Testing for Monotonicity in Credit Risk Modelling

by

Zhangqian Xie

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Mathematical Finance

Department of Mathematical and Statistical Sciences
University of Alberta

Abstract

Credit risk management, which deals with mitigating losses from lending activities, is crucial for financial institutions. Hence, credit risk modelling can be employed to reduce potential losses and avoid financial crises. There are sometimes monotonic relationships in credit risk models, which can simplify forms of models, reduce computational time, or be necessary to fulfill restrictions observed in reality. After reviewing commonly used credit risk models, several monotonicity testing methods are established and adapted to the situation for binary output of default indicators. Furthermore, we present a new test using the weighted sum of differences as the test statistic, with the weights optimized through combinations of their moments. Finally, we compare the performance of these tests on simulated data regarding the accuracy and power of the tests.

Acknowledgements

First, I wish to show my appreciation to my academic supervisor, Dr. Christoph Frei. On the one hand, his on-time professional guidance played a critical role in completing the thesis research. On the other hand, his industrial experience and connections were invaluable for my career development.

I would like to thank Kelly Keenan and the other members of the AI Labs Team of ATB Financial for the research collaboration and providing the idea of the thesis topic. Their industrial knowledge benefited me to a great extent during the internship at the ATB Financial. Financial support by Mitacs through the Accelerate Program and by the Social Sciences and Humanities Research Council of Canada through Insight Grant 435-2018-0049 is gratefully acknowledged.

Moreover, I wish to extend my special thanks to Dr. Abel Cadenillas, Dr. Adam Kashlak, and Dr. Alexander Melnikov for serving on my thesis examination committee.

Last but not least, I appreciate the support of my friends and family during the academic years of my master's studies.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Nowadays, loans have become an indispensable part of the modern financial system. From an individual to a corporation, it can be beneficial to acquire loans, which are negotiated to repay in the future, for making an investment or securing the present cash flow. Inevitably, there is some possibility or risk that the borrower cannot return the total money at the due date. It can occur that the lender will have a loss. If the amount of loss is large, the lender may go bankrupt. If the lender also owes money to another lender, it can result in a cascading effect, which may eventually cause a financial crisis. It is worthwhile to develop appropriate methods to quantify the risk of lending for risk management. In financial studies, *credit* refers to the money offered by a lender to a borrower and obligated to be repaid by the due date. *Credit risk* refers to the risk that a borrower is incapable of paying credit by the due date. Scholars developed models to quantify credit risk, which can be useful. Firstly, lenders can use credit risk models to reduce their financial losses. Secondly, it is helpful to apply the credit risk model to mitigate the impact of the financial crisis. Thirdly, by adjusting the credit risk models, they can be used for establishing a credit rating system for customers, or even constructing a highly interconnected network.

Since credit plays a crucial role in modern Finance, there have been various concepts and methodologies proposed by researchers over the decades. Initially, Altman (1968) came up with a numerical measure of credit risk through the Z-score. According to him, the bankruptcy probability of a firm can be predicted with financial data using multivariate discriminant analysis. Despite some objections to his method for using the historical values only, the Z-score is still popular among many practitioners thanks to its convenience (Benzschawel, 2012). The next well-known method is developed by Merton (1974) through more sophisticated techniques, such as stochastic processes and partial differential equations. With the help of the Black-Scholes framework, which had been used for option pricing (Black and Scholes, 1973), Merton modelled the debt and equity of the firm. Meanwhile, a new concept "distance-to-default" is proposed in his article. Based on the Merton model, there have been several advanced models to relax the restrictions of the model. These later credit risk models are known as structural models. Similar to the Merton model, the structural models require information about the firm's capital structure and asset value that may be unavailable for unlisted firms. To that end, reduced-form credit risk models are suggested. Referring to the views of Jarrow and Turnbull (1995), and Duffie and Singleton (1999), the reduced-form models regard the default process as a stopped jump process independent of the firm's asset value. More details about the structural and reduced-form credit risk models will be reviewed in Chapter 2.

Moreover, there are some studies on applying regression models to the estimation of credit risk. For example, Kruppa et al. (2013) compared the performance of several regression models (logistic regression, Random Forest, etc.) to estimate consumer credit risk. Along with the recent development of artificial intelligence, there have been gradually more experiments on applying machine learning models to credit risk studies. The machine learning models are similar to regressions to some extent and may have better performance, but it is more difficult to interpret the reasoning behind them. It should be noted that machine learning models are sensitive to the parameters. To put it more simply, it is likely to obtain a high accuracy in-sample by tuning the parameters deliberately while unable to perform well out of the sample. Hence, it is important to delve into the foundations of these models before implementing them. Moreover, it is necessary to explain meanings in the economic or financial senses for combining a complicated model with credit risk management.

In order to resolve the dilemma, this thesis is interested in seeking some constraints to simplify the models. These constraints could be based on theoretical grounds or reality restrictions. To exemplify, Dugas et al. (2000) stated that it is useful to incorporate some prior knowledge for reducing the complexity of the model and improving the performance of option pricing. According to some knowledge from the banking industry, monotonicity is used in the thesis. As evidence, the Merton model demonstrates that default risks are decreasing against the asset value returns. In addition, the Vasicek model indicates that the joint default probability is increasing against the correlation of the asset value returns. In light of such facts from financial studies, it is reasonable to impose monotonic constraints on credit risk models. Such constraints can make it easier to interpret complicated credit risk models. Therefore, the monotonic relationships in credit risk models will be explored in this thesis. The definition of *monotonicity* is described as follows.

Definition 1.1. Assume two real-valued random variables X and Y satisfy $Y = f(X)$ almost surely for a function f. Then

- Y is said to be *(strictly) increasing* against X if f is (strictly) increasing on the range of X,
- Y is said to be *(strictly) decreasing* against X if f is (strictly) decreasing on the range of X.

In practice, it is necessary to deal with the problem of how to know the monotonic relationship exactly from the data. A relatively direct method is to use the related graphs to determine the monotonicity. In other words, a variety of figures for the realizations of Y against X can be used to roughly determine whether they are monotonic or not. These figures can be line plots with ordering variables, scatter plots for smaller data, or box plots for categorical variables. Applying graphical methods provides a convenient way to view the relationship between two random variables. Nonetheless, the conclusion can be confused or even incorrect due to randomness in datasets.

To avoid these problems, the thesis prefers to use hypothesis testing, which is involved with developing hypotheses, as a numerical ground for the conclusion.

Without loss of generality, suppose that the null hypothesis H_0 is no monotonic relationship between X and Y, and the alternative hypothesis H_1 is that there exists monotonicity, which can be further divided into increasing and decreasing. There are two kinds of statistical tests, parametric tests, and non-parametric tests, which are both usually adjusted to the study of random series. Referring to the views of Esterby (1996), the later chapters only focus on non-parametric tests that are thought to be relatively more suitable for non-normally distributed data, which is fitted with the real data. To put it another way, non-parametric tests coincide with the practice in the financial industry. Three non-parametric tests using the ranks of random variables are studied in the thesis. In the case of applying these tests to credit risk modelling, the credit indicators, as Y in credit risk models, only have two values (1 and 0 represent a default or not, respectively). Since the range of Y is relatively simple, it is possible to make some improvements for three rank-based tests. That is to say, the issue is to improve and compare the performance of the monotonicity tests with the binary dependent variable Y. Despite numerous variants of statistical tests for monotonicity, there is still no specific test for the case of binary outputs. Hence, we also made an effort to develop a new test for the Vasicek model.

An overview of the later chapters is as follows. Chapter 2 reviews the structural and reduced-form credit risk models. For further research, the priority is put on the Vasicek model. Three types of tests based on correlation coefficients are discussed in Chapter 3. To improve the performance of these three tests for credit indicators with binary values, Chapter 4 adjusts these rank-based tests, and Chapter 5 creates a new test by optimizing the weighted sums. The performance of three rank-based tests and the new test is compared in Chapter 6. Chapter 7 draws conclusions for the thesis.

Chapter 2

Credit Risk Models

Credit risk models can be grouped into two types: structural models and reduced-form models. The former is based on the fact that the probability of default is computed through asset values and liabilities, while the latter estimates the probability of default by introducing an exogenous random event.

2.1 The Merton Model

One of the best known credit risk models is the Merton model that is developed by Merton (1974) through the Black-Scholes (BS) option pricing formula (Black and Scholes, 1973). In this model, the equity and debt of the firm are regarded as options on the asset value.

More precisely, let V_t be the asset value of the firm at time t, T be the maturity of loan with the notional amount of D. The model assumes that V_t equals to the sum of the values of equity and debt of the firm, denoted by E_t and D respectively, and we have $V_t = E_t + D$. There are also some assumptions required:

1. Assets can be fractionally and continuously traded. Short selling is allowed.

2. The interest rates of borrowing and lending are risk-free rate r, which is constant and deterministic.

3. No transaction fees or taxes. No dividends.

4. $(V_t)_{0\leq t\leq T}$ is a geometric Brownian motion, or in other words

$$\frac{dV_t}{V_t} = \mu\, dt + \sigma\, dW_t, \tag{2.1}$$

where μ and σ are the expected return and volatility of the firm's assets, and $(W_t)_{0\leq t\leq T}$ is a Brownian motion.

The first three assumptions are necessary for using the BS framework and the last is the evolution of the asset value V_t. Figure 2.1 depicts the path of V_t and the horizontal line is the constant loan value D. Naturally, we can observe that if the asset value is below the default point line, $V_T < D$, the firm will default, and hence the default can be represented by the shaded area. We assume that if the default occurs, the firm will compensate the bank for the total assets value V_T.

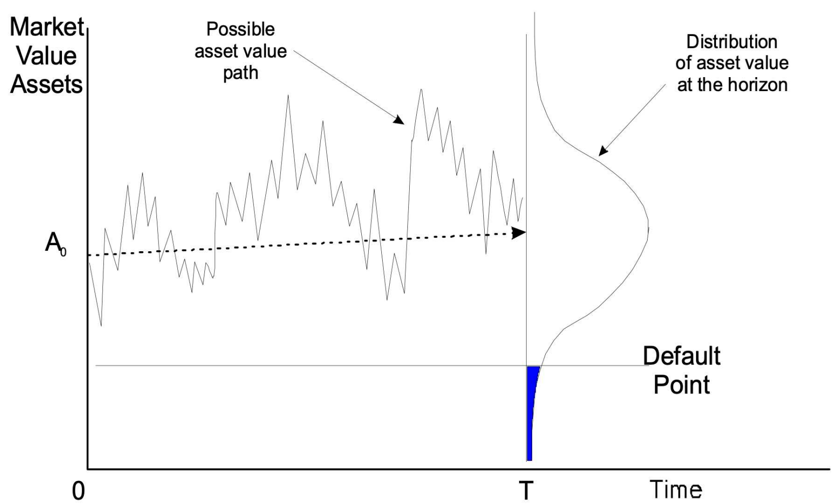

Figure 2.1: Dynamics of the Merton Model, Source: Vasicek (1984)

Since the amount paid at time T should be D or V_T, if the firm defaults or not, we know that the debt at the due date is

$$B_T = \min(V_T, D) = D - \max(D - V_T, 0).$$

According to (2.1), we can view the debt value as the payoff of the portfolio, which consists of a zero-coupon bond with a face value D while shorting a European put option on the firm's assets value with the exercise price D. Moreover, the equity payoff $E_T = \max(V_T - D, 0)$ can be viewed as the payoff of a European call option as well. Recall that we aim to calculate the shaded area in Figure 2.1 as the probability of default, and with the BS framework we have

$$P[V_T < D] = \Phi(-d_2),$$

where

$$d_2 = \frac{\ln \frac{V_0}{D} + \left(\mu - \frac{\sigma^2}{2}\right) T}{\sigma \sqrt{T}},$$

and $\Phi(\cdot)$ is the standard normal distribution function. d_2 is sometimes called the "distance-to-default".

It should be noted that Figure 2.1 shows the credit risk at the maturity date T, while Black and J. D. Cox (1976) enhanced the Merton model allowing that the default can be at any time between 0 and T.

2.2 The Vasicek Model

Following the Merton model, Vasicek (1984) studied the credit risk of the loan portfolio, rather than some financial contracts or a specific derivative. Hence, the Vasicek model still belongs to the structural models. The Vasicek model not only regards the joint loss distribution across multiple obligors but also divides risks into systematic and idiosyncratic parts.

Fix a complete probability space $(\Omega, \mathcal{F}, P)$. Considering a group of borrowers with the size of $n > 1$, the value of the loan portfolio for $j = 1, 2, \ldots, n$ at the maturity time T is

$$V_T^j = \rho W_T + \sqrt{1 - \rho^2} B_T^j, \tag{2.2}$$

where $(W_t)_{0\le t\le T}$ is a Brownian motion (BM) as a common factor of the model, $(B_t^j)_{0\le t\le T}$ is a BM specific to borrower j, and $\rho \neq \pm 1$ can be viewed as a correlation coefficient.

Assume that $(W_t)_{0\le t\le T}$ and $\{(B_t^j)_{0\le t\le T}\}_{j=1}^n$ are pairwise independent. Consider the default indicator for any borrower j, which should equal to one if j defaults,

$$Y_T^j = \begin{cases} 1 & \text{if } V_T^j < c^j, \\ 0 & \text{if } V_T^j \ge c^j, \end{cases} \tag{2.3}$$

where c^j is a capital criterion for borrower j. We assume that c is a deterministic variable with the discrete range of $\{c^j\}_{j=1}^n$. Taking an outcome $\omega \in \Omega$, we can think of (2.3) as a functional $V_T(\omega) : c \mapsto Y_T(\omega)$, where Y_T is a random variable depending on c. Take the realizations of Y_T as $\{y_T^j\}_{j=1}^n$. We want to test the monotonicity between $\{y_T^j\}_{j=1}^n$ and $\{c^j\}_{j=1}^n$.

If the common factor W_T is given, the conditional probability of default could be examined. Let $p_j = P[Y_T^j = 1|W_T]$ and $q_j = P[Y_T^j = 0|W_T] = 1 - p_j$. According to the definitions of p_j and q_j, we initially know that for any integer j between 1 and n

$$\begin{aligned} p_j(x) &= P\left[Y_T^j = 1|W_T = x\right] \\ &= P\left[V_T^j < c^j|W_T = x\right] \\ &= P\left[B_T^j < \frac{c^j - \rho x}{\sqrt{1-\rho^2}}\right] \\ &= \Phi\left(\frac{c^j - \rho x}{\sqrt{(1-\rho^2)T}}\right), \\ q_j(x) &= 1 - p_j(x), \end{aligned}$$

which suggests that p_j is increasing against j if $c^1 < c^2 < \cdots < c^n$.

A special case is that

$$\prod_{j=1}^{[\frac{n}{2}]} P[Y_T^j = 0|W_T = x] \prod_{j=[\frac{n}{2}]+1}^{n} P[Y_T^j = 1|W_T = x]$$
$$= \prod_{j=1}^{[\frac{n}{2}]} \left(1 - \Phi\left(\frac{c^j - \rho x}{\sqrt{(1-\rho^2)T}}\right)\right) \prod_{j=[\frac{n}{2}]+1}^{n} \Phi\left(\frac{c^j - \rho x}{\sqrt{(1-\rho^2)T}}\right).$$

Subsequently, we know that Y_T^j is conditionally independent of each other given the normal random variable W_T. Note that the conditional moments of Y_T^j and the covariance of Y_T^i and Y_T^j $(i \neq j)$ can be expressed as follows

$$\begin{aligned} \mathbf{E}[Y_T^j|W_T] &= p_j, \\ \mathrm{Var}(Y_T^j|W_T) &= p_j q_j, \\ \mathrm{Covar}(Y_T^i, Y_T^j|W_T) &= P[Y_T^i = 1, Y_T^j = 1|W_T] - p_i p_j = 0. \end{aligned} \tag{2.4}$$

Hence, the conditional probability of default will be $P[Y_T^j = 1|W_T]$ for each j, which is the focus of our study. Our main concern is to find the best methods for testing the monotonicity of Y_T against c through the realizations. We will compare different tests with the assumption that c is ascending. The main difficulty is that the relation of Y_T and c is hidden, which means that we cannot get the first-order derivative as a good measure of the monotonic relation. The randomness of data can hinder the testing of monotonicity as well. It is better to give up the traditional definitions of increasing and decreasing since the range of Y_T is $\{0, 1\}$. We would like to investigate whether the values or order of $\{c^j\}_{j=1}^n$ have a great impact on the performance of tests. Furthermore, the model's sensitivity to correlation ρ and maturity T would be examined.

2.3 Other Structural Models

In 1997, Credit Suisse Financial Products (CSFP) released an approach, CreditRisk+ (Credit Suisse First Boston, 1997). Its main idea comes from that the loss in actuarial science is determined by the probability of a disaster and the degree of loss or damage caused by it. Unlike the CreditMetrics model (Gupton, 1997), CreditRisk+'s credit risk metric does not include the case that the credit rating of the credit instrument is reduced. Moreover, an important assumption of the model is different from the CreditMetrics model. The CreditMetrics model generally measures the risk of portfolio loans and indirectly calculates the risk impact of individual loans on portfolio loans by calculating the marginal risk contribution of new loans. However, the CreditRisk+ model treats each loan as independent, and the probability of default on each loan is considered small. The default probability of each loan in the loan portfolio is random and constant, so it is consistent with the Poisson distribution. In addition, credit migration risk is not explicitly modelled in this analysis. Instead, CreditRisk+ allows for stochastic default rates which partially account for migration risk.

The advantages of this model are that it is assumed that the probability of a single asset default is subject to the Poisson distribution, only one case of default is considered, and the variables considered are small. All of those make the calculation process simple. The shortcomings of the model are that assuming the independence of each default may not be consistent with the actual situation. The model neglects the influence of market risk and credit rating decline and credit term changes.

At the end of the 20th century, KMV Corporation[1], a firm specializing in credit risk analysis, developed a credit risk methodology to evaluate default probability and the distribution related to the default. The KMV model is based on that the equity of the listed firm is regarded as the call option of the asset, while the liability is regarded as the put option of the asset (Crouhy et al., 2000). The default risk of the corporation is measured by estimating the probability that the future asset value of the corporation is lower than a certain value. The value cannot be directly observed but needs to be estimated based on the transaction data of the corporate capital market and the financial data. It is necessary to mention that KMV's model differs from CreditMetrics as it relies upon the Expected Default Frequency (EDF) for each issuer, rather than upon the average historical transition frequencies produced by the rating

[1]KMV Corporation was named after Stephen Kealhofer, John McQuown, and Oldrich Vasicek.

agencies for each credit class.

The advantages of the KMV model are that the calculation of the required data is easy to obtain, and it is convenient for the researcher to carry out the risk assessment of default. In addition, the daily update of the data can dynamically manage the default risk of the corporation. Nonetheless, the assumptions of the model may not be consistent with reality. For example, the asset value is subject to normal distribution, and the liabilities can be simply divided into long-term liabilities and short-term liabilities, which is questionable. Moreover, the model can only measure the default risk of listed companies.

2.4 Reduced-Form Models

The development of reduced-form models is mainly driven by the restrictions of structural models due to assumptions. The first breakthrough is to relax the limitation of deterministic default timing. For example, a structural model, proposed by Hull and White (1995), allows the default to happen stochastically at any time between 0 and T by taking the default probability as a process over a time interval. The model assumes that the joint default probability of different individuals is given by a multivariate normal distribution, which is consistent with the implied moments of the derivatives. Therefore, the default probability is highly correlated to the real information from the market. Some other structural models suggest some new concepts, which would not be discussed in this thesis.

Nonetheless, such models are still not independent of the asset values of the firms, which requires estimations through the financial data of the firms. In other words, it could cause inaccuracy since the estimation is based on historical data and even the financial data can be hard to access as some firms are not listed. For resolving these limitations, the reduced-form models are one of the recent solutions. The dependence on the asset value is neglected in these recent models and the arrival of default is not the most important to be captured in these models. The Merton model and its successors consider the default as an endogenous event and focus on computing the probability of default at an exact time. By contrast, the reduced-form models view the default as an exogenous event and imply the default probability over a small period via default intensity, which is the reason that they are also known as the intensity models. The grounds of the intensity models are stochastic calculus and risk-neutral pricing. The theory of risk-neutral pricing counts on that the present value of the future payoff of a derivative can be equivalent to the market value of the derivative.

A number of reduced-form models have been developed since the end of the last century. Two of them would be introduced here because they are well-known and have relatively greater impacts. One model is proposed by Jarrow and Turnbull (1995) during their studies of derivative pricing. The model assumes the default process as a stopped jump process and defines the intensity of default as the hazard rate. The hazard rate is a concept in probability theory, which is often used in survival analysis. Likewise, the survival probability is computed as the probability of the default process at a specific time. It is worthwhile to mention that the default intensity is regarded as a Poisson process while the default process depends on the intensity. This model is independent of the accounting data of the firm and extends the default arrival to a stochastic case. The other famous reduced-form model is developed by Duffie and Singleton (1999). Their research indicates that a defaultable zero-coupon bond can be equivalently replaced by a default-free bond with the interest rates modelled as default intensity. And then such a bond can be priced using the risk-neutral theory. Their work suggests a framework for converting the defaultable contract into a default-free one, which is crucial to dealing with gradually more complex financial contracts nowadays.

Chapter 3

Statistical Tests for Monotonicity

This section will introduce several testing for detecting monotonic relations. There are commonly used correlation-based tests and their variants for different situations, including the multivariate version.

3.1 Spearman Rank Test

The Spearman Rank (SR) test is a rank-based non-parametric statistical test that can be used to detect a monotonic trend in a time series (Lehmann and D'Abrera, 1975).

Given a sample data set $X_i, i = 1, 2, \ldots, n$ the null hypothesis H_0 of the SR test against trend tests is that all the X_i are no monotonic trend; the alternative hypothesis is that X_i increases or decreases with i. The test statistic is given by

$$D = 1 - \frac{6 \sum_{i=1}^{n} \left[R\left(X_i\right) - i\right]^2}{n\left(n^2 - 1\right)}, \tag{3.1}$$

where $R(X_i)$ is the rank of i-th observation X_i in the sample of size n.

There are some restrictions for the simplified formula (3.1): the data set should be without ties. If ties exist, we should assign the rank to each value, and then replace the rank with the arithmetic average of the tied values. Meanwhile, it is better to use the Pearson correlation coefficient in (3.2) as the test statistic to avoid incorrect results.

$$D = \frac{\mathrm{Cov}(R(X_i), i)}{\sigma_{X_i}\sigma_i}, \tag{3.2}$$

where σ_{X_i} and σ_i are the standard deviations of X_i and its rank, respectively.

Under the null hypothesis, the distribution of D is asymptotically normal with the mean and variance as follows (Lehmann and D'Abrera, 1975)

$$\mathbf{E}(D) = 0,$$
$$\mathrm{Var}(D) = \frac{1}{n-1}.$$

The p-value of the SR statistic (d) of the observed sample data is estimated using the normal cumulative distribution function (CDF) as its statistics are approximately normally distributed With the mean of zero and variance of $\mathrm{Var}(D)$ for the SR statistic. Using the following standardization

$$Z_{\mathrm{SR}} = \frac{D}{\sqrt{\mathrm{Var}(D)}},$$

the standardized statistic Z follows the standard normal distribution $Z \sim \mathrm{N}(0, 1)$ when $n > 4$.

Yue et al. (2002) compared the power of both the SR test and the following MK test to detect a trend. They also studied the influence of sample sizes and sample variations on the power of the tests.

3.2 Mann-Kendall Test

Another rank-based non-parametric test is the Mann-Kendall (MK) statistical test, which has been commonly used to assess the significance of trends in time series. The basic principle of MK tests for the trend is to examine the sign of all pairwise differences of observed values. Mann (1945) published the univariate form of such tests at first, while Hoeffding and Robbins (1948), Kendall (1955), and Dietz and Killeen (1981) expanded it to the multivariate situation.

The MK test is based on the test statistic S defined as follows:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sgn}(x_j - x_i), \tag{3.3}$$

where the x_j are the sequential data values, n is the length of the data set, and

$$\operatorname{sgn}(\theta) = \begin{cases} 1 & \text{if } \theta > 0, \\ 0 & \text{if } \theta = 0, \\ -1 & \text{if } \theta < 0. \end{cases}$$

Mann (1945) and Kendall (1955) have documented that when $n \geq 8$, the statistic S is approximately normally distributed with the mean and the variance as follows:

$$\mathbf{E}(S) = 0, \tag{3.4a}$$

$$\operatorname{Var}(S) = \frac{n(n-1)(2n+5) - \sum_p t_p(t_p - 1)(2t_p + 5)}{18}, \tag{3.4b}$$

where p indicates the p-th group of tied variables, and t_p is the size of this group. Such an adjustment for ties was proposed by Helsel et al. (2005). The standardized test statistic Z is computed by

$$Z_{\mathrm{MK}} = \begin{cases} \frac{S-1}{\sqrt{\operatorname{Var}(S)}} & \text{if } S > 0, \\ 0 & \text{if } S = 0, \\ \frac{S+1}{\sqrt{\operatorname{Var}(S)}} & \text{if } S < 0. \end{cases} \tag{3.5}$$

The standardized MK statistic Z follows the standard normal distribution with the mean of zero and variance of one when $n > 10$.

The p-value (probability value, p) of both the MK statistic (S) and the SR statistic (D) of sample data can be estimated using the normal CDF or simulation. If the p-value is small enough, the trend is quite unlikely to be caused by random sampling. For instance, at the significance level of 0.05, if $p \leq 0.05$, then the existing trend is considered to be statistically significant.

Although this test is widely used, a rather incomplete picture of the power of the MK test for the detection of the trend under various circumstances is the current state of the art.

There are several factors with impact on the performance of standard MK tests, such as serial correlation, missing data, and values below the detection limit. Due to the existence of auto-correlation, several seasonal MK tests have been developed (e.g., Robert M. Hirsch et al., 1982). From the practical point of view, missing data have been handled by ignoring the missing data and calculating the test statistic as if the sample is complete. We mainly discuss two directions of development for MK tests.

On the one hand, the variables can affect each other due to some hidden relations. When coping with practical data, we often face up to correlated variables. In order to apply the MK test, we need to get the covariance and use a general formula from Kendall (1955). Hipel and McLeod (1994) proposed a method for such a situation. He also extended the method to the multivariate case.

On the other hand, many factors are affecting the main studied response parameter, which can bias the trend results. To overcome this problem, the partial Mann-Kendall (PMK) tests have been developed. The PMK tests were first proposed by Clark and El-Shaarawi (1993), which can be derived from the general theory of multivariate MK tests by computing the conditional distribution of one MK statistic given a set of other MK statistics. Libiseller and Grimvall (2002) presented different variants of the PMK tests and compared their performance of them.

3.3 Cox-Stuart Test

D. Cox and Stuart (1955) improved the MK test and proposed new statistical testing. However, it requires several assumptions for ensuring linear regression. In other words, it requires that there is a linear relationship between the explanatory and response variables, which is not directly fitted with our case of a binary response variable.

Considering its limitation, we only give some brief review of the Cox-Stuart

(CS) test. Still, for the same denotation, the corresponding test statistic C is as follows:

$$C = \sum_{k=1}^{n/2} (n - 2k + 1) h_{k,n-k+1},$$

where

$$h_{ij} = \begin{cases} 1 & \text{if } y'_i > y'_j, \\ 0 & \text{if } y'_i \leq y'_j. \end{cases}$$

The CS statistic is proved to be asymptotically $C \sim \mathrm{N}(n^2/8, n(n^2-1)/24)$. However, the test is originally designed for considering a linear regression with an upward or downward trend.

3.4 Possible Extensions

In practice, the datasets typically contain several variables which have trends in different directions. This has led to the studies of multivariate techniques. Dietz and Killeen (1981) described an extension of the MK statistic for multivariate data where the trend for individual variables can be in different directions.

An example of such data is given by the matrix $Y_j = (y_{ijk})$ for season j, where the element y_{ijk} is the observation for variable k in season j of year i and $i = 1, 2, \ldots, n$ and $k = 1, 2, \ldots, m$. The test statistic is the quadratic form $S'V^-S$ of the vector, $S = (S_1, S_2, \ldots, S_m)$, of the MK statistics for the m parameters and V, the covariance matrix of S.

Lettenmaier (1988) proposed as a test statistic the sum of the squares of the S_k for $k = 1, 2, \ldots, m$, which he called the covariance eigenvalue (CE) method. To account for the additional dimension of season, the MK statistics are calculated separately for each season within the variable and then summed over the season for each variable.

The performance of these multivariate tests has been evaluated for some water monitoring programs (Lettenmaier, 1988; Ward et al., 1991). Assuming a linear trend, normal or log-normal error distributions, and between-variable and between-season correlations, the first two studies showed that the CE method was more powerful than the method of Dietz and Killeen (1981), called covariance inversion (CI), and, although both had empirical significance levels lower than the nominal, the CE method was less conservative than the CI method. Ward et al. (1991) evaluated modifications that retained dependence between variables but assumed independence between seasons, the latter assumption having been shown to be tenable for the datasets which motivated their study. The major result of the study was to show that these multivariate tests are more powerful than individual tests with nominal significance levels modified by a Bonferroni inequality.

The modified seasonal Kendall trend test given by Robert M Hirsch and Slack (1984) can also be viewed as a multivariate test and has been called the covariance sum test. It was evaluated in the above studies, but, as it does not allow for trends of different directions, it would not always be applicable.

Douglas et al. (2000) give a general framework for multivariate trend tests and show how the tests discussed here fit into this framework, as do tests based on the Spearman trend statistic. Further results are also given on testing the heterogeneity of trend, the null distribution of the CE statistic, and the identification of variables that are important contributors to the overall trend through canonical analysis.

Chapter 4

Modification of Classical Tests for Binary Outputs

We attempt to extend the application of the tests discussed in the last chapter to the case of the binary response variable. For the following simplification, we would like to discuss the realizations of variables, which would be deterministic.

4.1 Modification for Spearman Rank Test

Suppose that $n > 3$ and $\{c^j\}_{j=1}^n$ is ascending ordered. To apply the SR test, we need to obtain the ranks initially

$$R(y_T^j) = \begin{cases} \frac{n_0+1}{2} & \text{if } y_T^j = 0, \\ \frac{n+n_0+1}{2} & \text{if } y_T^j = 1, \end{cases}$$

where n_0 means the number of 0 in the realizations $\{y_T^j\}$. The mean and variance of the ranks can be computed

$$\overline{R(y)} = \overline{y} = \frac{n+1}{2},$$
$$\mathrm{Var}[R(y)] = \frac{(n-n_0)n_0}{4}.$$

Using (3.2), the observed SR statistic should be

$$\hat{D} = \frac{\sum_{j=1}^{n} \left[n+1-2R(y_T^j)\right](n+1-2j)}{n\sqrt{\frac{n_0(n-n_0)(n^2-1)}{3}}}$$

with the variance of $\mathrm{Var}(D) = 1/(n-1)$. Meanwhile, it can be standardized as follows

$$\hat{Z}_{\mathrm{SR}} = \frac{\hat{D}}{\sqrt{\mathrm{Var}(D)}} = \frac{\sum_{j=1}^{n} \left[n+1-2R(y_T^j)\right](n+1-2j)}{n\sqrt{\frac{n_0(n-n_0)(n+1)}{3}}}.$$

Using the same hypotheses in the previous chapter, we assume that H_0 will be rejected at the significance level α. As is mentioned in Section 3.1, Z_{SR} satisfies the standard normal distribution asymptotically. Hence, the p-value of the test should be

$$p = P\left[|Z_{\mathrm{SR}}| \geq |\hat{Z}_{\mathrm{SR}}|\right] = 2\left[1 - \Phi\left(|\hat{Z}_{\mathrm{SR}}|\right)\right].$$

If $p < \alpha$, we can reject H_0. Meanwhile, the critical values are $\Phi^{-1}(1-\alpha/2)$ and $\Phi^{-1}(\alpha/2)$ where $\Phi^{-1}(\cdot)$ is the quantile function of $\mathrm{N}(0,1)$. Supposing H_1 is true, we can obtain the power of the SR test as follows

$$B(\theta) = P\left[\theta\sqrt{n-1} > \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \text{ or } \theta\sqrt{n-1} < \Phi^{-1}\left(\frac{\alpha}{2}\right)\right],$$

where $\theta \neq 0$.

4.2 Modification for Mann-Kendall Test

Before presenting the details, it is necessary to introduce the definitions below.

Definition 4.1. Given a sequence $\{A_i\} = \{A_1, A_2, \dots\}$, we can define an increasing sequence $\{i_k\}$ such that

$$\begin{cases} i_0 = 0, \\ i_k = \min\{i > i_{k-1} : A_{i_{k-1}+1} = \cdots = A_i \neq A_{i+1}\}. \end{cases}$$

The integer $i_k,\ k \geq 1$ is called a *break-point* for the sequence $\{A_i\}$. If $A_{i_k} < A_{i_k+1},\ i_k$ is called a *upwards* break-point. If it is the reverse, i_k is called a *downwards* break-point.

Definition 4.2. Let $\{J_i^+\}$ be the set of upwards break-points for $\{y_T^j\}_{j=1}^n$, where $J_{i-1}^+ < J_i^+$. The size of $\{J_i^+\}$ is denoted by m^+.

Similarly, let $\{J_i^-\}$ be the set of downwards break-points for $\{y_T^j\}_{j=1}^n$, where $J_{i-1}^- < J_i^-$. The size of $\{J_i^-\}$ is denoted by m^-.

If there is no break-point or only one break-point, the observed MK test statistics can be computed using (3.3) as follows

$$\hat{S} = n_0 n_1.$$

Consider that there are two break-points: one is upwards (J^+), the other is downwards (J^-), which means $\hat{m}^+ = 1$ and $\hat{m}^- = 1$. According to the previous definitions, it implies that $y_T^1 = y_T^n$. Without loss of generality, assume that $y_T^1 = 0$. If $\hat{J}^+ > \hat{J}^-$, then $y_T^{\hat{J}^-} = y_T^1,\ y_T^{\hat{J}^-+1} < 0$. It is a contradiction and meanwhile $J^+ \neq J^-$ based on their definitions. Hence, $\hat{J}^+$ is less than $\hat{J}^-$. Because there is only one upwards break-point, $y_T^j = 0$ for all $\hat{J}^- < j \leq n$. In Table 4.1, we show what the realizations $\{y_T^j\}$ looks like.

$j =$	1	$\cdots$	$\hat{J}^+$	$(\hat{J}^+ + 1)$	$\cdots$	$\hat{J}^-$	$(\hat{J}^- + 1)$	$\cdots$	n
$y_T^j =$	0	$\cdots$	0	1	$\cdots$	1	0	$\cdots$	0

Table 4.1: Structure of the Realizations

On the contrary, consider the case that there are more break-points $\hat{J}_i$ where $i = 1, \ldots, m$. According to Table 4.1, we can express the observed test statistic as

$$\begin{aligned}
\hat{S} &= \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} \operatorname{sgn}\left(y_T^i - y_T^j\right) \\
&= \sum_{1 \le j < \hat{J}_1} \sum_{i=j+1}^{n} \operatorname{sgn}\left(y_T^i\right) + \sum_{\hat{J}_1 \le j \le \hat{J}_{\hat{m}}} \sum_{i=j+1}^{n} \operatorname{sgn}\left(y_T^i - y_T^j\right) \\
&\quad + \sum_{\hat{J}_{\hat{m}} < j < n} \sum_{i=j+1}^{n} \operatorname{sgn}\left(y_T^i - y_T^j\right) \\
&= n_1(\hat{J}_1 - 1) + \hat{S}_p + (n - \hat{J}_{\hat{m}})(n_0 - \hat{J}_1 + 2),
\end{aligned}$$

where $n_1 = n - n_0$ indicates the number of 1 in the observations, and

$$\hat{S}_p = \sum_{\hat{J}_1 \le j < i \le \hat{J}_{\hat{m}}} \operatorname{sgn}\left(y_T^i - y_T^j\right).$$

Meanwhile, based on (3.4b), the variance of this statistic should be

$$\operatorname{Var}(S) = \frac{n(n-1)(2n+5) - \sum_{i=0}^{1} n_i(n_i - 1)(2n_i + 5)}{18} = \frac{n_0 n_1 (n+1)}{3}.$$

Using (3.5), the standardized observed test statistic can be obtained

$$\hat{Z}_{\text{MK}} = \frac{\hat{S}}{\sqrt{\operatorname{Var}(S)}} = \frac{n_1(J_1 - 2) + \hat{S}_p + (n - J_m)(n_0 - J_1 + 2) - 1}{\sqrt{\frac{n_0 n_1 (n+1)}{3}}}$$

if $\hat{S} > 0$ that indicates the ascending order exists.

We can do similar procedures for the SR test and then obtain the p-value and power of the MK test as follows

$$\begin{aligned}
p &= 2\left(1 - \Phi(|\hat{Z}_{\text{MK}}|)\right), \\
B(\theta) &= P\left[\frac{\theta - 1}{\sqrt{\operatorname{Var}(S)}} > \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \text{ or } \frac{\theta + 1}{\sqrt{\operatorname{Var}(S)}} < \Phi^{-1}\left(\frac{\alpha}{2}\right)\right],
\end{aligned}$$

where $\theta \neq 0$.

4.3 Modification for Cox-Stuart Test

We will not discuss too many details about the CS test due to the limitation of its assumptions. However, we still give an idea for simplifying the test.

Since the CS test consists of a comparative indicator similar to the sign function, we can employ a similar method as the MK test. In exact, the CS test statistic can be computed as follows

$$\hat{C} = \sum_{k=1}^{n/2} (n - 2k + 1) h_{k,n-k+1},$$

where

$$h_{ij} = \begin{cases} 1 & \text{if } y'_i > y'_j, \\ 0 & \text{if } y'_i \leq y'_j. \end{cases}$$

It is possible to replace the sgn with h in the previous derivation, which is omitted here.

Chapter 5

A New Test for Binary Outputs

Fix the complete probability space $(\Omega, \mathcal{F}, P)$, which is the same as the space mentioned in Section 2.2. Without the loss of generality, assume that $n > 3$ and $\{c^j\}_{j=1}^n$ is ascending ordered, $c^1 < c^2 < \cdots < c^n$. We aim to create a different hypothesis testing with the better performance for the Vasicek model. The hypotheses are still the same as follows:

H_0 : There is no monotonicity of Y_T against c.

H_1 : Y_T monotonically increases as c, or Y_T monotonically decreases as c.

For an integer $1 \leq k < n$, we define the so-called *original monotonicity* as follows.

Definition 5.1. Let μ be a measure on natural numbers $\mathbb{N}$ and $\varepsilon, \eta > 0$ be positive numbers. We say that

- Y_T is (μ, ε, η)-originally increasing against c if

$$\mu(\{j \in \mathbb{N} : 1 \leq j \leq k, P[Y_T^j = 0] < P[Y_T^j = 1] + \eta\}) < \varepsilon$$

  and

$$\mu(\{j \in \mathbb{N} : k+1 \leq j \leq n, P[Y_T^j = 1] < P[Y_T^j = 0] + \eta\}) < \varepsilon,$$

- Y_T is (μ, ε, η)-originally decreasing against c if

$$\mu(\{j \in \mathbb{N} : 1 \leq j \leq k, P[Y_T^j = 1] < P[Y_T^j = 0] + \eta\}) < \varepsilon$$

and

$$\mu(\{j \in \mathbb{N} : k+1 \leq j \leq n, P[Y_T^j = 0] < P[Y_T^j = 1] + \eta\}) < \varepsilon.$$

Let $\mathbf{Y}_l = \{Y_T^j : 1 \leq j \leq k\}$ and $\mathbf{Y}_r = \{Y_T^j : k+1 \leq j \leq n\}$. In the case that the original monotonicity is increasing, some values of Y_T may decrease against c due to the randomness. In other words, it is possible that $Y_T^{j_1} > Y_T^{j_2}$ for $Y_T^{j_1} \in \mathbf{Y}_l$ and $Y_T^{j_2} \in \mathbf{Y}_r$. To examine the effect of randomness, we can employ a weighted sum of difference as the test statistic

$$\begin{aligned} \delta &= \sum_{j=1}^{k} w_j \mathbb{1}_{\{Y_T^j = 1\}} + \sum_{j=k+1}^{n} w_j \mathbb{1}_{\{Y_T^j = 0\}} \\ &= \sum_{j=1}^{k} w_j Y_T^j + \sum_{j=k+1}^{n} w_j (1 - Y_T^j), \end{aligned} \tag{5.1}$$

where $w_j \geq 0$ are weights summing up to 1, Y_T^j and $1 - Y_T^j$ can be viewed as the difference with ideal values. Similarly, the test statistic can be chosen as

$$\begin{aligned} \delta &= \sum_{j=1}^{k} w_j \mathbb{1}_{\{Y_T^j = 0\}} + \sum_{j=k+1}^{n} w_j \mathbb{1}_{\{Y_T^j = 1\}} \\ &= \sum_{j=1}^{k} w_j (1 - Y_T^j) + \sum_{j=k+1}^{n} w_j Y_T^j, \end{aligned}$$

in the case that the original monotonicity is decreasing.

Suppose that there is a positive number δ^* used for giving the result of test. If $\delta \leq \delta^*$, we say that the there is monotonicity, same as the original monotonicity, between Y_T and c and reject the H_0. Meanwhile, the hypotheses can be expressed as

$$H_0 : \delta > \delta^*, \qquad H_1 : \delta \leq \delta^*. \tag{5.2}$$

5.1 Important Parameters

There are two undetermined parameters: k and δ^*. On the one hand, given a relatively smaller number $\alpha_0 \in (0, 1)$, called an *acceptance level*, then consider the optimization problem as follows

$$\max_{\delta^*}\{P[\delta \leq \delta^* | W^T] \leq \alpha_0\} \tag{5.3}$$

or

$$\min_{\delta^*}\{P[\delta > \delta^* | W^T] \leq \alpha_0\}.$$

Note that we consider the conditional probability for a given W_T. We define δ^* as the solution to (5.3). According to this definition, δ^* depends on α_0 and the variables. Hence, its value would be obtained through the distribution of δ.

On the other hand, k is related to the variables and hence it is valued numerically. Let k change from 1 to $n-1$ and compare the result of the experiment. The value of k should have the best experimental performance, which maximizes (5.3). The reason is that as k varies, the result of δ would be different, which also impact on the result.

5.2 The Distribution of Test Statistic

According to the conditional moments of Y_T^j in (2.4), if the original monotonicity is increasing, the conditional expectation and variance of test statistic δ are

$$\mathbf{E}[\delta | W_T] = \sum_{j=1}^{k} w_j p_j + \sum_{j=k+1}^{n} w_j q_j, \tag{5.4a}$$

$$\mathrm{Var}(\delta | W_T) = \sum_{j=1}^{n} w_j^2 p_j q_j, \tag{5.4b}$$

where $p_j = P[Y_T^j = 1 | W_T]$ and $q_j = P[Y_T^j = 0 | W_T] = 1 - p_j$. On the contrary, if the original monotonicity is decreasing, the conditional expectation and variance of test statistic δ should be

$$\mathbf{E}[\delta|W_T] = \sum_{j=1}^{k} w_j q_j + \sum_{j=k+1}^{n} w_j p_j,$$
$$\mathrm{Var}(\delta|W_T) = \sum_{j=1}^{n} w_j^2 p_j q_j.$$

Since $Y_T^1, \ldots, Y_T^n$ are conditionally independent and Bernoulli distributed variables, we can apply the Central Limit Theorem and know that δ is asymptotically normal. This conclusion helps to evaluate the performance of the test for numerical experiments.

5.3 The Selection of Weights

This section provides some examples of weights at first. Then, the optimized weights are derived through the method similar as mean-variance optimization. The performance of these weights would be compared in details in the next chapter.

5.3.1 Uniform Weights

The first example is a natural form of weights. In other words, all values of weights are equal as follows

$$w_j = \frac{1}{n},$$

where $j = 1, 2, \ldots, n$.

5.3.2 Equivalent Weights to Mann-Kendall Test

If the original monotonicity of Y_T against c is increasing, it is possible to derive some form of weights to build a relationship between the new test and MK test. Consider that

$$w_j = \begin{cases} \frac{2(j-1)}{q} & \text{if } j \leq k, \\ \frac{2(n-j)}{q} & \text{if } j > k, \end{cases} \tag{5.5}$$

where $q = n^2 - (2k+1)n + 2k(k-1)$. With the weights in (5.5), we can get the corresponding statistic δ^{MK}. And we have the following relationship

$$S = q\delta^{\mathrm{MK}} + k(n-k) - (n-1)\left[\sum_{j=1}^{k} Y_T^j + \sum_{j=k+1}^{n} (1 - Y_T^j)\right],$$

where S is the test statistic of the MK test.

If the original monotonicity is decreasing, we get similar weights as follows

$$w_j = \begin{cases} \frac{2(n-j)}{q} & \text{if } j \le k, \\ \frac{2(j-1)}{q} & \text{if } j > k, \end{cases}$$

which satisfies

$$S = q\delta^{\mathrm{MK}} - k(n-k) - (n-1)\left[\sum_{j=1}^{k} Y_T^j + \sum_{j=k+1}^{n} (1 - Y_T^j)\right].$$

5.3.3 Probability-Related Weights

Assume that the capital criterion follows the form

$$c^j = f(j, M),$$

where M is not hidden but based on some given information. If the original monotonicity is increasing, we would like to find weights by estimating the range of M. If the original monotonicity is decreasing, we can convert it into the similar case by replacing Y_T^j with $1 - Y_T^j$.

Based on the definition, k should perform as a good splitting point such that

$$p_k(x) = P\left[Y_T^k = 1 \middle| W_T = x\right] < \frac{1}{2} \le P\left[Y_T^{k+1} = 1 \middle| W_T = x\right] = p_{k+1}(x).$$

In other words,

$$P\left[V_T^k < c^k \middle| W_T = x\right] < \frac{1}{2} \le P\left[V_T^{k+1} < c^{k+1} \middle| W_T = x\right].$$

Further,

$$\Phi\left(\frac{f(k,M)-\rho x}{\sqrt{(1-\rho^2)T}}\right) < \frac{1}{2} \le \Phi\left(\frac{f(k+1,M)-\rho x}{\sqrt{(1-\rho^2)T}}\right),$$
$$f(k,M) < \rho x \le f(k+1,M), \tag{5.6}$$

where Φ is the distribution function of N(0,1). Using (5.6), we can firstly estimate the range of M and then the range of $p_j(x)$. Let the range of $p_j(x)$ be (a_j, b_j) and $p_j^{mid} = (a_j + b_j)/2$.

Hence, the probability related weights are defined as follows:

$$v_j = \begin{cases} 1 - p_j^{mid} & \text{if } j \le k, \\ p_j^{mid} & \text{if } j > k, \end{cases}$$

and

$$w_j = \frac{v_j}{\sum_{\ell=1}^{n} v_\ell}.$$

5.3.4 The Optimized Weights

The goal is to find a better weights by optimizing some metrics of the test. Fix α_0 and remain the same hypotheses as (5.2) and regard δ as a function of inputs $\{c^j\}_{j=1}^n$. For robustness of testing, we attempt to interchange two of $\{c^j\}_{j=1}^n$, which is assumed to be an ascending sequence, and then the monotonicity should not hold. Ideally, the test statistic δ is expected to reveal the change of order. Put it more simply, we aim to solve the following minimization:

$$\min_{w_1,\ldots,w_n} \left\{\mathbf{E}^2[\delta|W_T] + a\,\mathrm{Var}(\delta|W_T)\right\} \tag{5.7}$$

subject to

$$\min_{\{\tilde{c}^j\}_{j=1}^n} \mathbf{E}[\delta(\{\tilde{c}^j\}_{j=1}^n)|W_T] \geq \alpha_0, \tag{5.8}$$

$$\sum_{j=1}^n w_j = 1, \tag{5.9}$$

$$w_1, \ldots, w_n \geq 0, \tag{5.10}$$

where $a > 0$ is a *tolerant factor*, and $\tilde{c}^j = c^j$ for all except for two different integers $j_1 \leq k < j_2$, such that $\tilde{c}^{j_2} = c^{j_1}$ and $\tilde{c}^{j_1} = c^{j_2}$.

In the case that the original monotonicity is increasing, come back to the optimization problem (5.7). Initially, we attempt to solve the optimization without the inequality constraints (5.8) and (5.10). Using the conditional moments of δ in (5.4) we can provide the Lagrange function

$$\begin{aligned} L(w_1, \ldots, w_n, \eta) =& \mathbf{E}^2[\delta|W_T] + a \operatorname{Var}(\delta|W_T) - \eta \left(\sum_{j=1}^n w_j - 1 \right) \\ =& \left(\sum_{j=1}^k w_j p_j + \sum_{j=k+1}^n w_j q_j \right)^2 + a \sum_{j=1}^n w_j^2 p_j q_j - \eta \left(\sum_{j=1}^n w_j - 1 \right), \end{aligned} \tag{5.11}$$

where w_j should be non-negative for all $1 \leq j \leq n$.

Subsequently, the first order derivative of L over w_j should be

$$\frac{\partial L}{\partial w_j} = \begin{cases} 2p_j S_\sigma + 2ap_j q_j w_j - \eta & \text{if } 1 \leq j \leq k, \\ 2q_j S_\sigma + 2ap_j q_j w_j - \eta & \text{if } j > k, \end{cases} \tag{5.12}$$

where $S_\sigma = \mathbf{E}[\delta|W_T] = \sum_{j=1}^k w_j p_j + \sum_{j=k+1}^n w_j q_j$. Using the first order condition, the linear system can be obtained as follows

$$\begin{cases} 2p_j S_\sigma + 2ap_j q_j w_j = \eta & \text{if } 1 \leq j \leq k, \\ 2q_j S_\sigma + 2aq_j p_j w_j = \eta & \text{if } j > k, \\ \sum_{j=1}^n w_j = 1. \end{cases} \tag{5.13}$$

According to (5.13), we have the following proposition.

Proposition 5.1. *The optimized weights are as follows*

If $j \le k$, then

$$w_j = \frac{1}{q_j} \cdot \frac{1}{\sum_{\ell=1}^{n} \frac{1}{p_\ell q_\ell} - \frac{U^2}{V+a}} \left(\frac{1}{p_j} - \frac{U}{V+a} \right).$$

If $j > k$, then

$$w_j = \frac{1}{p_j} \cdot \frac{1}{\sum_{\ell=1}^{n} \frac{1}{p_\ell q_\ell} - \frac{U^2}{V+a}} \left(\frac{1}{q_j} - \frac{U}{V+a} \right),$$

where

$$U = \sum_{\ell=1}^{k} \frac{1}{q_\ell} + \sum_{\ell=k+1}^{n} \frac{1}{p_\ell}, \qquad V = \sum_{\ell=1}^{k} \frac{p_\ell}{q_\ell} + \sum_{\ell=k+1}^{n} \frac{q_\ell}{p_\ell}.$$

The proof can be seen in Appendix A.1. Meanwhile, the sum of weights should be

$$S_\sigma = \frac{a}{\sum_{j=1}^{n} \frac{1}{p_j q_j} - \frac{U^2}{V+a}} \cdot \frac{U}{V+a}.$$

The next is to consider a sequence $\{\tilde{c}^j\}_{j=1}^n$ such that $\tilde{c}^j = c^j$ except for $\tilde{c}^{j_1} = c^{j_2}$ and $\tilde{c}^{j_2} = c^{j_1}$. Similarly, the conditional probabilities can be denoted by $\tilde{p}_j$ and $\tilde{q}_j$, where $\tilde{p}_j + \tilde{q}_j = 1$. Assume that k becomes $\tilde{k}$ as the order of $\{c_j\}$ changes. In this context, the values and sum of weights would be $\tilde{w}_j$ and $\tilde{S}$, respectively. Furthermore, the constraint (5.8) is equivalent to that for all

$j_1 \leq k < j_2$

$$\begin{aligned}
\mathbf{E}[\delta(\{\tilde{c}^j\}_{j=1}^n)|W_T] &= \sum_{j=1}^{\tilde{k}} \tilde{w}_j \tilde{p}_j + \sum_{j=\tilde{k}+1}^{n} \tilde{w}_j \tilde{q}_j \\
&= \tilde{S}_\sigma \\
&= \frac{a}{\sum\limits_{j=1}^{n} \frac{1}{\tilde{p}_j \tilde{q}_j} - \frac{\tilde{U}^2}{\tilde{V}+a}} \cdot \frac{\tilde{U}}{\tilde{V}+a} \\
&\geq \alpha_0,
\end{aligned}$$

where

$$\tilde{U} = \sum_{\ell=1}^{\tilde{k}} \frac{1}{\tilde{q}_\ell} + \sum_{\ell=\tilde{k}+1}^{n} \frac{1}{\tilde{p}_\ell}, \qquad \tilde{V} = \sum_{\ell=1}^{\tilde{k}} \frac{\tilde{p}_\ell}{\tilde{q}_\ell} + \sum_{\ell=\tilde{k}+1}^{n} \frac{\tilde{q}_\ell}{\tilde{p}_\ell}.$$

Finally, combining the above inequality with (5.14), there are several restrictions about a for the rightness of constraints as the below proposition states.

Proposition 5.2. *The tolerant factor a and acceptance level α_0 should satisfy that*

for any $1 \leq j \leq k$,

$$\frac{1}{q_j} \cdot \frac{1}{\sum\limits_{\ell=1}^{n} \frac{1}{p_\ell q_\ell} - \frac{U^2}{V+a}} \left(\frac{1}{p_j} - \frac{U}{V+a} \right) \geq 0; \tag{5.15a}$$

for any $k < j \leq n$,

$$\frac{1}{p_j} \cdot \frac{1}{\sum\limits_{\ell=1}^{n} \frac{1}{p_\ell q_\ell} - \frac{U^2}{V+a}} \left(\frac{1}{q_j} - \frac{U}{V+a} \right) \geq 0; \tag{5.15b}$$

for any $1 \le j_1 \le k < j_2 \le n$*,*

$$\frac{a}{\sum_{j=1}^{n} \frac{1}{\tilde{p}_j \tilde{q}_j} - \frac{\tilde{U}^2}{\tilde{V}+a}} \cdot \frac{\tilde{U}}{\tilde{V}+a} \ge \alpha_0. \tag{5.15c}$$

Note that $U > 1$, $\tilde{U} > 1$, $U > V$ and $\tilde{U} > \tilde{V}$. Despite that we may calculate the range of a, we want to focus on the existence of the best weights rather than the complete selection of a. Therefore, we only provide a sufficient condition.

Proposition 5.3. *There is a sufficient condition of Proposition 5.2 that*

$$a \ge \max(p_n, q_1) \cdot U - V \tag{5.16}$$

and

$$\alpha_0 < \min(p_1, q_n). \tag{5.17}$$

Proof. We know that (5.15a) is equivalent to

$$\begin{cases} \frac{1}{p_j} \ge \frac{U}{V+a} \\ \sum_{\ell=1}^{n} \frac{1}{p_\ell q_\ell} > \frac{U^2}{V+a} \end{cases} \quad \text{or} \quad \begin{cases} \frac{1}{p_j} \le \frac{U}{V+a} \\ \sum_{\ell=1}^{n} \frac{1}{p_\ell q_\ell} < \frac{U^2}{V+a} \end{cases}$$

for all $1 \le j \le k$. Since p_j is increasing, we have

$$\begin{cases} \frac{1}{p_k} \ge \frac{U}{V+a} \\ \sum_{j=1}^{n} \frac{1}{p_j q_j} > \frac{U^2}{V+a} \end{cases} \quad \text{or} \quad \begin{cases} \frac{1}{p_1} \le \frac{U}{V+a} \\ \sum_{j=1}^{n} \frac{1}{p_j q_j} < \frac{U^2}{V+a} \end{cases}$$

Then

$$\begin{cases} a \ge p_k U - V \\ a > \frac{U^2}{\sum_{j=1}^{n} \frac{1}{p_j q_j}} - V \end{cases} \quad \text{or} \quad \begin{cases} 0 < a \le p_1 U - V \\ a < \frac{U^2}{\sum_{j=1}^{n} \frac{1}{p_j q_j}} - V \end{cases}$$

According to the Cauchy–Schwarz inequality, we can get that

$$
\begin{aligned}
V\sum_{j=1}^{n}\frac{1}{p_jq_j} &= \left(\sum_{j=1}^{k}\frac{p_j}{q_j}+\sum_{j=k+1}^{n}\frac{q_j}{p_j}\right)\sum_{j=1}^{n}\frac{1}{p_jq_j} \\
&= \left[\sum_{j=1}^{k}\left(\sqrt{\frac{p_j}{q_j}}\right)^2+\sum_{j=k+1}^{n}\left(\sqrt{\frac{q_j}{p_j}}\right)^2\right]\sum_{j=1}^{n}\left(\sqrt{\frac{1}{p_jq_j}}\right)^2 \\
&\geq \left(\sum_{j=1}^{k}\frac{1}{q_j}+\sum_{j=k+1}^{n}\frac{1}{p_j}\right)^2 = U^2,
\end{aligned}
$$

which means that $U^2/[\sum_{j=1}^{n} 1/(p_jq_j)] - V \leq 0$. Consequently, the set of available a should be

$$
\begin{cases} a \geq p_kU - V, \\ a > 0. \end{cases} \tag{5.18}
$$

Similarly, the set of all solutions to (5.15b) is

$$
\begin{cases} a \geq q_kU - V, \\ a > 0. \end{cases} \tag{5.19}
$$

Remark 5.3.1. What is necessary to mention is that $p_kU - V$ and $q_kU - V$ can be any real number. To illustrate, suppose that $n = 3, k = 2$ and $p_1 = 1/16, p_2 = 1/8, p_3 = 1/4$. At this point,

$$
p_kU - V = \frac{1}{8}\left(\frac{16}{15}+\frac{8}{7}+4\right)-\left(\frac{1}{15}+\frac{1}{7}+3\right) = -\frac{73}{30} < 0.
$$

Conversely, for the same n and k let $p_1 = 1/2, p_2 = 3/4, p_3 = 7/8$, resulting in

$$
p_kU - V = \frac{3}{4}\left(2+4+\frac{8}{7}\right)-\left(1+3+\frac{1}{7}\right) = \frac{17}{14} > 0.
$$

Note that $p_n > p_k$, $q_1 > q_k$ and

$$
\max(p_n, q_1)\cdot U - V \geq \left(\sum_{j=1}^{k}\frac{p_n}{q_j}+\sum_{j=k+1}^{n}\frac{q_1}{p_j}\right)-\sum_{j=1}^{k}\frac{p_j}{q_j}-\sum_{j=k+1}^{n}\frac{q_j}{p_j} > 0.
$$

Therefore, we can know that both (5.18) and (5.19) are satisfied.

Before considering (5.15c), note that for any $j \neq j_1, j_2$ and $j_1 \leq k < j_2$

$$\tilde{q}_j = 1 - \tilde{p}_j, \qquad \tilde{p}_j = p_j, \qquad \tilde{p}_{j_1} = p_{j_2}, \qquad \tilde{p}_{j_2} = p_{j_1}.$$

As a result,

$$\sum_{j=1}^{n} \frac{1}{\tilde{p}_j \tilde{q}_j} = \sum_{j=1}^{n} \frac{1}{p_j q_j}.$$

Return to (5.15c), it implies that for any $j_1 \leq k < j_2$

$$\begin{cases} a\tilde{U} - \alpha_0(\tilde{V} + a)\sum_{j=1}^{n} \frac{1}{\tilde{p}_j \tilde{q}_j} + \alpha_0 \tilde{U}^2 \geq 0, \\ \sum_{j=1}^{n} \frac{1}{\tilde{p}_j \tilde{q}_j} > \frac{\tilde{U}^2}{\tilde{V}+a}, \end{cases}$$

or

$$\begin{cases} a\tilde{U} - \alpha_0(\tilde{V} + a)\sum_{j=1}^{n} \frac{1}{\tilde{p}_j \tilde{q}_j} + \alpha_0 \tilde{U}^2 \leq 0, \\ \sum_{j=1}^{n} \frac{1}{\tilde{p}_j \tilde{q}_j} < \frac{\tilde{U}^2}{\tilde{V}+a}. \end{cases}$$

We still employ the Cauchy-Schwarz inequality to get $\tilde{U}^2/[\sum_{j=1}^{n} 1/(\tilde{p}_j \tilde{q}_j)] \leq \tilde{V}$. Then, we obtain that for any $j_1 \leq k < j_2$

$$\begin{cases} \left(\tilde{U} - \alpha_0 \sum_{j=1}^{n} \frac{1}{p_j q_j}\right) a \geq \alpha_0 \left(\tilde{V} \sum_{j=1}^{n} \frac{1}{p_j q_j} - \tilde{U}^2\right), \\ a > 0. \end{cases}$$

Note that $\tilde{U} < \sum_{j=1}^{n} 1/(\tilde{p}_j \tilde{q}_j) = \sum_{j=1}^{n} 1/(p_j q_j)$. To assure the existence of a, for each pair of j_1 and j_2, $\alpha_0 > 0$ should be small enough to satisfy

$$\tilde{U} - \alpha_0 \sum_{j=1}^{n} \frac{1}{p_j q_j} > 0. \tag{5.20}$$

According to (5.17), we know that $\alpha_0 < p_j$ and $\alpha_0 < q_j$ for all $1 \leq j \leq n$. Furthermore, the order of sequence is the only difference between $\{\tilde{c}^j\}_{j=1}^n$ and

$\{c^j\}_{j=1}^n$ so that $\alpha_0 < \tilde{p}_j$ and $\alpha_0 < \tilde{q}_j$ for all $1 \le j \le n$. Then we know that

$$\begin{aligned}\tilde{U} - \alpha_0 \sum_{j=1}^{n} \frac{1}{p_j q_j} &= \sum_{j=1}^{\tilde{k}} \frac{1}{\tilde{q}_j} + \sum_{j=\tilde{k}+1}^{n} \frac{1}{\tilde{p}_j} - \alpha_0 \sum_{j=1}^{n} \frac{1}{\tilde{p}_j \tilde{q}_j} \\ &= \sum_{j=1}^{\tilde{k}} \frac{\tilde{p}_j - \alpha_0}{\tilde{p}_j \tilde{q}_j} + \sum_{j=\tilde{k}+1}^{n} \frac{\tilde{q}_j - \alpha_0}{\tilde{p}_j \tilde{q}_j} > 0.\end{aligned}$$

Hence, the last step is to show that for any possible j_1 and j_2, (5.16) is a subset of solutions to

$$a \ge \alpha_0 \frac{\tilde{V} \sum\limits_{j=1}^{n} \frac{1}{p_j q_j} - \tilde{U}^2}{\tilde{U} - \alpha_0 \sum\limits_{j=1}^{n} \frac{1}{p_j q_j}}. \tag{5.21}$$

Put it another way, rearranging terms in (5.21), we can obtain that

$$\begin{aligned} a &\ge \frac{\alpha_0 \tilde{V} \sum\limits_{j=1}^{n} \frac{1}{p_j q_j} - \alpha_0 \tilde{U}^2}{\tilde{U} - \alpha_0 \sum\limits_{j=1}^{n} \frac{1}{p_j q_j}} \\ &= \frac{\tilde{V} \left(\alpha_0 \sum\limits_{j=1}^{n} \frac{1}{p_j q_j} - \tilde{U} \right) + \tilde{U}\tilde{V} - \alpha_0 \tilde{U}^2}{\tilde{U} - \alpha_0 \sum\limits_{j=1}^{n} \frac{1}{p_j q_j}} \\ &= \frac{\tilde{V} - \alpha_0 \tilde{U}}{\tilde{U} - \alpha_0 \sum\limits_{j=1}^{n} \frac{1}{p_j q_j}} \tilde{U} - \tilde{V}. \end{aligned} \tag{5.22}$$

Then based on (5.17), we can obtain that

$$\frac{\tilde{V} - \alpha_0 \tilde{U}}{\tilde{U} - \alpha_0 \sum\limits_{j=1}^{n} \frac{1}{p_j q_j}} = \frac{\sum\limits_{j=1}^{\tilde{k}} \tilde{p}_j \frac{\tilde{p}_j - \alpha_0}{\tilde{p}_j \tilde{q}_j} + \sum\limits_{j=\tilde{k}+1}^{n} \tilde{q}_j \frac{\tilde{q}_j - \alpha_0}{\tilde{p}_j \tilde{q}_j}}{\sum\limits_{j=1}^{\tilde{k}} \frac{\tilde{p}_j - \alpha_0}{\tilde{p}_j \tilde{q}_j} + \sum\limits_{j=\tilde{k}+1}^{n} \frac{\tilde{q}_j - \alpha_0}{\tilde{p}_j \tilde{q}_j}} < \frac{p_n \sum\limits_{j=1}^{\tilde{k}} \frac{\tilde{p}_j - \alpha_0}{\tilde{p}_j \tilde{q}_j} + q_1 \sum\limits_{j=\tilde{k}+1}^{n} \frac{\tilde{q}_j - \alpha_0}{\tilde{p}_j \tilde{q}_j}}{\sum\limits_{j=1}^{\tilde{k}} \frac{\tilde{p}_j - \alpha_0}{\tilde{p}_j \tilde{q}_j} + \sum\limits_{j=\tilde{k}+1}^{n} \frac{\tilde{q}_j - \alpha_0}{\tilde{p}_j \tilde{q}_j}},$$

and hence

$$\frac{\tilde{V} - \alpha_0 \tilde{U}}{\tilde{U} - \alpha_0 \sum_{j=1}^{n} \frac{1}{p_j q_j}} \leq \max(p_n, q_1).$$

$\square$

Appendix A.2 demonstrates some ideas for extending this proposition.

Chapter 6

Numerical Experiments

6.1 Setups of Simulation

Now that we have a new test for detecting monotonicity, we would like to use Monte Carlo simulation to generate the data for examining its performance. Let N be the number of simulations with the value of 1000. For each simulation, the criteria $\{c^j\}$ has a size of $n = 100$ and is ascending ordered based on our assumption. Meanwhile, the observations of $\{y_T^j\}$ are obtained by generating individual random variables $\{B_T^j\}$ in (2.2) with the maturity $T = 1$ and the correlation ρ varying between 0 and 1. According to the previous discussion, we mainly consider conditional probability with respect to the Vasicek model, and hence, we set the condition to be $W_T = 0.1$, which equals to the risk-free rate.

To employ a well-rounded experiment, the total data $\{(c^j, y_T^j)\}$ are equally divided into four groups: increasing, almost equal (most of y_T^j are the same), nearly no monotonicity (weakly increasing), and no monotonicity. For each group, we choose a different range of $\{c^j\}$ to control the monotonicity between y_T^j and c^j. Figure 6.1 illustrates 4 scatter plots of the corresponding situations.

In Figure 6.1, we could observe that the upper left plot has sparse points for $c^j < 0, y_T^j = 1$ and $c^j \geq 0, y_T^j = 0$, which implies an increasing relationship between y_T^j and c^j. The upper right one is relatively a special case, due to fewer points with $y_T^j = 1$, that indicates that is not strictly increasing. According

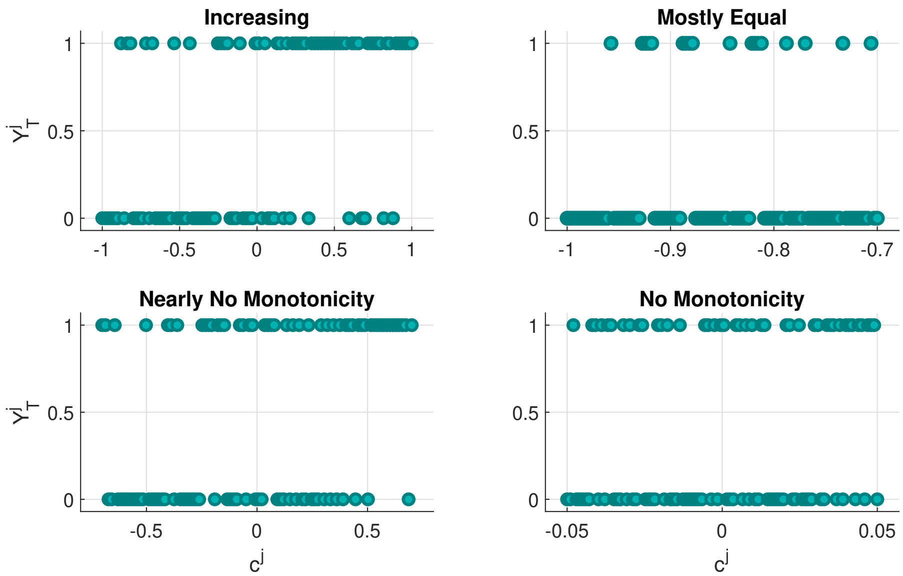

Figure 6.1: Examples of Four Groups in Total Data

to our expectation, the testing method should detect the monotonicity in this case. Increasing the number of points for $c^j < 0, y^j_T = 1$ and $c^j \geq 0, y^j_T = 0$, we got the other two plots that are considered to be weakly increasing and not monotonic, respectively. In these situations, the monotonicity could be more sensitive due to the randomness, and hence, we labelled the monotonicity for each simulation. Out of the same consideration, we only choose the part of data showing nearly no monotonicity when comparing the power of different tests.

6.2 Estimation of Parameters

For our new test, there are three parameters to be determined, which is discussed in Section 5.1. The values of a and α_0 are restricted by Proposition 5.3, which suggests that $a = \max(p_n, q_1) \cdot U - V + 1$ and $\alpha_0 = \min(p_1, q_n)/2$. The parameter k is determined by minimizing the objective function in (5.7). The acceptance level α_0 controls the critical value of the new test, while the tolerant factor a affect the derivation of weights. When changing $\{c^j\}$, the conditional probabilities p_j and q_j can be so different that the effect of a vary. We found that eliminating the monotonicity has an impact on the effect of a. To illustrate, we explored the conditional moments of the test statistic δ, the objective function for finding weights, and the values of k appearing in (5.1).

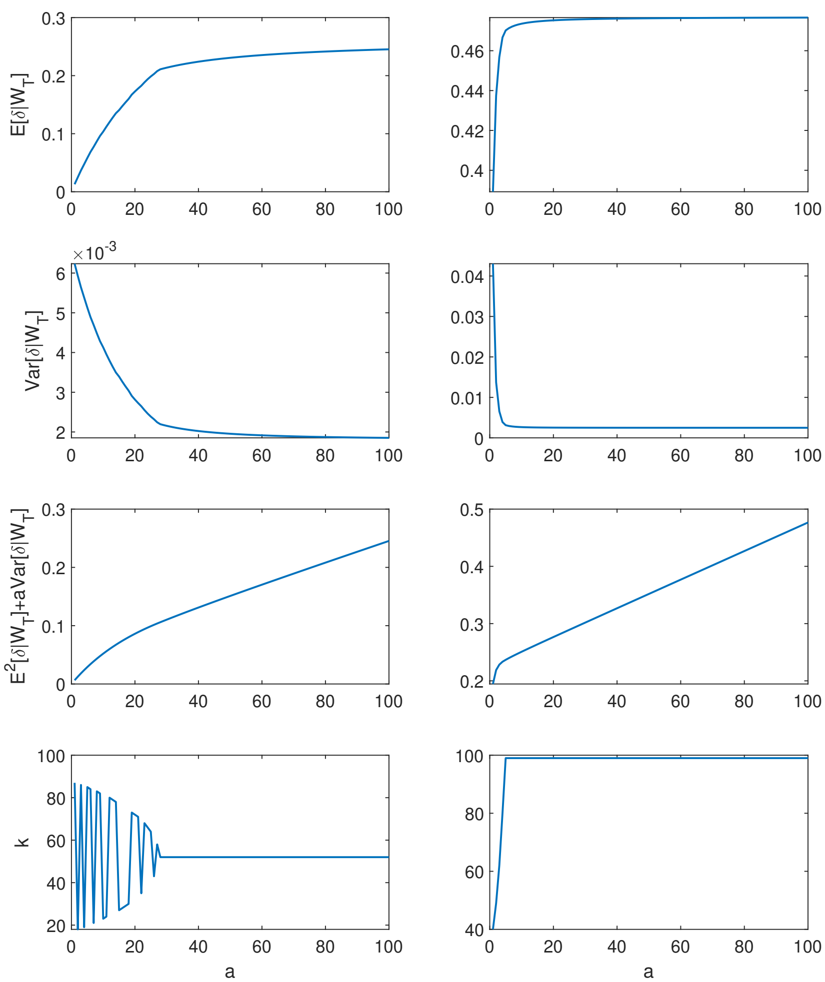

Figure 6.2: Performance of New Test against a

Figure 6.2 depicts the performance of the new test when changing a in the situations of increasing and no monotonicity. The left four line plots are for the increasing case, while the right four line plots are for the case of no monotonicity. At first glance, we found that as a increases, δ's conditional expectation and the objective function increases, δ's conditional variance decreases, and k fluctuates initially and stays the same for larger a. Especially, the conditional moments of the test statistics show a stable trend when a is increasing. On the other hand, the stabilization of these values over a are magnificent when there is no monotonicity. Meanwhile, the stable value of k are different and k even takes the maximum of its range $(n-1)$ when there is no monotonicity, which seems to be reasonable referring to our definition of δ. The findings suggests that we should change a for different $\{c^j\}$, and it is worthwhile to take a smaller a for minimizing the objective function. In other words, it is better to find the value of a numerically rather than using the sufficient condition in Proposition 5.3.

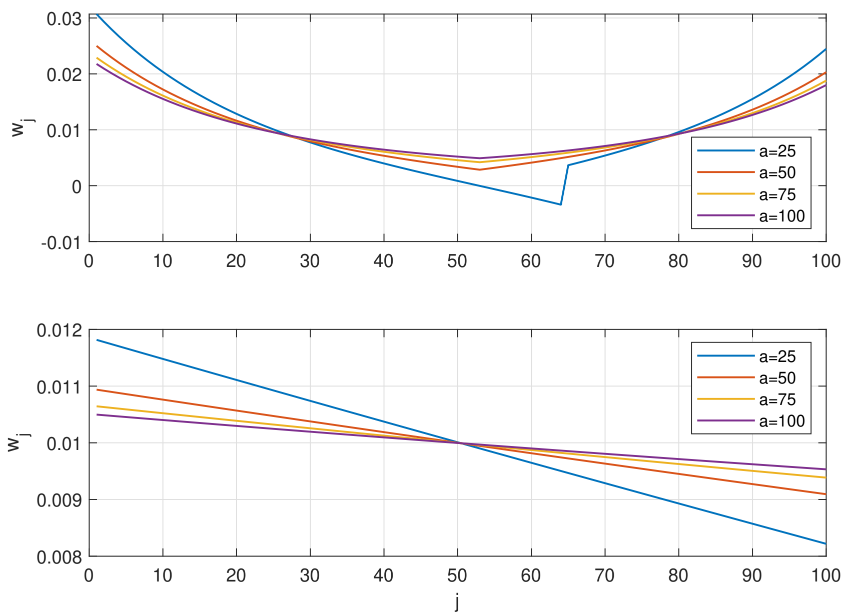

Figure 6.3: Optimal Weights for Different a

To delve into the underlying of the effect of a, we compared the best weights for several values of a. The illustration are shown in Figure 6.3, where the top plot is for increasing and the bottom one is for no monotonicity. On the one hand, the best weight decreases for $j \leq k$ and then increases for $j > k$. Meanwhile, the magnitude of these changes are much slighter as a becomes larger. Further, it should be noted that the best weight for increasing data goes negative if $a = 25$, which could cause the fluctuations of k in Figure 6.2. On the other hand, the optimal weights for no monotonicity show a decline over j as a result of $k = n - 1$. That could explain our findings from Figure 6.2 since the absolute value of gradients are bigger when the monotonicity is diminishing. Lastly, it indicates that we could use the minimal a that makes all of the weights positive to improve the performance of the new test.

6.3 Results

Now that we had methodologies to get all of the parameters. The next step is to examine the effect of ρ, which is a correlation in the Vasicek model. From the theoretical perspective, the correlation ρ would affect the conditional probabilities p_j, q_j, and δ. In a financial sense, the correlation determines how the individuals impact on each other. Therefore, it is worthwhile to study the performance of the test against ρ. Figure 6.4 depicts the line plots of δ's conditional expectation and variance over ρ. The left two plots are for increasing and the right two are for no monotonicity.The other two situations are not presented since they look similar to these plots. On the one hand,

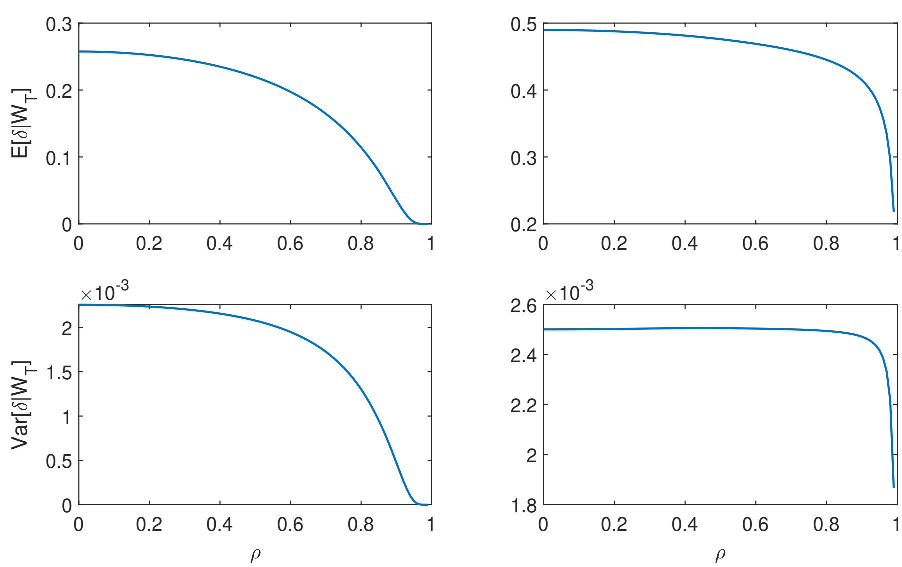

Figure 6.4: Conditional Moments of δ against Correlation ρ

the values of the conditional moments decreases as ρ grows. Especially, the decreasing speed (the absolute value of the first derivative) are larger as ρ comes closer to 1. On the other hand, the magnitude of declining for no monotonicity is greater than that of the increasing case.

The monotonicity shown in Figure 6.4 seems to originate from the relationship between the optimal weights and ρ and the formula of δ. For the former one, we apply the similar methods used for analyzing the effect of a to it. The line plots of the optimal weights for different values of ρ are illustrated in Figure 6.5, where the top plot has increasing data and the bottom one has no monotonic data. Observing the upper graph, we can find the pattern of weights still be that the value decreases for $j \leq k$ and increases for $j > k$, which seems similar to the plot of the optimal weights over tolerant factor a. However, the second plot demonstrates difference that the value of weights has the fair same

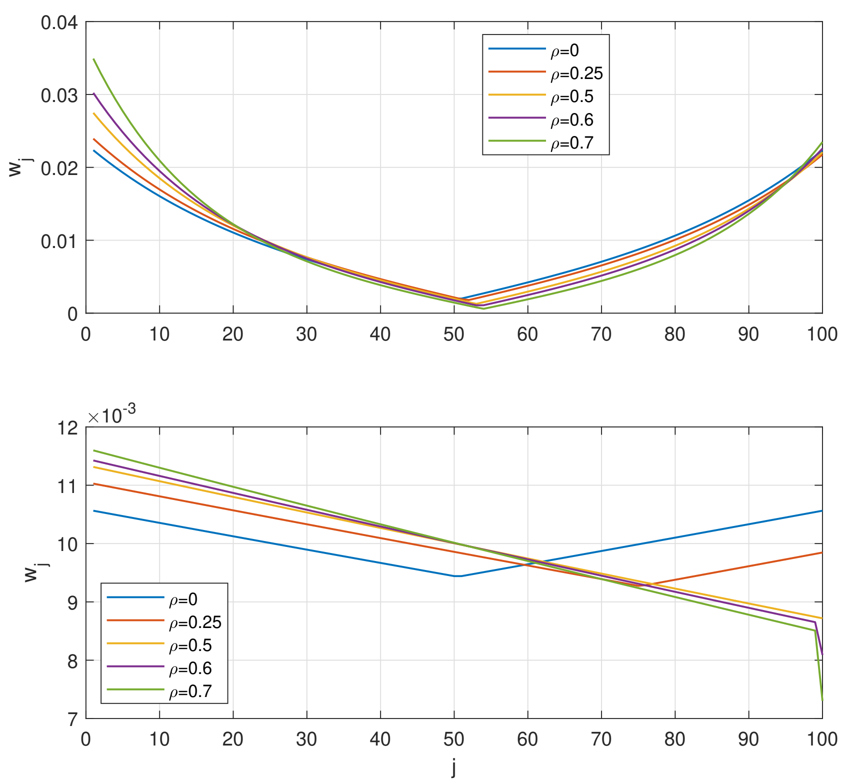

Figure 6.5: Optimal Weights for Different ρ

pattern as the increasing data plot if ρ is small enough. Furthermore, that the value of k is shifting towards right seems to be the grounds that drives the decreasing speed of the moments faster. However, it is notable that the magnitude of increasing and no monotonicity are different, which ensures the correctness of the test.

From the previous results, it seems the shape of the weights curve (the relationship between w_j and j) plays a role in detecting monotonicity. Hence, we attempt to compare the performance of the new test with some different weights introduced in Section 5.3. Table 6.1 lists the value of test statistic δ and testing results p-value of three methods for constructing the new test. The table does not contain the MK-equivalent weights that would be examined later. To exclude the influence of biased results due to different $\{c^j\}$, we compare the performance on the nearly-no-monotonicity data as a representative group. Meanwhile, we choose an appropriate a to make k fixed based on our previous findings. On the other hand, we would like to mitigate the impact of correlation ρ, which is examined before. To illustrate, we can observe that the difference caused by changing ρ is slight for the value of δ or p-value. Therefore, we could focus on studying the factor of the weights curve. To combine the testing results for each simulation, we reported the medians of test statistic δ and then employed the bootstrapping method to generate sample data for calculating empirical p-value. We did not use the analytical p-value out of the consideration that the new test depends on some parameters, such as k, that are sensitive to data. The data generating process is still based on Monte Carlo method with the same $\{c^j\}$. Now that we obtain all values in Table 6.1, we could say that

		$\rho = 0$	$\rho = 0.25$	$\rho = 0.5$	$\rho = 0.75$
Optimal	δ	0.1410	0.1407	0.1399	0.1382
	p-Value	0.0160	0.0160	0.0120	0.0120
Uniform	δ	0.1531	0.1531	0.1531	0.1531
	p-Value	0.0240	0.0240	0.0240	0.0240
Probability Related	δ	0.1531	0.1529	0.1572	0.1523
	p-Value	0.0240	0.0240	0.0240	0.0240

Table 6.1: Summary for Tests with Different Weights

the weights curves bring about the difference of testing results to some extent. According to Table 6.1, the testing with the optimal weights outperformed the testings with the other two weights. It is interesting that the probability-related weights have a similar performance as the uniform distributed weights. That may be because the setting of $c^j = f(j, M)$ are not very complicated, which deserves some further discussions.

To figure out the comparative power of the new test, we attempt to apply three hypothesis testings on the simulation data. Table 6.2 demonstrates the statistical significance, power and effort (used time) of new test, the SR test and the MK test. There is no result of the CS test because of two reasons. One reason is that the assumption of the CS test is not fitted with the situation, which needs to be a linear relationship. Another reason is that the performance of CS test is not comparable with other testing methods even if applying some transformations in Chapter 4.

		$\rho = 0$	$\rho = 0.25$	$\rho = 0.5$	$\rho = 0.75$
New	p-Value	0.0144	0.0168	0.0128	0.0068
	Power	0.9856	0.9832	0.9872	0.9932
	Elapsed Time	0.0159	0.0165	0.0161	0.0155
SR	p-Value	0.1352	0.1236	0.0792	0.0124
	Power	0.8648	0.8764	0.9208	0.9876
	Elapsed Time	0.7275	0.7391	0.7269	0.7335
MK	p-Value	0.1364	0.1244	0.0792	0.0132
	Power	0.8636	0.8756	0.9208	0.9868
	Elapsed Time	0.1956	0.1974	0.1926	0.1832

Table 6.2: Comparative Power of Different Tests

According to Table 6.2, we can see that new test outperforms the other two tests with lower computation time for the total 10,000 simulations. The SR test cost the most time for running the program, whereas the fastest new test only cost approximately 1/50 of its time. That follows the previous analysis in Chapter 3 that the SR test computes the covariance and the MK test compares pair-wise values, while the new test can mainly use arithmetic operators. On the other hand, the significance of the three tests are relatively similar. Under the circumstance that y_T^j is mostly equal or nearly not monotonic against c^j, the new test brings about a lower p-value and a higher power of the test. On the contrary, the testing significance p-values and powers are not as well as the SR and MK tests.

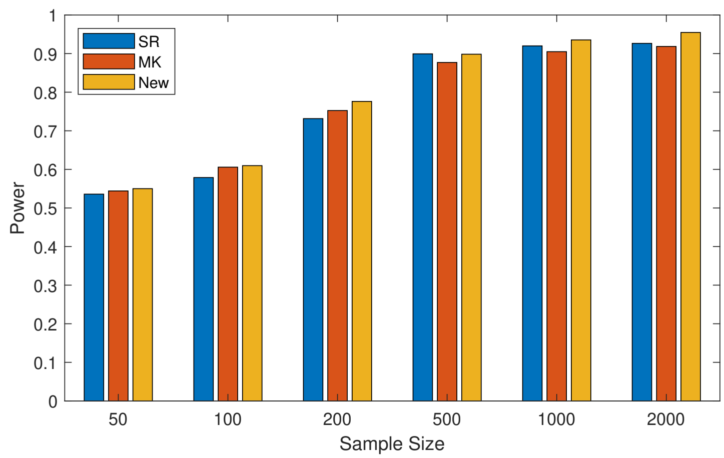

Figure 6.6: Power of Different Tests against Sample Size

To get some further information about the three testing methods, we randomly sample the data to compute the power of the test, which is shown in Figure 6.6. The supposing alternative hypothesis is that y_T^j is nearly not monotonic against c^j, which account for 25% of the total simulation. Observing the figure, we could view that all three tests have a growing curve of the power and the difference actually is not large. Further, it demonstrates that all powers exceed 0.8 when the sample size is 500 or above, which is an evidence of the ability to detect the monotonicity. Meanwhile, all values of the power almost stay stable as the sample size reach at 1000. The findings offer us some insights for future application.

Chapter 7

Conclusion

In credit risk modelling, practitioners maintain monotonic relationships regularly when dealing with banking data whose sizes are much more considerable nowadays. The subsequent issue, lacking enough specific studies, is how to precisely detect the monotonicity, particularly for variables in a huge dataset. Even though there are a few widely-used hypothesis testing methods, we still attempt to pursue simplified statistical tests for the binary response variable that frequently occurs in banking data. Using the Vasicek model framework, we modified three existing monotonicity tests for the binary credit risk indicator. We introduced crucial assumptions and pros and cons of these testing methods in details. In addition, we briefly discussed their computational complexity as a performance metrics.

Motivated by the MK test and cross-entropy, we propose a new hypothesis test using a weighted sum of difference as the statistics with flexible parameters. Via a similar methodology for the mean-variance optimization, we analytically solved for the optimal weights, and restricted the parameters in some regions. There are also other options for weights that we supplement. In particular, we employ a weight specification to convert the new test into the MK test. Besides, the asymptomatic distribution of the test statistic δ is derived from central limit theorem and moment generating function under suitable assumptions.

After laying these theoretical foundations for the topic, we implement the statistical simulation with four scenarios to examine the impact of parameters from the Vasicek model and our new test. The analysis on the results of the

10,000 simulations demonstrates that the conditional moments of the statistics δ of the optimal new test reduce as the correlation ρ in the Vasicek model grows from 0 to 1. This is because most of the conditional moments except for the conditional variance are positively related to the tolerance factor a. Furthermore, both parameters have dominant influence on the optimal weights. Based on this conclusion, we investigated the effect of the shape of the weights curve by comparing several weights. It indicates that the different options for weights can affect the performance of our new test to some extent. Last but not least, by comparing all mentioned tests, we find that the new test can perform as well as the existing with a remarkably lower computation time. In some extreme case, our new test outperforms the other tests. In general, not only does our new test improve the performance of monotonicity testing for some extreme situations, but it also notably reduces the computation time for monotonicity testing.

There are still some potential topics for further discussion. On the one hand, we can enhance the methodology of credit risk modelling or create a monotonicity testing based on variants of the SR and MK tests. We mainly studied the conditional probability under the Vasicek model framework. It is possible to expand it to either marginal probability for the Vasicek model or another credit risk model. In addition, some variants are likely to perform well for the binary response variable. On the other hand, we may improve the design of the new test. For the optimization of weights, it is possible to employ some other objective function from Mathematical Finance, such as log-utility. Moreover, the formula of the test statistics can be more complicated for flexibility by adding more parameters.

Bibliography

Altman, Edward I. "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy". In: *The Journal of Finance* 23.4 (1968), pp. 589–609.

Benzschawel, Terry. *Credit Risk Modelling-Facts, Theory and Applications*. Risk books, 2012.

Black, Fischer and John D. Cox. "Valuing corporate securities: Some effects of bond indenture provisions". In: *The Journal of Finance* 31.2 (1976), pp. 351–367.

Black, Fischer and Myron Scholes. "The Pricing of Options and Corporate Liabilities". In: *Journal of Political Economy* 81.3 (1973), pp. 637–654.

Clark, James A and AH El-Shaarawi. "Evaluation of commercial presence-absence test kits for detection of total coliforms, Escherichia coli, and other indicator bacteria". In: *Applied and Environmental Microbiology* 59.2 (1993), pp. 380–388.

Cox, D. and A. Stuart. "Some Quick Sign Tests for Trend in Location and Dispersion". In: *Biometrika* 42.1/2 (1955), p. 80.

Credit Suisse First Boston. *CreditRisk+: A credit risk management framework*. Tech. rep. Technical report, Credit Suisse First Boston, 1997.

Crouhy, Michel, Dan Galai, and Robert Mark. "A comparative analysis of current credit risk models". In: *Journal of Banking and Finance* 24.1-2 (2000), pp. 59–117.

Dietz, E. Jacquelin and Timothy J. Killeen. "A Nonparametric Multivariate Test for Monotone Trend with Pharamaceutical Applications". In: *Journal of the American Statistical Association* 76.373 (1981), pp. 169–174.

Douglas, EM, RM Vogel, and CN Kroll. "Trends in floods and low flows in the United States: impact of spatial correlation". In: *Journal of Hydrology* 240.1-2 (2000), pp. 90–105.

Duffie, Darrell and Kenneth J Singleton. "Modeling term structures of defaultable bonds". In: *The Review of Financial Studies* 12.4 (1999), pp. 687–720.

Dugas, Charles et al. "Incorporating second-order functional knowledge for better option pricing". In: *Advances in Neural Information Processing Systems* 13 (2000).

Esterby, S. "Review of Methods for the Detection and Estimation of Trends with Emphasis on Water Quality Applications". In: *Hydrological Processes* 10.2 (1996), pp. 127–149.

Gupton, Gred M. *Creditmetrics: technical document*. JP Morgan & Company, 1997.

Helsel, Dennis R et al. *Nondetects and Data Analysis. Statistics for Censored Environmental Data*. Wiley-Interscience, 2005.

Hipel, K. W. and A. I. McLeod. *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier Science & Techn., 1994.

Hirsch, Robert M and James R Slack. "A nonparametric trend test for seasonal data with serial dependence". In: *Water Resources Research* 20.6 (1984), pp. 727–732.

Hirsch, Robert M., James R. Slack, and Richard A. Smith. "Techniques of Trend Analysis for Monthly Water Quality Data". In: *Water Resources Research* 18.1 (1982), pp. 107–121.

Hoeffding, Wassily and Herbert Robbins. "The central limit theorem for dependent random variables". In: *Duke Mathematical Journal* 15.3 (1948), pp. 773–780.

Hull, John and Alan White. "The impact of default risk on the prices of options and other derivative securities". In: *Journal of Banking and Finance* 19.2 (1995), pp. 299–322.

Jarrow, Robert A and Stuart M Turnbull. "Pricing derivatives on financial securities subject to credit risk". In: *The Journal of Finance* 50.1 (1995), pp. 53–85.

Kendall, M. *Rank Correlation Methods*. C. Griffin, 1955.

Kruppa, Jochen et al. "Consumer credit risk: Individual probability estimates using machine learning". In: *Expert Systems with Applications* 40.13 (2013), pp. 5125–5131.

Lehmann, Erich Leo and Howard J D'Abrera. *Nonparametrics: Statistical Methods Based on Ranks*. Holden-day, 1975.

Lettenmaier, Dennis P. "Multivariate nonparametric tests for trend in water quality 1". In: *JAWRA Journal of the American Water Resources Association* 24.3 (1988), pp. 505–512.

Libiseller, Claudia and Anders Grimvall. "Performance of Partial Mann-Kendall Tests for Trend Detection in the Presence of Covariates". In: *Environmetrics* 13.1 (2002), pp. 71–84.

Mann, H. "Nonparametric Tests Against Trend". In: *Econometrica* 13.3 (1945), pp. 245–259.

Merton, Robert C. "On the pricing of corporate debt: The risk structure of interest rates". In: *The Journal of Finance* 29.2 (1974), pp. 449–470.

Vasicek, Oldrich A. *The Philosophy of Credit Valuation: The Credit Valuation Model*. 1984.

Ward, Robert C, Jim C Loftis, and Graham B McBride. *Design of Water Quality Monitoring Systems*. John Wiley & Sons, 1991.

Yue, S., P. Pilon, and G. Cavadias. "Power of the Mann–Kendall and Spearman's Rho Tests for Detecting Monotonic Trends in Hydrological Series". In: *Journal of Hydrology* 259.1-4 (2002), pp. 254–271.

Appendix A

Derivations and Proofs in Section 5.3

A.1 Proof of Proposition 5.1

For (5.13), assume that none of p_j and q_j are zeros at first. That is for avoiding some extreme cases. In the meanwhile, it can be achieved by choosing appropriate values for ρ and each c^j. Hence, the following proof will be in the context that all of p_j and q_j are non-zero.

Since the sum of the weights is 1, we can get the expression of η as follows

$$\eta = \frac{2a}{\sum_{j=1}^{n} \frac{1}{p_j q_j} - \frac{\left(\sum_{j=1}^{k} \frac{1}{q_j} + \sum_{j=k+1}^{n} \frac{1}{p_j}\right)^2}{\sum_{j=1}^{k} \frac{p_j}{q_j} + \sum_{j=k+1}^{n} \frac{q_j}{p_j} + a}},$$

where

$$\sum_{j=1}^{n} \frac{1}{p_j q_j} \cdot \left(\sum_{j=1}^{k} \frac{p_j}{q_j} + \sum_{j=k+1}^{n} \frac{q_j}{p_j} + a\right) \neq \left(\sum_{j=1}^{k} \frac{1}{q_j} + \sum_{j=k+1}^{n} \frac{1}{p_j}\right)^2.$$

Therefore, the values and sum of weights can be derived by replacing the η.

If $j \leq k$, then

$$w_j = \frac{1}{q_j} \cdot \frac{1}{\sum\limits_{\ell=1}^{n} \frac{1}{p_\ell q_\ell} - \frac{\left(\sum_{\ell=1}^{k} \frac{1}{q_\ell} + \sum_{\ell=k+1}^{n} \frac{1}{p_\ell}\right)^2}{\sum_{\ell=1}^{k} \frac{p_\ell}{q_\ell} + \sum_{\ell=k+1}^{n} \frac{q_\ell}{p_\ell} + a}} \left(\frac{1}{p_j} - \frac{\sum\limits_{\ell=1}^{k} \frac{1}{q_\ell} + \sum\limits_{\ell=k+1}^{n} \frac{1}{p_\ell}}{\sum\limits_{\ell=1}^{k} \frac{p_\ell}{q_\ell} + \sum\limits_{\ell=k+1}^{n} \frac{q_\ell}{p_\ell} + a} \right).$$

If $j > k$, then

$$w_j = \frac{1}{p_j} \cdot \frac{1}{\sum\limits_{\ell=1}^{n} \frac{1}{p_\ell q_\ell} - \frac{\left(\sum_{\ell=1}^{k} \frac{1}{q_\ell} + \sum_{\ell=k+1}^{n} \frac{1}{p_\ell}\right)^2}{\sum_{\ell=1}^{k} \frac{p_\ell}{q_\ell} + \sum_{\ell=k+1}^{n} \frac{q_\ell}{p_\ell} + a}} \left(\frac{1}{q_j} - \frac{\sum\limits_{\ell=1}^{k} \frac{1}{q_\ell} + \sum\limits_{\ell=k+1}^{n} \frac{1}{p_\ell}}{\sum\limits_{\ell=1}^{k} \frac{p_\ell}{q_\ell} + \sum\limits_{\ell=k+1}^{n} \frac{q_\ell}{p_\ell} + a} \right)$$

and

$$S_\sigma = \frac{a}{\sum\limits_{j=1}^{n} \frac{1}{p_j q_j} - \frac{\left(\sum_{j=1}^{k} \frac{1}{q_j} + \sum_{j=k+1}^{n} \frac{1}{p_j}\right)^2}{\sum_{j=1}^{k} \frac{p_j}{q_j} + \sum_{j=k+1}^{n} \frac{q_j}{p_j} + a}} \cdot \frac{\sum\limits_{j=1}^{k} \frac{1}{q_j} + \sum\limits_{j=k+1}^{n} \frac{1}{p_j}}{\sum\limits_{j=1}^{k} \frac{p_j}{q_j} + \sum\limits_{j=k+1}^{n} \frac{q_j}{p_j} + a}.$$

A.2 Extension of Proposition 5.3

We have showed that sufficient conditions for a and α_0 in Section 5.3. However, we would like to offer some ideas for finding out the complete available sets of α_0 and a. The goal is to find the maximum of the right-hand side of (5.21). If using the parameter $\tilde{k}$, which is used after the order of $\{c^j\}_{j=1}^n$ varies, the situations can be divided into three: $\tilde{k} \in [1, j_1)$, $\tilde{k} \in [j_2, n)$ and $\tilde{k} \in [j_1, j_2)$. Albeit that $\tilde{k}$ can be useful for categorizing, it is still difficult to establish the analytic relationship between $\tilde{k}$ and $\{c^j\}_{j=1}^n$. For this reason, we will find a way to represent $\tilde{U}$ and $\tilde{V}$ with U and V, which is computed by k and the conditional probabilities that is used before rearranging $\{c^j\}_{j=1}^n$.

Let $\Delta\tilde{U} = \tilde{U} - U$ and $\Delta\tilde{V} = \tilde{V} - V$. Put $\tilde{U} = U + \Delta\tilde{U}$ and $\tilde{V} = V + \Delta\tilde{V}$

into (5.20). It provides that U and $\sum_{j=1}^{n} 1/(p_j q_j)$ are independent of $\{\tilde{c}^j\}_{j=1}^{n}$. On the one hand, the total range of the acceptance level should be

$$\alpha_0 < \frac{U}{\sum\limits_{j=1}^{n} \frac{1}{p_j q_j}} + \frac{\Delta\tilde{U}_{\min}}{\sum\limits_{j=1}^{n} \frac{1}{p_j q_j}}, \tag{A.1}$$

where $\Delta\tilde{U}_{\min}$ is the minimum of $\Delta\tilde{U}$ over all possible values of j_1 and j_2.

On the other hand, by replacing $\tilde{U}, \tilde{V}$ with U, V, we can get that

$$\begin{aligned}
a &\geq \alpha_0 \frac{\left(V + \Delta\tilde{V}\right)\Sigma - \left(U + \Delta\tilde{U}\right)^2}{U + \Delta\tilde{U} - \alpha_0\Sigma} \\
&= -\alpha_0 \left(U + \Delta\tilde{U} + \alpha_0\Sigma - \frac{V + \Delta\tilde{V} - \alpha_0^2\Sigma}{U + \Delta\tilde{U} - \alpha_0\Sigma}\Sigma\right) \\
&= -2\alpha_0^2\Sigma - \alpha_0 \left[\left(U + \Delta\tilde{U} - \alpha_0\Sigma\right) - \frac{\left(V + \Delta\tilde{V}\right)\Sigma - \alpha_0^2\Sigma^2}{U + \Delta\tilde{U} - \alpha_0\Sigma}\right],
\end{aligned}$$

where $\Sigma = \sum_{j=1}^{n} 1/(p_j q_j)$. There are three cases that need handling.

Firstly, if $\tilde{k} < j_1 < j_2$, then

$$\begin{aligned}
\tilde{U} &= U + \frac{1}{p_{j_1}} - \frac{1}{q_{j_1}} + \mathrm{error}(k, \tilde{k}), & \Delta\tilde{U} &= \frac{1}{p_{j_1}} - \frac{1}{q_{j_1}} + \mathrm{error}(k, \tilde{k}), \\
\tilde{V} &= V + \frac{q_{j_1}}{p_{j_1}} - \frac{p_{j_1}}{q_{j_1}} + \mathrm{error}(k, \tilde{k}), & \Delta\tilde{V} &= \frac{q_{j_1}}{p_{j_1}} - \frac{p_{j_1}}{q_{j_1}} + \mathrm{error}(k, \tilde{k}),
\end{aligned}$$

where the term $\mathrm{error}(k, \tilde{k})$ can be eliminating through the optimization of the value of k. To obtain the minimum of $\Delta\tilde{U}$, consider a function defined on the interval $(0, 1)$ as follows

$$g(x) = \frac{1}{x} - \frac{1}{1-x}, \quad x \in (0, 1).$$

Provided that its first derivative is

$$g'(x) = -\frac{1}{x^2} - \frac{1}{(1-x)^2} < 0,$$

$g(x)$ is strictly decreasing against x. In fact, we observe that $\Delta\tilde{U} = g(p_{j_1})$ so that $\Delta\tilde{U}$ arrives at its minimum when p_{j_1} is largest. That is, $j_1 = k$ and

$$\Delta\tilde{U}_{\min} = \frac{1}{p_k} - \frac{1}{q_k}.$$

Subsequently, we can get an upper threshold of the acceptance level in (A.1)

$$\alpha_0 < \frac{U}{\sum\limits_{j=1}^{n} \frac{1}{p_j q_j}} + \frac{\frac{1}{p_k} - \frac{1}{q_k}}{\sum\limits_{j=1}^{n} \frac{1}{p_j q_j}} = \frac{\sum\limits_{j=1}^{k-1} \frac{1}{q_j} + \sum\limits_{j=k}^{n} \frac{1}{p_j}}{\sum\limits_{j=1}^{n} \frac{1}{p_j q_j}}.$$

Secondly, if $j_1 < j_2 \leq \tilde{k}$, then

$$\tilde{U} = U + \frac{1}{q_{j_2}} - \frac{1}{p_{j_2}} + \mathrm{error}(k, \tilde{k}), \qquad \Delta\tilde{U} = \frac{1}{q_{j_2}} - \frac{1}{p_{j_2}} + \mathrm{error}(k, \tilde{k}),$$
$$\tilde{V} = V + \frac{p_{j_2}}{q_{j_2}} - \frac{q_{j_2}}{p_{j_2}} + \mathrm{error}(k, \tilde{k}), \qquad \Delta\tilde{V} = \frac{p_{j_2}}{q_{j_2}} - \frac{q_{j_2}}{p_{j_2}} + \mathrm{error}(k, \tilde{k}).$$

With the help of $g(x)$ defined before, we can derive that $\Delta\tilde{U} = -g(p_{j_2})$. As $-g(x)$ increases strictly over x, $\Delta\tilde{U}$ arrives at its minimum when p_{j_2} is smallest. That is, $j_2 = k+1$ and

$$\Delta\tilde{U}_{\min} = \frac{1}{q_{k+1}} - \frac{1}{p_{k+1}}.$$

Accordingly, the acceptance level should has an upper boundary in (A.1)

$$\alpha_0 < \frac{\sum\limits_{j=1}^{k+1} \frac{1}{q_j} + \sum\limits_{j=k+2}^{n} \frac{1}{p_j}}{\sum\limits_{j=1}^{n} \frac{1}{p_j q_j}}$$

Thirdly, if $j_1 \leq \tilde{k} < j_2$, then

$$\tilde{U} = U + \frac{1}{q_{j_2}} + \frac{1}{p_{j_1}} - \frac{1}{q_{j_1}} - \frac{1}{p_{j_2}} + \mathrm{error}(k, \tilde{k}),$$

$$
\begin{aligned}
\Delta\tilde{U} &= \frac{1}{q_{j_2}} + \frac{1}{p_{j_1}} - \frac{1}{q_{j_1}} - \frac{1}{p_{j_2}} + \mathrm{error}(k,\tilde{k}),\\
\tilde{V} &= V + \frac{p_{j_2}}{q_{j_2}} + \frac{q_{j_1}}{p_{j_1}} - \frac{p_{j_1}}{q_{j_1}} - \frac{q_{j_2}}{p_{j_2}} + \mathrm{error}(k,\tilde{k}),\\
\Delta\tilde{V} &= \frac{p_{j_2}}{q_{j_2}} + \frac{q_{j_1}}{p_{j_1}} - \frac{p_{j_1}}{q_{j_1}} - \frac{q_{j_2}}{p_{j_2}} + \mathrm{error}(k,\tilde{k}).
\end{aligned}
$$

In this case, we can find that $\Delta\tilde{U} > 0$ and $\Delta\tilde{V} > 0$. For seeking the minimum of the $\Delta\tilde{U}$, we rearrange the terms of $\Delta\tilde{U}$ and get that

$$
\Delta\tilde{U} = \left(\frac{1}{p_{j_1}} - \frac{1}{q_{j_1}}\right) - \left(\frac{1}{p_{j_2}} - \frac{1}{q_{j_2}}\right) = g(p_{j_1}) - g(p_{j_2}),
$$

where $g(x) = 1/x - 1/(1-x)$ is a function on the interval $(0,1)$.