constraint on the expected quality of the final analysis.

Most recent literature on Bayesian sample size determination follows the framework outlined in Section 1.2 (see, e.g., Spiegelhalter and Freedman 1986; Joseph, Wolfson, and Berger 1995; Adcock 1997; Weiss 1997; and Pham-Gia and Turkkan 2003). We address a different issue. Despite philosophical differences between Bayesians and frequentists, could they agree when determining sample size? The answer seems obvious given the parallels between the two at the inferential stage when Bayesians use noninformative priors. But, at the design stage, data are yet to be observed. We illustrate that substantial agreement can occur and that it is not restricted to the class of noninformative priors.

In the remainder of the article we illustrate important parallels between frequentist and Bayesian sample size determination in simple and commonly encountered applications, as summarized in Table 1. We will explore two types of parallels between B and F: (a) parallels based on B and F that are derived from the same underlying goals, to highlight differences induced by the type of approach chosen; and (b) parallels based on B and F that are derived from goals that are apparently different.

In Section 2 we assume that $G_{\mathrm{F}}$ is power and $G_{\mathrm{B}}$ is the rate of correct classification. These are both related to controlling error rates in repeated sampling. In the normal case, for all three scenarios considered, we identify a simple relationship between the prior distribution in a Bayesian analysis and the target alternative value in a frequentist analysis.

In Section 3 we compare frequentist power to Bayesian information. In 3.1 we introduce the notion of Lindley information and use it in Sections 3.2 and 3.3 for Bayesian sample size determination. In Section 4 we provide a final discussion.

## 2. CONTROLLING CLASSIFICATION ERROR

Consider a random sample $x_1, \ldots, x_n$ from $N(\theta, \sigma^2)$, the normal distribution with unknown mean $\theta$ and known variance $\sigma^2$. The sample mean, $\bar{x}$, is a sufficient statistic. Using conventional notation, let $H_0$ and $H_1$ denote, respectively, the null and alternative hypotheses. Type I and Type II error rates are, respec-

*Table 2. Loss Function: No Loss is Incurred With a Correct Decision, but a Loss of 1 is Incurred if $H_0$ is not Rejected When in Fact $H_1$ is True, and a Loss of K is Incurred if $H_0$ is Rejected When in Fact $H_0$ is True*

| | $H_0$ is True | $H_1$ is True |
|---|---|---|
| Do not Reject $H_0$ | 0 | 1 |
| Reject $H_0$ | $K$ | 0 |

tively, $\alpha$ and $\beta$ and $z_\alpha$ is the $\alpha$-quantile of the standard normal distribution, that is, $\Phi(z_\alpha) = P(Z \le z_\alpha) = \alpha$.

A standard frequentist approach to sample size determination in many biomedical applications is to specify $\alpha$ and $\beta$ and a "relevant difference," $\delta$, between the null and the alternative hypotheses. In testing a simple hypothesis against a simple alternative, $\delta$ is the difference of the hypothesized values. For composite alternative hypotheses, $\delta$ is chosen to lie somewhere within the alternative set. We take this to be the reference frequentist approach throughout this section.

### 2.1 Simple Versus Simple Hypotheses

We want to test the hypothesis $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$, where $\theta_1 > \theta_0$. Under the frequentist approach, when testing $H_0$ versus $H_1$ at level of significance $\alpha$, the critical region is $\bar{x} \ge \theta_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$. When using power as the frequentist goal function, we find the sample size $n_F$ so that the test has power $1 - \beta$ at the alternative hypothesis and obtain

$$n_F = (z_\alpha + z_\beta)^2 \left(\frac{\sigma}{\delta}\right)^2, \tag{5}$$

where $\delta = \theta_1 - \theta_0$. In terms of the notation of Section 1.2, the frequentist input for sample size determination is $\mathbf{u} = (\alpha, \delta, \sigma)$.

Under the Bayesian approach, uncertainty about $\theta$ is represented by its prior distribution. Assume a priori that $P(\theta_0) = 1 - P(\theta_1) = \pi$. A Bayesian decision between $H_0$ and $H_1$ is based on their posterior probabilities. Suppose that the null hypothesis $H_0$ is not rejected if the posterior probability of the null hypothesis is at least $1/(1 + K)$. This cutoff for the posterior probability is consistent with a $0 - 1 - K$ loss function shown in Table 2 and minimizes the posterior expected loss (Berger 1985). Moreover, this cutoff implies that the null hypothesis is not rejected if

$$\bar{x} \le \frac{\sigma^2 \log(K\frac{\pi}{1-\pi})}{n\delta} + \frac{\theta_1 + \theta_0}{2}.$$

In this example we choose the Bayesian goal function to be the rate of correctly classifying a hypothesis as true or false. Using again the notation of Section 1.2, the Bayesian input for sample size determination is $\mathbf{v} = (\pi, K, \delta, \sigma)$ and the Bayesian goal function can be formally calculated as

$$\begin{aligned} & G_{\mathrm{B}}(n, \mathbf{v}) \\ &= KP(H_0)P(\text{correct decision}|H_0) \\ &\quad + P(H_1)P(\text{correct decision}|H_1) \\ &= K\pi P_{\theta_0}\left(\bar{x} \le \frac{\sigma^2 \log(K\frac{\pi}{1-\pi})}{n\delta} + \frac{\theta_1 + \theta_0}{2}\right) \\ &\quad + (1-\pi)P_{\theta_1}\left(\bar{x} > \frac{\sigma^2 \log(K\frac{\pi}{1-\pi})}{n\delta} + \frac{\theta_1 + \theta_0}{2}\right) \\ &= K\pi P\left(Z \le \frac{\sigma \log(K\frac{\pi}{1-\pi})}{\sqrt{n}\delta} + \frac{\delta\sqrt{n}}{2\sigma}\right) \\ &\quad + (1-\pi)P\left(Z \ge \frac{\sigma \log(K\frac{\pi}{1-\pi})}{\sqrt{n}\delta} - \frac{\delta\sqrt{n}}{2\sigma}\right), \end{aligned} \tag{6}$$

where $Z$ is the standard normal.

The Bayesian goal function in Equation (6) relates to the Bayes risk (Berger 1985) and weighs the probability of making a correct decision by the prior probabilities of reaching each decision. Moreover, the utility of correctly accepting $H_0$ is not necessarily the same as that of correctly rejecting it as this depends on the value of $K$.

A Bayesian finds the sample size to ensure a minimum rate $r^*$ of correct classification. Note that there are common elements between this approach and the frequentist approach of controlling Type II error given Type I error. However, instead of using frequentist error rates $\alpha$ and $\beta$, the Bayesian goal function uses a prior probability $\pi$ for the null hypothesis, a cutoff determined by $K$ for testing the hypotheses and a cutoff $r^*$ for ensuring some desired rate of correct classification.

To illustrate sample size determination under both approaches, suppose that $\sigma^2 = 1$ and that $\delta = .10$. Assuming $\alpha = .05, \beta = .10$, a frequentist obtains $n_{\mathrm{F}} = 857$. A Bayesian assuming that $\pi = .5, K = 1$ and $r^* = .9283$ finds $n_{\mathrm{B}} = 857$. Computationally, this is obtained by using Equation (6) for varying increasing sample sizes until Equation (6) is equal to $r^* = .9283$. Similarly, suppose that $\delta = .05$. The sample sizes are equal to $n_{\mathrm{B}} = n_{\mathrm{F}} = 3{,}426$. In fact, the equality $n_{\mathrm{F}} = n_{\mathrm{B}}$ holds for all values of $\delta$. To see this, note that when setting $n = n_F$ in Equation (6), we obtain

$$\begin{aligned} G_{\mathrm{B}}(n = n_F, \mathbf{v}) &= K\pi P\left(Z \le \frac{\log(K\frac{\pi}{1-\pi})}{|z_\alpha + z_\beta|} + \frac{|z_\alpha + z_\beta|}{2}\right) \\ &\quad + (1-\pi)P\left(Z \ge \frac{\log(K\frac{\pi}{1-\pi})}{|z_\alpha + z_\beta|} - \frac{|z_\alpha + z_\beta|}{2}\right), \end{aligned} \tag{7}$$

that is, $G_{\mathrm{B}}(n = n_{\mathrm{F}}, \mathbf{v})$ is algebraically constant as a function of $\delta$. This is an example of a formal correspondence between frequentist and Bayesian inputs.

### 2.2 Composite Hypotheses

We now consider composite hypotheses $H_0 : \theta \le \theta_0$ versus $H_1 : \theta > \theta_0$. The frequentist sample size is given by Equation (5) with $\delta$ equal to the minimum difference to be detected with power $1 - \beta$. In a Bayesian analysis it is common to adopt priors that are symmetric around $\theta_0$, as they assign the same prior probability to the null and alternative regions. If we assume $\theta$ is a priori $N(\theta_0, \tau^2)$, it facilitates the illustration of the the relationship between the two approaches. The posterior distribution of $\theta$ is $N(m_1, v_1^2)$, where $m_1 = \frac{\theta_0\sigma^2 + \bar{x}n\tau^2}{\sigma^2 + n\tau^2}$ and $v_1^2 = \frac{\sigma^2\tau^2}{\sigma^2 + n\tau^2}$ (DeGroot 1970). A Bayesian does not reject the null hypothesis when the posterior probability is at least $1/(1 + K)$. This means that $H_0$ would not be rejected when $\bar{x} \le y_n$ where $y_n = \theta_0 + z_{\frac{K}{1+K}} \frac{\sigma}{\sqrt{n}} \sqrt{1 + \frac{\sigma^2}{n\tau^2}}$. The rate of correct classification

*Table 3. Sample Size and Correct Classification Rate $G_B$ as Functions of the Difference $\delta$. The frequentist sample size ($n_F$) is determined with $\alpha$ = .05, $\beta$ = .10, and $\sigma$ = 1 up to the nearest integer. Rates $G_B(n = n_F, \mathbf{v} = \mathbf{v}_1)$, $G_B(n = n_F, \mathbf{v} = \mathbf{v}_{C\delta})$ for C = .5, 1, 2 assume n = $n_F$ and that, respectively, $\tau$ = 1 and $\tau = C\delta$, for C = .5, 1, 2.*

| $\delta$ | $n_F$ | $G_B(n_F, \mathbf{v} = \mathbf{v}_1)$ | $G_B(n = n_F, \mathbf{v} = \mathbf{v}_{.5\delta})$ | $G_B(n = n_F, \mathbf{v} = \mathbf{v}_\delta)$ | $G_B(n, \mathbf{v} = \mathbf{v}_{2\delta})$ |
|---|---|---|---|---|---|
| .01 | 85638 | .999 | .857 | .895 | .925 |
| .02 | 21410 | .998 | .857 | .895 | .925 |
| .03 | 9515 | .997 | .857 | .895 | .925 |
| .04 | 5352 | .996 | .857 | .895 | .925 |
| .05 | 3426 | .995 | .857 | .895 | .925 |
| .06 | 2379 | .993 | .857 | .895 | .925 |
| .07 | 1748 | .992 | .857 | .895 | .925 |
| .08 | 1338 | .991 | .857 | .895 | .925 |
| .09 | 1057 | .990 | .857 | .895 | .925 |
| .10 | 856 | .989 | .857 | .895 | .925 |
| .20 | 214 | .978 | .857 | .895 | .925 |
| .30 | 95 | .967 | .857 | .895 | .925 |
| .40 | 54 | .957 | .857 | .895 | .925 |
| .50 | 34 | .946 | .857 | .895 | .925 |



is

$$G_{\mathrm{B}}(n, \mathbf{v}) = K \int_{-\infty}^{\theta_0} P_\theta(\bar{x} \le y_n)\pi(\theta)d\theta + \int_{\theta_0}^{\infty} P_\theta(\bar{x} > y_n)\pi(\theta)d\theta. \quad (8)$$

It is shown in Appendix A.1 that,

$$G_{\mathrm{B}}(n, \mathbf{v}) = \frac{1}{2} + K\Phi\left(z_{\frac{K}{1+K}} \frac{\sigma}{\sqrt{n}\tau}\right) - (1+K)\int_0^{\infty} \Phi\left(-\frac{u\sqrt{n}}{\sigma} + z_{\frac{K}{1+K}}\sqrt{1+\frac{\sigma^2}{n\tau^2}}\right) \times \frac{1}{\sqrt{2\pi}\tau} e^{-\frac{u^2}{2\tau^2}} du. \quad (9)$$

Again, as in Section 2.1, $\mathbf{u} = (\alpha, \delta, \sigma)$ is the frequentist input. Fixing the values of $\alpha$, $\beta$, and $\sigma$, the frequentist approach would produce different sample sizes for different values of $\delta$. This is illustrated in Table 3 with the sample size $n_{\mathrm{F}}$ calculated under the frequentist approach with $\alpha = .05$ and $\beta = .10$ for some values of $\delta$ and assuming that $\sigma = 1$.

Suppose a Bayesian wants to find the sample size to ensure a minimum rate $r^*$ of correct decisions. The Bayesian input set is $\mathbf{v} = (\tau, K, \sigma)$, as seen from Equation (9) which is constant in $\delta$. When setting $n = n_{\mathrm{F}}$, the Bayesian goal function is no longer constant in $\delta$, because $n_{\mathrm{F}}$ depends on $\delta$. This is shown in the third column of Table 3 assuming that $\tau = 1$ and $K = 1$. When $\delta$ changes, $G_{\mathrm{B}}(n = n_{\mathrm{F}}, \mathbf{v} = \mathbf{v}_1)$ changes accordingly.

To establish a correspondence between the two approaches with $n = n_{\mathrm{F}} = n_{\mathrm{B}}$, then just as the frequentist power is constant over all values $\delta$, the Bayesian goal function should also be constant over $\delta$. This is achieved by taking $\tau^2 = C\delta^2$ (where $C$ is a positive constant), that is, by assuming that the prior variance depends on $\delta$. By setting $n = n_{\mathrm{F}}$ from Equation (5) along with $\tau^2 = C\delta^2$ in the goal function, it can be shown (see Appendix A.2) that

$$G_{\mathrm{B}}(n = n_{\mathrm{F}}, \mathbf{v}) = \frac{1}{2} + K\Phi\left(\frac{z_{\frac{K}{1+K}}}{\sqrt{C}|z_\alpha + z_\beta|}\right) - (1+K)\int_0^{\infty} \Phi\left(-t|z_\alpha + z_\beta| + z_{\frac{K}{1+K}}\sqrt{1+\frac{1}{C(z_\alpha+z_\beta)^2}}\right)\frac{1}{\sqrt{2\pi C}} e^{-\frac{t^2}{2C}} dt. \quad (10)$$

The above expression brings out a parallels between the two approaches by establishing a relationship between the (Bayesian) prior variance and the (frequentist) target difference. The last three columns of Table 3 give the rates of correct classification $G_{\mathrm{B}}(n = n_{\mathrm{F}}, \mathbf{v} = \mathbf{v}_\tau)$ calculated with $n_{\mathrm{F}}$ and $\tau = C\delta$ for $C = .5, 1, 2$. We see from Table 3 that when $\tau$ is a function of $\delta$, the rate of correct classification is constant as a function of $\delta$.

### 2.3 Simple Versus Two-Sided Composite Hypothesis

Consider now the simple null hypothesis $H_0 : \theta = \theta_0$ versus the two-sided composite hypothesis $H_1 : \theta \neq \theta_0$. The minimum frequentist sample size to detect a difference $\delta$ with fixed probabilities of Types I and II errors is

$$n_{\mathrm{F}} = (z_{\alpha/2} + z_\beta)^2 \left(\frac{\sigma}{\delta}\right)^2, \quad (11)$$

only slightly different from Equation (5). Define $P(H_0) = 1 - P(H_1) = \pi$ and assume that $\theta \sim N(\theta_0, \tau^2)$ under $H_1$. A discussion of the choice of priors for testing two-sided hypotheses was provided by Berger (1985). Accepting the null hypothesis when the posterior probability is at least $1/(1+K)$ means accepting when $\bar{x} \in I = (-y_n + \theta_0, y_n + \theta_0)$ with

$$y_n = \frac{\sigma}{\sqrt{n}}\sqrt{2\left(1+\frac{\sigma^2}{n\tau^2}\right)\left(\log\left(K\frac{\pi}{1-\pi}\right) + \frac{1}{2}\log\left(1+\frac{n\tau^2}{\sigma^2}\right)\right)}.$$

The rate of correct classification is

$$G_{\mathrm{B}}(n, \mathbf{v}) = K\pi P_{\theta_0}(\bar{x} \in I) + (1-\pi)\int_{-\infty}^{\infty} P_\theta(\bar{x} \notin I)\pi(\theta)d\theta. \quad (12)$$

Expression (12) is equivalent to

$$G_{\mathrm{B}}(n, \mathbf{v}) = [1 + (K-1)\pi] - 2K\pi\Phi\left(-\frac{y_n\sqrt{n}}{\sigma}\right)$$

$$+ \ (1-\pi)\int_{-\infty}^{\infty}\left[\Phi\left(\frac{(-y_n-u)\sqrt{n}}{\sigma}\right) - \Phi\left(\frac{(y_n-u)\sqrt{n}}{\sigma}\right)\right]\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{u^2}{2\tau^2}}du. \quad (13)$$

(see Appendix A.3).

The frequentist input set is again $\mathbf{u} = (\alpha, \delta, \sigma)$. A Bayesian seeking the minimum sample size with rate $r^*$ of correct decisions requires, from (13), the input $\mathbf{v} = (\pi, \tau, K, \sigma)$. If we consider again the case $\tau^2 = C\delta^2$ with $n = n_F$ in Equation (11),

$$\begin{aligned} G_{\mathrm{B}}(n &= n_F, \mathbf{v}) \\ &= [1 + (K-1)\pi] - 2K\pi\Phi(-A) + (1-\pi) \\ &\quad \int_{-\infty}^{\infty}[\Phi(-A-tB) - \Phi(A-tB)]\frac{1}{\sqrt{2\pi C}}e^{-\frac{t^2}{2C}}dt, \end{aligned} \quad (14)$$

where

$$A = \sqrt{2\left(1+\frac{1}{CB^2}\right)\left(\log\left(K\frac{\pi}{1-\pi}\right) + \frac{1}{2}\log(1+CB^2)\right)}$$

and $B = |z_{\alpha/2} + z_\beta|$ (see Appendix A.3 for details). Again, this is constant as a function of $\delta$ and depends on $z_{\alpha/2}$ and $z_\beta$.

### 2.4 Simple Versus Simple Hypothesis—Bernoulli Data

Consider a random sample $x_1, x_2, \ldots, x_n$ from the Bernoulli($\theta$) distribution and let $y = \sum_{i=1}^n x_i$. We want to test $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$. Desu and Raghavarao (1990) derived the frequentist sample size by taking the arc sin transformation of the sample proportion $y/n$ and the fact that $2\sqrt{n}(\arcsin\sqrt{y/n} - \arcsin\sqrt{\theta})$ has asymptotically a $N(0,1)$ distribution. They obtained

$$n_{\mathrm{F}} = \left(\frac{z_\alpha + z_\beta}{2(\arcsin\sqrt{\theta_1} - \arcsin\sqrt{\theta_0})}\right)^2. \quad (15)$$

Suppose a priori that $P(\theta_0) = 1 - P(\theta_1) = \pi$. Accepting the null hypothesis when the posterior probability is at least $1/(1+K)$ is equivalent to accepting it when $y \le y_c$ with the cutoff point

$$y_c = \left(\frac{\log\left(K\frac{\pi}{1-\pi}\right) - n\log\left(\frac{1-\theta_1}{1-\theta_0}\right)}{\log\left(\frac{\theta_1}{\theta_0}\right) - \log\left(\frac{1-\theta_1}{1-\theta_0}\right)}\right).$$

The rate of correct decisions is

$$G_{\mathrm{B}}(n, \mathbf{v}) = K\pi P_{\theta_0}(Y \le y_c) + (1-\pi)P_{\theta_1}(Y > y_c). \quad (16)$$

Under the frequentist power-based approach to sample size determination the input is $\mathbf{u} = (\alpha, \theta_0, \theta_1)$. When sample size is determined on the basis of the rate $r^*$ of correct decisions, the Bayesian input is $\mathbf{v} = (\pi, K, \theta_0, \theta_1)$. Table 4 shows the frequentist sample size $n_F$ determined with $\alpha = .05, \beta = .10, \theta_0 = .01$ and $\theta_1 = \delta + \theta_0$. The values $G_{\mathrm{B}}(n = n_F, \mathbf{v})$ correspond to the rates of correct decisions assuming $\pi = .5$ and with $n = n_{\mathrm{F}}$. The function $G_{\mathrm{B}}(n = n_{\mathrm{F}}, \mathbf{v})$ is not constant as a function of $\delta$. Similar to the approach taken in Section 2.2, we can further establish a relationship between the prior probability $\pi$ and the value of $\delta$ being tested to obtain a constant goal function. In the above example, the prior probabilities that are necessary to achieve a constant goal function with the Bayesian approach can be calculated numerically. Figure 1 shows the prior probability $\pi$ as a function of $\delta$. Under these prior probabilities the Bayesian goal function is approximately constant and equal to $r^* = .93$. The figure shows that the prior probabilities decreases as the target difference increases.

*Table 4. Sample Size in the Bernoulli Example. The frequentist sample size $n_F$ is determined with $\alpha = .05$, $\beta = .10$, $\theta_0 = .01$, $\theta_1 = \delta + \theta_0$ and rounded to the nearest integer. The values of $G_B(n = n_F, \mathbf{v} = \mathbf{v}_{.5})$ correspond to the rate of making the correct decisions computed assuming $\pi = .5$ and with $n = n_F$.*

| $\delta$ | $n_{\mathrm{F}}$ | $G_{\mathrm{B}}(n = n_{\mathrm{F}}, \mathbf{v} = \mathbf{v}_{\pi=.5})$ |
|---|---|---|
| .001 | 88927 | .928 |
| .002 | 23244 | .928 |
| .003 | 10771 | .928 |
| .004 | 6302 | .928 |
| .005 | 4187 | .929 |
| .006 | 3012 | .929 |
| .007 | 2289 | .929 |
| .008 | 1809 | .929 |
| .009 | 1474 | .929 |
| .010 | 1230 | .929 |
| .020 | 392 | .929 |
| .030 | 210 | .932 |
| .040 | 137 | .933 |
| .050 | 99 | .932 |
| .060 | 77 | .935 |
| .070 | 62 | .929 |
| .080 | 52 | .929 |
| .090 | 44 | .936 |
| .100 | 38 | .938 |

## 3. COMPARING ERROR RATES AND INFORMATION

### 3.1 Lindley Information

Consider observing $X$ from the distribution $p(x|\theta)$. Lindley (1956) proposed to measure the amount of information about $\theta$ provided by the outcome of the experiment using the quantity:

$$\begin{aligned} \mathcal{I} &\equiv E\{\log\pi(\theta|X)\} - E\{\log\pi(\theta)\} \\ &= E\left\{\log\left(\frac{\pi(\theta|X)}{\pi(\theta)}\right)\right\}, \end{aligned} \quad (17)$$

where expectations are taken with respect to the joint distribution of $\theta$ and $X$. Expression (17) quantifies the difference between the knowledge available on $\theta$ before and after the experiment, represented by the prior distribution $\pi(\theta)$ and by the posterior distribution $\pi(\theta|X)$, respectively.

There are two standard justifications for (17). One is decision-theoretic: $\mathcal{I}$ is the expected utility of reporting a density on the unknown quantity $\theta$, when the utility function is the logarithm of the density. See Good (1952), DeGroot (1984) for additional details, and DeGroot (1962) for the information derived with different utility functions. The other is information-theoretic: $\mathcal{I}$

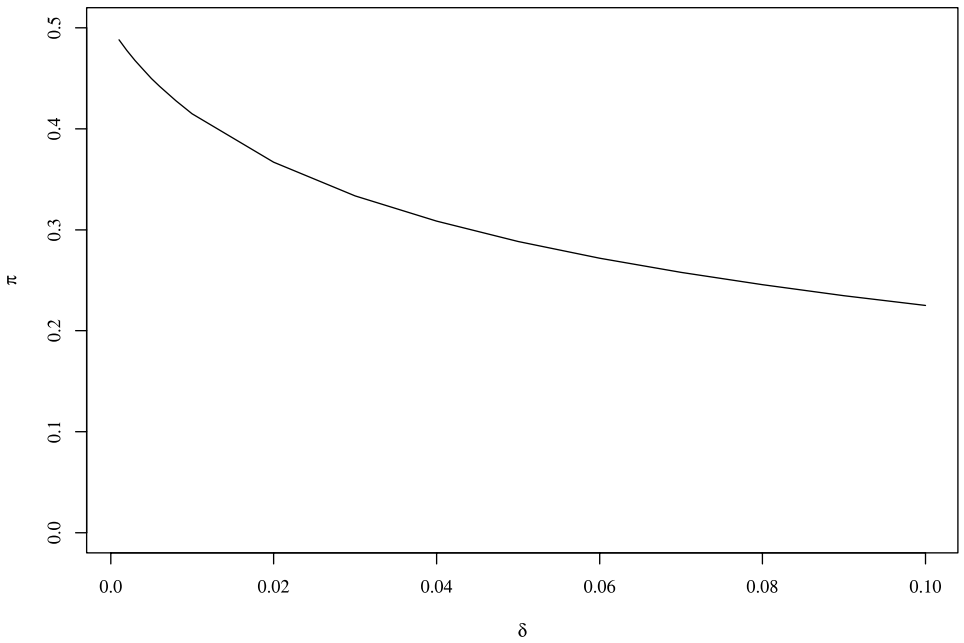


*Figure 1. Prior probability $\pi$ as a function of $\delta = \theta_1 - \theta_0$ in the Bernoulli example. The prior probabilities are numerically calculated to keep a rate $r^* = .93$ of correct decisions when $n = n_F$ and using the normal approximation to the binomial probabilities in Equation (16).*



is the difference in entropies between the prior and the posterior distribution. Both interpretations are consistent with viewing $\mathcal{I}$ as quantifying the information provided by an experiment.

Lindley (1956) considered the sequential sampling problem and suggested that experimentation should continue until obtaining enough information $\mathcal{I}$ about the parameter of interest. Bernardo (1979, 1997) provided asymptotic results on information, and applied them to determining sample size. Parmigiani and Berry (1994) developed general first-order conditions to obtain optimal designs. In another application, Verdinelli and Kadane (1992) derived designs for certain hierarchical linear models by maximizing the expected utility expressed as a linear combination of expected information and expected value of some function of the outcomes of the experiment.

In this section we use information to investigate parallels between Bayesian and frequentist approaches to sample size determination under different goals. As in the previous section, the frequentist goal function $G_{\mathrm{F}}$ is power. Now the Bayesian goal function $G_{\mathrm{B}}$ is Lindley information. With these rather different goals, one can identify close correspondences between the frequentist and Bayesian inputs which lead to the same sample sizes.

### 3.2 Information and Sample Size—Normal Data

When a random sample of size $n$ is taken from a $N(\theta, \sigma^2)$ distribution and $\theta$ has a priori $N(\mu, \tau^2)$ distribution, the expected information for the experiment (Lindley 1956) is

$$\mathcal{I} = \frac{1}{2} \log \left( 1 + n \frac{\tau^2}{\sigma^2} \right). \tag{18}$$

From Equation (18), the Bayesian sample size $n_{\mathrm{B}}$ to achieve a fixed amount of expected information $\mathcal{I}^*$ is

$$n_{\mathrm{B}} = (\exp(2\mathcal{I}^*) - 1) \left( \frac{\sigma}{\tau} \right)^2, \tag{19}$$

and the Bayesian input is $\mathbf{v} = (\tau, \sigma)$.

Note that the expected information $\mathcal{I}$ is not constant in $\delta$ when setting $n = n_{\mathrm{F}}$ in Equation (18), where $n_{\mathrm{F}}$ is the frequentist power-based sample size based on the input $\mathbf{u} = (\alpha, \delta, \sigma)$. When $n = n_{\mathrm{F}}$, assuming that $\tau^2 = C\delta^2$ (for $C > 0$) makes $\mathcal{I}$ constant in $\delta$ and it follows from Equations (5) and (19), $n_{\mathrm{B}} = n_{\mathrm{F}}$ if and only if $(\exp(2\mathcal{I}^*) - 1) = C \left( z_\alpha + z_\beta \right)^2$, which implies

$$\mathcal{I}^* = \frac{1}{2} \log \left( 1 + C(z_\alpha + z_\beta)^2 \right). \tag{20}$$

### 3.3 Information and Sample Size—Bernoulli Data

Consider independent and identical $\mathrm{Bernoulli}(\theta)$ observations and assume that $\theta$ takes only two values, $\theta_0$ and $\theta_1$, with $P(\theta_0) = 1 - P(\theta_1) = \pi$. The frequentist sample size $n_{\mathrm{F}}$ is determined as in Section 2.4 with input $\mathbf{u} = (\alpha, \theta_0, \theta_1)$.

If a Bayesian seeks the sample size giving expected information $\mathcal{I}^*$, the input is $\mathbf{v} = (\pi, \theta_0, \theta_1)$. Using $n = n_{\mathrm{F}}$ (with $n_{\mathrm{F}}$ calculated with $\alpha = .05, \beta = .10, \theta_0 = .01$ and $\theta_1 = \delta + \theta_0$), Figure 2 shows that the expected information is nearly constant as a function of $\delta$ for a broad range of prior probabilities. Again, if a frequentist and a Bayesian agree on a sample size for one $\delta$ then they will be close to agreeing for all $\delta$.

## 4. DISCUSSION

In this article we provided a simple but general framework for identifying mappings between the frequentist and Bayesian approaches to sample size determination. We also provided examples of how this correspondence can be used in standard situations. By way of illustration we focused on the correspondence between the "target difference" $\delta$ in the classical approach and the prior standard deviation in a Bayesian approach. Showing that one can translate a sample size calculation made according to one philosophy into the other philosophy will help convey a

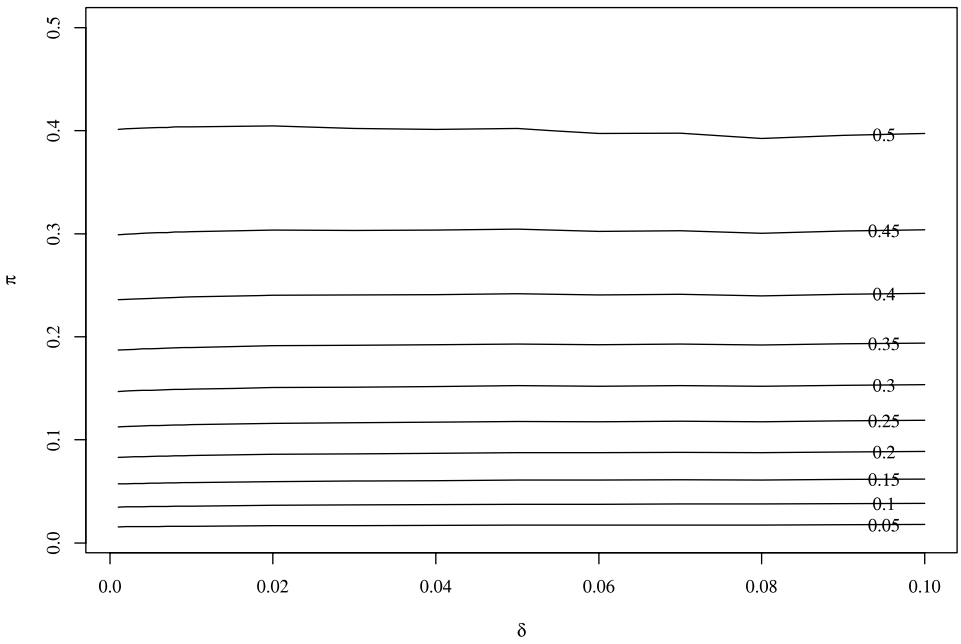


*Figure 2. Contour plot of the expected information as a function of $\delta = \theta_1 - \theta_0$ and $\pi$. The expected information is calculated with $n = n_F$. The frequentist sample size $n_F$ is calculated with $\alpha = .05$, $\beta = .10$, $\theta_0 = .01$, $\theta_1 = \delta + \theta_0$.*



greater understanding of the appropriateness of any particular sample size choice and for the method that derived it.

Bayesian and frequentist approaches to sample size determination differ in philosophy and goals. However, both provide algorithms for determining a sample size from specified criteria and inputs. Sample size determination is a critical aspect of the evaluation of proposed research funding and regulatory evaluation of proposed drug and medical device trials. Reviewers can reasonably criticize the inputs. But in view of our simple observation, there is no philosophical basis for criticizing a proposed sample size. If a different philosophical approach is deemed to be more appropriate, it is possible to derive a set of inputs in that alternative approach in which the sample size is the same as the one proposed.

An increasing number of marketing approval applications for drugs and medical devices are being submitted for regulatory approval using a Bayesian approach. Many reviewers object to mixing frequentist and Bayesian philosophies in the application process. One of the problems in Bayesian submissions is that there is no standard method for calculating sample size from a Bayesian perspective. Another is that reviewers do not understand Bayesian sample size calculations as well as they understood Bayesian analyses. Our results lend credence to frequentist sample size calculations from a Bayesian perspective and allow companies to use standard frequentist methods while still claiming to be Bayesian.

We have focused on a special design consideration and explored correspondences between frequentist and Bayesian approaches. A related question is, once data have been collected, how to perform the analysis? There are correspondences at the analysis stage when the Bayesian uses a noninformative prior. Using conditional frequentist hypothesis testing as discussed by Berger, Brown, and Wolpert (1994) and Berger, Boukai, and Wang (1997) leads to equivalent inferences under the Bayesian approach. Our article indicates that there are correspondences between the two approaches at the design phase as well as at the inferential stage.

## APPENDIX

### A.1 DERIVATION OF $G_{\mathrm{B}}$ FROM SECTION 2.2

Let $y_n = \theta_0 + z_{\frac{K}{1+K}} \frac{\sigma}{\sqrt{n}} \sqrt{1 + \frac{\sigma^2}{n\tau^2}}$ and suppose that $\theta \sim N(\theta_0, \tau^2)$. The probability of making the correct decision is given by

$$\begin{aligned} G_{\mathrm{B}}(n, \mathbf{v}) &= K \int_{-\infty}^{\theta_0} P_\theta(\bar{x} \le y_n)\pi(\theta)d\theta \\ &\quad + \int_{\theta_0}^{\infty} P_\theta(\bar{x} > y_n)\pi(\theta)d\theta \\ &= K \int_{-\infty}^{\theta_0} P_\theta(\bar{x} \le y_n)\pi(\theta)d\theta \\ &\quad + \int_{\theta_0}^{\infty} (1 - P_\theta(\bar{x} \le y_n))\pi(\theta)d\theta \\ &= P(\theta \ge \theta_0) + K \int_{-\infty}^{\theta_0} P_\theta(\bar{x} \le y_n)\pi(\theta)d\theta \\ &\quad - \int_{\theta_0}^{\infty} P_\theta(\bar{x} \le y_n)\pi(\theta)d\theta \end{aligned}$$

Adding and subtracting $K \int_{\theta_0}^{\infty} P_\theta(\bar{x} \le y_n)\pi(\theta)d\theta$

$$\begin{aligned} &= P(\theta \ge \theta_0) + K \int_{-\infty}^{\infty} P_\theta(\bar{x} \le y_n)\pi(\theta)d\theta \\ &\quad -(1 + K) \int_{\theta_0}^{\infty} P_\theta(\bar{x} \le y_n)\pi(\theta)d\theta \\ &= P(\theta \ge \theta_0) + KP(\bar{x} \le y_n) \\ &\quad -(1 + K) \int_{\theta_0}^{\infty} P_\theta(\bar{x} \le y_n)\pi(\theta)d\theta. \quad \text{(A.1)} \end{aligned}$$

We have $P(\theta \geq \theta_0) = 1/2$. The predictive distribution of $\bar{x}$ is $N(\theta_0, \sigma^2/n + \tau^2)$. Thus,

$$P(\bar{x} \leq y_n) \;=\; \Phi\left(\frac{y_n - \theta_0}{\sqrt{\sigma^2/n + \tau^2}}\right) = \Phi\left(z_{\frac{K}{1+K}} \frac{\sigma}{\sqrt{n}\tau}\right). \tag{A.2}$$

Furthermore,

$$\begin{aligned}
&\int_{\theta_0}^{\infty} P_\theta(\bar{x} \leq y_n)\pi(\theta)d\theta \\
&= \int_{\theta_0}^{\infty} \Phi\left(\frac{y_n - \theta}{\sigma/\sqrt{n}}\right)\pi(\theta)d\theta \\
&= \int_{\theta_0}^{\infty} \Phi\left(\frac{\theta_0 - \theta}{\sigma/\sqrt{n}} + z_{\frac{K}{1+K}}\sqrt{1 + \frac{\sigma^2}{n\tau^2}}\right)\pi(\theta)d\theta.
\end{aligned}$$

Let $u = \theta - \theta_0$

$$= \int_0^{\infty} \Phi\left(-\frac{u\sqrt{n}}{\sigma} + z_{\frac{K}{1+K}}\sqrt{1 + \frac{\sigma^2}{n\tau^2}}\right)\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{u^2}{2\tau^2}}du. \tag{A.3}$$

Thus,

$$\begin{aligned}
G_{\mathrm{B}}(n, \mathbf{v}) \;=\; & \frac{1}{2} + K\Phi\left(z_{\frac{K}{1+K}}\frac{\sigma}{\sqrt{n}\tau}\right) \\
& -(1+K)\int_0^{\infty}\Phi\left(-\frac{u\sqrt{n}}{\sigma}\right. \\
& \left. + z_{\frac{K}{1+K}}\sqrt{1 + \frac{\sigma^2}{n\tau^2}}\right)\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{u^2}{2\tau^2}}du.
\end{aligned}$$

### A.2 DERIVATION OF $G_{\mathrm{B}}$ FROM SECTION 2.2 ASSUMING $\tau^2 = C\delta^2$ AND $N = N_{\mathrm{F}}$

As in A.1 suppose that $y_n = \theta_0 + z_{\frac{K}{1+K}}\frac{\sigma}{\sqrt{n}}\sqrt{1 + \frac{\sigma^2}{n\tau^2}}$ and $\theta \sim N(\theta_0, \tau^2)$. Furthermore, assume that $\tau^2 = C\delta^2$ and $n = n_F$. This implies from Equation (A.2) that

$$P(\bar{x} \leq y_n) \;=\; \Phi\left(\frac{z_{\frac{K}{1+K}}}{\sqrt{C}|z_\alpha + z_\beta|}\right). \tag{A.4}$$

From Equation (A.3)

$$\begin{aligned}
&\int_{\theta_0}^{\infty} P_\theta(\bar{x} \leq y_n)\pi(\theta)d\theta \\
&= \int_0^{\infty}\Phi\left(-\frac{u\sqrt{n}}{\sigma} + z_{\frac{K}{1+K}}\sqrt{1 + \frac{\sigma^2}{n\tau^2}}\right)\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{u^2}{2\tau^2}}du \\
&= \int_0^{\infty}\Phi\left(-\frac{u|z_\alpha + z_\beta|}{\delta} + z_{\frac{K}{1+K}}\sqrt{1 + \frac{1}{C(z_\alpha + z_\beta)^2}}\right) \\
&\quad \times\frac{1}{\sqrt{2\pi C}\delta}e^{-\frac{u^2}{2C\delta^2}}du
\end{aligned}$$

Let $t = \frac{u}{\delta}$

$$\begin{aligned}
&= \int_0^{\infty}\Phi\left(-t|z_\alpha + z_\beta| + z_{\frac{K}{1+K}}\sqrt{1 + \frac{1}{C(z_\alpha + z_\beta)^2}}\right) \\
&\quad \frac{1}{\sqrt{2\pi C}}e^{-\frac{t^2}{2C}}dt.
\end{aligned} \tag{A.5}$$

By combining Equations (A.4) and (A.5) into (A.1)

$$\begin{aligned}
G_{\mathrm{B}}(n = n_F, \mathbf{v}) &= \frac{1}{2} + K\Phi\left(\frac{z_{\frac{K}{1+K}}}{\sqrt{C}|z_\alpha + z_\beta|}\right) \\
&\quad -(1+K)\int_0^{\infty}\Phi\left(-t|z_\alpha + z_\beta|\right. \\
&\quad \left. + z_{\frac{K}{1+K}}\sqrt{1 + \frac{1}{C(z_\alpha + z_\beta)^2}}\right)\frac{1}{\sqrt{2\pi C}}e^{-\frac{t^2}{2C}}dt.
\end{aligned}$$

### A.3 DERIVATION OF $G_{\mathrm{B}}$ FROM SECTION 2.3

Let $\bar{x} \in I = (-y_n + \theta_0, y_n + \theta_0)$ with $y_n = \frac{\sigma}{\sqrt{n}}\sqrt{2\left(1 + \frac{\sigma^2}{n\tau^2}\right)\left(\log\left(K\frac{\pi}{1-\pi}\right) + \frac{1}{2}\log\left(1 + \frac{n\tau^2}{\sigma^2}\right)\right)}$ define the acceptance region for $H_0$. The probability of making the correct decision is

$$\begin{aligned}
G_{\mathrm{B}}(n, \mathbf{v}) &= K\pi P_{\theta_0}(\bar{x} \in I) \\
&\quad +(1-\pi)\int_{-\infty}^{\infty} P_\theta(\bar{x} \notin I)\pi(\theta)d\theta
\end{aligned} \tag{A.6}$$

Observe that

$$P_{\theta_0}(\bar{x} \in I) \;=\; 1 - 2\Phi\left(-\frac{y_n\sqrt{n}}{\sigma}\right). \tag{A.7}$$

Furthermore,

$$\begin{aligned}
&\int_{-\infty}^{\infty} P_\theta(\bar{x} \notin I)\pi(\theta)d\theta \\
&= 1 + \int_{-\infty}^{\infty}[P_\theta(\bar{x} \leq -y_n + \theta_0) - P_\theta(\bar{x} \leq y_n + \theta_0)]\pi(\theta)d\theta \\
&= 1 + \int_{-\infty}^{\infty}\left[\Phi\left((-y_n - (\theta - \theta_0))\frac{\sqrt{n}}{\sigma}\right)\right. \\
&\quad \left. -\Phi\left((y_n - (\theta - \theta_0))\frac{\sqrt{n}}{\sigma}\right)\right]\pi(\theta)d\theta
\end{aligned}$$

Let $u = \theta - \theta_0$

$$\begin{aligned}
&= 1 + \int_{-\infty}^{\infty}\left[\Phi\left(\frac{(-y_n - u)\sqrt{n}}{\sigma}\right)\right. \\
&\quad \left. -\Phi\left(\frac{(y_n - u)\sqrt{n}}{\sigma}\right)\right]\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{u^2}{2\tau^2}}du.
\end{aligned} \tag{A.8}$$

By combining Equations (A.7) and (A.8) into (A.6)

$$\begin{aligned}
G_{\mathrm{B}}(n, \mathbf{v}) \;=\; & [1 + (K-1)\pi] - 2K\pi\Phi\left(-\frac{y_n\sqrt{n}}{\sigma}\right) \\
& +(1-\pi)\int_{-\infty}^{\infty}\left[\Phi\left(\frac{(-y_n - u)\sqrt{n}}{\sigma}\right)\right. \\
& \left. -\Phi\left(\frac{(y_n - u)\sqrt{n}}{\sigma}\right)\right]\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{u^2}{2\tau^2}}du.
\end{aligned} \tag{A.9}$$

Assuming that $\tau^2 = C\delta^2$, setting $n = n_{\mathrm{F}}$ as in Equation (11) and using the transformation $t = u/\delta$ in the integrand,

$$\begin{aligned}
&G_{\mathrm{B}}(n = n_{\mathrm{F}}, \mathbf{v}) = [1 + (K-1)\pi] - 2K\pi\Phi(-A) \\
&+(1-\pi)\int_{-\infty}^{\infty}[\Phi(-A - tB) - \Phi(A - tB)]\frac{1}{\sqrt{2\pi C}}e^{-\frac{t^2}{2C}}dt,
\end{aligned}$$

where

$$A = \sqrt{2\left(1+\frac{1}{CB^2}\right)\left(\log\left(K\frac{\pi}{1-\pi}\right)+\frac{1}{2}\log\left(1+CB^2\right)\right)}$$

and

$$B = \quad |z_{\alpha/2} + z_{\beta}|.$$

*[Received April 2002. Revised September 2004.]*

## REFERENCES


Adcock, C. J. (1997), "Sample Size Determination: A Review," *The Statistician*, 46, 261–283.

Berger, J. O. (1985), *Statistical Decision Theory and Bayesian Analysis* (2nd ed.), New York: Springer-Verlag.

Berger, J., Boukai, B., and Wang, Y. (1997), "Unified Frequentist and Bayesian Testing of a Precise Hypothesis," *Statistical Science*, 12, 133–160.

Berger, J., Brown, L., and Wolpert, R. (1994), "A Unified Conditional Frequentist and Bayesian Test for Fixed and Sequential Simple Hypothesis Testing," *The Annals of Statistics*, 22, 1787–1807.

Bernardo, J. M. (1979), "Asymptotic Behaviour of the Information Provided by an Experiment" (in Spanish), *Revista de la Real Academia di Ciencias Exactas, Fisicas y Naturales*, 73, 491–502.

——— (1997), "Statistical Inference as a Decision Problem: The Choice of Sample Size," *The Statistician*, 46, 151–153.

DeGroot, M. H. (1962), "Uncertainty, Information, and Sequential Experiments," *The Annals of Mathematical Statistics*, 33, 404–419.

——— (1970), *Optimal Statistical Decisions*, New York: McGraw-Hill.

——— (1984), "Changes in Utility as Information," *Theory and Decision*, 17, 287–303.

Desu, M., and Raghavarao, D. (1990), *Sample Size Methodology*, New York: Academic Press.

Good, I. J. (1952), "Rational Decisions," *Journal of the Royal Statistical Society*, Ser. B, 14, 107–114.

Grundy, P., Healy, M., and Rees, D. (1956), "Economic Choice of the Amount of Experimentation," *Journal of The Royal Statistical Society*, Ser. B., 18, 32–55.

Joseph, L., Wolfson, D. B., and Berger, R. D. (1995), "Sample Size Calculations for Binomial Proportions via Highest Posterior Density Intervals," *The Statistician*, 44, 143–154.

Lindley, D. V. (1956), "On a Measure of the Information Provided by an Experiment," *Annals of Mathematical Statistics*, 27, 986–1005.

——— (1997), "The Choice of Sample Size," *The Statistician*, 46, 129–138.

Neyman, J., and Pearson, E. S. (1933), "On the Problem of the Most Efficient Test of Statistical Hypotheses," *Philosophical Transactions of the Royal Society*, Ser. A, 231, 286–337.

Parmigiani, G., and Berry, D. A. (1994), "Applications of Lindley Information Measure to the Design of Clinical Experiments," in *Aspects of Uncertainty. A Tribute to D. V. Lindley*, Chichester: Wiley, pp. 329–348.

Pham-Gia, T., and Turkkan, N. (2003), "Determination of Exact Sample Sizes in the Bayesian Estimation of the Difference of Two Proportions," *The Statistician*, 52, 131–150.

Raiffa, H., and Schlaifer, R. (1961), *Applied Statistical Decision Theory*, Boston: Harvard University Press.

Spiegelhalter, D. J., and Freedman, L. S. (1986), "A Predictive Approach to Selecting the Size of a Clinical Trial, Based on Subjective Clinical Opinion," *Statistics in Medicine*, 5, 1–13.

Verdinelli, I., and Kadane, J. B. (1992), "Bayesian Designs for Maximizing Information and Outcome," *Journal of the American Statistical Association*, 87, 510–515.

Wald, A. (1950), *Statistical Decision Functions*, New York: Wiley.

Weiss, R. (1997), "Bayesian Sample Size Calculations for Hypothesis Testing," *The Statistician*, 46, 185–191.