SEQUENTIAL DESIGNS WITH APPLICATION IN SOFTWARE ENGINEERING

A DISSERTATION IN
Mathematics
and
Telecommunication and Computer Networking

Presented to the Faculty of the University
of Missouri-Kansas City in partial fulfillment of
the requirements for the degree

DOCTOR OF PHILOSOPHY

by
WEI WU

B.S. Nanjing University of Technology, 2005
M.S. University of Missouri-Kansas City, 2010

Kansas City, Missouri
2013

SEQUENTIAL DESIGNS WITH APPLICATION
IN SOFTWARE ENGINEERING

Wei Wu, Candidate for the Doctor of Philosophy Degree
University of Missouri-Kansas City, 2013

ABSTRACT

Presented here is a Bayesian approach to test case allocation in the software reliability estimation. Bayesian analysis allows us to update our beliefs about the reliability of a particular partition as we test, and thus, dynamically refine our allocation of test cases during the reliability testing process. We started with a fully sequential sampling scheme to estimate the reliability of a software system using partition testing. We have shown both theoretically and through simulation that the proposed scheme always performs at least as well as fixed sampling approaches where test case allocation is predetermined, and in all but the most unlikely circumstances, outperform them. Based on the sequential allocation, a multistage sampling scheme is established, which is less time consuming and more efficient. Meanwhile, an efficient sampling scheme is also developed to accommodate more situations. In the last chapter, we extend our study from parallel systems to series systems. We again use a Bayesian approach to allocate test cases to estimate the reliability of a series system with two components. A second-order lower bound

for the incurred Bayes risk is established theoretically and Monte Carlo simulations with several proposed sequential designs are implemented to achieve this second-order lower bound.

APPROVAL PAGE

The faculty listed below, appointed by the Dean of the College of Arts and Sciences have examined a dissertation titled "Sequential Designs with Application in Software Engineering," presented by Wei Wu, candidate for the Doctor of Philosophy, and certify that in their opinion it is worthy of acceptance.

Supervisory Committee

Kamel Rekab, Ph.D., Committe Chair
Department of Mathematics and Statistics

Yong Zeng, Ph.D.
Department of Mathematics and Statistics

Majid Bani-Yaghoub, Ph.D.
Department of Mathematics and Statistics

Deep Medhi, Ph.D.
Department of Computer Science and Electrical Engineering

Baek-Young Choi, Ph.D.
Department of Computer Science and Electrical Engineering

CONTENTS

LIST OF TABLES

LIST OF NOTATIONS

Notations for Chapter 1 to Chapter 4

R	Software reliability
D_i	The ith partition
p_i	Usage probability of the ith partition
R_i	Reliability of the ith partition
M	Total number of test cases to be allocated
k	Total number of partitions
m_i	Number of test cases allocated to the ith partition
$m_{i,t}$	Number of test cases allocated to the ith partition after t test cases have been executed
$x_{i,j}$	Outcome of the jth test taken from ith partition
$\hat{R}_i$	Estimator of R_i
$\hat{R}$	Estimator of R
$p.d.f.$	Probability Density Function
$\alpha_i^0, \beta_i^0, r_i$	Initial beta parameters for the ith partition
$\Re(F)$	Bayes risk of the fixed sampling scheme
$\Re(F_o)$	Bayes risk of the optimal fixed sampling scheme
$\Re(O)$	Bayes risk of the optimal sampling scheme
$\hat{C}_i(t), \hat{C}_{i,j}(t)$	Allocation ratios after t test cases have been executed
$\Re(S_p)$	Bayes risk of the purely sequential sampling scheme
$\Re(S_m)$	Bayes risk of the multistage sampling scheme
$\Re(S_a)$	Bayes risk of the accelerated sampling scheme

Notations for Chapter 5

R	Software reliability
$\mathcal{P}_i$	The ith component
t	The total number of test cases to be allocated
M	Number of test cases allocated to the first component
N	Number of test cases allocated to the second component
θ	Reliability of the first component
ω	Reliability of the second component
X_i	The outcome of ith test taken from the ith component
a, b, r	Initial beta parameters for the first component
c, d, s	Initial beta parameters for the second component
$\Re(\Delta)$	Bayes risk of the sequential design Δ

ACKNOWLEDGEMENT

I would like to thank Kamel Rekab for his help, support and guidance that made this dissertation possible. I would like to thank my wife for unending patience in the face of tremendous adversity. Professor Yong Zeng, Professor Deep Medhi, Associate Professor Baek-Young Choi and Assistant Professor Majid Bani-Yaghoub kindly agreed to be on my doctoral committee, and I would like to express my gratitude to them.

CHAPTER 1

INTRODUCTION

Testing software to estimate reliability is most often done late in the product development cycle, when functionality is in place, and the application is relatively stable. It is at this phase when decision makers are most concerned about estimating field reliability. Understandably, such an estimate greatly impacts the decision to release. In critical systems, this impact is especially important, where the tolerance for failure can be in the order of 10^{-3} or smaller [17]. Overestimating reliability in this instance could have disastrous consequences, both for those who depend on the system and the company that produces it. Alternatively, if management has given some reliability guarantee to users, either explicitly or implicitly, and in-house reliability estimates inaccurately fall below these benchmarks, then release of the software may be delayed unnecessarily, sometimes at great expense to the organization. Thus, estimates that closely approximate reality can reduce risk and decrease the cost of software development. The goal of this paper is to introduce a sequential sampling method to test for software reliability.

Our approach requires that we break up the domain of possible test cases into partitions. These partitions should be non-overlapping such that if test case i belongs to partition j then no partition other than j will contain test case i. Software testing using samples of test cases drawn from partitions such as this is referred to as partition testing [1], [34], [39]. There are several possible criteria we can use to partition the software, by component for example. There are many general approaches in the literature that address the mechanics, such as the usage model [34] and data flow testing criteria [39], among others [10], [40].

We begin formally by partitioning the test domain into k subdomains, denoted here by $D_1, D_2, \ldots, D_k$. For each subdomain D_i, we have two associated values: p_i, which is the probability that a given post-delivery use of the software will be of partition i ($p_1, p_2, \ldots, p_k$ are known parameters of the operational profile [1], [38]); and R_i, which is the conditional reliability of a use case or a test case [22], on condition that it was randomly chosen from within partition i. Within partition i, each test case has an equal chance of being selected. There are several approaches for estimating the operational profile and the corresponding usage probabilities, such as the usage model [38] and markov chain model [41].

The definition of reliability used here will be the one described by Poore et. al. [22]: Reliability is the probability that the software will give the correct result for a single randomly chosen [according to the operational profile] use. Using this definition, software reliability, R, is represented by

$$R = \sum_{i=1}^{k} p_i R_i, \tag{1.0.1}$$

see for example [1], [34]. The impossibility of complete testing of any software system of non-trivial size precludes us from knowing the conditional reliability R_i of each subdomain exactly [14]. Instead, we must distribute the M test cases allocated for reliability estimation among these k partitions, and use the results to estimate each R_i. Specifically, sample sizes $m_1, m_2, \ldots, m_k$ are taken from subdomains $D_1, D_2, \ldots, D_k$ respectively, where $m_1 + m_2 + \ldots + m_k = M$. Several approaches have been taken in the past to allocate test cases under the criteria of minimizing the variance of the maximum likelihood estimator of R (see for example [34]). These allocation decisions, however, rely solely on the information obtained from samples taken during reliability testing. To take advantage of an initial belief

about the conditional reliability within a particular subdomain, we adopt a fully Bayesian approach which allows us to refine these beliefs as we sample to improve future allocation decisions.

Because the total number of test cases is fixed, it is possible to determine the optimal allocation of test cases among the partitions before testing begins based only on the prior distributions within the subdomains and usage probabilities $p_1, \ldots, p_k$. These allocations, which are made before sampling begins, are referred to as fixed sampling schemes. We will show that making allocation decisions as we sample, based on the prior distributions updated by sampling results, will outperform the best fixed sampling scheme in terms of minimizing the expected loss incurred by the Bayes estimator of the overall reliability.

Bayesian-based allocations were sought by [3], [13], and [18], although the focus here is on estimating software reliability instead of software predictive reliability [3], p-values estimation [13], or estimating the probability of failure [18]. Other criteria such as maximizing the payoff in expected reliability [37] have been studied. For related work on Bayesian-based optimal test allocations in software reliability, see [31].

CHAPTER 2

A FULLY SEQUENTIAL TEST ALLOCATION FOR SOFTWARE RELIABILITY ESTIMATION

In this section we propose a method to determine how to sequentially allocate test cases among partitions of the software to minimize the expected loss incurred by the Bayes estimator of the overall reliability when the total number of software test cases is fixed. In contrast to fixed sampling schemes, where the proportion of test cases taken from each partition is determined before reliability testing begins, we make allocation decisions dynamically throughout the testing process. Using a fully Bayesian approach, we can take advantage of information from previous functional testing and insights from developers. We then refine these estimates in an iterative manner as we sample. We also compare the results from a multistage sampling scheme with the optimal fixed sampling scheme, and demonstrate its superiority in terms of the expected loss incurred when the overall reliability is estimated by its Bayes estimator both theoretically and through Monte Carlo simulations.

2.1 The Baysian Model

We model the outcome of the jth test taken from the ith partition as a Bernoulli random variable $x_{i,j}$ such that

$$x_{i,j} = \begin{cases} 1, \text{ if test } j \text{ taken from } D_i \text{ is processed correctly.} \\ 0, \text{ otherwise,} \end{cases}$$

where $x_{i,j}$ has a Bernoulli distribution with parameter R_i. We define $\hat{R}_i$ as estimate of R_i after m_i tests have been allocated to partition i, namely $\hat{R}_i = E[R_i|x_{i,1}, \ldots, x_{i,m_i}]$. Based on this allocation, our estimate of R, denoted as $\hat{R}$, can thus be defined as $\hat{R} = \sum_{i=1}^{k} p_i \hat{R}_i$.

Our objective is to seek the most accurate measure of reliability we can obtain by allocating the M total test cases amongst the k partitions. To do this allocation, there are several alternatives including minimizing the expected loss [8], and maximizing expected utility [16], [42], among others [4], [15]. Since we are using a Bayesian method, we choose to define the accuracy of our prediction in terms of minimizing the average loss incurred by estimating R by $\hat{R}$, its Bayes estimator. Instead of determining optimal allocations, Littlewood and Wright [17] proposed Bayesian stopping rules to meet some specified reliability.

There are several potential choices to define a loss function, such as the absolute value of the difference between R and $\hat{R}$ [4], the squared difference [8], [33], and several others. The most common selection [8], [11], [15], [33] is the squared error loss because of its natural correspondence to Euclidian distance. For software reliability estimation, the squared error loss has been used often [8], [11], [15], [33], and it is perhaps the most natural way to think of a positive distance between two parameters. Therefore, we choose the squared error loss as our measure of distance, and thus define the loss function as $\ell(R, \hat{R}) = (R - \hat{R})^2$. Note that the squared error loss function can be extended to a more general loss function [36] as $\ell(R, \hat{R}) = s_o(\hat{R} - R)^r I_{\{\hat{R} \geq R\}} + s_u(R - \hat{R})^r I_{\{\hat{R} < R\}}$, where r is the order of the function, and s_o and s_u are the loss coefficients for overestimation and underestimation, respectively. The general loss function reduces to the squared error loss function when we set $s_u = s_o = 1$, and $r = 2$.

By choosing the squared error loss, the Bayes estimator of R is the posterior mean (see [4] pp.161). We thus seek to measure the expected loss incurred by estimating R by its Bayes estimator $\hat{R}$. This expected loss, $\Re(P)$, which is also referred to as the Bayes Risk $\Re(P) = E^{R^J}[(R - \hat{R})^2]$, is the expected loss with respect to R^J, the joint density of R, and the observed data. Because $\hat{R}$ is the posterior mean, we can rewrite the expected loss as

$$\Re(P) = E[Var[R|x_{1,1}, \ldots, x_{1,m_1}, \ldots, x_{k,1}, \ldots, x_{k,m_k}]], \tag{2.1.1}$$

which is the expectation with respect to the marginal $p.d.f.$ of $x_{1,1}, \ldots, x_{1,m_1}, \ldots, x_{k,1}, \ldots, x_{k,m_k}$ of the posterior variance of R given that the results of the M total test cases have been observed (see [15], [16], [32]).

Now we seek a form of (2.1.1) that is useful in making allocation decisions. We assume that the prior distributions for the partitions are independent. The independence of the priors is a reasonable assumption when partitioning is done by component. In these cases, the reliability of one component may be completely unrelated to that of another, especially if one is a recent addition, and the other is well tested; or the components were created by different groups that may use alternate tools, methodologies, and standards internally. Under the assumption of independence of the priors, and because $R = \sum_{i=1}^{k} p_i R_i$, and $p_1, \ldots, p_k$ are fixed, then

$$\Re(P) = \sum_{i=1}^{k} p_i^2 E[Var[R_i|x_{i,1}, \ldots, x_{i,m_i}]]. \tag{2.1.2}$$

Although many distributions are supported on interval $(0, 1)$ such as Uniform distribution on $(0, 1)$ and Logit-normal distribution, we choose the *Beta* probability distribution as our prior distribution. The *Beta* distribution can take a wide

variety of shapes to accommodate many diverse situations through the choice of the prior parameters. The *Beta* distribution was also found to be the most conservative choice of prior distributions under uncertainty [9], and has a rich history in probability theory [16]. For each subdomain D_i, we assume that the corresponding conditional reliability follows a *Beta* distribution such that $R_i \sim Beta(\alpha_i^0, \beta_i^0)$. One choice of α_i^0 and β_i^0 can be made based on what we believe to be the reliability of the subdomain i prior to reliability testing. We refer to this estimate as μ_i. Our degree of certainty in this prediction is expressed through the standard deviation σ_i. The expected value of R_i before testing begins, assuming a *Beta* distribution with parameters α_i^0 and β_i^0 is

$$E[R_i] = \frac{\alpha_i^0}{\alpha_i^0 + \beta_i^0},$$

and similarly, the variance is

$$Var(R_i) = \frac{\alpha_i^0 \beta_i^0}{(\alpha_i^0 + \beta_i^0)^2(\alpha_i^0 + \beta_i^0 + 1)}.$$

In search of the parameters of the Beta distribution, we set $E[R_i] = \mu_i$ and $Var[R_i] = \sigma_i^2$; and by combining the two equations, and through algebraic manipulation [18], we find

$$\alpha_i^0 = \frac{(\mu_i^2(1-\mu_i) - \sigma_i^2\mu_i)}{\sigma_i^2},$$

and

$$\beta_i^0 = \frac{(\mu_i^2(1-\mu_i) - \sigma_i^2\mu_i)(1-\mu_i)}{\sigma_i^2\mu_i}.$$

This result is true given the restriction that

$$0 < \mu_i^0 < 1,$$

and

$$0 < \sigma_i^2 < \mu_i^0(1-\mu_i^0).$$

This result offers some guidance to the practitioner in that α_i^0 and β_i^0 can be computed directly based on the values of μ_i and σ_i. The following theorem gives the overall software reliability.

Theorem 2.1.1

Given the independence of the priors, the expected loss incurred by estimating the overall reliability R by its Bayes estimator $\hat{R}$ is

$$\Re(P) = \sum_{i=1}^{k} p_i^2 E\left[\frac{R_i(1-R_i)}{m_i + r_i^0}\right], \tag{2.1.3}$$

where

$$r_i^0 = \alpha_i^0 + \beta_i^0.$$

Equation (2.1.3) will serve to guide the sequential allocation scheme.

Proof of theorem (2.1.1)

We will need show that

$$E[Var[R_i|x_{i,1},\ldots,x_{i,m_i}]] = E\left[\frac{R_i(1-R_i)}{m_i + r_i^0}\right]. \tag{2.1.4}$$

For the left hand side of (2.1.4), we first evaluate [4]

$$Var[R_i|x_{i,1},\ldots,x_{i,m_i}] = \frac{\alpha_i^{m_i}\beta_i^{m_i}}{(m_i + r_i^0)^2(m_i + r_i^0 + 1)},$$

then

$$E\left[Var[R_i|x_{i,1},\ldots,x_{i,m_i}]\right] = E\left[\frac{\alpha_i^{m_i}\beta_i^{m_i}}{(m_i + r_i^0)^2(m_i + r_i^0 + 1)}\right].$$

Next, we evaluate the R.H.S. of (2.1.4) yielding

$$E\left[\frac{R_i(1-R_i)}{m_i+r_i^0}\right] = E\left[E\left[\frac{R_i(1-R_i)}{m_i+r_i^0}|x_{i,1},\ldots,x_{i,m_i}\right]\right]$$

by application of the iterated expectation theorem (see [6] page 481). Because the inner expectation is conditional on the data, m_i can be taken out of the inner expectation, yielding

$$E\left[\frac{R_i(1-R_i)}{m_i+r_i^0}|x_{i,1},\ldots,x_{i,m_i}\right] = \frac{E[R_i(1-R_i)|x_{i,1},\ldots,x_{i,m_i}]}{m_i+r_i^0}.$$

Now, to show that the theorem is true, all that remains is to show that

$$E[R_i(1-R_i)|x_{i,1},\ldots,x_{i,m_i}] = \frac{\alpha_i^{m_i}\beta_i^{m_i}}{(m_i+r_i^0)(m_i+r_i^0+1)}. \tag{2.1.5}$$

For the L.H.S. of (2.1.5), by definition of the expectation,

$$E[R_i(1-R_i)|x_{i,1},\ldots,x_{i,m_i}] = \textstyle\int_0^1 R_i(1-R_i)\pi(R_i|x_1,\ldots,x_{m_i})dR_i,$$

where $\pi(\bullet)$ denotes the *p.d.f.* It then follows that

$$E[R_i(1-R_i)|x_{i,1},\ldots,x_{i,m_i}] = \int_0^1 \frac{R_i(1-R_i)R_i^{\alpha_i^{m_i}-1}(1-R_i)^{\beta_i^{m_i}-1}}{\beta(\alpha_i^{m_i},\beta_i^{m_i})}dR_i, \tag{2.1.6}$$

where

$$\beta(\alpha_i^{m_i},\beta_i^{m_i}) = \frac{\Gamma(\alpha_i^{m_i})\Gamma(\beta_i^{m_i})}{\Gamma(\alpha_i^{m_i}+\beta_i^{m_i})}, \tag{2.1.7}$$

which is a constant because the expectation on the LHS of (2.1.6) is conditional on the data. We can rewrite (2.1.6) as

$$E[R_i(1-R_i)|x_{i,1},\ldots,x_{i,m_i}] =$$
$$\frac{\beta(\alpha_i^{m_i}+1,\beta_i^{m_i}+1)}{\beta(\alpha_i^{m_i},\beta_i^{m_i})}\int_0^1 \frac{R_i^{(\alpha_i^{m_i}+1)-1}(1-R_i)^{(\beta_i^{m_i}+1)-1}}{\beta(\alpha_i^{m_i}+1,\beta_i^{m_i}+1)}dR_i.$$

Because

$$\int_0^1 \frac{R_i^{(\alpha_i^{m_i}+1)-1}(1-R_i)^{(\beta_i^{m_i}+1)-1}}{\beta(\alpha_i^{m_i}+1,\beta_i^{m_i}+1)}dR_i = 1.$$

We have now shown that

$$E[R_i(1-R_i)|x_{i,1},\ldots,x_{i,m_i}] = \frac{\beta(\alpha_i^{m_i}+1,\beta_i^{m_i}+1)}{\beta(\alpha_i^{m_i},\beta_i^{m_i})}.$$

By expanding the above through the use of (2.1.7), and using the properties of the Gamma function, it is now trivial to show that

$$E[R_i(1-R_i)|x_{i,1},\ldots,x_{i,m_i}] = \frac{\alpha_i^{m_i}\beta_i^{m_i}}{(m_i+r_i^0)(m_i+r_i^0+1)}. \qquad \blacksquare$$

2.2 Estimating Software Reliability

It is our contention that operating with the benefit of past knowledge gives us a distinct advantage over approaches in which test cases are distributed based solely on usage probabilities. This knowledge, as represented by the *Beta* distributions, is dynamic in that we update our beliefs about the softwares reliability periodically during the reliability testing process. The prior *p.d.f.*, which is denoted as $\pi(R_i)$, is given by $Beta(\alpha_i^0,\beta_i^0)$ [6] such that

$$\pi(R_i) = \frac{R_i^{\alpha_i^0-1}(1-R_i)^{\beta_i^0-1}}{\beta(\alpha_i^0,\beta_i^0)},$$

where the normalizing constant $\beta(\alpha_i^0,\beta_i^0)$ is the complete *Beta* function

$$\beta(\alpha_i^0,\beta_i^0) = \int_0^1 R_i^{\alpha_i^0-1}(1-R_i)^{\beta_i^0-1}dR_i.$$

We can determine the *updated* prior, referred to as the *posterior* distribution, of R_i given the results of the tests executed, to reflect our improved knowledge of the system. We define $m_{i,t}$ as the number of tests allocated to partition i after t total tests have been allocated. It can be shown that the posterior distribution $\pi(R_i|x_{i,1},\cdots,x_{i,m_{i,t}})$ is given as $Beta(\alpha_i^{m_{i,t}},\beta_i^{m_{i,t}})$ with *p.d.f.*

$$\pi(R_i|x_{i,1},\cdots,x_{i,m_{i,t}}) = \frac{R_i^{\alpha_i^{m_{i,t}}-1}(1-R_i)^{\beta_i^{m_{i,t}}-1}}{\beta(\alpha_i^{m_{i,t}},\beta_i^{m_{i,t}})},$$

where

$$\beta(\alpha_i^{m_{i,t}}, \beta_i^{m_{i,t}}) = \int_0^1 R_i^{(\alpha_i^{m_{i,t}}-1)}(1-R_i)^{(\beta_i^{m_{i,t}}-1)}dR_i,\ \alpha_i^{m_{i,t}} = \alpha_i^0 + \sum_{j=1}^{m_{i,t}} x_{i,j},$$

and

$$\beta_i^{m_{i,t}} = \beta_i^0 + m_{i,t} - \sum_{j=1}^{m_{i,t}} x_{i,j}.$$

Using this distribution, we can estimate the reliability of subdomain i by calculating its expected value as

$$\hat{R}_i = E[R_i | x_{i,1}, \cdots, x_{i,m_{i,t}}] = \frac{\alpha_i^0 + \sum_{j=1}^{m_{i,t}} x_{i,j}}{\alpha_i^0 + \beta_i^0 + m_{i,t}}.$$

What makes this method appealing is that our starting point is a *Beta* distribution. We adjust that distribution based on the results of tests, and the result is another *Beta* distribution. This property is essential for dynamic decision making. As we decide where to test next, we need only to look at these updated distributions to select the subdomain that would most improve the accuracy of our overall reliability estimate

$$\hat{R} = \sum_{i=1}^{k} p_i \left(\frac{\alpha_i^0 + \sum_{j=1}^{m_{i,t}} x_{i,j}}{\alpha_i^0 + \beta_i^0 + m_i} \right). \tag{2.2.1}$$

2.3 Optimal Fixed Sampling Scheme

In a fixed sampling scheme, test cases are allocated before reliability testing begins based solely on prior information and usage probabilities [34]. We outline this procedure below. Recall Theorem (2.1.1). Because for fixed sampling schemes

each m_i is fixed, then equation (2.1.3) becomes

$$\Re(F) = \sum_{i=1}^{k} \frac{p_i^2 E[R_i(1-R_i)]}{m_i + r_i^0}, \tag{2.3.1}$$

where

$$r_i = \alpha_i^0 + \beta_i^0,$$

and $\Re(F)$ is the Bayes risk incurred using a fixed sampling scheme.

Theorem 2.3.1

Let $\Re(F)$ and $\Re(F_o)$ be the Bayes risks incurred using a fixed sampling scheme and the optimal fixed sampling scheme, respectively. Then

$$\Re(F) \geq \Re(F_o)$$

with equality iff

$$m_i = \frac{\left(M + \sum_{\ell=1}^{k} r_\ell^0\right) p_i \sqrt{\frac{\alpha_i^0 \beta_i^0}{r_i^0(r_i^0+1)}}}{\sum_{\ell=1}^{k} p_\ell \sqrt{\frac{\alpha_\ell^0 \beta_\ell^0}{r_\ell^0(r_\ell^0+1)}}} - r_i^0, \tag{2.3.2}$$

where

$$\Re(F_o) = \frac{\left(\sum_{i=1}^{k} p_i \sqrt{\frac{\alpha_i^0 \beta_i^0}{r_i^0(r_i^0+1)}}\right)^2}{M + \sum_{i=1}^{k} r_i^0}$$

and $\Re(F)$ is given in (2.3.1).

Proof of theorem (2.3.1)

Equation (2.3.1) can be rewritten as [23]

$$\Re(F) = \frac{\left(\sum_{i=1}^{k} p_i \sqrt{E[R_i(1-R_i)]}\right)^2}{M + \sum_{i=1}^{k} r_i^0} + \frac{1}{M + \sum_{i=1}^{k} r_i^0} \cdot$$
$$\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \left(\frac{(m_i + r_i^0) p_j \sqrt{E[R_j(1-R_j)]} - (m_j + r_j^0) p_i \sqrt{E[R_i(1-R_i)]}}{(m_i + r_i^0)(m_j + r_j^0)} \right)^2,$$

which is minimized by choosing

$$\frac{m_i + r_i^0}{m_j + r_j^0} = \frac{p_i \sqrt{E[R_i(1-R_i)]}}{p_j \sqrt{E[R_j(1-R_j)]}}.$$

Combining the above with the fact that

$$E[R_i(1-R_i)] = \frac{\alpha_i^0 \beta_i^0}{(m_i + r_i^0)(m_i + r_i^0 + 1)},$$

$\Re(F)$ is minimized iff

$$m_i = \frac{\left(M + \sum_{\ell=1}^{k} r_\ell^0\right) p_i \sqrt{\frac{\alpha_i^0 \beta_i^0}{r_i^0 (r_i^0 + 1)}}}{\sum_{\ell=1}^{k} p_\ell \sqrt{\frac{\alpha_\ell^0 \beta_\ell^0}{r_\ell^0 (r_\ell^0 + 1)}}} - r_i^0$$

for $i = 1, \cdots, k-1$, and $m_k = M - \sum_{i=1}^{k-1} m_i$ with an associated Bayes risk of

$$\Re(F_o) = \frac{\left(\sum_{i=1}^{k} p_i \sqrt{\frac{\alpha_i^0 \beta_i^0}{r_i^0 (r_i^0 + 1)}}\right)^2}{M + \sum_{i=1}^{k} r_i^0}. \quad \blacksquare$$

It is interesting to note from equation (2.3.2) that the best fixed allocation is very sensitive to the choice of prior parameters. Small inaccuracies in α_i^0 and β_i^0 can have a large impact on the proportion of test cases allocated to a particular subdomain. Because with fixed allocation these proportions are determined in advance, we have no opportunity to recover from our mistakes in choosing the parameters in light of testing results. It is this shortcoming that motivates a dynamic allocation approach which will be discussed next.

2.4 Sequential Allocation

We seek a sampling approach in which we can refine our allocation of test cases dynamically among subdomains as we learn more about the software. Our objective is still to minimize the expected loss, but here we make allocation decisions at various intervals during the testing process. Consider again (from Theorem 2.1.1) the expected loss

$$\Re(P) = \sum_{i=1}^{k} p_i^2 E\left[\frac{R_i(1-R_i)}{m_i + r_i^0}\right] \tag{2.4.1}$$

where

$$r_i^0 = \alpha_i^0 + \beta_i^0.$$

In contrast to (2.3.1), m_i cannot be taken out of the expectation because it is a random variable, and therefore not fixed. Because each R_i is unknown, our objective is to determine the values of $m_1, \ldots, m_k$ that will minimize the expected loss based on reliability estimates updated periodically during the testing process. For any sampling rule, the expected loss of (2.4.1) can be rewritten as

$$\begin{aligned}\Re(P) = &\frac{E\left[\left(\sum_{i=1}^{k} p_i\sqrt{R_i(1-R_i)}\right)^2\right]}{M + \sum_{i=1}^{k} r_i^0} + \frac{1}{\left(M + \sum_{i=1}^{k} r_i^0\right)}\cdot \\ &E\left[\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} \frac{\left((m_i + r_i^0)p_j\sqrt{R_j(1-R_j)} - (m_j + r_j^0)p_i\sqrt{R_i(1-R_i)}\right)^2}{(m_i + r_i^0)(m_j + r_j^0)}\right].\end{aligned} \tag{2.4.2}$$

Our objective is to minimize $\Re(P)$ through the choice of $m_1, \ldots, m_k$. Because the sum $m_1 + m_2 + \ldots + m_k = M$ is fixed, the first term in the sum above

is also fixed, and beyond our control. Thus, $\Re(P)$ is bounded below by

$$\Re(P) \geq \frac{E\left[\left(\sum_{i=1}^{k} p_i\sqrt{R_i(1-R_i)}\right)^2\right]}{M + \sum_{i=1}^{k} r_i^0}$$

with equality iff

$$\frac{m_i + r_i^0}{m_j + r_j^0} = \frac{p_i\sqrt{R_i(1-R_i)}}{p_j\sqrt{R_j(1-R_j)}} \tag{2.4.3}$$

for each i, j. Since the actual conditional reliability R_i for each subdomain is unknown, the optimal sampling scheme is not practical. Instead, we estimate its value at intervals during the testing process, and choose which subdomain to sample from to adjust the ratio such that

$$\frac{m_{i,t} + r_i^0}{m_{j,t} + r_j^0}$$

is close to $\hat{C}_{i,j}(t)$, where

$$\hat{C}_{i,j}(t) = \frac{p_i}{p_j}\frac{E[\sqrt{R_i(1-R_i)}|\mathscr{F}_t]}{E[\sqrt{R_j(1-R_j)}|\mathscr{F}_t]}, \tag{2.4.4}$$

and $\mathscr{F}_t$ is the data collected after t test cases have been executed.

2.5 A Purely Sequential Sampling Scheme

In this section we present a method for allocating test cases to partitions of the software based on the results of previous tests. In a fully sequential procedure individual test cases must be executed in series because allocation decisions are made after each test execution. This approach would be appropriate for software that runs on platforms where accuracy of the reliability prediction outweighs the time saved by executing tests on multiple machines or where obtaining time on

multiple machines is expensive, as is the case for mainframes or custom test hardware. We proceed to test as follows:

Step 1:

Test one unit from each partition D_i.

Step 2:

After ℓ tests have been allocated, where $\ell \geq k$, we take test $\ell + 1$ from partition i if for all j:

$$\frac{m_{i,\ell} + r_i^0}{m_{j,\ell} + r_j^0} < \hat{C}_{i,j}(\ell).$$

Step 3:

This approach is applied sequentially until all M tests are allocated.

We will show that this approach yields a smaller expected loss (Bayes risk) than the best fixed sampling scheme. The disadvantage here is that $M - k$ decisions are necessary to distribute the test cases and attain our estimate. This process can easily be automated however and is especially attractive when tests are time consuming and expensive to run. The time needed to execute a single test often increases with the products maturity, as most easy to find failures have been weeded out. It is during this latter phase of the development process when decision makers are most interested in accurate reliability estimates. Since reliability is likely to be closer to release levels and other milestones only in the latter stages of development, it is only here when reliability becomes worth measuring and measuring accurately.

2.6 Sampling Scheme Comparisons

When compared purely sequential sampling scheme with the optimal fixed design, the purely sequential sampling scheme is expected to be more effective in

terms of minimizing the Bayes risk, especially when M is large. It is expected to be more effective because we use information gathered during testing in our allocation procedures as opposed to predetermining allocation based solely on the prior *Beta* distribution. We can also determine the Bayes risk of the theoretical optimal sampling scheme. Recall Equation (2.4.2) and we see that the first term does not depend upon the individual m_i values. By choosing $m_1, m_2, \ldots, m_k$ such that Equation (2.4.3) becomes equality, we see that the Bayes risk of the fully sequential design, $\Re(O)$, becomes:

$$\Re(P) \geq \frac{E\left[\left(\sum_{i=1}^{k} p_i \sqrt{R_i(1-R_i)}\right)^2\right]}{M + \sum_{i=1}^{k} r_i^0}.$$

We will present theoretical results comparing the Bayes risk of the sequential sampling schemes presented here with both the fully sequential and best fixed designs. Section (2.9) shows the experimental results obtained through Monte Carlo simulations comparing the fully sequential sampling scheme to the best fixed sampling scheme.

2.7 Theoretical Results

We now compare the performance of the sampling schemes presented here when M is large. One method of comparison is to determine the cost, in terms of the Bayes risk, that is incurred over the Bayes risk of the optimal sampling scheme. Specifically, we seek to determine $\Re(S_p) - \Re(O)$, where $\Re(S_p)$ is the Bayes risk of purely sequential sampling scheme and $\Re(O)$ is the Bayes risk of the optimal sampling scheme. The following theorem quantifies the order of $\Re(S_p) - \Re(O)$ for large M.

Theorem 2.7.1

The excess Bayes risk incurred by the purely sequential sampling scheme over the Bayes risk incurred by the optimal sampling scheme is of order $1/M$.

Proof of theorem (2.7.1)

We need to show that $M[\Re(S_p) - \Re(O)] \to 0$ as $M \to \infty$, where

$$\Re(S_p) = \frac{E\left[\left(\sum_{i=1}^{k} p_i\sqrt{R_i(1-R_i)}\right)^2\right]}{M + \sum_{i=1}^{k} r_i^0} + \frac{1}{\left(M + \sum_{i=1}^{k} r_i^0\right)} \cdot$$

$$E\left[\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\left(\frac{(m_i + r_i^0)p_j\sqrt{R_j(1-R_j)} - (m_j + r_j^0)p_i\sqrt{R_i(1-R_i)}}{(m_i + r_i^0)(m_j + r_j^0)}\right)^2\right],$$

and

$$\Re(O) = \frac{E\left[\left(\sum_{i=1}^{k} p_i\sqrt{R_i(1-R_i)}\right)^2\right]}{M + \sum_{i=1}^{k} r_i^0}.$$

The proof of the theorem is established if we show that in the purely sequential sampling scheme

$$E\left[\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\left(\frac{(m_i + r_i^0)p_j\sqrt{R_j(1-R_j)} - (m_j + r_j^0)p_i\sqrt{R_i(1-R_i)}}{(m_i + r_i^0)(m_j + r_j^0)}\right)^2\right] \to 0$$

as $M \to \infty$.

The proof of the theorem follows if one shows

$$\frac{m_i}{m_j} \to \frac{p_i\sqrt{R_i(1-R_i)}}{p_j\sqrt{R_j(1-R_j)}}$$

as $M \to \infty$.

The proof of the theorem is similar to Rekab [25] although in this work we are using a Bayesian methodology to estimate software reliability instead of classical statistics to estimate the reliability of a series system. For ℓ large enough there exists

$$\ell^{(i)} = \sup\{q < \ell : \frac{m_{i,q}}{m_{j,q}} < C_{i,j}(q) \text{ for all } j \neq i\}$$

and $\ell^{(i)} \to \infty$ as $\ell \to \infty$. Then

$$\frac{m_{i,\ell}}{m_{j,\ell}} \leq \frac{m_{i,\ell^{(i)}} + 1}{m_{j,\ell^{(i)}}} \leq C_{i,j}(\ell^{(i)}) + \frac{1}{m_{j,\ell^{(i)}}}.$$

On the other hand,

$$\frac{m_{i,\ell}}{m_{j,\ell}} \geq \frac{m_{i,\ell^{(i)}} - 1}{m_{j,\ell^{(i)}}} \geq C_{i,j}(\ell^{(i)}) \left(1 - \frac{1}{m_{j,\ell^{(i)}}}\right).$$

By the optional stopping theorem [2], $C_{i,j}(\ell^{(i)})$ is a martingale. The rest of the proof follows from the martingale convergence theorem [2] [7]. ∎

Theorem (2.7.1) has several interesting implications. As M gets large, we can see that the purely sequential sampling scheme approaches the optimal, in terms of the Bayes risk. This seems reasonable because in each scheme we use the knowledge gained from test results to improve our allocations during the testing process.

Theorem 2.7.2

For large M,

$$\frac{\Re(S_p)}{\Re(F_o)} \leq 1$$

where $\Re(S_p)$ is the Bayes risk incurred by the purely sequential sampling scheme and $\Re(F_o)$ is the Bayes risk incurred by the best fixed design.

Proof of theorem (2.7.2)

For large M,

$$\left(M + \sum_{i=1}^{k} r_i^0\right) \Re(S_p) \to E\left[\left(\sum_{i=1}^{k} p_i \sqrt{R_i(1-R_i)}\right)^2\right],$$

and

$$\left(M + \sum_{i=1}^{k} r_i^0\right) \Re(F) \to \left(\sum_{i=1}^{k} p_i \sqrt{R_i(1-R_i)}\right)^2.$$

The proof is established by showing

$$\frac{E\left[\left(\sum_{i=1}^{k} p_i \sqrt{R_i(1-R_i)}\right)^2\right]}{\left(\sum_{i=1}^{k} p_i \sqrt{E[R_i(1-R_i)]}\right)^2} \leq 1.$$

We can rewrite the ratio above as

$$E\left[\left(\sum_{i=1}^{k} p_i \sqrt{R_i(1-R_i)}\right)^2\right] = T + T_1,$$

and

$$\left(\sum_{i=1}^{k} p_i \sqrt{E[R_i(1-R_i)]}\right)^2 = T + T_2,$$

where

$$T = \sum_{i=1}^{k} p_i^2 E[R_i(1-R_i)],$$
$$T_1 = 2\sum_{i<j} p_i p_j E[\sqrt{R_i(1-R_i)R_j(1-R_j)}],$$

and

$$T_2 = 2\sum_{i<j} p_i p_j \sqrt{E[R_i(1-R_i)R_j(1-R_j)]}.$$

Hence,

$$\frac{E\left[\left(\sum_{i=1}^{k} p_i \sqrt{R_i(1-R_i)}\right)^2\right]}{\left(\sum_{i=1}^{k} p_i \sqrt{E[R_i(1-R_i)]}\right)^2} = \frac{T + T_1}{T + T_2}.$$

Because the square root is concave, by Jensen's Inequality $T_1 \leq T_2$, the proof easily follows. ∎

Theorem (2.7.2) shows that for a large number of test cases, allocating tests through one of the sequential schemes presented here will always yield estimates of reliability that are at least as accurate as those computed by predetermining allocations before testing begins. For large M, if the parameters chosen for the prior correctly describe the distribution of R_i then we have equality. It is unlikely that such parameters would be chosen correctly however, and to assume that they would negate the need for reliability testing altogether.

2.8 Monte Carlo Simulations

Below we present the results of experimental comparisons between the purely sequential sampling scheme described above and the best fixed sampling scheme. Table 1 in page 64 shows the ratios of the Bayes risk of the purely sequential sampling scheme to the Bayes risk of the best fixed scheme.

We consider the case where the test domain is divided into two disjoint subdomains, D_1 and D_2, with reliability R_1 and R_2 respectively. In each subdomain we assume that these unknown reliability statistics follow a *Beta* distribution such that $R_1 \sim Beta(\alpha_1, \beta_1)$ and $R_2 \sim Beta(\alpha_2, \beta_2)$. The parameters of these distributions are chosen to be small relative to the number of test cases executed, which indicates reasonable uncertainty as expressed through a high standard deviation. This is by far the dominant situation in practice. For large parameter values, the posterior distribution is dominated by the prior when the number of test cases is small, which means the initial guess will govern our prediction. This is obviously undesirable, as our goal is for the prior to serve as a guide to help in allocation decisions. If we were to assume that these guesses were exceptionally accurate, this would negate the need for reliability testing.

Associated with each domain D_i is also a usage probability p_i. The tables present results for various total sample sizes, M, and the last column indicates the limiting value of the ratio $\Re(S_p)/\Re(F_o)$ as M becomes very large. The ratio in parenthesis below is the percentage of tests that were allocated to subdomain D_1 by the particular sequential sampling scheme.

In all cases, even for small sample sizes, the purely sequential sampling scheme outperformed the best fixed sampling scheme. Intuitively, and by Theorem (2.7.2), this is what we would expect given that the sequential sampling schemes use information contained in the Beta priors as well as information gained during the testing process.

Rows 1 and 2 in the table show scenarios with uniform priors where $\alpha_1 = \beta_1 = \alpha_2 = \beta_2 = 1$. This is the most commonly used prior parameter configuration, and from a practical point of view, this indicates that very little is known about the reliability of each partition. In these cases, we still gain by using a sequential approach over the best fixed allocation as is evident by ratios less than 1. As the total sample size M increases we approach the limiting value given in the last column. Rows 3 and 4 have identical prior distributions for the two subdomains such that $\alpha_1 = \alpha_2 = 0.5$ and $\beta_1 = \beta_2 = 0.01$. This choice represents a high expected reliability in each partition since α is much larger than β. This is a very plausible scenario in practice and indicates that reliability in each subdomain is assumed to be high. Under these conditions, and where the probability of use for each subdomain is equal ($p_1 = p_2$), the sequential sampling schemes performed exceptionally well compared to the best fixed allocation. This is consistent with the results found separately by Shapiro [35] and Hardwick and Stout [12] when using a fully sequential approach with equally weighted subdomains.

Perhaps the most interesting comparison results come from rows 5 and 6 of the table. Here, in D_1, reliability is assumed to be high as expressed through the prior parameters $\alpha_1 = 0.1$ and $\beta_1 = 0.001$, and a uniform prior for D_2 indicates very little information about the reliability of that partition. The usage probabilities $p_1 = .9$ and $p_2 = .1$ indicates that the usage probability of D_2, is far less than that of D_1. In this instance, D_1 could represent test cases that traverse the majority of functionality of the product which has been thoroughly tested and is assumed to be fairly reliable. Similarly, D_2 may contain test paths that exercise a recently added component for which little is known in terms of its reliability. In this case purely sequential sampling scheme shows a marked improvement over the best fixed scheme. This is one of the most encouraging results from a practitioners point of view.

While these results are compelling, minimizing the variance is not the only objective that a tester may pursue. Reliability testing may in fact just be used as a gauge to judge the product reaching a mile stone or its fitness for release. If system reliability is clearly insufficient in these circumstances, knowing accurately how unreliable the system is may be useless. For circumstances where reliability is close to meaningful levels, however, the methods presented here show significant promise.

CHAPTER 3

A MULTISTAGE SEQUENTIAL TEST ALLOCATION FOR SOFTWARE RELIABILITY ESTIMATION

Consider the same problem in chapter 2 where our purpose is to determine how to sequentially allocate test cases among partitions of the software to minimize the expected loss incurred by the Bayes estimator of the overall reliability when the total number of software test cases is fixed. In a fully sequential sampling scheme, individual test cases must be executed one by one because allocation decisions are made after each test execution. Therefore using a fully sequential sampling scheme could be very costly in terms of both efficiency and affordability. In this section we will propose a multistage sampling scheme that is more efficient. We will also compare the results from a multistage sampling scheme with the optimal fixed sampling scheme, and demonstrate its superiority in terms of the expected loss incurred when the overall reliability is estimated by its Bayes estimator both theoretically and through Monte Carlo simulations. A real case study will also be introduced to demonstrate the superiority of the proposed method.

3.1 Sequential Allocation

Recall section (2.5), $\Re(P)$ is bounded below by

$$\Re(P) \geq \frac{E\left[\left(\sum_{i=1}^{k} p_i \sqrt{R_i(1-R_i)}\right)^2\right]}{M + \sum_{i=1}^{k} r_i^0}$$

with equality iff

$$\frac{m_i + r_i^0}{m_j + r_j^0} = \frac{p_i\sqrt{R_i(1-R_i)}}{p_j\sqrt{R_j(1-R_j)}} \tag{3.1.1}$$

for each i, j. Note that (3.1.1) is equivalent to

$$\frac{m_i + r_i^0}{M + \sum_{i=1}^{k} r_i^0} = \frac{p_i\sqrt{R_i(1-R_i)}}{\sum_{j=1}^{k} p_j\sqrt{R_j(1-R_j)}}.$$

Because the actual conditional reliability R_i for each subdomain is unknown, the optimal sampling scheme is not practical. Instead, we estimate its value at intervals during the testing process, and choose which subdomain to sample from to adjust the ratio such that

$$\frac{m_{i,t} + r_i^0}{M + \sum_{j=1}^{k} r_j^0}$$

is close to $\hat{C}_i(t)$, where

$$\hat{C}_i(t) = \frac{p_i E\left[\sqrt{R_i(1-R_i)}|\mathscr{F}_t\right]}{\sum_{j=1}^{k} p_j E\left[\sqrt{R_j(1-R_j)}|\mathscr{F}_t\right]}, \tag{3.1.2}$$

and $\mathscr{F}_t$ is the data collected after t test cases have been executed.

3.2 Sequential Multistage Sampling

As opposed to purely sequential sampling [25] where we make allocation decisions one test at a time, multistage sampling allocates groups of test cases among the partitions based on previous test results. This approach may be more appropriate than the purely sequential sampling when M is large. If this is the case, it may become cost effective to test in parallel given the potential time savings.

Another consideration is the possible cost of switching between partitions during the testing process. In [34], Sayre and Poore discuss a partitioning scheme that separates test cases into two groups based on whether or not a specific risky input was part of the test path. In such a case, in general, it can be assumed that executing a test from one subdomain followed by another test from another subdomain would incur little overhead in switching between partitions. In other circumstances, however, when test cases are partitioned by component for instance, it could be significantly more efficient to execute a group of test cases from one partition before executing tests from another. In these instances, the accuracy lost by choosing a multistage sampling approach over a fully sequential scheme may be outweighed by the time saved executing tests. In the multistage scheme, we distribute the M test cases in L stages, which are fixed before testing begins, such that at stage j there are S_j test cases distributed during that stage. We define $S_{j,i}$ to be the *cumulative* test cases allocated to partition i up to stage j. Here is an outline of the process.

Stage 1:

Traditional approaches to multistage sampling [23], [24] use a balanced allocation of the S_1 test cases in the first stage, which may lead to over-sampling from a particular subdomain. To eliminate this concern, we propose allocating test cases at the first stage according to the best fixed scheme with S_1 total test cases. Thus, by (2.3.2), we allocate test cases in the first stage by the optimal fixed sampling scheme such that

$$S_{1,i} = \frac{\left(S_1 + \sum_{\ell=1}^{k} r_\ell^0\right) p_i \sqrt{\frac{\alpha_i^0 \beta_i^0}{(r_i^0)(r_i^0+1)}}}{\sum_{\ell=1}^{k} p_\ell \sqrt{\frac{\alpha_\ell^0 \beta_\ell^0}{(r_i^0)(r_i^0+1)}}} - r_i^0 \tag{3.2.1}$$

for $i = 1, 2, \ldots, k-1$ where $S_{1,i}$ is rounded to the nearest integer, and

$$S_{1,k} = S_1 - \sum_{i=1}^{k-1} S_{1,i}.$$

Unlike the purely sequential sampling scheme, once allocations have been determined within a stage, we are no longer restricted by having to execute these test cases in series. This advantage can be beneficial if multiple machines are available to execute tests, or if it is expensive to switch between partitions.

Stage 2 through L:

For Stages 2 through L, we distribute the test cases to the partitions based on the prior distribution updated by testing results from previous stages, and determine $S_{j,1}, S_{j,2}, \ldots, S_{j,k-1}$ sequentially such that, at each stage j, $2 \leq j \leq L$, for partition i,

$$S_{j,i} = \left(\left(\sum_{\ell=1}^{j} S_\ell \right) + \left(\sum_{\ell=1}^{k} r_\ell^0 \right) \right) \hat{C}_i(\bar{S}_{j-1}) - r_i^0, \tag{3.2.2}$$

and

$$S_{j,k} = \sum_{\ell=1}^{j} S_\ell - \sum_{\ell=1}^{k-1} S_{j,\ell},$$

where

$$\bar{S}_{j-1} = \sum_{y=1}^{j-1} S_y.$$

Therefore, at stage L, the total number of tests allocated to partition i is

$$M_i = \min\{M - \sum_{\substack{j=1 \\ j \neq i}}^{k} S_{L-1,j}, \max((M + \sum_{\ell=1}^{k} r_\ell^0)\hat{C}_i(\bar{S}_{L-1}) - r_i^0, S_{L-1,i})\}. \tag{3.2.3}$$

Note that (3.2.3) means M_i should be at least the total number of tests allocated up to stage $L-1$, i.e., $S_{L-1.i}$. In addition, M_i has to be at most $M - \sum_{\substack{j=1 \\ j \neq i}}^{k} S_{L-1,j}$

because there are only this many tests left. However, in the last stage, we are still allocating the tests sequentially, which leads to the max part, and thus (3.2.3).

There are several potential approaches to choosing the number of test cases allocated to each sampling stage. The simplest is to pick

$$S_1 = S_2 = \cdots = S_L = M/L,$$

and thus distribute the tests equally among the stages. There is much to be gained by choosing these stage sizes analytically, however. As a result, extensive work has been done on determining such allocations for two and three stage sampling procedures. In [23], Rekab shows that for a two stage procedure we should choose S_1 such that

$$\lim_{M\to\infty} \tfrac{S_1}{M} = 0, \text{ and } \lim_{M\to\infty} S_1 = \infty.$$

Although such criteria do not offer practical guidance toward selecting S_1, this result seems to indicate a choice of S_1 of the form M^λ, where $0 < \lambda < 1$. Hardwick and Stout [11] offer more specific guidance for some particular cases.

3.3 Theoretical Results

When choosing a sampling strategy, it is important to be aware of the tradeoffs involved in selecting one particular strategy over another. In this section, we will compare the Bayes risk incurred by the multistage to the Bayes risk incurred by the optimal sampling and to the Bayes risk incurred by the best fixed design when M is large. Specifically, we seek to determine $\Re(S_m) - \Re(O)$, where $\Re(S_m)$ is the Bayes risk of sequential sampling scheme S, and $\Re(O)$ is the Bayes risk of the optimal sampling scheme. Below, we state the theoretical results of this difference for multistage sampling schemes.

Theorem 3.3.1

The excess Bayes risk incurred by the multistage sampling schemes over the Bayes risk incurred by the optimal sampling scheme is of the order of $1/M$.

Proof of theorem (3.3.1)

We need to show that $M[\Re(S_m) - \Re(O)] \to 0$ as $M \to \infty$, where

$$\Re(S_m) = \frac{E\left[\left(\sum_{i=1}^{k} p_i\sqrt{R_i(1-R_i)}\right)^2\right]}{M + \sum_{i=1}^{k} r_i^0} + \frac{1}{\left(M + \sum_{i=1}^{k} r_i^0\right)}$$
$$E\left[\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\left(\frac{(m_i + r_i^0)p_j\sqrt{R_j(1-R_j)} - (m_j + r_j^0)p_i\sqrt{R_i(1-R_i)}}{(m_i + r_i^0)(m_j + r_j^0)}\right)^2\right],$$

and

$$\Re(O) = \frac{E\left[\left(\sum_{i=1}^{k} p_i\sqrt{R_i(1-R_i)}\right)^2\right]}{M + \sum_{i=1}^{k} r_i^0}.$$

The proof of the theorem is established if we show that

$$E\left[\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\left(\frac{(m_i + r_i^0)p_j\sqrt{R_j(1-R_j)} - (m_j + r_j^0)p_i\sqrt{R_i(1-R_i)}}{(m_i + r_i^0)(m_j + r_j^0)}\right)^2\right] \to 0$$

as $M \to \infty$.

Thus it suffices to show that in the multistage scheme

$$\frac{m_i}{M} \to \frac{p_i\sqrt{R_i(1-R_i)}}{\sum_{j=1}^{k} p_j\sqrt{R_j(1-R_j)}}$$

with probability one as $M \to \infty$. The rest of the proof is similar to Rekab and Tahir [26]. A sketch of it is as follows.

Let $U_{i,L-1} = E[\sqrt{R_i(1-R_i)}|\Im_{\bar{S}_{L-1}}]$, because $E[\sqrt{R_i(1-R_i)}]$ is finite then $U_{i,L-1}$ is a uniformly integrable martingale and $\bar{S}_{L-1} \to \infty$ as $M \to \infty$. It follows from the martingale convergence theorem that

$$U_{i,L-1} \to \sqrt{R_i(1-R_i)}$$

with probability one as $M \to \infty$. Then

$$\hat{C}_i(\bar{S}_{L-1}) \to \frac{p_i\sqrt{R_i(1-R_i)}}{\sum_{j=1}^{k} p_j\sqrt{R_j(1-R_j)}}$$

with probability one as $M \to \infty$. Divide m_i in (3.2.3) by M, take the limit as $M \to \infty$, the proof follows. ∎

Theorem 3.3.2

For large M, we have $\Re(S_m)/\Re(F_o) \leq 1$, where $\Re(S_m)$ is the Bayes risk incurred by the multistage sampling scheme, and $\Re(F_o)$ is the Bayes risk incurred by the best fixed design.

The proof is similar to the proof of theorem (2.7.2) and thus be skipped here.

3.4 Monte Carlo Simulations

We consider a real application which is the administration feature in an operations support system developed at AT&T, which dealt with the maintenance and scheduling of maintenance tasks performed by on-site work personnel [20]. To test the application, we divide the test domain into 64 sub-domains based on grouping factors identified by the requirements of the screen field, and the software development process. As shown in Table (2) on page 65, there are $4 \times 2 \times 2 \times 4 = 64$ possible combinations with each combination corresponding to a partition of the input domain. In each subdomain, we assume that these unknown reliability statistics follow a *Beta* distribution such that $R_i \sim Beta(\alpha_i, \beta_i)$, as shown in Table (3) on page 66, which also shows the usage probability p_i. The

parameters of these distributions are chosen to be small relative to the number of test cases executed, which indicates reasonable uncertainty as expressed through a high standard deviation. This type of result is by far the dominant situation in practice.

Table (4) on page 67 presents the simulation results for various total sample sizes, and the limiting value of the ratio $\Re(S_m)/\Re(F_o)$ as M becomes very large. It is obvious from the simulation results that the sequential sampling scheme outperformed the best fixed sampling scheme, even with relatively small sample sizes. Intuitively, and by Theorem (3.3.2), this result is what we would expect given that the sequential sampling schemes use information contained in the *Beta* priors, as well as information gained during the testing process.

Table (5) on page 68 shows the percentage allocated to each subdomain after the sequential sampling scheme in the case when $M = 50,000$. Subdomains D1-D10 show a uniform setting with different usage probabilities. Subsequently, D11-D30 illustrate a choice of prior parameters where the conditional reliabilities have equal means but different variances, while D31-D50 illustrate the situation when the conditional reliabilities have the same variance but different means. What is interesting to note is that the allocations among the partitions closely track with the usage probabilities, and appear to be uninfluenced by the differences in the means of the conditional reliabilities. This appearance is reasonable, however, because allocation decisions are based on the ratio of (3.1.2), which is closely linked to the ratio of the standard deviations of the conditional reliabilities.

While these results are compelling, minimizing the variance is not the only objective that a tester may pursue. Reliability testing may in fact just be used as a gauge to judge the product reaching a milestone, or its fitness for release. If

system reliability is clearly insufficient in these circumstances, knowing accurately how unreliable the system is may be useless. For circumstances where reliability is close to meaningful levels, however, the methods presented here show significant promise.

CHAPTER 4

AN EFFICIENT TEST ALLOCATION FOR SOFTWARE RELIABILITY ESTIMATION

In this section we will propose an efficient sampling scheme in the software reliability estimation. Unlike fully sequential sampling schemes where individual test cases must be executed one by one and multistage sampling schemes where test cases are executed in batches, an accelerated sampling scheme is less time-consuming than the fully sequential sampling scheme but achieves more accuracy than the multistage sampling scheme. The superiority of the accelerated sampling scheme over best fixed sampling scheme is presented theoretically. Meanwhile, comparisons between the accelerated sampling scheme and other sampling schemes such as balanced design, best fixed design and multistage sampling scheme will be drawn through Monte Carlo simulations to demonstrate its merit. A real case study is presented as well to better illustrate the proposed method.

4.1 Sequential Allocation

Recall section (2.5), $\Re(P)$ is bounded below by

$$\Re(P) \geq \frac{E\left[\left(\sum_{i=1}^{k} p_i\sqrt{R_i(1-R_i)}\right)^2\right]}{M+\sum_{i=1}^{k} r_i^0}$$

with equality if

$$\frac{m_i + r_i^0}{m_j + r_j^0} = \frac{p_i\sqrt{R_i(1-R_i)}}{p_j\sqrt{R_j(1-R_j)}} \tag{4.1.1}$$

for each i, j. Since the actual conditional reliability R_i for each subdomain is unknown, the optimal sampling scheme is not practical. Instead, we estimate its value at intervals during the testing process, and choose which subdomain to sample from to adjust the ratio such that

$$\frac{m_{i,t} + r_i^0}{m_{j,t} + r_j^0}$$

is close to $\hat{C}_{i,j}(t)$, where

$$\hat{C}_{i,j}(t) = \frac{p_i}{p_j} \frac{E[\sqrt{R_i(1-R_i)}|\mathscr{F}_t]}{E[\sqrt{R_j(1-R_j)}|\mathscr{F}_t]}. \tag{4.1.2}$$

Or equivalently,

$$\frac{m_{i,t} + r_i^0}{M + \sum_{j=1}^{k} r_j^0}$$

is close to $\hat{C}_i(t)$, where

$$\hat{C}_i(t) = \frac{p_i E\left[\sqrt{R_i(1-R_i)}|\mathscr{F}_t\right]}{\sum_{j=1}^{k} p_j E\left[\sqrt{R_j(1-R_j)}|\mathscr{F}_t\right]}, \tag{4.1.3}$$

and $\mathscr{F}_t$ is the data collected after t test cases have been executed.

4.2 Accelerated Sampling Schemes for Estimating Reliability

Below we discuss the accelerated sampling approach. The method is particularly appropriate in the case of automated testing, where test cases may be generated randomly. If the application is partitioned by component for instance and one test case is defined as a random path through a particular software component from invocation to termination or entry into a particular function or com-

ponent to exit. In these cases, time for generation is relatively small [41]. With the goal of reliability estimation, the information which would be most helpful is "which component do I select for the next test case?". There are several other scenarios such as these that the methods proposed in this section are reasonable in a practical implementation sense.

In the accelerated scheme, we distribute the M test cases in L stages, which are fixed before testing begins, such that at stage j there are S_j test cases distributed during that stage. We define $S_{j,i}$ to be the *distinct* test cases allocated to partition i during stage j. We outline the accelerated sampling approach as below.

Stage 1:

We allocate test cases in the first stage according to the best-fixed scheme with S_1 total test cases where

$$\lim_{M\to\infty} \tfrac{S_1}{M} = 0 \text{ and } \lim_{M\to\infty} S_1 = \infty \text{ [23]}.$$

Thus by equation (2.3.2) we allocate test cases in the first stage such that:

$$S_{1,i} = \frac{\left(S_1 + \sum_{\ell=1}^{k} r_\ell^0\right) p_i \sqrt{\frac{\alpha_i^0 \beta_i^0}{(r_i^0)(r_i^0+1)}}}{\sum_{\ell=1}^{k} p_\ell \sqrt{\frac{\alpha_\ell^0 \beta_\ell^0}{(r_i^0)(r_i^0+1)}}} - r_i^0 \tag{4.2.1}$$

for $i = 1, 2, \ldots, k-1$ where $S_{1,i}$ is rounded to the nearest integer, and

$$S_{1,k} = S_1 - \sum_{i=1}^{k-1} S_{1,i}.$$

Stage 2 through $L-1$:

We continue with the method for stages 2 to $L-1$ and distribute tests within these stages to the k partitions based on the prior distribution updated by testing

results from previous stages and determine $S_{j,1}, S_{j,2}, \ldots, S_{j,k-1}$ sequentially such that at each stage $j, 2 \leq j \leq L-1$, for partition i, by equation (4.1.3),

$$S_{j,i} = \left(S_j + \left(\sum_{\ell=1}^{k} r_\ell^0 \right) \right) \hat{C}_i(\bar{S}_{j-1}) - r_i^0, \tag{4.2.2}$$

and

$$S_{j,k} = S_j - \sum_{\ell=1}^{k-1} S_{j,\ell},$$

where

$$\bar{S}_{j-1} = \sum_{y=1}^{j-1} S_y.$$

Note here for $j = 2, \ldots, L-1, S_j$ are chosen such that

$$\lim_{M\to\infty} \tfrac{S_j}{M} = 0 \text{ and } \lim_{M\to\infty} S_j = \infty.$$

In Stage 2 to Stage $L-1$, instead of making allocation decisions one at a time, we allocate groups of test cases among the partitions based on the previous stage results, which may become cost effective to test in parallel given the potential time savings.

Stage L:

By (4.1.2), we allocate one test from partition i if

$$\frac{m_{i,\ell} + r_i^0}{m_{j,\ell} + r_j^0} < \hat{C}_{i,j}(\ell)$$

for all $j \neq i$ where $m_{i,\ell}$ is the *cumulative* test cases allocated to partition i after a total of ℓ test cases has been allocated, where

$$\sum_{j=1}^{L-1} S_{j,i} \leq \ell \leq \sum_{j=1}^{L-1} S_{j,i} + M - \sum_{i=1}^{L-1} S_i.$$

We proceed sequentially in this manner until all the remaining S_L test cases have been allocated, where

$$S_L = M - \sum_{j=1}^{L-1} S_j.$$

At this final stage, all test cases are allocated one at a time to achieve the best accuracy for reliability estimation.

4.3 Theoretical Results

In this section the comparison between the Bayes risk incurred by the accelerated sampling scheme to the Bayes risk incurred by the optimal sampling scheme as well as the best fixed design is demonstrated in theory.

Theorem 4.3.1

The excess Bayes risk incurred by the accelerated sampling schemes over the Bayes risk incurred by the optimal sampling scheme is of the order of $1/M$.

The proof is similar to the proof of theorem (2.7.1) and thus be skipped here.

Theorem 4.3.2

For large M, *we have* $\Re(S_a)/\Re(F_o) \leq 1$, *where* $\Re(S_a)$ *is the Bayes risk incurred by the accelerated sampling scheme, and* $\Re(F_o)$ *is the Bayes risk incurred by the best fixed design.*

The proof is similar to the proof of theorem (2.7.2) and thus be skipped here.

4.4 Monte Carlo Simulations

In this section the accelerated sampling scheme is compared to multiple sampling schemes in terms of incurred Bayes risk through Monte Carlo simulations. We consider the case where the test domain is divided into two disjoint

subdomains, D_1 and D_2, with reliability R_1 and R_2 respectively. In each subdomain we assume that these unknown reliability statistics follow a *Beta* distribution such that $R_1 \sim Beta(\alpha_1, \beta_1)$ and $R_2 \sim Beta(\alpha_2, \beta_2)$. Associated with each domain D_i is also a usage probability p_i.

Table (6) on page 69 presents results for comparison between accelerated sampling scheme and equal allocation sampling scheme, which allocates equal test cases to each domain. The results show that Bayes risk ratios under various initial parameter settings and different total sample sizes. With no surprise, even with equal usage probability accelerated sampling scheme outperforms equal allocation scheme in a significant way.

Table (7) on page 70 presents results for comparison between the accelerated sampling scheme and the best fixed sampling scheme. As shown in Table (7), even for small sample sizes, the accelerated sequential sampling scheme outperforms the best fixed sampling scheme. Intuitively, and by Theorem (4.3.2), this is what we would expect given that the sequential sampling scheme uses information contained in the *Beta* priors as well as information gained during the testing process.

In Table (8) on page 71 a comparison is drawn between accelerated sampling scheme and two-stage sampling scheme. A two-stage sampling scheme is also a multistage sampling allocation scheme with only 2 stages. As observed in Table (8), accelerated scheme outperforms two-stage scheme even for small sample sizes, which makes sense because allocating test cases one by one in the last stage really gives a better estimation accuracy in terms of Bayes risk. It is interesting to note that as M becomes very large they have the same performance, which also makes sense because as M becomes very large allocating test cases one by one does not

play a role any more. However, in practice accelerated scheme will definitely have a better accuracy then two-stage scheme.

4.5 Real Case Study

We consider a real application [19] [21] that is a data driven telephone billing system where the operations are classified by two types of service (residential or business), usage of three discount calling plans (none, national or international) and two types of payment status (paid of delinquent). The profile thus results in $K = 2 \times 3 \times 2 = 12$ operations (subdomains) with occurrence probabilities (usage probabilities) given in Table (9) on page 72. *Beta* distributions are assumed for every subdomain with different parameters. As we can see in Table (10) on page 73, it is obvious that the accelerated sampling scheme performs better than the best fixed sampling scheme in every setting.

CHAPTER 5

TEST ALLOCATION FOR ESTIMATING RELIABILITY OF SERIES SYSTEMS WITH TWO COMPONENTS

In the previous chapters, we are dealing with parallel systems where the reliability is the weighted sum of reliabilities of each component (subdomain), as in (1.0.1). The reliabillity of a series system, however, is defined as the product of reliabilities of each component. In this chapter, we adopted a Bayesian approach to estimate the reliability of a series system with two components under the squared error loss. We derived an asymptotic second-order lower bound for the Bayes risk of a sequential procedure that allocates M test cases to one component and $t-M$ to the other component, where M is determined according to a sequential design and t denotes the total number of test cases.

5.1 Introduction

Let $\mathcal{P}_1, \mathcal{P}_2$ be two components in a series system with associated values θ and ω, respectively, where θ and ω is the conditional reliability of a use case or a test case [22], on condition that it was randomly chosen within two components. Within each component, each test case has an equal chance to be selected. Note that θ and ω are unknown parameters [14]. Then the reliability of the series system with two components is defined as

$$R = \theta \cdot \omega. \tag{5.1.1}$$

The problem is to distribute t test cases into two components and use the

results to estimate θ and ω. Specifically, sample sizes M and N are taken from components $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. Rekab [25] has studied this problem and proposed a fully sequential sampling scheme that is nearly optimal and performs better than the balanced design. However, the fully sequential sampling scheme only reaches the first-order optimality. It is desired to derive an asymptotic second-order lower bound for the Bayes risk of a sequential design.

5.2 The Bayesian Model

We model the outcome of the jth test taken from the first component as a Bernoulli random variable X_j such that

$$X_j = \begin{cases} 1, \text{ if test } j \text{ taken from } \mathcal{P}_1 \text{ is processed correctly.} \\ 0, \text{ otherwise,} \end{cases}$$

where X_j has a Bernoulli distribution with parameter θ.
Similarly, the outcome of the jth test taken from the second component is modeled as a Bernoulli random variable Y_j such that

$$Y_j = \begin{cases} 1, \text{ if test } j \text{ taken from } \mathcal{P}_2 \text{ is processed correctly.} \\ 0, \text{ otherwise,} \end{cases}$$

where Y_j has a Bernoulli distribution with parameter ω.

Within each component, we assume that the corresponding conditional reliability follows a *Beta* prior distribution such that $\theta \sim Beta(a, b)$ and $\omega \sim Beta(c, d)$. Then the posterior density of θ when M observations are sampled from population $\mathcal{P}_1$ is

$$d\Pi_M(\theta) = \frac{\prod_{i=1}^{M} d\Pi(X_i|\theta) d\Pi(\theta)}{\int \prod_{i=1}^{M} d\Pi(X_i|\theta) d\Pi(\theta) d\theta}$$

$$= \frac{(\prod_{i=1}^{M} \theta^{X_i}(1-\theta)^{1-X_i}) \cdot \theta^{a-1}(1-\theta)^{b-1} d\theta}{c(a_M, b_M)}$$
$$= \frac{\theta^{a+\sum_{i=1}^{M} X_i - 1}(1-\theta)^{b+M-\sum_{i=1}^{M} X_i - 1} d\theta}{c(a_M, b_M)},$$

where

$$c(a_M, b_M) = \int \theta^{a+\sum_{i=1}^{M} X_i - 1}(1-\theta)^{b+M-\sum_{i=1}^{M} X_i - 1} d\theta.$$

Therefore,

$$d\Pi_M(\theta) \sim Beta(a + \sum_{i=1}^{M} X_i, b + M - \sum_{i=1}^{M} X_i). \tag{5.2.1}$$

Similarly,

$$d\Pi_N(\omega) \sim Beta(a + \sum_{i=1}^{N} Y_i, b + N - \sum_{i=1}^{N} Y_i). \tag{5.2.2}$$

Theorem 5.2.1

The following equations hold.

$$Var[\theta|X_1, \ldots, X_j] = E[\frac{\theta(1-\theta)}{a+b+j}|X_1, \ldots, X_j], \tag{5.2.3}$$

and

$$Var[\omega|Y_1, \ldots, Y_k] = E[\frac{\omega(1-\omega)}{a+b+k}|Y_1, \ldots, Y_k]. \tag{5.2.4}$$

Proof of theorem (5.2.1)

By equation (5.2.1),

$$d\Pi_j(\theta) \sim Beta(a + \sum_{i=1}^{j} X_i, b + j - \sum_{i=1}^{j} X_i),$$

Then, we have

$$Var[\theta|X_1,\ldots,X_j] = \frac{(a+\sum_{i=1}^{j} X_i)(b+j-\sum_{i=1}^{j} X_i)}{(a+b+j)^2(a+b+j+1)}.$$

Let $B(x,y)$ stands for the *Beta* function of x and y, then

$$E[\frac{\theta(1-\theta)}{a+b+j}|X_1,\ldots,X_j] =$$

$$\int \frac{\theta(1-\theta)\cdot\theta^{a+\sum_{i=1}^{j} X_i-1}(1-\theta)^{b+j-\sum_{i=1}^{j} X_i-1}}{(a+b+j)\cdot B(a+\sum_{i=1}^{j} X_i, b+j-\sum_{i=1}^{j} X_i)}d\theta =$$

$$\int \frac{\theta^{a+\sum_{i=1}^{j} X_i}(1-\theta)^{b+j-\sum_{i=1}^{j} X_i}}{(a+b+j)\cdot B(a+\sum_{i=1}^{j} X_i, b+j-\sum_{i=1}^{j} X_i)}d\theta =$$

$$\frac{B(a+\sum_{i=1}^{j} X_i+1, b+j-\sum_{i=1}^{j} X_i+1)}{(a+b+j)\cdot B(a+\sum_{i=1}^{j} X_i, b+j-\sum_{i=1}^{j} X_i)}\cdot$$

$$\int \frac{\theta^{a+\sum_{i=1}^{j} X_i+1-1}(1-\theta)^{b+j-\sum_{i=1}^{j} X_i+1-1}}{B(a+\sum_{i=1}^{j} X_i+1, b+j-\sum_{i=1}^{j} X_i+1)}d\theta =$$

$$\frac{B(a+\sum_{i=1}^{j} X_i+1, b+j-\sum_{i=1}^{j} X_i+1)}{(a+b+j)\cdot B(a+\sum_{i=1}^{j} X_i, b+j-\sum_{i=1}^{j} X_i)}\cdot 1 =$$

$$\frac{\frac{\Gamma(a+\sum_{i=1}^{j} X_i+1)\Gamma(b+j-\sum_{i=1}^{j} X_i+1)}{\Gamma(a+b+j+2)}}{\frac{(a+b+j)\Gamma(a+\sum_{i=1}^{j} X_i)\Gamma(b+j-\sum_{i=1}^{j} X_i)}{\Gamma(a+b+j)}} =$$

$$\frac{(a+\sum_{i=1}^{j} X_i)(b+j-\sum_{i=1}^{j} X_i)}{(a+b+j)^2(a+b+j+1)}.$$

This proves equation (5.2.3). A similar argument can be used to show equation (5.2.4). This concludes the proof. ∎

Our objective is to estimate the reliability of the series system $R(\theta, \omega) = \theta\omega$ with squared error loss

$$L^*(R, \hat{R}) = (R - \hat{R})^2, \tag{5.2.5}$$

where $\hat{R}$ is the Bayes estimate, i.e., the estimate that minimizes the expected posterior loss, namely

$$f(\hat{R}) = E[(R - \hat{R})^2] = E(R^2) - 2E(R\hat{R}) + 2E(\hat{R}^2)$$
$$f(\hat{R}) = \hat{R}^2 - 2E(R)\hat{R} + E(R^2)$$
$$f^{'}(\hat{R}) = 2\hat{R} - 2E(R),$$
$$\text{set } f^{'}(\hat{R}) = 0 \text{ to get } \hat{R} = E(R).$$
$$f^{''}(\hat{R}) = 2 > 0.$$

Therefore $f(\hat{R})$ is minimized at $\hat{R} = E(R)$, and the value of the minimum of $f(\hat{R})$ is

$$E(R - E(R))^2 = Var(R).$$

Let $\mathscr{F}_{m,n}$ be the sigma algebra generated by $X_1, \ldots, X_m, Y_2, \ldots, Y_n$ for $m, n = 1, 2, \ldots$. A sequential desgin Δ is a sequence of indicator variables $\Delta_1, \cdots, \Delta_t$, each of which takes the values 0 or 1 (0 if the test case is allocated to $\mathcal{P}_2$ and 1 if the test case is allocated $\mathcal{P}_1$) and such that Δ_{k+1} is $\mathscr{F}_{m_k,n_k}$-measurable for all $k = 2, \cdots, t-1$ where

$$n_k = \Delta_1 + \ldots + \Delta_k \text{ and } m_k = k - n_k$$

for $k = 1, \ldots, t$.

Theorem 5.2.2

Let $m_t = M$ and $n_t = N$. If $M, N \geq 1$, then the posterior expected loss, given

$\mathscr{F}_{M,N}$ *is*

$$L^*(M,N) = \frac{A}{M+r} + \frac{B}{N+s} - \frac{C}{(M+r)(N+s)}, \tag{5.2.6}$$

where $r = a+b, s = c+d$ *and*

$$\begin{aligned} A &= E_t[\theta(1-\theta)\omega^2], \\ B &= E_t[\omega(1-\omega)\theta^2], \\ C &= E_t[\theta(1-\theta)\omega(1-\omega)]. \end{aligned} \tag{5.2.7}$$

Here E_t *represents the posterior expectation given the sigma algebra* $\mathscr{F}_{M,N}$.

Proof of theorem (5.2.2)

Throughout the proof, we will skip the phrase given $\mathscr{F}_{M,N}$ in all the expressions involving E_t and Var_t:

$$\begin{aligned} E_t(R-\hat{R})^2 &= Var_t(R) \\ Var_t(R) &= E_t(\theta\omega)^2 - E_t^2(\theta\omega), \end{aligned}$$

by definition of conditional variance.

$$Var_t(R) = E_t(\theta^2\omega^2) - E_t^2(\theta)E_t^2(\omega),$$

by independence. Using the fact:

$$Var_t(\theta) = E_t(\theta^2) - E_t^2(\theta),$$

the latter equality can be rewritten as

$$Var_t(R) = E_t(\theta^2)E_t(\omega^2) - [E_t(\theta^2) - Var_t(\theta)][E_t(\omega^2) - Var_t(\omega)].$$

Multiplying the expressions inside the brackets gives:

$$\begin{aligned} Var_t(R) &= E_t(\theta^2)E_t(\omega^2) - E_t(\theta^2)E_t(\omega^2) \\ &+E_t(\theta^2)Var_t(\omega) + E_t(\omega^2)Var_t(\theta) - Var_t(\theta)Var_t(\omega). \end{aligned}$$

In the latter equality, the first two terms cancel. By theorem (5.2.1),

$$Var_t(\theta) = \frac{1}{M+r} E_t(\theta(1-\theta)),$$

and

$$Var_t(\omega) = \frac{1}{N+s} E_t(\omega(1-\omega)).$$

We write the latter equality as

$$\begin{aligned}
Var_t(R) &= E_t(\theta^2)\frac{1}{N+s}E_t(\omega(1-\omega)) + E_t(\omega^2)\frac{1}{M+r}E_t(\theta(1-\theta)) \\
&\quad -\frac{1}{M+r}E_t(\theta(1-\theta))\frac{1}{N+s}E_t(\omega(1-\omega)) \\
&= \frac{1}{N+s}E_t(\theta^2)E_t(\omega(1-\omega)) + \frac{1}{M+r}E_t(\omega^2)E_t(\theta(1-\theta)) \\
&\quad -\frac{1}{M+r}\frac{1}{N+s}E_t(\theta(1-\theta))E_t(\omega(1-\omega)) \\
&= \frac{1}{N+s}E_t(\theta^2\omega(1-\omega)) + \frac{1}{M+r}E_t(\omega^2\theta(1-\theta)) \\
&\quad -\frac{1}{M+r}\frac{1}{N+s}E_t(\theta(1-\theta)\omega(1-\omega)) \\
&= \frac{A}{M+r} + \frac{B}{N+s} - \frac{C}{(M+r)(N+s)}.
\end{aligned}$$

This concludes the proof. ∎

5.3 Second Order Lower Bound

Recall the Bayes risk incurred in the sequential design is

$$\Re(\Delta) = E_{\theta,\omega}[L^*(M,N)] = E_{\theta,\omega}\{\frac{A}{M+r} + \frac{B}{N+s} - \frac{C}{(M+r)(N+s)}\}, \quad (5.3.1)$$

where A, B and C are of the form in (5.2.7).

Theorem 5.3.1

Suppose that

$$a > 0, b > 1 \text{ and } c > 0, d > 1. \quad (5.3.2)$$

Then for any sequential design Δ

$$\Re(\Delta) \geq \frac{1}{t+r+s} E_{\theta,\omega}[(\omega\sqrt{\theta(1-\theta)} + \theta\sqrt{\omega(1-\omega)})^2] +$$
$$\frac{1}{(t+r+s)^2} E_{\theta,\omega}\left[\frac{(3-4\theta)^2\sqrt{\omega(1-\omega)}}{4(1-\theta)}(\omega\sqrt{\theta(1-\theta)} + \theta\sqrt{\omega(1-\omega)}) + \right.$$
$$\left.\frac{(3-4\omega)^2\sqrt{\theta(1-\theta)}}{4(1-\omega)}(\omega\sqrt{\theta(1-\theta)} + \theta\sqrt{\omega(1-\omega)})\right] - \quad (5.3.3)$$
$$\frac{1}{(t+r+s)^2} E_{\theta,\omega}\left[\theta(1-\theta)\omega(1-\omega)\left(2 + \frac{\omega\sqrt{\theta(1-\theta)}}{\theta\sqrt{\omega(1-\omega)}} + \frac{\theta\sqrt{\omega(1-\omega)}}{\omega\sqrt{\theta(1-\theta)}}\right)\right]$$
$$+o(1/t^2)$$

where $r = a + b$ *and* $s = c + d$.

To prove theorem (5.3.1), we need to break up the Bayes risk into two parts as follows.

$$\begin{aligned}\Re(\Delta) &= E_{\theta,\omega}\left\{\frac{A}{M+r} + \frac{B}{N+s} - \frac{C}{(M+r)(N+s)}\right\} \\ &= E_{\theta,\omega}\left\{\frac{A}{M+r} + \frac{B}{N+s}\} - E\{\frac{C}{(M+r)(N+s)}\right\} \\ &= \Re_1 - \Re_2\end{aligned}$$

where

$$\Re_1 = E_{\theta,\omega}\left\{\frac{A}{M+r} + \frac{B}{N+s}\right\}, \quad (5.3.4)$$

and

$$\Re_2 = E_{\theta,\omega}\left\{\frac{C}{(M+r)(N+s)}\right\}, \quad (5.3.5)$$

where A, B and C are of the form in (5.2.7).

The following propositions are needed for the proof of theorem (5.3.1).

Proposition 5.3.1.1

$$\begin{aligned}\Re_1 \geq \frac{1}{t+r+s}E_{\theta,\omega}[(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})^2]+\\ \frac{1}{(t+r+s)^2}E_{\theta,\omega}\left[\frac{(3-4\theta)^2\sqrt{\omega(1-\omega)}}{4(1-\theta)}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})+\right.\\ \left.\frac{(3-4\omega)^2\sqrt{\theta(1-\theta)}}{4(1-\omega)}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})\right]+o(1/t^2).\end{aligned} \tag{5.3.6}$$

The following lemmas are required for the proof of the proposition (5.3.1.1).

Lemma 5.3.1.1

Given α, let $X_1,\ldots,X_k$ be i.i.d. random variables that follow a distribution in the form of one-parameter exponential family, that is

$$dF_\alpha(x)=\exp\{\alpha x-\psi(\alpha)\}dx, \tag{5.3.7}$$

where α has the conjugate prior distribution in the form of

$$d\Pi(\alpha)=\frac{\exp\{\lambda\mu\alpha-\lambda\psi(\alpha)\}}{c(\lambda,\mu)}d\alpha. \tag{5.3.8}$$

Let $g(\alpha)=\psi^{'}(\alpha)\sqrt{\psi^{''}(\alpha)}$, if

$$\int\frac{[g^{'}(\alpha)]^2}{\psi^{''}(\alpha)}d\Pi(\alpha)<\infty, \tag{5.3.9}$$

then

$$Var\,[g(\alpha)|\mathscr{F}_k]\leq E\left[\frac{[g^{'}(\alpha)]^2}{(k+\lambda)\psi^{''}(\alpha)}|\mathscr{F}_k\right] \tag{5.3.10}$$

where $\mathscr{F}_k$ is the sigma algebra generated by $X_1,\ldots,X_k$.

The proof of lemma (5.3.1.1) is established directly from the proof of lemma 1 in [27] by changing $g(\alpha)$ from $\sqrt{\psi''(\alpha)}$ to $\psi'(\alpha)\sqrt{\psi''(\alpha)}$.

Lemma 5.3.1.2

Suppose condition (5.3.2) holds, then the following inequalities hold.

$$Var_t\left[\theta\sqrt{\theta(1-\theta)}\right] \leq E_t\left[\frac{\theta^2(3-4\theta)^2}{4(M+r)}\right]$$

and

$$Var_t\left[\omega\sqrt{\omega(1-\omega)}\right] \leq E_t\left[\frac{\omega^2(3-4\omega)^2}{4(N+s)}\right].$$

Proof of lemma (5.3.1.2)

Since $X_1, \ldots, X_M$ are i.i.d. Bernoulli distributed with parameter θ and that θ has Beta distribution that is the conjugate prior of Bernoulli distribution, it follows that each X_i has the distribution in the form of (5.3.7) with

$$\alpha = \log\frac{p}{1-p},$$
$$\psi(\alpha) = \log(1+e^{\theta})$$

and that α has the conjugate prior distribution in the form of (5.3.8). Since

$$g(\alpha) = \psi'(\alpha)\sqrt{\psi''(\alpha)} = \theta^{\frac{3}{2}}(1-\theta)^{\frac{1}{2}}$$

and

$$g'(\alpha) = \tfrac{1}{2}\theta^{\frac{3}{2}}(3-4\theta)(1-\theta)^{\frac{1}{2}},$$

then

$$\int \frac{[g'(\alpha)]^2}{\psi''(\alpha)} d\Pi(\alpha) = \int \frac{\theta(3-4\theta)^2}{4(1-\theta)} d\theta$$
$$= \int \frac{9\theta}{4(1-\theta)} - \frac{6\theta^2}{1-\theta} + \frac{4\theta^3}{1-\theta} d\theta$$
$$= \frac{9B(a+1,b-1)}{4B(a,b)} - \frac{4B(a+2,b-1)}{B(a,b)} + \frac{4B(a+3,b-1)}{B(a,b)}.$$

By condition (5.3.2)

$$\int \frac{[g^{'}(\alpha)]^2}{\psi^{''}(\alpha)} d\Pi(\alpha) < \infty.$$

Therefore, by lemma (5.3.1.1), the first inequality follows. A similar argument can be used to show the second inequality. This concludes the proof. ∎

Lemma 5.3.1.3

Let $Z_t = Var_t\{\theta^{3/2}(1-\theta)^{1/2}\omega^{3/2}(1-\omega)^{1/2}\}$,

and suppose that condition (5.3.2) holds, then

$$Z_t \leq E_t\{\frac{\theta^2(3-4\theta)^2\omega^3(1-\omega)}{4(M+r)} + \frac{\omega^2(3-4\omega)^2\theta^3(1-\theta)}{4(N+s)}\}.$$

Proof of lemma (5.3.1.3) $Z_t = Var_t\{\theta^{3/2}(1-\theta)^{1/2}\omega^{3/2}(1-\omega)^{1/2}\}$

$$= E_t\{\theta^3(1-\theta)\omega^3(1-\omega)\} - E_t\{\theta^{3/2}(1-\theta)^{1/2}\omega^{3/2}(1-\omega)^{1/2}\}^2$$

$$= E_t\{\theta^3(1-\theta)\}E_t\{\omega^3(1-\omega)\}$$

$$-E_t\{\theta^{3/2}(1-\theta)^{1/2}\}^2 E_t\{\omega^{3/2}(1-\omega)^{1/2}\}^2$$

$$= (E_t\{\theta^3(1-\theta)\} - E_t\{\theta^{3/2}(1-\theta)^{1/2}\}^2)E_t\{\omega^3(1-\omega)\}$$

$$+(E_t\{\omega^3(1-\omega)\} - E_t\{\omega^{3/2}(1-\omega)^{1/2}\}^2)E_t\{\theta^{3/2}(1-\theta)^{1/2}\}^2$$

$$= Var_t\{\theta^{3/2}(1-\theta)^{1/2}\}E_t\{\omega^3(1-\omega)\}$$

$$+Var_t\{\omega^{3/2}(1-\omega)^{1/2}\}E_t\{\theta^{3/2}(1-\theta)^{1/2}\}^2$$

$$\leq Var_t\{\theta^{3/2}(1-\theta)^{1/2}\}E_t\{\omega^3(1-\omega)\}$$

$$+Var_t\{\omega^{3/2}(1-\omega)^{1/2}\}E_t\{\theta^3(1-\theta)\}.$$

By Lemma (5.3.1.2) the proof follows. ∎

Lemma 5.3.1.4

Suppose that condition (5.3.2) holds, and

$$\frac{M+r}{t+r+s} \to \frac{\omega\sqrt{\theta(1-\theta)}}{\omega\sqrt{\theta(1-\theta)} + \theta\sqrt{\omega(1-\omega)}} \tag{5.3.11}$$

with probability one, then

$$(t+r+s)Z_t \to \frac{\theta^{3/2}(3-4\theta)^2\omega^2(1-\omega)}{4\sqrt{1-\theta}}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})+$$
$$\frac{\omega^{3/2}(3-4\omega)^2\theta^2(1-\theta)}{4\sqrt{1-\omega}}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})$$

with probability one as $t \to +\infty$.

Proof of lemma (5.3.1.4)

By lemma (5.3.1.3), it follows that

$$limsup(t+r+s)Z_t \leq$$
$$limsup E_t\{\frac{t+r+s}{M+r}\cdot\frac{1}{4}\theta^2(3-4\theta)^2\omega^3(1-\omega)+\frac{t+r+s}{N+s}\cdot\frac{1}{4}\omega^2(3-4\omega)^2\theta^3(1-\theta)\}.$$

Now, observe that

$$E_t\{\frac{t+r+s}{M+r}\cdot\frac{1}{4}\theta^2(3-4\theta)^2\omega^3(1-\omega)+\frac{t+r+s}{N+s}\cdot\frac{1}{4}\omega^2(3-4\omega)^2\theta^3(1-\theta)\}$$
$$\to \frac{\theta^{3/2}(3-4\theta)^2\omega^2(1-\omega)}{4\sqrt{1-\theta}}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})+$$
$$\frac{\omega^{3/2}(3-4\omega)^2\theta^2(1-\theta)}{4\sqrt{1-\omega}}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})$$

with probability one as $t \to +\infty$, by the martingale convergence theorem. So

$$limsup(t+r+s)Z_t \leq \frac{\theta^{3/2}(3-4\theta)^2\omega^2(1-\omega)}{4\sqrt{1-\theta}}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})+$$
$$\frac{\omega^{3/2}(3-4\omega)^2\theta^2(1-\theta)}{4\sqrt{1-\omega}}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)}).$$

The reverse inequality will follow by showing that

$$liminf(t+r+s)Var_t\{\theta^{3/2}(1-\theta)^{1/2}\}E_t\{\omega^3(1-\omega)\} \geq$$
$$\frac{\theta^{3/2}(3-4\theta)^2\omega^2(1-\omega)}{4\sqrt{1-\theta}}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)}) \tag{5.3.12}$$

and

$$liminf(t+r+s)Var_t\{\omega^{3/2}(1-\omega)^{1/2}\}E_t\{\theta^{3/2}(1-\theta)^{1/2}\}^2 \geq \frac{\omega^{3/2}(3-4\omega)^2\theta^2(1-\theta)}{4\sqrt{1-\omega}}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)}). \tag{5.3.13}$$

Observe that the left hand side of (5.3.12) can be written as

$$\tfrac{(t+r+s)}{M+r}(M+r)Var_t\{\theta^{3/2}(1-\theta)^{1/2}\}E_t\{\omega^3(1-\omega)\}.$$

Now recall the exponential family notation in lemma (5.3.10),

$$Var_t\{\theta^{3/2}(1-\theta)^{1/2}\} = Var_t\{g(\alpha)\}$$

where $g(\alpha) = \psi^{'}(\alpha)\sqrt{\psi^{''}(\alpha)}$ and $\psi(\alpha) = \log(1+e^{\theta})$. By using the fact that the posterior distribution of α is asymptotically normal with mean $\hat{\alpha}_m$ and variance $1/((M+r)\psi^{''}(\hat{\alpha}_m))$ [5] and the fact that

$$Var[f(X)] \approx f^{'}(E[X])^2 Var[X],$$

we have

$$Var_t\{g(\alpha)\} \approx \frac{g^{'}(\hat{\alpha}_m)^2}{(M+r)\psi^{''}(\hat{\alpha}_m)}.$$

Now, by conditions (5.3.2) and (5.3.11) and Fatou's Lemma, we have

$$\begin{aligned}\liminf \frac{(t+r+s)}{M+r}(M+r)Var_t\{g(\alpha)\}E_t\{g^2(\beta)\} \geq \\ \frac{\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)}}{\omega\sqrt{\theta(1-\theta)}} \cdot \tfrac{g^{'}(\alpha)^2}{\psi^{''}(\alpha)} \cdot \omega^3(1-\omega) \\ = \frac{\theta^{3/2}(3-4\theta)^2\omega^2(1-\omega)}{4\sqrt{1-\theta}}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)}).\end{aligned}$$

Therefore we proved the inequality (5.3.12). A similar argument can be used to show that the inequality (5.3.13) holds. This concludes the proof of the lemma (5.3.1.4). ∎

Now we can prove the proposition (5.3.1.1). Note that the proof is similar to [27] in spirit although we are estimating product of means instead of sum of means.

Proof of Proposition (5.3.1.1)

Recall $\Re_1$ in equation (5.3.4),

$$\begin{aligned}\Re_1 &= E_{\theta,\omega}[\frac{A}{M+r}+\frac{B}{N+s}]\\ &= E_{\theta,\omega}\{\frac{(\sqrt{A}+\sqrt{B})^2}{t+r+s}+\frac{[(N+s)\sqrt{A}-(M+r)\sqrt{B}]^2}{(t+r+s)(M+r)(N+s)}\}\\ &\geq E_{\theta,\omega}[\frac{(\sqrt{A}+\sqrt{B})^2}{t+r+s}] = E_{\theta,\omega}\{\frac{A+B+2\sqrt{AB}}{t+r+s}\}\\ &= E_{\theta,\omega}[\frac{A+B}{t+r+s}]+2E_{\theta,\omega}[\frac{\sqrt{AB}}{t+r+s}]\end{aligned}$$

where $A = E_t[\omega^2\theta(1-\theta)]$ and $B = E_t[\theta^2\omega(1-\omega)]$.

Let $W_t = \sqrt{AB} = E_t[\theta\sqrt{\theta(1-\theta)}\omega\sqrt{\omega(1-\omega)}]$, therefore

$$Z_t = AB - W_t^2$$

where Z_t is defined in lemma (5.3.1.3).

Notice that W_t is uniformly integrable martingale, then by martingale convergence theorem,

$$E_{\theta,\omega}[W_t] = W_o = \int \theta\sqrt{\theta(1-\theta)}\cdot\omega\sqrt{\omega(1-\omega)}d\Pi_{\theta,\omega}.$$

Note that

$$\sqrt{AB} = W_t + \frac{Z_t}{\sqrt{AB}+W_t},$$

so

$$E_{\theta,\omega}[\sqrt{AB}] = W_o + E_{\theta,\omega}[\frac{Z_t}{\sqrt{AB}+W_t}].$$

Then by Lemma (5.3.1.4), it follows that

$$\frac{(t+r+s)Z_t}{\sqrt{AB}+W_t} \to \frac{1}{2}\frac{(3-4\theta)^2\sqrt{\omega(1-\omega)}}{4(1-\theta)}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})$$
$$+\frac{1}{2}\frac{(3-4\omega)^2\sqrt{\theta(1-\theta)}}{4(1-\omega)}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})$$

with probability one as $t \to +\infty$. Then by Fatou's Lemma, we have

$$liminf E_{\theta,\omega}\{(t+r+s)\frac{Z_t}{\sqrt{AB}+W_t}\} \geq$$
$$E_{\theta,\omega}\{\frac{(3-4\theta)^2\sqrt{\omega(1-\omega)}}{8(1-\theta)}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})$$
$$+\frac{(3-4\omega)^2\sqrt{\theta(1-\theta)}}{8(1-\omega)}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})\}.$$

Therefore

$$\Re_1 = E_{\theta,\omega}[\frac{A+B}{t+r+s}] + 2E_{\theta,\omega}[\frac{\sqrt{AB}}{t+r+s}]$$
$$\geq E_{\theta,\omega}[\frac{A+B}{t+r+s}] + \frac{2W_o}{t+r+s}$$
$$+\frac{1}{(t+r+s)^2}\left\{\frac{(3-4\theta)^2\sqrt{\omega(1-\omega)}}{4(1-\theta)}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})\right.$$
$$\left.+\frac{(3-4\omega)^2\sqrt{\theta(1-\theta)}}{4(1-\omega)}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})\right\}$$
$$= \frac{1}{t+r+s}E_{\theta,\omega}[(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})^2]$$
$$+\frac{1}{(t+r+s)^2}\left\{\frac{(3-4\theta)^2\sqrt{\omega(1-\omega)}}{4(1-\theta)}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})\right.$$
$$\left.+\frac{(3-4\omega)^2\sqrt{\theta(1-\theta)}}{4(1-\omega)}(\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)})\right\}.$$

This concludes the proof of proposition (5.3.1.1). ∎

Proposition 5.3.1.2

Suppose the condition (5.3.11) holds

$$\Re_2 \to E_{\theta,\omega}\left\{\frac{\theta(1-\theta)\omega(1-\omega)}{(t+r+s)^2}\left(2+\frac{\omega\sqrt{\theta(1-\theta)}}{\theta\sqrt{\omega(1-\omega)}}+\frac{\theta\sqrt{\omega(1-\omega)}}{\omega\sqrt{\theta(1-\theta)}}\right)\right\} \quad (5.3.14)$$

with probability one.

Proof of proposition (5.3.1.2)

Recall $\Re_2$ in equation (5.3.5),

$$\Re_2 = E_{\theta,\omega}\{\frac{C}{(M+r)(N+s)}\} = E_{\theta,\omega}\{\frac{E_t[\theta(1-\theta)\omega(1-\omega)]}{(M+r)(N+s)}\}$$
$$= \frac{1}{(t+r+s)^2}E_{\theta,\omega}\{E_t[\frac{t+r+s}{(M+r)}\frac{t+r+s}{(N+s)}\theta(1-\theta)\omega(1-\omega)]\}.$$

Then by condition (5.3.11),

$$\frac{t+r+s}{(M+r)}\frac{t+r+s}{(N+s)}\theta(1-\theta)\omega(1-\omega) \to \theta(1-\theta)\omega(1-\omega)\cdot$$
$$\frac{\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)}}{\omega\sqrt{\theta(1-\theta)}}\cdot\frac{\omega\sqrt{\theta(1-\theta)}+\theta\sqrt{\omega(1-\omega)}}{\theta\sqrt{\omega(1-\omega)}}$$

so that

$$\frac{t+r+s}{(M+r)}\frac{t+r+s}{(N+s)}\theta(1-\theta)\omega(1-\omega) \to$$
$$\frac{1}{(t+r+s)^2}\theta(1-\theta)\omega(1-\omega)\left(2+\frac{\omega\sqrt{\theta(1-\theta)}}{\theta\sqrt{\omega(1-\omega)}}+\frac{\theta\sqrt{\omega(1-\omega)}}{\omega\sqrt{\theta(1-\theta)}}\right)$$

almost surely.

The proof will follow if we establish that $\frac{t^2}{MN}\theta(1-\theta)\omega(1-\omega)$ is uniformly integrable. To show uniform integrability, we must show that there exsits $p>1$ such that

$$\sup_t E_{\theta,\omega}[\frac{t^2}{MN}\theta(1-\theta)\omega(1-\omega)]^p < \infty.$$

Let

$$A_t = \{min(M,N) < \epsilon t\}.$$

Then

$$E[\frac{t^2}{MN}\theta(1-\theta)\omega(1-\omega)]^p \leq \frac{1}{\epsilon^{2p}}E[\theta(1-\theta)\omega(1-\omega)]^p$$
$$+(2t)^pE[\theta(1-\theta)\omega(1-\omega)]^pI_{AN}.$$

This completes the proof of the proposition (5.3.1.2). ∎

Now we can prove our second-order lower bound for the Bayes risk incurred by the sequential design under the squared loss.

Proof of Theorem (5.3.1)

Since $\Re(\Delta) = \Re_1 - \Re_2$,

by proposition (5.3.1.1) and (5.3.1.2), theorem (5.3.1) is proved. ∎

5.4 Sequential Allocation

Recall equation (5.3.3), the second order lower bound for $\Re(\Delta)$ is

$$\begin{aligned}\Re(\Delta) \geq \frac{1}{t+r+s} E_{\theta,\omega}[(\omega\sqrt{\theta(1-\theta)} + \theta\sqrt{\omega(1-\omega)})^2] + \\ \frac{1}{(t+r+s)^2} E_{\theta,\omega}\left[\frac{(3-4\theta)^2\sqrt{\omega(1-\omega)}}{4(1-\theta)}(\omega\sqrt{\theta(1-\theta)} + \theta\sqrt{\omega(1-\omega)}) + \right. \\ \left. \frac{(3-4\omega)^2\sqrt{\theta(1-\theta)}}{4(1-\omega)}(\omega\sqrt{\theta(1-\theta)} + \theta\sqrt{\omega(1-\omega)})\right] - \\ \frac{1}{(t+r+s)^2} E_{\theta,\omega}\left[\theta(1-\theta)\omega(1-\omega)\left(2 + \frac{\omega\sqrt{\theta(1-\theta)}}{\theta\sqrt{\omega(1-\omega)}} + \frac{\theta\sqrt{\omega(1-\omega)}}{\omega\sqrt{\theta(1-\theta)}}\right)\right] \\ + o(1/t^2)\end{aligned}$$

with the condition in equation (5.3.11)

$$\frac{M+r}{t+r+s} \to \frac{\omega\sqrt{\theta(1-\theta)}}{\omega\sqrt{\theta(1-\theta)} + \theta\sqrt{\omega(1-\omega)}}$$

with probability one. Since the actual value of θ and ω is unknown, sequential allocation using this condition is not practical. Instead, we estimate the ratio at intervals during the test allocation process, and choose which component to sample from to adjust the ratio such that

$$\frac{M+r}{t+r+s}$$

is close to $\hat{C}(t)$, where

$$\hat{C}(t) = \frac{E[\omega\sqrt{\theta(1-\theta)}|\mathscr{F}_t]}{E[\omega\sqrt{\theta(1-\theta)} + \theta\sqrt{\omega(1-\omega)}|\mathscr{F}_t]}, \tag{5.4.1}$$

and $\mathscr{F}_t$ is the data collected after t test cases have been allocated.

5.5 Fully Sequential Sampling Scheme

In this section we present a fully sequential sampling scheme where individual test case must be allocated in series because allocation decisions are made after each test case is executed. We proceed to test as follows:

Step 1:

Allocate one test case to each component $\mathcal{P}_1$ and $\mathcal{P}_2$.

Step 2:

After ℓ test cases have been allocated, where $\ell \geq 2$, we allocate test case $\ell + 1$ to component $\mathcal{P}_i$ if

$$\frac{m_{i,\ell} + r}{t + r + s} < \hat{C}(\ell),$$

where $m_{i,\ell}$ is the *cumulative* test cases allocated to component $\mathcal{P}_i$.

This approach is applied sequentially until all t test cases are allocated.

5.6 Multistage Sampling Scheme

In this section the multistage scheme is presented where we allocate t test cases in L stages, which are fixed before testing begins, such that at stage j there are S_j test cases distributed during that stage. We define $S_{j,i}$ to be the *cumulative* test cases allocated to component $\mathcal{P}_i$ up to stage j. Here is an outline of the

process.

Stage 1:

We allocate test cases in the first stage with S_1 total test cases such that

$$\lim_{t\to\infty} \frac{S_1}{t} = 0 \text{ and } \lim_{t\to\infty} S_1 = \infty \text{ [23]}.$$

Thus by equation (5.4.1) we allocate test cases in the first stage such that:

$$S_{1,1} = (S_1 + r + s)E\left[\frac{\omega\sqrt{\theta(1-\theta)}}{\omega\sqrt{\theta(1-\theta)} + \theta\sqrt{\omega(1-\omega)}}\right] - r \tag{5.6.1}$$

where the expectation is calculated based on the prior parameters and $S_{1,1}$ is rounded to the nearest integer, and

$$S_{1,2} = S_1 - S_{1,1}.$$

Stage 2 through L:

For Stages 2 through L, we allocate test cases to two components based on the prior distribution updated by test results from previous stages, and determine $S_{j,1}, S_{j,2}$ sequentially such that, at each stage j, $2 \le j \le L$,

$$S_{j,1} = \left(\left(\sum_{\ell=1}^{j} S_\ell\right) + r + s\right)\hat{C}_i(\bar{S}_{j-1}) - r, \tag{5.6.2}$$

and

$$S_{j,2} = \sum_{\ell=1}^{j} S_\ell - S_{j,1},$$

where

$$\bar{S}_{j-1} = \sum_{y=1}^{j-1} S_y.$$

Therefore, at stage L, the total number of test cases allocated to component P_1 is

$$M = \min\{t - S_{L-1,2}, \max((t + r + s)\hat{C}_i(\bar{S}_{L-1}) - r, S_{L-1,i})\}. \tag{5.6.3}$$

Note that (5.6.3) means M should be at least the total number of test cases allocated up to stage $L-1$, i.e., $S_{L-1,1}$. In addition, M has to be at most $M-S_{L-1,2}$ because there are only this many test cases left. However, in the last stage, we are still allocate test cases sequentially, which leads to the max part, and thus (5.6.3).

5.7 Accelerated Sampling Scheme

In this section the accelerated sampling scheme is proposed where we allocate t test cases in L stages, which are fixed before testing begins, such that at stage j there are S_j test cases distributed during that stage. We define $S_{j,1}$ and $S_{j,2}$ to be the *distinct* test cases allocated to component $\mathcal{P}_1$ and $\mathcal{P}_2$ during stage j. We outline the accelerated sampling approach as below.

Stage 1:

We allocate test cases in the first stage with S_1 total test cases such that

$$\lim_{t\to\infty} \tfrac{S_1}{t} = 0 \text{ and } \lim_{t\to\infty} S_1 = \infty \text{ [23]}.$$

Thus by equation (5.4.1) we allocate test cases in the first stage such that:

$$S_{1,1} = (S_1 + r + s)E\left[\frac{\omega\sqrt{\theta(1-\theta)}}{\omega\sqrt{\theta(1-\theta)} + \theta\sqrt{\omega(1-\omega)}}\right] - r \tag{5.7.1}$$

where the expectation is calculated based on the prior parameters and $S_{1,1}$ is rounded to the nearest integer, and

$$S_{1,2} = S_1 - S_{1,1}.$$

Stage 2 through $L-1$:

We continue with the method for stages 2 to $L-1$ and allocate test cases within these stages to two components based on the prior distribution updated by observations from previous stages and determine $S_{j,1}, S_{j,2}$ sequentially such that at each stage $j, 2 \leq j \leq L-1$, by equation (5.4.1),

$$S_{j,1} = \left(\left(\sum_{\ell=1}^{j} S_\ell \right) + r + s \right) \hat{C}_i(\bar{S}_{j-1}) - r, \tag{5.7.2}$$

and

$$S_{j,2} = S_j - S_{j,1},$$

where

$$\bar{S}_{j-1} = \sum_{y=1}^{j-1} S_y.$$

Note here for $j = 2, \ldots, L-1, S_j$ are chosen such that

$$\lim_{t\to\infty} \tfrac{S_j}{t} = 0 \text{ and } \lim_{t\to\infty} S_j = \infty.$$

Stage L:

By (5.4.1), we allocate one test case to component $\mathcal{P}_i$ if

$$\frac{m_{i,\ell} + r}{t + r + s} < \hat{C}(\ell)$$

for $i = 1, 2$ where $m_{i,\ell}$ is the *cumulative* test cases allocated to component $\mathcal{P}_i$ after ℓ total test cases have been allocated, where

$$\sum_{j=1}^{L-1} S_{i,j} \leq m_{i,\ell} \leq \sum_{j=1}^{L-1} S_{i,j} + t - \sum_{i=1}^{L-1} S_i.$$

We proceed sequentially in this manner until all the remaining S_L test cases have been allocated, where

$$S_L = t - \sum_{j=1}^{L-1} S_j.$$

5.8 Monte Carlo Simulations

Numerical results are presented in this section to illustrate the proposed sampling schemes to achieve the second-order optimality of Bayes risk when estimating the reliability of a series system with two components.

Table (11) on page (74) present the results for a fully sequential design with various prior settings. The results show the trend of the difference between incurred Bayes risk and optimal Bayes risk as the total number of test cases t increases. It is indicated that as t increases, the incurred Bayes risk is approaching the optimal Bayes risk in the order of t^2, which is expected.

The results of the multistage design are illustrated in Table (12) on page (75). Without surprises, the difference also converges to zero in the order of t^2 which agrees with our theory. However, when compared with the results from the fully sequential design, the difference is bigger than the fully sequential design in almost every situation. This is because fully sequential design allocates test cases one by one and thus has more accuracy over the multistage design, which saves execution time while maintaining an acceptable accuracy.

We present the results of the accelerated design in Table (13) on page (76). As revealed in the table, the difference decreases in the order of t^2 as in the other two sequential designs. We noticed by comparing results with other two designs that the accuracy of the accelerated design is just in between, which also makes sense because it only allocate test cases in series at the last stage. We could utilize this more balanced sequential design when both accuracy and efficiency are pursued under certain circumstances.

SUMMARY AND CONCLUSION

Reliability estimation for software is rarely undertaken blindly. There is almost always some notion by those responsible for this task about the reliability of a system before reliability testing begins. Traditional approaches to partition based reliability estimation use this information alone to establish fixed test allocations before reliability testing begins. Therefore, once testing begins we do not use the intermediate results of test execution to refine the allocation of test cases among the partitions. If the prior parameters are chosen poorly, the accuracy in estimating reliability using a fixed test case allocation could suffer greatly.

In this dissertation we took a Bayesian approach to test case allocation. Bayesian analysis allows us to update our beliefs about the reliability of a particular partition as we test, and thus, dynamically refine our allocation of test cases during the reliability testing process.

After introducing some background in the first chapter, we presented purely sequential sampling scheme in chapter 2 to estimate the reliability of a software system using partition testing. We have shown both theoretically and through simulation that the purely sequential sampling scheme always performs at least as well as fixed sampling approaches where test case allocation is predetermined. The content of this chapter is published in Journal of Applied Statistical Science [29].

In the third chapter, the multistage sampling scheme is presented. Unlike the purely sequential sampling scheme, the multistage sampling scheme allocate test cases in batches instead of one at a time. It is thus less costly and highly implementable while maintaining trustworthy results. The proposed sampling scheme

is applied to a real world application with 64 subdomains and shows superiority over fixed sampling schemes. The content of this chapter is published in IEEE Transactions in Reliability [30].

To balance accuracy and efficiency, we proposed the accelerated sampling scheme in chapter 4. Before the last stage, the accelerated sampling scheme will allocate test cases in batches just like the multistage sampling scheme. In the last stage, however, test cases will be allocated one by one in order to achieve a better accuracy. Comparisons have been made through Monte Carlo simulations among accelerated sampling scheme, equal allocation, multistage sampling scheme and purely sequential sampling scheme. It is indicated that the acclerated sampling scheme outperforms all the other sampling schemes in terms of incurred Bayes risk except purely sequential sampling scheme, which is much more time-consuming than the proposed sampling scheme. The content of this chapter is published in Applied Mathematics and Computation [28].

We extend our studies to series systems in chapter 5. We start with a series system with two components. A second-order lower bound is established for the incurred Bayes risk. Purely sequential, multistage and accelerated sampling schemes are then proposed in order to achieve this second-order lower bound. The simulation result is very promising. The problem could be more general if we study finite components instead of two components, which will be our focus of the future researches.

APPENDIX

TABLES

Table 1
Bayes Risk ratios of fully sequential vs. the best fixed ($\Re(S_p)/\Re(F_o)$)

α_1, β_1	α_2, β_2	p_1, p_2	$M = 40$ (% *to* D_1)	$M = 100$ (% *to* D_1)	$M = 1000$ (% *to* D_1)	$M \to \infty$
1,1	1,1	.5,.5	.975 (50%)	.968 (50%)	.964 (50%)	.963
1,1	1,1	.1,.9	.989 (8%)	.986 (10%)	.986 (11%)	.986
.5,.01	.5,.01	.5,.5	.730 (50%)	.637 (50%)	.552 (50%)	.529
.5,.01	.5,.01	.1,.9	.941 (18%)	.903 (20%)	.841 (21%)	.830
.1,.001	1,1	.5,.5	.934 (6%)	.922 (5%)	.896 (2%)	.884
.1,.001	1,1	.1,.9	.723 (22%)	.665 (17%)	.602 (9%)	.563
1,.05	.1,.005	.5,.5	.783 (69%)	.676 (70%)	.641 (73%)	.618
1,.05	.1,.005	.1,.9	.828 (36%)	.772 (39%)	.738 (43%)	.704
1,2	2,1	.5,.5	.979 (50%)	.969 (50%)	.965 (50%)	.963
1,2	2,1	.1,.9	.997 (6%)	.991 (9%)	.989 (11%)	.987

Table 2
Partitioning of the input domain obtained by grouping the factors

Factor A: Unique Tasks	Factor B: Replicates	Factor C: Replicate Type	Factor D: User Entry
A1: Single Task	B1: Single	C1: Defined in Data Base	D1: Null
A2: 2-5 Tasks	B2: two or more	C2: Defined by User	D2: Invalid
A3: 6-10 Tasks			D3: Partial
A4: $\geq$ 11 Taks			D4: Complete

Table 3
Initial Parameter Settings

Domain	$D1-D5$	$D6-D10$	$D11-D15$	$D16-D20$	$D21-D25$	$D26-D30$
(α_i, β_i)	(1,1)	(1,1)	(1,.05)	(1,.05)	(.1,.005)	(.1,.005)
p_i	.01	.02	.01	.02	.01	.02
Domain	$D31-D35$	$D36-D40$	$D41-D45$	$D46-D50$	$D51-D55$	$D56-D60$
(α_i, β_i)	(1,2)	(1,2)	(2,1)	(2,1)	(.5,.01)	(.5,.01)
p_i	.01	.02	.01	.02	.01	.02
Domain	$D61-D62$	$D63-D64$				
(α_i, β_i)	(.1,.001)	(.1,.001)				
p_i	.01	.04				

Table 4
Bayes risk ratios of two-stage vs. the best-fixed ($\Re(S_m)/\Re(F_o)$)

M	$\Re(S_m)/\Re(F_o)$
1000	94.18%
10000	85.9%
50000	79.47%
∞	69.36%

Table 5
Percentage to each subdomain after simulation (M=50,000)

Domain	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
% to D_i	2.04%	2.04%	2.05%	2.05%	2.04%	4.09%	4.09%	4.1%	4.09%	4.11%
Domain	D11	D12	D13	D14	D15	D16	D17	D18	D19	D20
% to D_i	.36%	.36%	.36%	.36%	.35%	0.75%	0.75%	0.74%	0.72%	0.74%
Domain	D21	D22	D23	D24	D25	D26	D27	D28	D29	D30
% to D_i	.08%	.09%	.08%	.09%	.08%	.16%	.17%	.17%	.17%	.18%
Domain	D31	D32	D33	D34	D35	D36	D37	D38	D39	D40
% to D_i	2.03%	2.04%	2.04%	2.05%	2.05%	4.09%	4.08%	4.07%	4.09%	4.08%
Domain	D41	D42	D43	D44	D45	D46	D47	D48	D49	D50
% to D_i	2.04%	2.04%	2.05%	2.03%	2.05%	4.09%	4.08%	4.1%	4.07%	3.69%
Domain	D51	D52	D53	D54	D55	D56	D57	D58	D59	D60
% to D_i	.1%	.09%	.1%	.09%	.09%	.2%	.19%	.2%	.2%	.19%
Domain	D61	D62	D63	D64						
% to D_i	.02%	.02%	.1%	.09%						

Table 6
Bayes risk ratios of accelerated vs. equal allocation ($\Re(S_a)/\Re(F_e)$)

α_1, β_1	α_2, β_2	p_1, p_2	$M = 40$ (% *to* D_1)	$M = 100$ (% *to* D_1)	$M = 1000$ (% *to* D_1)	$M \to \infty$
1,1	1,1	.5,.5	.976 (50%)	.966 (50%)	.963 (50%)	.963
1,1	1,1	.1,.9	.608 (8%)	.604 (10%)	.601 (11%)	.601
.5,.01	.5,.01	.5,.5	.696 (50%)	.648 (50%)	.555 (50%)	.529
.5,.01	.5,.01	.1,.9	.605 (18%)	.533 (19%)	.524 (21%)	.506
.1,.001	1,1	.5,.5	.568 (6%)	.545 (5%)	.522 (2%)	.507
.1,.001	1,1	.1,.9	.415 (28%)	.348 (21%)	.309 (11%)	.286
1,.05	.1,.005	.5,.5	.666 (68%)	.578 (70%)	.552 (72%)	.534
1,.05	.1,.005	.1,.9	.637 (33%)	.542 (36%)	.522 (42%)	.522
1,2	2,1	.5,.5	.985 (50%)	.976 (50%)	.963 (50%)	.963
1,2	2,1	.1,.9	.608 (6%)	.603 (9%)	.602 (11%)	.602

Table 7
Bayes risk ratios of accelerated vs. the best fixed ($\Re(S_a)/\Re(F_o)$)

α_1, β_1	α_2, β_2	p_1, p_2	$M = 40$ (% *to* D_1)	$M = 100$ (% *to* D_1)	$M = 1000$ (% *to* D_1)	$M \to \infty$
1,1	1,1	.5,.5	.978 (50%)	.969 (50%)	.964 (50%)	.963
1,1	1,1	.1,.9	.991 (8%)	.988 (10%)	.986 (11%)	.986
.5,.01	.5,.01	.5,.5	.710 (50%)	.665 (50%)	.559 (50%)	.529
.5,.01	.5,.01	.1,.9	.949 (18%)	.896 (19%)	.843 (21%)	.830
.1,.001	1,1	.5,.5	.949 (6%)	.931 (5%)	.907 (2%)	.884
.1,.001	1,1	.1,.9	.751 (28%)	.671 (21%)	.609 (11%)	.563
1,.05	.1,.005	.5,.5	.771 (68%)	.680 (70%)	.656 (72%)	.618
1,.05	.1,.005	.1,.9	.866 (33%)	.778 (36%)	.749 (42%)	.704
1,2	2,1	.5,.5	.989 (50%)	.981 (50%)	.967 (50%)	.963
1,2	2,1	.1,.9	.994 (6%)	.989 (9%)	.988 (11%)	.987

Table 8
Bayes risk ratios of accelerated vs. two-stage ($\Re(S_a)/\Re(S_m)$)

α_1, β_1	α_2, β_2	p_1, p_2	$M = 40$ (% *to* D_1)	$M = 100$ (% *to* D_1)	$M = 1000$ (% *to* D_1)	$M \to \infty$
1,1	1,1	.5,.5	.997 (50%)	.991 (50%)	.993 (50%)	1
1,1	1,1	.1,.9	.993 (8%)	.992 (10%)	.998 (11%)	1
.5,.01	.5,.01	.5,.5	.985 (50%)	.974 (50%)	.986 (50%)	1
.5,.01	.5,.01	.1,.9	.992 (18%)	.973 (19%)	.998 (21%)	1
.1,.001	1,1	.5,.5	.995 (6%)	.999 (5%)	.992 (2%)	1
.1,.001	1,1	.1,.9	.949 (28%)	.968 (21%)	.940 (11%)	1
1,.05	.1,.005	.5,.5	.983 (68%)	.955 (70%)	.986 (72%)	1
1,.05	.1,.005	.1,.9	.997 (33%)	.948 (36%)	.971 (42%)	1
1,2	2,1	.5,.5	.990 (50%)	.999 (50%)	.997 (50%)	1
1,2	2,1	.1,.9	.999 (6%)	.998 (9%)	.990 (11%)	1

Table 9
Operational profile and usage probabilities for billing system

Subdomain	Usage Probability	(α, β)
Residential, no calling plan, paid	0.5940	(1,1)
Residential, national calling plan, paid	0.1580	(1,1)
Business, no calling plan, paid	0.1485	(.5,.01)
Business, national calling plan, paid	0.0396	(.5,.01)
Residential, international calling plan, paid	0.0396	(.1,.005)
Business, international calling plan, paid	0.0099	(.1,.005)
Residential, no calling plan, delinquent	0.0060	(1,.05)
Residential, national calling plan, delinquent	0.0016	(1,.05)
Business, no calling plan, delinquent	0.0015	(1,2)
Business, national calling plan, delinquent	0.0004	(1,2)
Residential, international calling plan, delinquent	0.0004	(2,1)
Business, international calling plan, delinquent	0.0001	(2,1)

Table 10
Bayes risk ratios of accelerated vs. the best-fixed ($\Re(S_a)/\Re(F_o)$)

M	$\Re(S_a)/\Re(F_o)$
5000	97.63%
10000	93.78%
50000	91.22%
∞	89.02%

Table 11
$t^2 * (\Re(\Delta) - \Re(o))$ By Fully Sequential Design

(a, b)	(c, d)	$t = 100$	$t = 300$	$t = 500$	$t = 800$	$t = 1000$
(2,2)	(2,2)	1.00	0.97	1.12	0.792	0.42
(4,3)	(4,3)	1.19	1.57	1.77	1.42	1.03
(4,3)	(3,4)	0.78	0.84	0.63	0.65	0.49
(0.1,2)	(1,20)	0.007	0.009	0.005	0.004	0.001
(5,2)	(4,3)	1.19	1.23	1.11	0.98	0.83

Table 12
$t^2 * (\Re(\Delta) - \Re(o))$ By Multistage Design

(a, b)	(c, d)	$t = 100$	$t = 300$	$t = 500$	$t = 800$	$t = 1000$
(2,2)	(2,2)	1.56	1.44	1.77	1.23	1.11
(4,3)	(4,3)	1.76	1.98	1.83	1.56	1.22
(4,3)	(3,4)	1.33	1.04	0.97	0.84	0.81
(0.1,2)	(1,20)	0.012	0.008	0.013	0.008	0.005
(5,2)	(4,3)	1.45	1.72	1.56	1.44	1.32

Table 13
$t^2 * (\Re(\Delta) - \Re(o))$ By Accelerated Design

(a, b)	(c, d)	$t = 100$	$t = 300$	$t = 500$	$t = 800$	$t = 1000$
(2,2)	(2,2)	1.23	1.11	1.24	1.02	0.95
(4,3)	(4,3)	1.34	1.75	1.88	1.53	1.11
(4,3)	(3,4)	0.88	0.93	0.72	0.74	0.63
(0.1,2)	(1,20)	0.009	0.007	0.009	0.006	0.004
(5,2)	(4,3)	1.23	1.65	1.44	1.32	1.11

REFERENCE

[1] Adams, T. (1996). Total variance approach to software reliability estimation. *IEEE Transactions on Software Engineering 22*(9), 687–688.

[2] Ash, R. (1972). *Real Analysis and Probability*. New York: Academic Press.

[3] Bai, C., C. Jiang, and K. Cai (2010, July 29-31). A reliability improvement predictive approach to software testing with bayesian method. Proceedings of the 29th Chinese Control Conference, Beijing, China.

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

[5] Bickel, P. and J. Yahav (1969). Some contributions to the asymptotic theory of Bayes solutions. *Wahrscheinlichkeitstheorie und verw. Gebietr. 11*, 257–276.

[6] Bickel, P. J. and K. A. Doksum (2001). *Mathematical Statistics: Basic Ideas and Selected Topics* (2nd ed.). New Jersey: Prentice Hall.

[7] Breiman, L. (1968). *Probability*. Addison-Wesley Publishing.

[8] Chen, T., A. Mayrhauser, A. Hajjar, C. Anderson, and M. Sahinoglu (1999). How much testing is enough? applying stopping rules to behavioral model testing. Proceedings of the 4th IEEE International Symposium on High-Assurance Systems Engineering, pp. 249–256.

[9] Dyer, D. and P. Chiou (1984). An information theoretic approach to incorporating prior information in binomial sampling. *Communications in Statistics-Theory and Methods 13*(17), 2051–2083.

[10] Gutjahr, W. J. (1995). Optimal test distributions for software failure cost estimation. *IEEE Transactions on Software Engineering 21*(3), 219–228.

[11] Hardwick, J. and Q. Stout (1995). Determining optimal few-stage allocation procedures. *Computing Science and Statistics 27*.

[12] Hardwick, J. and Q. Stout (1996). Optimal allocation for estimating the mean of a bivariate polynomial. *Sequential Analysis 15*(2& 3), 71–90.

[13] Jensen, S., S. Soi, and L. Wang (2009). A bayesian approach to efficient different allocation for resampling-based significance testing. *BMC Bioinformatics 10*, 198.

[14] Kaner, C. (1997). The impossibility of complete testing. *Software QA 4*(4), 28.

[15] Lehmann, E. L. (1983). *Theory of Point Estimation*. New York: John Wiley & Sons.

[16] Lindgren, B. (1976). *Statistical Theory* (3rd ed.). New York: Macmillan.

[17] Littlewood, B. and D. Wright (1997). Some conservative stopping rules for the operational testing of safety critical software. *IEEE Transactions on Software Engineering 23*(11), 673–683.

[18] Miller, K. W., L. J. Morell, R. E. Noonan, S. K. Park, D. M. Nicol, B. W. Murrill, and J. Voas (1992). Estimating the probability of failure when testing reveals no failures. *IEEE Transactions on Software Engineering 18*(1), 33–43.

[19] Musa, J. (1993). Operational profiles in software - reliability engineering. *IEEE Software 10*(2), 14–32.

[20] Nair, V. N., D. A. James, W. K. Ehrlich, and J. Zevallos (1998). A statistical assessment of some software testing strategies and application of experimental design techniques. *Statistica Sinica 8*, 165–184.

[21] Özekici, S. and R. Soyer (2003). Reliability of software with an operational profile. *European Journal of Operational Research 149*(2), 459–474.

[22] Poore, J. H., H. D. Mills, and D. Mutchler (1993). Planning and certifying software system reliability. *IEEE Software 10*(1), 88–99.

[23] Rekab, K. (1992a). A nearly optimal 2-stage procedure. *Communications in Statistics Theory and Methods 21*(1), 197–201.

[24] Rekab, K. (1992b). Sequential allocation for estimation problem. *Stochastic Analysis and Applications 10*(4), 465–470.

[25] Rekab, K. (1993). A sampling scheme for estimating the reliability of a series system. *IEEE Transactions On Reliability 42*(2), 287–290.

[26] Rekab, K. and M. Tahir (2000). A two-stage sequential allocation scheme for estimating the product of several means. *Stochastic Analysis and Applications 18*(2), 289–298.

[27] Rekab, K. and M. Tahir (2004). An asymptotic second-order lower bound for the bayes risk of a sequential procedure. *Sequential Analysis 23*, 451–464.

[28] Rekab, K., H. Thompson, and W. Wu (2013a). An efficient test allocation for software reliability estimation. *Applied Mathematics and Computation 220*(1), 94–103.

[29] Rekab, K., H. Thompson, and W. Wu (2013b). A fully sequential test allocation for software reliability estimation. *Journal of Applied Statistical Science 20*(1), 63–78.

[30] Rekab, K., H. Thompson, and W. Wu (2013c). A multistage sequential test allocation for software reliability estimation. *IEEE Transactions on Reliability 62*(2), 424–433.

[31] Rekab, K., X. Yan, and H. Thompson (2008). An efficient sampling scheme for estimating software reliability with associated cost. *Stochastic Analysis and Applications 28*(2), 353–366.

[32] Rohatgi, V. K. (1976). *An Introduction to Probability Theory and Mathematical Statistics*. New York: John Wiley and Sons.

[33] Sahinoglu, M. and E. Chow (1999). Empirical-bayesian availability indices of safety and time critical software systems with corrective maintenance. Proceedings of the 1999 Pacific Rim International Symposium on Dependable Computing, pp. 84–92.

[34] Sayre, K. and J. H. Poore (1999, May). Partition testing with usage models. Proceedings of Science and Engineering for Software Development, Los Angeles.

[35] Shapiro, C. (1985). Allocation schemes for estimating the product of positive parameters. *Journal of the American Statistical Association 80*(390), 449–454.

[36] Stüger, H. P. (2006). Asymmetric loss functions and sample size determination: A bayesian approach. *Austrian Journal Of Statistics 35*(1), 57–66.

[37] Turner, R. and A. Corporation (2007, Jan). Optimized allocation of testing budget for missile defense vehicle. Reliability and Maintainability Symposium, Orlando.

[38] Walton, G. H., J. H. Poore, and C. J. Trammel (1995). Statistical testing of software based on a usage model. *Software Practice and Experience 25*(1), 97–108.

[39] Weyuker, E. J. and B. Jeng (1991). Analyzing partition testing strategies. *IEEE Transactions on Software Engineering 17*(7), 703–711.

[40] Weyuker, E. J. and T. J. Ostrand (1980). Theories of program testing and the application of revealing sub-domains. *IEEE Transactions on Software Engineering SE-6*(3), 236–246.

[41] Whittaker, J. A. and M. G. Thomason (1994). A markov chain model for statistical software testing. *IEEE Transactions on Software Engineering 20*(10), 812–824.

[42] Zahedi, Z. and N. Ashrafi (1991). Software reliability allocation based on structure, utility, price, and cost. *IEEE Transactions on Software Engineering 17*(4), 345–356.

VITA

Wei Wu was born on July 20th, 1983, in Wuhan, Hubei Province, China. In 2001 he graduated from No.3 High School of WISCO. He received a B.S. degree from Nanjing University of Technology in Geographic Information System 2005 and an M.S. degree in University of Missouri-Kansas City in Mathematics and Statistics 2010. He has been a business analyst in Sprint Corporation since 2011. In 2013, the contents of this dissertation were published in IEEE transactions on Reliability, Applied Mathematics and Computation and Journal of Applied Statistical Science.