应用概率统计 第 36 卷
第 5 期 2020 年 10 月

Chinese Journal of Applied Probability and Statistics
Oct., 2020, Vol. 36, No. 5, pp. 523-535
doi: 10.3969/j.issn.1001-4268.2020.05.007

A New Expectation Identity and Its Application in the Calculations of Predictive Powers Assuming Normality *

ZHANG Yingying* RONG Tengzhong LI Manman

(*Department of Statistics and Actuarial Science, College of Mathematics and Statistics, Chongqing University, Chongqing, 401331, China*)

Abstract: For calculating the predictive powers, we suggest an elegant expectation identity to directly calculate the expectations. We calculate the predictive powers of the hypotheses with a nonzero threshold for five different categories, which are non-sequential trials with classical power and Bayesian power, and sequential trials with hybrid predictions, Bayesian predictions, and classical predictions. Moreover, the calculations of the various predictive powers are illustrated through three examples. Finally, when calculating the average success probability in [9], it is tricky to find the predictive distribution for the predictive power, whereas, it is straightforward to utilize the expectation identity for the calculation.

Keywords: expectation identity; predictive power; normal model; one-sided hypothesis; average success probability (ASP)

2010 Mathematics Subject Classification: 62F03; 62F15

Citation: ZHANG Y Y, RONG T Z, LI M M. A new expectation identity and its application in the calculations of predictive powers assuming normality [J]. Chinese J Appl Probab Statist, 2020, 36(5): 523–535.

§1. Introduction

For financial and ethical reasons, an increasingly utilized feature in clinical trial practice is to allow a study to stop early for futility or efficacy. The researches in the literature often assume normality for the prior and the likelihood. Many methods for addressing futility or efficacy have been described in the literature, including means based on conditional power [1–4], sequential monitoring [5–7], expected or predictive power [8–10], beta

*The research was supported by the Fundamental Research Funds for the Central Universities (Grant Nos. 2019CDXYST0016; 2018CDXYST0024), the China Scholarship Council (Grant No. 201606055028), the National Natural Science Foundation of China (Grant No. 11671060), the MOE Project of Humanities and Social Sciences on the West and the Border Area (Grant Nos. 20XJC910001; 14XJC910001) and Chongqing Key Laboratory of Analytic Mathematics and Applications.

*Corresponding author, E-mail: robertzhangyying@qq.com or robertzhang@cqu.edu.cn.

Received December 3, 2018. Revised January 12, 2020.

spending functions, and others. Some statisticians consider the determination of the sample sizes [11–13]. Other clinicians investigate the type I and II error probabilities [14, 15]. Many medical researchers exploit Bayesian approaches for futility or efficacy [9, 10, 16, 17]. A few faculty members discuss some optimal adaptive designs in clinical trials [18].

Spiegelhalter et al. [8] have calculated the rejection region, the power or the conditional power, and the predictive power or the conditional predictive power of the hypotheses $H_0 : \theta < 0$ versus $H_1 : \theta > 0$ for five different categories, which are non-sequential trials with classical power and Bayesian power, and sequential trials with hybrid predictions, Bayesian predictions, and classical predictions in Sections 6.5 and 6.6. In this paper, we will calculate the above quantities of the hypotheses $H_0 : \theta \leqslant \theta_0$ versus $H_1 : \theta > \theta_0$ (henceforth, hypotheses A) and $H_0 : \theta \geqslant \theta_0$ versus $H_1 : \theta < \theta_0$ (henceforth, hypotheses B) for the five different categories, where θ_0 is the threshold value of the hypotheses. Note that $\theta_0 \neq 0$ corresponds to a non-inferiority trial and $\theta_0 = 0$ corresponds to a superiority trial. A detailed discussion of the non-inferiority issue of the hypotheses can be found in the supplement of [13].

There are two ways to calculate the predictive powers: One way is to calculate the predictive powers by using the predictive distributions, and the other way is to calculate the expectations which are very involved and are always circumvented by Spiegelhalter et al. [8] and other researchers. We prove and utilize an elegant new expectation identity to calculate the predictive powers of the five different categories by directly calculating the expectations without circumvention.

For the average success probability (ASP) in [9], we can also calculate it in two ways. One way is to calculate the ASP by using the predictive distribution, and the other way is to calculate the expectation by utilizing the new expectation identity.

The rest of the paper is organized as follows. In Section 2, we prove an elegant new expectation identity and calculate the predictive powers of the hypotheses with a nonzero threshold for five different categories. Moreover, we utilize the new expectation identity to analytically calculate the ASP in [9]. Section 3 illustrates the calculations of the predictive powers through three examples. Some conclusions and discussions are provided in Section 4.

§2. The Calculations of Predictive Powers Assuming Normality

There are two ways to calculate the predictive powers: One way is to calculate the predictive powers by using the predictive distributions, and the other way is to calculate

the expectations which are very involved and are always circumvented by Spiegelhalter et al.[8] and other researchers.

We have the following elegant expectation identity (1) which is very useful for the calculation of the predictive power by directly calculating the expectation. The proof of the theorem can be found in the supplement.

Theorem 1 Assume $Z \sim \mathrm{N}(0,1)$ with cumulative distribution function $\Phi(x)$, and let a and b be real constants. Then

$$\mathsf{E}[\Phi(aZ+b)] = \Phi\Big(\frac{b}{\sqrt{1+a^2}}\Big). \tag{1}$$

As mentioned in the introduction section, Spiegelhalter et al.[8] have calculated the rejection region, the power or the conditional power, and the predictive power or the conditional predictive power of the hypotheses $H_0 : \theta < 0$ versus $H_1 : \theta > 0$ for five different categories, which are non-sequential trials with classical power and Bayesian power, and sequential trials with hybrid predictions, Bayesian predictions, and classical predictions in Sections 6.5 and 6.6. In the later part of this section, we will utilize the expectation identity (1) to calculate various predictive powers of hypotheses A and B with a nonzero threshold by directly calculating the expectations.

2.1 Non-Sequential Trials

Suppose we have a normal prior $\theta \sim \mathrm{N}(\mu, \sigma^2/n_0)$ and our future data Y_n given θ have distribution $Y_n \,|\, \theta \sim \mathrm{N}(\theta, \sigma^2/n)$. We wish to calculate the predictive probability of obtaining a "significant" result, when testing the hypotheses A and B.

2.1.1 Classical Power: Hybrid Classical-Bayesian Methods

By utilizing the expectation identity (1) for testing the hypotheses A, we have the following corollary in which we have shown that the hybrid predictive power can be calculated in two ways: One way is to calculate the hybrid predictive power by using the predictive distribution, and the other way is to directly calculate an expectation. The proof of the corollary can be found in the supplement.

Corollary 2 The hybrid predictive power is

$$\begin{aligned} P(S^C_{\epsilon,\theta_0}) &= \int P(S^C_{\epsilon,\theta_0} \,|\, \theta)\pi(\theta)\mathrm{d}\theta = \mathsf{E}_\theta[P(S^C_{\epsilon,\theta_0} \,|\, \theta)] \\ &= \Phi\Big[\sqrt{\frac{n_0}{n_0+n}}\Big(\frac{\mu-\theta_0}{\sigma/\sqrt{n}} + z_\epsilon\Big)\Big], \end{aligned} \tag{2}$$

where E_θ takes expectation with respect to the random variable $\theta \sim \pi(\theta)$.

Similarly, by using the expectation identity for testing the hypotheses B, the calculation of the hybrid predictive power

$$P(S^{C-}_{\epsilon,\theta_0}) = \int P(S^{C-}_{\epsilon,\theta_0} \,|\, \theta)\pi(\theta)\mathrm{d}\theta = \mathsf{E}_\theta[P(S^{C-}_{\epsilon,\theta_0} \,|\, \theta)]$$

can be found in the supplement. The expressions of the hybrid predictive powers for "Classical power" are summarized in Table 1.

Table 1 The expected or predictive powers for non-sequential trials

	$H_0 : \theta \leqslant \theta_0$, $H_1 : \theta > \theta_0$ (Hypotheses A)	$H_0 : \theta \geqslant \theta_0$, $H_1 : \theta < \theta_0$ (Hypotheses B)
Hybrid predictive power	$P(S^{C}_{\epsilon,\theta_0}) = \Phi\left[\sqrt{\frac{n_0}{n_0+n}}\left(\frac{\mu-\theta_0}{\sigma/\sqrt{n}} + z_\epsilon\right)\right]$	$P(S^{C-}_{\epsilon,\theta_0}) = \Phi\left[\sqrt{\frac{n_0}{n_0+n}}\left(-\frac{\mu-\theta_0}{\sigma/\sqrt{n}} + z_\epsilon\right)\right]$
Bayesian predictive power	$P(S^{B}_{\epsilon,\theta_0}) = \Phi\left[\sqrt{\frac{(n_0+n)n_0}{n}}\frac{\mu-\theta_0}{\sigma} + \sqrt{\frac{n_0}{n}}z_\epsilon\right]$	$P(S^{B-}_{\epsilon,\theta_0}) = \Phi\left[-\sqrt{\frac{(n_0+n)n_0}{n}}\frac{\mu-\theta_0}{\sigma} + \sqrt{\frac{n_0}{n}}z_\epsilon\right]$

2.1.2 Bayesian Power

We wish to calculate the predictive probability of obtaining a "significant" Bayesian result, and we shall denote such "Bayesian significance" as $S^{B}_{\epsilon,\theta_0} \equiv \{\mathsf{P}(\theta \leqslant \theta_0 \,|\, \text{data}) < \epsilon\}$ or $S^{B-}_{\epsilon,\theta_0} \equiv \{\mathsf{P}(\theta \geqslant \theta_0 \,|\, \text{data}) < \epsilon\}$. By using the expectation identity for testing the hypotheses A and B, the calculations of the Bayesian predictive, expected, or average powers

$$P(S^{B}_{\epsilon,\theta_0}) = \int P(S^{B}_{\epsilon,\theta_0} \,|\, \theta)\pi(\theta)\mathrm{d}\theta = \mathsf{E}_\theta[P(S^{B}_{\epsilon,\theta_0} \,|\, \theta)],$$
$$P(S^{B-}_{\epsilon,\theta_0}) = \int P(S^{B-}_{\epsilon,\theta_0} \,|\, \theta)\pi(\theta)\mathrm{d}\theta = \mathsf{E}_\theta[P(S^{B-}_{\epsilon,\theta_0} \,|\, \theta)],$$

respectively can be found in the supplement. The expressions of the Bayesian predictive powers for "Bayesian power" are also summarized in Table 1.

It is worthy to mention that the hybrid predictive power and the Bayesian predictive power in Table 1 only utilize the historical data through the prior $\pi(\theta)$.

2.2 Monitoring Sequential Trials Using Predictions: Conditional Power

This subsection deals with the concept of "futility" (see [8]), that is, given the data so far, what is the chance of getting a "significant" result? Suppose we have a normal prior $\theta \sim \mathrm{N}(\mu, \sigma^2/n_0)$, our current data y_m given θ have distribution $y_m \,|\, \theta \sim \mathrm{N}(\theta, \sigma^2/m)$, our future data Y_n given θ have distribution $Y_n \,|\, \theta \sim \mathrm{N}(\theta, \sigma^2/n)$ where σ^2 is assumed known.

We wish to calculate the predictive probability of obtaining a "significant" result, when testing the hypotheses A and B.

2.2.1 Hybrid Predictions: Using a Prior and Current Data to Predict a Future Classical Analysis

By utilizing the expectation identity (1) for testing the hypotheses A, we also have the following corollary in which we have shown that the hybrid conditional predictive power (HCPP) can be calculated in two ways: One way is to calculate the HCPP by using the predictive distribution, and the other way is to directly calculate an expectation. The proof of the corollary can be found in the supplement.

Corollary 3 The HCPP is

$$\begin{aligned}
&P(S^C_{\epsilon,\theta_0} \,|\, y_m, \text{prior}) \\
&= \int P(S^C_{\epsilon,\theta_0} \,|\, y_m, \theta)\pi(\theta \,|\, y_m)\mathrm{d}\theta = \mathsf{E}_{\theta|y_m}[P(S^C_{\epsilon,\theta_0} \,|\, y_m, \theta)] \\
&= \Phi\Big[\sqrt{\frac{n_0 n}{(n_0+m)(n_0+m+n)}}\frac{\sqrt{n_0}(\mu-\theta_0)}{\sigma} + \sqrt{\frac{m(n_0+m+n)}{n(n_0+m)}}\frac{\sqrt{m}(y_m-\theta_0)}{\sigma} \\
&\quad + \sqrt{\frac{(m+n)(n_0+m)}{n(n_0+m+n)}}z_\epsilon\Big],
\end{aligned} \tag{3}$$

where $\mathsf{E}_{\theta|y_m}$ takes expectation with respect to the random variable $\theta \,|\, y_m \sim \pi(\theta \,|\, y_m)$.

The word "prior" in $P(S^C_{\epsilon,\theta_0} \,|\, y_m, \text{prior})$ means that the prior distribution $\pi(\theta) \sim \mathrm{N}(\mu, \sigma^2/n_0)$ where $n_0 > 0$ is incorporated in the calculation of the HCPP through the posterior distribution $\pi(\theta \,|\, y_m)$. Therefore, the mathematical expression of the HCPP involves the hyper parameters n_0 and μ of the prior distribution $\pi(\theta)$. The word "prior" in $P(S^C_{\epsilon,\theta_0} \,|\, y_m, \text{prior})$ can be omitted without much confusion, with the understanding that the prior $\pi(\theta)$ is used to derive the posterior $\pi(\theta \,|\, y_m)$. Note that the symbol $P(S^C_\epsilon \,|\, y_m, \text{prior})$ has been used in [8] in the special case of $\theta_0 = 0$.

By using the expectation identity for testing the hypotheses B, the calculation of the HCPP

$$P(S^{C-}_{\epsilon,\theta_0} \,|\, y_m, \text{prior}) = \int P(S^{C-}_{\epsilon,\theta_0} \,|\, y_m, \theta)\pi(\theta \,|\, y_m)\mathrm{d}\theta = \mathsf{E}_{\theta|y_m}[P(S^{C-}_{\epsilon,\theta_0} \,|\, y_m, \theta)]$$

can be found in the supplement. The expressions of the hybrid conditional predictive powers (HCPPs) for "Hybrid predictions" are summarized in Table 2.

Table 2 The conditional expected or predictive powers for sequential trials

	$H_0:\theta\leqslant\theta_0,\ H_1:\theta>\theta_0$ (Hypotheses A)	$H_0:\theta\geqslant\theta_0,\ H_1:\theta<\theta_0$ (Hypotheses B)
Hybrid conditional predictive power (HCPP)	$P(S^{C}_{\epsilon,\theta_0}\,\vert\, y_m,\text{prior})$ $=\Phi\left[\begin{array}{l}\sqrt{\frac{n_0 n}{(n_0+m)(n_0+m+n)}}\frac{\sqrt{n_0}(\mu-\theta_0)}{\sigma}\\ +\sqrt{\frac{m(n_0+m+n)}{n(n_0+m)}}\frac{\sqrt{m}(y_m-\theta_0)}{\sigma}\\ +\sqrt{\frac{(m+n)(n_0+m)}{n(n_0+m+n)}}z_\epsilon\end{array}\right]$	$P(S^{C-}_{\epsilon,\theta_0}\,\vert\, y_m,\text{prior})$ $=\Phi\left[\begin{array}{l}-\sqrt{\frac{n_0 n}{(n_0+m)(n_0+m+n)}}\frac{\sqrt{n_0}(\mu-\theta_0)}{\sigma}\\ -\sqrt{\frac{m(n_0+m+n)}{n(n_0+m)}}\frac{\sqrt{m}(y_m-\theta_0)}{\sigma}\\ +\sqrt{\frac{(m+n)(n_0+m)}{n(n_0+m+n)}}z_\epsilon\end{array}\right]$
Bayesian conditional predictive power (BCPP)	$P(S^{B}_{\epsilon,\theta_0}\,\vert\, y_m,\text{prior})$ $=\Phi\left[\begin{array}{l}\sqrt{\frac{n_0+m+n}{(n_0+m)n}}\frac{n_0(\mu-\theta_0)+m(y_m-\theta_0)}{\sigma}\\ +\sqrt{\frac{n_0+m}{n}}z_\epsilon\end{array}\right]$	$P(S^{B-}_{\epsilon,\theta_0}\,\vert\, y_m,\text{prior})$ $=\Phi\left[\begin{array}{l}-\sqrt{\frac{n_0+m+n}{(n_0+m)n}}\frac{n_0(\mu-\theta_0)+m(y_m-\theta_0)}{\sigma}\\ +\sqrt{\frac{n_0+m}{n}}z_\epsilon\end{array}\right]$
Classical conditional predictive power (CCPP)	$P(S^{C}_{\epsilon,\theta_0}\,\vert\, y_m,n_0=0)$ $=\Phi\left[\sqrt{\frac{m+n}{n}}\frac{\sqrt{m}(y_m-\theta_0)}{\sigma}+\sqrt{\frac{m}{n}}z_\epsilon\right]$	$P(S^{C-}_{\epsilon,\theta_0}\,\vert\, y_m,n_0=0)$ $=\Phi\left[-\sqrt{\frac{m+n}{n}}\frac{\sqrt{m}(y_m-\theta_0)}{\sigma}+\sqrt{\frac{m}{n}}z_\epsilon\right]$

2.2.2 Bayesian Predictions: Using a Prior and Current Data to Predict a Future Bayesian Analysis

We wish to calculate the predictive probability of obtaining a "significant" Bayesian result, and we shall denote such "Bayesian significance" as $S^{B}_{\epsilon,\theta_0}\equiv\{\mathsf{P}(\theta\leqslant\theta_0\,|\,\text{data})<\epsilon\}$ or $S^{B-}_{\epsilon,\theta_0}$ $equiv\{\mathsf{P}(\theta\geqslant\theta_0\,|\,\text{data})<\epsilon\}$. By using the expectation identity for testing the hypotheses A and B, the calculations of the Bayesian conditional predictive powers (BCPPs)

$$P(S^{B}_{\epsilon,\theta_0}\,|\,y_m,\text{prior})=\int P(S^{B}_{\epsilon,\theta_0}\,|\,y_m,\theta)\pi(\theta\,|\,y_m)\mathrm{d}\theta=\mathsf{E}_{\theta|y_m}[P(S^{B}_{\epsilon,\theta_0}\,|\,y_m,\theta)],$$
$$P(S^{B-}_{\epsilon,\theta_0}\,|\,y_m,\text{prior})=\int P(S^{B-}_{\epsilon,\theta_0}\,|\,y_m,\theta)\pi(\theta\,|\,y_m)\mathrm{d}\theta=\mathsf{E}_{\theta|y_m}[P(S^{B-}_{\epsilon,\theta_0}\,|\,y_m,\theta)],$$

respectively can be found in the supplement. The expressions of the BCPPs for "Bayesian predictions" are also summarized in Table 2.

2.2.3 Classical Predictions: Using Only Current Data to Predict a Future Classical Analysis

Classical predictions means that we ignore prior opinion both in the prediction and in the reporting. By using the expectation identity for testing the hypotheses A and B, the calculations of the classical conditional predictive powers (CCPPs)

$$P(S^{C}_{\epsilon,\theta_0}\,|\,y_m,n_0=0)=\int P(S^{C}_{\epsilon,\theta_0}\,|\,y_m,\theta)\pi(\theta\,|\,y_m,n_0=0)\mathrm{d}\theta=\mathsf{E}_{\theta|y_m,n_0=0}[P(S^{C}_{\epsilon,\theta_0}\,|\,y_m,\theta)],$$

$$P(S^{C-}_{\epsilon,\theta_0} \,|\, y_m, n_0 = 0) = \int P(S^{C-}_{\epsilon,\theta_0} \,|\, y_m, \theta)\pi(\theta \,|\, y_m, n_0 = 0)\mathrm{d}\theta = \mathsf{E}_{\theta|y_m,n_0=0}[P(S^{C-}_{\epsilon,\theta_0} \,|\, y_m, \theta)],$$

respectively can be found in the supplement. The expressions of the CCPPs for "Classical predictions" are also summarized in Table 2.

It is worthy to mention that the HCPP and the BCPP in Table 2 utilize both the historical data and the interim data, while the CCPP only utilizes the interim data.

2.3 The Average Success Probability in [9]

In [9], the average success probability (ASP) is defined by

$$\mathrm{ASP} = \int_{-\infty}^{\infty} P(S^{C,d_2}_{\alpha,\delta_0} \,|\, \delta)\pi(\delta \,|\, d_1)\mathrm{d}\delta,$$

where $S^{C,d_2}_{\alpha,\delta_0}$ is the classical rejection region of the hypotheses $H_0 : \delta \leqslant \delta_0$ versus $H_1 : \delta > \delta_0$, $P(S^{C,d_2}_{\alpha,\delta_0} \,|\, \delta)$ is the classical power of the confirmatory trial, α is the significance level, δ_0 is the threshold value of the hypotheses, δ is the unknown true treatment effect of the early and confirmatory trials, d_1 and d_2 are the observed treatment differences in the treatment group and the control (or placebo) group means of the early and confirmatory trials respectively.

There are two ways to analytically calculate the ASP. One way is to calculate the ASP by using the predictive distribution,

$$d_2 \,|\, d_1 \sim \mathrm{N}\Big(d_1, 2\sigma^2\Big(\frac{1}{m_1} + \frac{1}{m_2}\Big)\Big),$$

where σ^2 is a common known variance of the observations from the treatment and control groups, m_1 and m_2 are the per group number of patients of the early and confirmatory trials respectively. Note that when deriving the predictive distribution $d_2 \,|\, d_1$ of the ASP, the tricky part is

$$\pi(d_2 \,|\, \delta) = \pi(d_2 \,|\, \delta, d_1),$$

since $d_1 \,|\, \delta$ and $d_2 \,|\, \delta$ are assumed independent, and thus the marginal distribution of $d_2 \,|\, \delta$ is equal to the conditional distribution of $d_2 \,|\, \delta, d_1$. More details of the derivation of the predictive distribution $d_2 \,|\, d_1$ of the ASP can be found in the supplement.

Alternatively, the ASP can be rewritten as

$$\mathrm{ASP} = \mathsf{E}_{\delta|d_1}[P(S^{C,d_2}_{\alpha,\delta_0} \,|\, \delta)],$$

where $\mathsf{E}_{\delta|d_1}$ takes expectation with respect to the random variable $\delta \,|\, d_1 \sim \pi(\delta \,|\, d_1)$, and we can utilize the new expectation identity to analytically calculate the expectation. The

analytically formula of the ASP is found to be

$$\mathrm{ASP} = \Phi\Big[\Big(\frac{d_1 - \delta_0}{\sqrt{2/m_2}\sigma} - Z_\alpha\Big)\sqrt{\frac{m_1}{m_1 + m_2}}\Big]. \tag{4}$$

The analytically calculation of the ASP (4) by the two ways can be found in the supplement.

Note that in [9], she used the one-dimensional numerical integration to calculate the ASP. Although the numerical integration is accurate, it is time consuming. With the analytical formula of the ASP given by the equation (4), we can calculate the number of patients of the confirmatory trial m_2, which is the solution of the equation (4), for a given ASP. The solution of the equation (4) can be obtained by using the R function uniroot() very quickly and very accurately [19].

§3. Applications

In this section, we will illustrate the calculations of the predictive powers through examples.

Example 4 (Examples 2.6, 6.2, and 6.3 in [8]) Suppose we are designing a trial for a new cancer treatment which it is hoped will raise 5-year survival from 20% to 40%. This is equivalent to a hazard ratio of $\ln(0.40)/\ln(0.20) = 0.57$, when assuming proportional hazards, or a $\ln$(hazard ratio) of $\theta_A = -0.56$. We can take $\theta_A = 0.56$, which is equivalent to redefining the hazard ratio as control hazard divided by new intervention hazard instead of its inverse. Therefore, the hypotheses are $H_0 : \theta = 0$ versus $H_1 : \theta = \theta_A = 0.56 > 0$. Taking $\sigma = 2$ and assuming $\epsilon = 0.025$, 80% power is achieved at $n = 7.85 \times 2^2/(0.56)^2 = 100$.

Consider an archetypal enthusiastic prior centred on the alternative hypothesis and with 5% prior probability that $\theta < 0$. Hence $\theta \sim \mathrm{N}(\mu, \sigma^2/n_0)$, where $\mu = 0.56$, $\sigma = 2$, and $\mu - 1.645\sigma/\sqrt{n_0} = 0$ hold, such that $n_0 = 1.645^2\sigma^2/\mu^2 = 34.5$. The classical power evaluated at the prior mean is 80% as designed, *the expected power (the hybrid predictive power in Table 1) averaging over the entire prior distribution is* 0.66, showing the decline from the conditional value of 0.80. *Moreover, the expected Bayesian power (the Bayesian predictive power in Table 1), averaged with respect to the prior distribution, is* 0.78.

Example 5 ([20]; Example 6.7 in [8]) Long-term tamoxifen therapy is used for prevention of recurrence of breast cancer. The aim of the study is to estimate disease-free survival benefit from tamoxifen over placebo, in patients who already have had 5 years of taking tamoxifen without a recurrence. To detect a 40% reduction in annual risk associated

with tamoxifen (hazard ratio $= 0.6$), with 85% power and a one-sided tail area of 5%, 115 events were required. The statistical model is the proportional hazards regression model, with summary using the approximate hazard ratio analysis. If there are O_T events on treatment, and O_C events on control, then $y_m = 2(O_T - O_C)/m$ is an approximate estimate of the $\ln$(hazard ratio) θ, with mean θ and variance $4/m$. Prior distributions: An 'enthusiastic' (or optimistic) prior was centred on a 40% hazard reduction and a 5% chance of a negative effect (i.e., $\mathrm{HR} > 1$), equivalent on the $\ln(\mathrm{HR})$ scale to a normal prior with mean $\mu_o = \ln(0.6) = -0.51$ and standard deviation 0.31 ($\sigma = 2$, $n_0 = 41.4$). Also a sceptical prior was adopted with the same standard deviation as the enthusiastic prior but centred on $\mu_s = 0$. The estimated $\ln(\mathrm{HR})$ after the first interim analysis in 1993 is $y_m = 0.435$, at that time $m = 46$ events have been observed, and a further $n = 115 - 46 = 69$ events are to be observed.

Under the prior assumption Reference (the prior is neither used in the analysis nor in the prediction), the three probabilities are

$$\begin{aligned}
\mathrm{CCPP}_T &= P(S^{C-}_{\epsilon,\theta_0} \,|\, y_m, n_0 = 0),\\
\mathrm{CCPP}_C &= P(S^{C}_{\epsilon,\theta_0} \,|\, y_m, n_0 = 0),\\
\mathrm{CCPP}_E &= 1 - \mathrm{CCPP}_T - \mathrm{CCPP}_C,
\end{aligned}$$

for 'tamoxifen superior', 'control superior', and 'equivocal', respectively. Similarly, under the prior assumption 'When using prior in analysis' (the prior is used both in the analysis as well as the prediction), the three probabilities are

$$\begin{aligned}
\mathrm{BCPP}_T &= P(S^{B-}_{\epsilon,\theta_0} \,|\, y_m, \mathrm{prior}),\\
\mathrm{BCPP}_C &= P(S^{B}_{\epsilon,\theta_0} \,|\, y_m, \mathrm{prior}),\\
\mathrm{BCPP}_E &= 1 - \mathrm{BCPP}_T - \mathrm{BCPP}_C.
\end{aligned}$$

Under the prior assumption 'When not using prior in analysis' (the prior is not used in the analysis, but it is used in the prediction), the three probabilities are

$$\begin{aligned}
\mathrm{HCPP}_T &= P(S^{C-}_{\epsilon,\theta_0} \,|\, y_m, \mathrm{prior}),\\
\mathrm{HCPP}_C &= P(S^{C}_{\epsilon,\theta_0} \,|\, y_m, \mathrm{prior}),\\
\mathrm{HCPP}_E &= 1 - \mathrm{HCPP}_T - \mathrm{HCPP}_C.
\end{aligned}$$

The analytical forms of the probabilities CCPP_T, CCPP_C, BCPP_T, BCPP_C, HCPP_T, and HCPP_C can be obtained as in Table 2.

Table 3 illustrates the probabilities of eventual conclusions for the B-14 trial after the first interim analysis in 1993. From Table 3, we observe that the sceptical analysis and

the optimistic analysis both firmly predict an equivocal result at the end of the trial, and the observations are reflected in the predictive powers. The chance of finding in favour of tamoxifen is less than 0.017 in all cases, and thus we should stop the trial for futility.

Table 3 Probabilities of eventual conclusions for the B-14 trial after the first interim analysis in 1993

Final conclusion	Reference (CCPP)	When using prior in analysis (BCPP)		When not using prior in analysis (HCPP)	
		Sceptical	'Optimistic'	Sceptical	'Optimistic'
'Tamoxifen superior'	0.000	0.000	0.017	0.000	0.003
'Equivocal'	0.380	0.724	0.972	0.610	0.846
'Control superior'	0.619	0.276	0.011	0.390*	0.151

Example 6 ([9]) The ASPs when there are 128 and 172 patients per group when the posterior distribution of δ given $d_1 = 2.5$ is $\mathrm{N}(2.5, (2/m_1)7.14^2)$ are given in Table 4. Comparing Table 4 with Table I of [9], we find that the ASPs are the same. Though the one-dimensional numerical integration used in [9] is accurate, it is time consuming.

Table 4 The ASP when there are 128 and 172 patients per group when the posterior distribution of δ given $d_1 = 2.5$ is $\mathrm{N}(2.5, (2/m_1)7.14^2)$

Sample size in the future trial	$m_1 = 25$	$m_1 = 70$
$m_2 = 128$/group (80% power)	0.633	0.692
$m_2 = 172$/group (90% power)	0.677	0.756

Now we consider the inverse problem: Given an ASP and the number of patients of the early trial m_1, find the number of patients of the confirmatory trial m_2. Table 5 displays the m_2 for a given ASP assuming the normal ($m_1 = 25$ or 70) treatment effect. Note that there is an NA in Table 5. The reason is that for the normal treatment effect with $m_1 = 25$, the limiting ASP is 0.892 as $m_2 \to \infty$. That is, beyond the limiting ASP (0.892), we cannot find m_2. For the normal treatment effect with $m_1 = 70$, the limiting ASP is 0.981 as $m_2 \to \infty$, and thus for given ASPs equal to 0.8 and 0.9, we can find the corresponding m_2.

Table 5 The m_2 for a given ASP assuming the normal treatment effect

	$m_1 = 25$	$m_1 = 70$
ASP $= 0.8$	664	221
ASP $= 0.9$	NA	536

§4. Conclusions and Discussions

There are two ways to calculate the predictive powers: One way is to calculate the predictive powers by using the predictive distributions, and the other way is to calculate the expectations which are very involved and are always circumvented by [8] and other researchers. We prove and utilize an elegant expectation identity (1) to calculate the predictive powers by directly calculating the expectations without circumvention. The calculations of the predictive powers of the hypotheses A and B with a threshold θ_0 are divided into five different categories. That is, non-sequential trials with classical power and Bayesian power, and sequential trials with hybrid predictions, Bayesian predictions, and classical predictions. For the ASP in [9], we can also calculate it in two ways: By using the predictive distribution, and by utilizing the new expectation identity. Moreover, the calculations of the various predictive powers are illustrated through three examples. Finally, it indicates that the expectation identity can be utilized for the calculations of other predictive powers for normal prior and likelihood. It is also worthy to mention that in some cases such as the ASP in [9], it is tricky to find the predictive distribution for the predictive power, whereas, it is straightforward to utilize the expectation identity for the calculation of the predictive power.

Note that throughout the paper, we always assume that the variance σ^2 is known. This maybe not the truth especially for the early phase trial. However, in real applications and literature (see for instance [8, 9, 13, 21–24]), it is common practice to assume that the variance σ^2 is known to obtain analytical solutions, such as $\Phi(\cdot)$ for powers and average powers. For the unknown variance case, one might use the historical data to specify a sampling prior for σ^2 (see [25]). Alternatively, one might exploit a t statistic. As stated in [21], the sampling distribution of t is a non-central t distribution (which only becomes an ordinary Student t distribution if $\delta = 0$). However, based on publications or early phase trials, the estimate of σ^2 is good enough, so that it provides some assurance to the practioners that probably there is no need to have a prior for σ^2 when designing the confirmative trial.

Acknowledgments We would like to thank the referees for their constructive comments, which have led a substantial improvement of the paper. Moreover, we would like to thank Professor Ming-Hui Chen at University of Connecticut for providing a proof of the expectation identity.

Supporting Information Additional information for this article is available.

Supplement: Some proofs of the article.

R folder: R codes used in the article.

References

[1] LACHIN J M. A review of methods for futility stopping based on conditional power [J]. *Stat Med*, 2005, **24(18)**: 2747–2764.

[2] ZHANG Y, CLARKE W R. A flexible futility monitoring method with time-varying conditional power boundary [J]. *Clin Trials*, 2010, **7(3)**: 209–218.

[3] NEUENSCHWANDER B, ROYCHOUDHURY S, SCHMIDLI H. On the use of co-data in clinical trials [J]. *Stat Biopharm Res*, 2016, **8(3)**: 345–354.

[4] DENG Q Q, ZHANG Y Y, ROY D, et al. Superiority of combining two independent trials in interim futility analysis [J]. *Stat Methods Med Res*, 2020, **29(2)**: 522-540.

[5] LAN K K G, ZUCKER D M. Sequential monitoring of clinical trials: the role of information and brownian motion [J]. *Stat Med*, 1993, **12(8)**: 753–765.

[6] WANG Y N, LAN K K G, LI G, et al. A group sequential procedure for interim treatment selection [J]. *Stat Biopharm Res*, 2011, **3(1)**: 1–13.

[7] CHENG Y S, RAY S, CHANG M, et al. Statistical monitoring of clinical trials with multiple co-primary endpoints using multivariate B-value [J]. *Stat Biopharm Res*, 2014, **6(3)**: 241–250.

[8] SPIEGELHALTER D J, ABRAMS K R, MYLES J P. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation* [M]. Chichester: Wiley, 2004.

[9] CHUANG-STEIN C. Sample size and the probability of a successful trial [J]. *Pharm Stat*, 2006, **5(4)**: 305–309.

[10] IBRAHIM J G, CHEN M H, LAKSHMINARAYANAN M, et al. Bayesian probability of success for clinical trials using historical data [J]. *Stat Med*, 2015, **34(2)**: 249–264.

[11] LAN K K G, HU P, PROSCHAN M A. A conditional power approach to the evaluation of predictive power [J]. *Stat Biopharm Res*, 2009, **1(2)**: 131–136.

[12] CHUANG-STEIN C, YANG R Y. A revisit of sample size decisions in confirmatory trials [J]. *Stat Biopharm Res*, 2010, **2(2)**: 239–248.

[13] ZHANG Y Y, TING N. Bayesian sample size determination for a Phase III clinical trial with diluted treatment effect [J]. *J Biopharm Statist*, 2018, **28(6)**: 1119–1142.

[14] CHANG W H, CHUANG-STEIN C. Type I error and power in trials with one interim futility analysis [J]. *Pharm Stat*, 2004, **3(1)**: 51–59.

[15] LACHIN J M. Futility interim monitoring with control of type I and II error probabilities using the interim Z-value or confidence limit [J]. *Clin Trials*, 2009, **6(6)**: 565–573.

[16] DMITRIENKO A, WANG M D. Bayesian predictive approach to interim monitoring in clinical trials [J]. *Stat Med*, 2006, **25(13)**: 2178–2195.

[17] SCHMIDLI H, BRETZ F, RACINE-POON A. Bayesian predictive power for interim adaptation in seamless phase II/III trials where the endpoint is survival up to some specified timepoint [J]. *Stat Med*, 2007, **26(27)**: 4925–4938.

[18] ZHU C H, CHEN G J. Some optimal adaptive designs in clinical trials [J]. *Chinese J Appl Probab Statist*, 2005, **21(1)**: 67–75.

[19] R CORE TEAM. R: A Language and Environment for Statistical Computing [CP]. Vienna, Austria: R Foundation for Statistical Computing, 2019.

[20] DIGNAM J J, BRYANT J, WIEAND H S, et al. Early stopping of a clinical trial when there is evidence of no treatment benefit: protocol B-14 of the National Surgical Adjuvant Breast and Bowel Project [J]. *Control Clin Trials*, 1998, **19(6)**: 575–588.

[21] O'HAGAN A, STEVENS J W, CAMPBELL M J. Assurance in clinical trial design [J]. *Pharm Stat*, 2005, **4(3)**: 187–201.

[22] WANG S J, HUNG H M J, O'NEILL R T. Adapting the sample size planning of a phase III trial based on phase II data [J]. *Pharm Stat*, 2006, **5(2)**: 85–97.

[23] KIRBY S, BURKE J, CHUANG-STEIN C, et al. Discounting phase 2 results when planning phase 3 clinical trials [J]. *Pharm Stat*, 2012, **11(5)**: 373–385.

[24] LAN K K G, WITTES J T. Some thoughts on sample size: a Bayesian-frequentist hybrid approach [J]. *Clin Trials*, 2012, **9(5)**: 561–569.

[25] CHEN M H, IBRAHIM J G, LAM P, et al. Bayesian design of noninferiority trials for medical devices using historical data [J]. *Biometrics*, 2011, **67(3)**: 1163–1170.

一个新的期望恒等式及其在正态预测势计算中的应用

张应应　荣腾中　李曼曼

(重庆大学数学与统计学院统计与精算学系, 重庆, 401331)

摘　要: 为了计算预测势, 我们建议使用一个简洁的期望恒等式来直接计算期望值. 我们计算了具有非零阈值的假设对五种不同类型的预测势, 即具有经典势和贝叶斯势的非序贯试验, 以及混合预测、贝叶斯预测和经典预测的序贯试验. 此外, 通过三个例子说明了各种预测势的计算. 最后, 在计算文献 [9] 中的平均成功概率时, 很难找到预测势的预测分布, 而利用期望恒等式进行计算是很简单的.

关键词: 期望等式; 预测势; 正态模型; 单边假设; 平均成功概率

中图分类号: O212; O213; R9