



141
062
THS

LIBRARY
Michigan State
University

This is to certify that the
dissertation entitled

ON SOME INFERENCE PROBLEMS FOR CURRENT
STATUS DATA

presented by

DEEPA AGGARWAL

has been accepted towards fulfillment
of the requirements for the

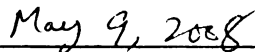
Ph.D.

degree in

STATISTICS AND
PROBABILITY



Major Professor's Signature



Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

On Some Inference Problems For Current Status Data

By

Deepa Aggarwal

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics and Probability

2008

ABSTRACT

On Some Inference Problems For Current Status Data

By

Deepa Aggarwal

In the current status or interval censored case 1 data, one does not observe the event occurrence time but only the inspection time and whether the event has occurred prior to the inspection time or not. This thesis consists of two parts. The first part pertains to fitting a parametric model to the distribution function of the event occurrence time in the one sample set up with current status data. In this part, we first discuss two analogous minimum distance inference procedures for fitting a regression function in the classical regression set up. These distances are based on squared deviations of a nonparametric regression function estimator and the model being fitted. In the one distance the integrating measure is σ -finite and in the second, it is data dependent. The thesis establishes asymptotic normality of the proposed empirical minimum distance statistic and that of the corresponding estimator under the fitted model in a general regression set up. Then, these results for empirical minimum distance test are adapted to fit a parametric model to the distribution of the event occurrence time based on current status data. It also contains a finite

sample comparison of the proposed test with Koul and Yi test and the one sample Cramér-von Mises test based on nonparametric maximum likelihood estimator of the distribution function of the event occurrence time.

The second part of the thesis pertains to testing for the equality of the two event occurrence time distribution functions in the two sample setting when the data is interval censored case 1 from both samples. It derives the asymptotic distribution of the underlying test statistic both under the null hypothesis and under local alternatives. It also contains a finite sample comparison of the proposed test with the two sample Cramér-von Mises test based on nonparametric maximum likelihood estimators of the time to event distribution functions.

ACKNOWLEDGMENTS

I wish to express my deepest regard to my advisor Professor Hira L. Koul for his invaluable guidance, generous support. I have enjoyed our interactions tremendously. I thank Dr Koul for providing me with countless opportunities to grow both personally and professionally. This dissertation could not have been completed without his help and support.

I would also like to thank Professors Sarat Dass, Dennis Gilliland and Habib Salehi for serving on my guidance committee. My special thanks go to Professors Connie Page, Dennis Gilliland and Sandra Herman for their advice when I was at the consulting service.

I would like to thank my family especially my mother, my husband and my daughter for all the support and encouragement provided by them during my graduate education.

This research was partly supported by the NSF grant DMS 0704130 with Professor Koul as P.I.

TABLE OF CONTENTS

1	Introduction	1
2	Empirical Minimum Distance	
	Lack-Of-Fit Testing	
	In Regression Model	8
2.1	Introduction	8
2.2	Assumptions	12
2.3	Consistency of θ_n^* and $\hat{\theta}_n$	14
2.4	Asymptotic distribution of $\hat{\theta}_n$	28
2.5	Asymptotic normality of $M_n(\hat{\theta}_n)$	41
3	Minimum Distance Goodness-Of-Fit Tests	
	For Current Status Data	60
3.1	Introduction	60
3.2	Minimum Distance Statistics and Tests	63
3.3	Empirical Minimum Distance Statistic	67
3.4	Simulations	68
4	Testing the equality of two distributions with Current Status Data	75
4.1	Introduction	75
4.2	Asymptotic behavior under the null hypothesis and local alternatives	78
4.3	Simulations	88
	BIBLIOGRAPHY	93

LIST OF TABLES

3.1	Mean and MSE of $\hat{\theta}_n$, $X, T \sim \exp(1)$, $\theta_0 = 1$	72
3.2	Empirical sizes of $M_n(\hat{\theta}_n)$, $X, T \sim \exp(1)$	72
3.3	Power of $M_n(\hat{\theta}_n)$, $T \sim \exp(1)$, $(c_1, c_2) = (.9, 1)$	73
3.4	Simulated percentiles of CV_1 , $X, T \sim \exp(1)$	73
3.5	Empirical sizes of CV_1 , $X, T \sim \exp(1)$	73
3.6	Mean and MSE of $\hat{\zeta}_n$, $X, T \sim \exp(1)$, $\theta_0 = 1$	74
3.7	Power of CV_1 , $T \sim \exp(1)$	74
3.8	Empirical sizes, $X, T \sim \exp(1)$, $(c_1, c_2) = (.9, .8)$	74
4.1	Empirical sizes of $\hat{T}$, $X, Y \sim \exp(1)$, $S, T \sim \exp(1)$	90
4.2	Empirical sizes of $\hat{T}$, $X, Y \sim \exp(1)$, $(n_1, n_2) = (180, 200)$	90
4.3	Power of $\hat{T}$, $S, T \sim \exp(1)$, $X \sim \exp(1)$, $n_1 = n_2 = 50$	91
4.4	Simulated percentiles of CV_2 , $X, Y \sim \exp(1)$, $S, T \sim \exp(1.5)$	91
4.5	Empirical sizes of CV_2 , $X, Y \sim \exp(1)$, $S, T \sim \exp(1.5)$	91
4.6	Simulated 95th percentile of CV_2 , $X, Y \sim \exp(1)$	92
4.7	Empirical sizes, $X, Y \sim \exp(1)$, $(c_1, c_2) = (.25, 1)$	92
4.8	Power, $S, T \sim \exp(1)$, $X \sim \exp(1)$, $(c_1, c_2) = (.2, .9)$	92

CHAPTER 1

Introduction

In recent years there has been a considerable research on the analysis of interval-censored data. In the case I interval censored data, an event occurrence time X is unobservable, but one observes an inspection time T and whether an event has occurred prior to this time or not. This type of data is also known as current status data. It is different from the right censored data where one observes true life time in the case of no censoring and only censoring time when life time is censored.

Current status data often arises in epidemiology, demography and economics. For example, as mentioned in Jewell and Van der Laan (2004), in the study of infectious disease Human Immunodeficiency Virus (HIV), in particular, the partner studies of HIV infection. These partnerships are assumed to include an index case who has been infected via some external source, and a susceptible partner who has no other means of infection except the contact with the index case. Suppose X denotes the time from infection of the index case to infection of the susceptible partner and T is the time the susceptible partner is examined after infection of the index case. Then the

infection status of the susceptible partner provides current status data. For some more applications for current status data, see Hoel and Walburg (1972), Finkelstein and Wolfe (1985), Finkelstein (1986), Diamond, McDonald and Shah (1986), Diamond and McDonald (1991), Keiding (1991) and Jewell and Van der Laan (2004).

This thesis is concerned with the following two problems. The first problem pertains to fitting a parametric model to the distribution function of the event occurrence time in the one sample set up with current status data. The second problem is concerned with testing for the equality of the two event occurrence times distribution functions in the two sample setting when the data is interval censored case 1 from both samples.

We shall now focus on the first problem for the moment. To describe this problem a bit more precisely, let F denote the distribution function (d.f.) of the event occurrence time X , Θ be a subset of q -dimensional Euclidean space $\mathbb{R}^q$, and $\{F_\theta, \theta \in \Theta\}$ be a known parametric family of d.f.'s on $[0, \infty)$. Let T be the inspection time and $\delta := I(X \leq T)$ and $\mathcal{I}$ be a compact interval of $[0, \infty)$, where $I[A]$ denotes the indicator function of the event A . We assume X to be independent of T . The problem of interest here is to test the hypothesis

$$H_{01} : F(t) = F_{\theta_0}(t), \text{ for all } t \in \mathcal{I}, \text{ for some } \theta_0 \in \Theta,$$

against the alternative

$$H_{11} : H_{01} \text{ is not true.}$$

It is natural to base tests of H_{01} on a distance between the nonparametric maximum likelihood estimates $\hat{F}$ of F and $\{F_\theta, \theta \in \Theta\}$. One such test statistic is the

Cramér-von-Mises statistic

$$CV_1 = \inf_{\theta} \int \left(\hat{F}(x) - F_{\theta}(x) \right)^2 d\hat{F}(x).$$

But unfortunately neither the finite sample nor the asymptotic null distribution of this statistic is known because of the complicated nature of the distribution of $\hat{F}$. Even asymptotic distribution of a suitably standardized $\hat{F}$ is intractable, cf., Groeneboom and Wellner (1992).

An alternative way to proceed is to use the well known regression relationship between δ and $F(T)$, i.e., $E(\delta|T) = F(T)$, and the fact that this regression is heteroscedastic. In this context then the problem of testing H_{01} is equivalent to testing the lack-of-fit of the parametric regression model $\{F_{\theta}, \theta \in \Theta\}$.

There is a vast literature on the problem of testing for the lack-of-fit of a parametric regression model. The monograph of Hart (1997) provides a nice overview on the subject till 1997. Using the ideas of Khmaladze (1979), Stute, Thies and Zhu (1998) proposed an asymptotically distribution free test for this problem based on a martingale transform of a certain marked empirical process of the residuals. Koul and Ni (2005) used the minimum distance methodology to propose a class of tests for the same problem. In all this literature the data is completely observable.

Using the above mentioned equivalence between testing H_{01} and the corresponding lack-of-fit testing of a regression model, Koul and Yi (2006) adapted the Stute-Thies-Zhu test to test for H_{01} . They provide sufficient conditions for consistency of their test at a fixed alternative and derive an expression for its asymptotic power against local alternatives.

Koul and Ni (2004) used the integrated square distance between a kernel type nonparametric estimator of the regression function and the model being fitted, where the integrating measure is a σ -finite measure. A practical problem that arises in using these statistics is the choice of the integrating measure. Although one may choose this by using some optimality criteria, such a measure will invariably depend on the model being fitted and the design distribution.

In this thesis we first discuss two analogous minimum distance inference procedures in the classical regression set up, first when the integrating measure is σ -finite and second when the integrating measure is data dependent-viz, the empirical d.f. of the design variable. We prove asymptotic normality of the proposed empirical minimum distance statistic and that of the corresponding estimator under the fitted model in a general regression set up. Then, these results for empirical minimum distance test are adapted to fit a parametric model to the distribution of the event occurrence times based on current status data. We also show consistency of the proposed minimum distance tests against a fixed alternative and obtain asymptotic power against a class of local alternatives for current status data.

We now describe the second problem of this thesis. To describe it more precisely, let F_1 (F_2) denote the d.f. of event occurrence time X (Y) from the first (second) population, and let S (T) be the corresponding inspection time. In the two sample current status data set up, one observes (δ, S) and (η, T) , where $\delta = I[X \leq S]$ and $\eta = I[Y \leq T]$. The problem of interest here is to test the null hypothesis that the

two event occurrence distributions are the same, i.e.

$$H_{02} : F_1(x) = F_2(x), \quad \text{for all } x \in \mathcal{I},$$

against the alternative

$$H_{12} : F_1(x) \neq F_2(x), \quad \text{for some } x \in \mathcal{I}.$$

Similar to the one sample set up discussed above, it is natural to base tests of H_{02} on nonparametric maximum likelihood estimates $\hat{F}_1$ and $\hat{F}_2$ of F_1 and F_2 . One such test is based on the Cramér-von-Mises statistic

$$\begin{aligned} CV_2 = & \frac{n_1}{n_1 + n_2} \int_{\mathcal{I}} \left(\hat{F}_1(x) - \hat{F}_2(x) \right)^2 d\hat{F}_1(x) \\ & + \frac{n_2}{n_1 + n_2} \int_{\mathcal{I}} \left(\hat{F}_1(x) - \hat{F}_2(x) \right)^2 d\hat{F}_2(x). \end{aligned}$$

Again for the same reasons given above asymptotic null distribution of such a statistic is not currently tractable.

An alternative way to proceed is to use the well known regression relationship between δ and $F_1(S)$, and η and $F_2(T)$, i.e., $E(\delta|S) = F_1(S)$ and $E(\eta|T) = F_2(T)$, and the fact that these two regressions are heteroscedastic. In this context then the problem of testing H_{02} is equivalent to testing the equality of the two regression functions under heteroscedasticity.

The problem of comparing the two regression functions has been discussed by several authors. In general, see, e.g., Hall and Hart (1990), King, Hart and Wehrly (1991), Carroll and Hall (1992), Delgado (1993), Kulasekera (1995), Koul and Schick (1997, 2003), Neumeyer and Dette (2003), among others. The data is completely observable in the above mentioned literature.

Koul and Schick (2003) proposed a test using covariate matching for the same problem in a general regression set up. In this thesis, we adapt this test to the two sample current status data and discuss its asymptotic normality under a general set of assumptions.

This thesis is organized as follows. Chapter 2 studies empirical minimum distance tests of lack of fit in classical regression set up. Corollary 2.3.1 and Theorem 2.3.1 state and prove consistency of empirical minimum distance estimates of the underlying parameters of the model being fitted. Theorem 2.4.1 and Theorem 2.5.1 give asymptotic distribution of the parameter estimator and the empirical minimum distance statistic under the null hypothesis.

In chapter 3, section 2, we apply the results of Koul and Ni (2004) for minimum distance tests of goodness of fit hypothesis based on current status data. After that, we discuss consistency of these tests against a fixed alternative and obtain asymptotic power against a class of local alternatives. Section 3.3 uses the results of Chapter 2 for empirical minimum distance test to fit a parametric model to the distribution of the event occurrence times based on current status data.

Section 3.4 reports the numerical results of the three simulation studies in the one sample set up. The first one assesses the finite sample level and power behavior of the empirical minimum distance test. The simulation results of empirical minimum distance statistic are consistent with asymptotic theory. Also, simulation results show little bias in the estimator of the best fitted parameter for all the chosen sample sizes. The second simulation study investigates Monte Carlo size and power behavior of the Cramér-Von-Mises test CV_1 . The finite sample level of this test approximates

the nominal level well for all the chosen sample sizes. The third simulation study investigates Monte Carlo size comparison of empirical minimum distance test, CV_1 , and Koul and Yi (2006) test. Simulation results show that empirical sizes are better for CV_1 and Koul and Yi (2006) test as compared to empirical minimum distance test, when sample size is less than 200. But when sample size is 200 or large, empirical sizes are comparable in all the three tests. In our simulations, $\hat{F}$ is obtained by the one step procedure for the calculation of the nonparametric maximum likelihood estimator, based on isotonic regression, cf. Groeneboom and Wellner (1992).

Chapter 4 deals with the problem of testing the equality of two distribution functions against the two sided alternative based on the current status data. Proposition 4.2.1 discuss asymptotic normality of the underlying test statistic under a general set of assumptions. Section 4.3 reports the numerical results of the two simulation studies. The first one assesses the finite sample level and power behavior of the proposed test statistic. The simulation results of the proposed test statistic are consistent with asymptotic theory.

In the second simulation study, the finite sample comparison of the proposed test statistic with the two sample Cramér-Von Mises test is made. Simulation results show that for all the chosen alternatives, bandwidths and sample sizes, significance level and power of the proposed and CV_2 tests are comparable. Again, in our simulations, $\hat{F}_1$ and $\hat{F}_2$ are obtained by the one step procedure for the calculation of the nonparametric maximum likelihood estimator, based on isotonic regression.

CHAPTER 2

Empirical Minimum Distance

Lack-Of-Fit Testing

In Regression Model

2.1 Introduction

This chapter discusses an empirical minimum distance method for fitting a parametric model to the regression function $\mu(x) := E(Y|X = x)$, $x \in \mathbb{R}^d$, $d \geq 1$, assuming it exists, where Y is the one dimensional response variable and X is a d dimensional design variable. Let $\{m_\theta(x) : x \in \mathbb{R}^d, \theta \in \Theta \subset \mathbb{R}^q, q \geq 1\}$ be a given parametric family of regression functions and let $\mathcal{I}$ be a compact subset of $\mathbb{R}^d$. The problem of interest is that of model checking, i.e., to test the hypothesis

$$\mathcal{H}_0 : \mu(x) = m_{\theta_0}(x), \quad \text{for some } \theta_0 \in \Theta \quad \text{and for all } x \in \mathcal{I};$$

$$\mathcal{H}_1 : \mathcal{H}_0 \text{ is not true,}$$

based on a random sample (X_i, Y_i) , $1 \leq i \leq n$, from the distribution of (X, Y) .

Several authors have addressed the problem of regression model checking: see, Hart (1997) and references there in. The recent paper of Koul and Ni (2004) (K-N) uses the minimum distance ideas of Wolfowitz (1953, 1954, 1957) and Beran (1977, 1978) to propose tests of lack-of-fit for the regression model with heteroscedastic errors. In a finite sample comparison of these tests with some other existing tests, they noted that a member of this class preserves the asymptotic level and has very high power against some alternatives when compared to some other existing lack-of-fit tests. The distance used in their paper is the integrated square deviation between a nonparametric estimator of the regression function and the parametric model being fitted with respect to a general integrating measure.

To be specific, K-N considered the following tests of $\mathcal{H}_0$ where the design is random and observable, and the errors are heteroscedastic. Let Φ be a sigma finite measure on $\mathbb{R}^d$, G denote the d.f of the design variable X , and G_n be the empirical d.f. based on X_i , $1 \leq i \leq n$. For any density kernel K , let $K_h(x) := K(x/h)/h^d$, $h > 0$, $x \in \mathbb{R}^d$. Define, as in K-N,

$$\begin{aligned}\hat{g}_h(x) &:= \frac{1}{n} \sum_{i=1}^n K_h(x - X_i), \\ g_w^*(x) &:= \frac{1}{n} \sum_{i=1}^n K_w^*(x - X_i), \quad w = w_n \sim (\log n/n)^{1/(d+4)}, \\ \mathcal{T}_n(\theta) &:= \int_{\mathcal{I}} \left[\frac{1}{n} \sum_{i=1}^n K_h(x - X_i)(Y_i - m_\theta(X_i)) \right]^2 \frac{d\Phi(x)}{g_w^{*2}(x)},\end{aligned}$$

and $\hat{\vartheta}_n := \operatorname{argmin}_{\vartheta \in \Theta} \mathcal{T}_n(\vartheta)$, where K, K^* are kernel density functions, possibly different, $h = h_n$ and $w = w_n$ are the window widths, depending on the sam-

ple size n . K-N gave some sufficient conditions on the underlying entities for consistency and asymptotic normality of $\hat{v}_n$ under H_0 , and asymptotic normality of $\mathcal{D}_n := nh_n^{d/2}(\mathcal{I}_n(\hat{v}_n) - \hat{R}_n)/\hat{\mathcal{G}}_n^{1/2}$ under H_0 , where

$$\begin{aligned}\hat{R}_n &:= \frac{1}{n^2} \sum_{i=1}^n \int_{\mathcal{I}} \frac{K_h^2(x - X_i) \hat{\varepsilon}_i^2}{g_w^{*2}(x)} d\Phi(x), \\ \hat{\varepsilon}_i &= Y_i - m_{\hat{\theta}_n}(X_i), \\ \hat{\mathcal{G}}_n &:= 2n^{-2} h^d \sum_{i \neq j} \left(\int_{\mathcal{I}} \frac{K_h(x - X_i) K_h(x - X_j) \hat{\varepsilon}_i \hat{\varepsilon}_j}{\hat{g}_h^2(x)} d\Phi(x) \right)^2.\end{aligned}\tag{2.1}$$

A practical problem that arises in using these statistics is the choice of the integrating measure Φ . Although one may choose Φ using some optimality criteria, such a Φ will invariably depend on the model being fitted and the design distribution. One way to simplify the choice of Φ is to use the empirical d.f. of design in the above entities.

We are thus motivated to propose empirical minimum distance tests of lack-of-fit in the classical regression model. Accordingly, let $\mathcal{I}_j = I(X_j \in \mathcal{I})$ and define

$$M_n(\theta) := n^{-1} \sum_{j=1}^n \left[\frac{1}{n} \sum_{i=1}^n \frac{K_h(X_j - X_i)(Y_i - m_\theta(X_i))}{g_w^*(X_j)} \mathcal{I}_j \right]^2,$$

and $\hat{\theta}_n := \operatorname{argmin}_{\theta \in \Theta} M_n(\theta)$.

We also need the following entities:

$$\begin{aligned}\hat{m}_n(x) &:= \frac{1}{n} \sum_{i=1}^n \frac{K_h(x - X_i) Y_i}{g_w^*(x)}, \quad x \in \mathbb{R}^d, \\ M_n^*(\theta) &:= n^{-1} \sum_{j=1}^n (\hat{m}_n(X_j) - m_\theta(X_j))^2 \mathcal{I}_j, \quad \theta \in \mathbb{R}^q, \\ \theta_n^* &:= \operatorname{argmin}_{\theta \in \Theta} M_n^*(\theta).\end{aligned}$$

In this thesis we prove the consistency of θ_n^* and $\hat{\theta}_n$. We also prove asymptotic normality of $\sqrt{n}(\hat{\theta}_n - \theta_0)$, and $nh^{d/2}(\tilde{M}_n(\theta_0) - \tilde{C}_n)$ under H_0 , where $\tilde{C}_n$ is given below at (2.2). Then, similar to K-N, sequences of estimators $\hat{C}_n$ and $\hat{\Gamma}_n$ are provided such that $\hat{C}_n$ is $nh^{d/2}$ -consistent for $\tilde{C}_n$ and $\hat{\Gamma}_n$ is consistent for Γ , and under some sufficient conditions on the underlying entities, asymptotic null distribution of $nh^{d/2}\hat{\Gamma}_n^{-1/2}(M_n(\hat{\theta}_n) - \hat{C}_n)$, is shown to be standard normal. These results are similar in nature to Theorems 4.1 and 5.1 of K-N. Here,

$$\begin{aligned}
\tilde{C}_n &:= n^{-3} \sum_{j=1}^n \sum_{i=1}^n \frac{K_h^2(X_j - X_i) \varepsilon_i^2}{g^2(X_j)} \mathcal{I}_j, \\
\hat{C}_n &:= n^{-3} \sum_{j=1}^n \sum_{i=1}^n \frac{K_h^2(X_j - X_i) \hat{\varepsilon}_i^2}{g_w^{*2}(X_j)} \mathcal{I}_j, \\
\hat{\varepsilon}_i &:= Y_i - m_{\hat{\theta}_n}(X_i), \quad 1 \leq i \leq n, \\
\Gamma &:= 2 \int_{\mathcal{I}} \sigma^4(x) dx \int \left(\int K(u) K(v+u) du \right)^2 dv, \\
\hat{\Gamma}_n &:= 2h^d n^{-4} \sum_{i \neq j} \left(\sum_l \frac{K_h(X_l - X_i) K_h(X_l - X_j) \hat{\varepsilon}_i \hat{\varepsilon}_j}{\hat{g}_h^2(X_l)} \mathcal{I}_l \right)^2,
\end{aligned} \tag{2.2}$$

where $\sigma^2(x) := E[(Y - m_{\theta_0}(x))^2 | X = x]$, $x \in \mathbb{R}^d$, and g is Lebesgue density of G .

This chapter is organized as follows. Section 2.2 states the needed assumptions. In the beginning of section 2.3, we summarize some of the results of K-N and Koul and Song (2006) (K-S) for the sake of completeness. Section 2.3 contains the proofs of consistency of θ_n^* and $\hat{\theta}_n$, while section 2.4 and 2.5 contains the proofs of asymptotic normality of $\hat{\theta}_n$ and that of the proposed empirical minimum distance test statistic, respectively.

2.2 Assumptions

Here we shall state the assumptions that are in K-N for reference where theorems and lemmas are proved. Throughout the thesis θ_0 denotes the true parameter value under $\mathcal{H}_0$ assumed to be in the interior of Θ . About the errors, the underlying design and the σ -finite measure Φ on $\mathbb{R}^d$ we assume the following:

- (e1) The random variables $\{(X_i, Y_i) : X_i \in \mathbb{R}^d, Y_i \in \mathbb{R}, i = 1, 2, \dots, n\}$ are i.i.d., $E|Y| < \infty$, and the conditional expectation $\mu(x) := E(Y|X = x)$ satisfying $\int \mu^2(x) d\Phi(x) < \infty$.
- (e2) $E(Y - \mu(X))^2 < \infty$, and the function $\sigma^2(x) = E[(Y - \mu(X))^2|X = x]$ is a.s. (Φ) continuous on $\mathcal{I}$ and $\sigma_{\theta_0}^2(x)$ is continuous on $\mathbb{R}^d$.
- (e3) $E|Y - \mu(X)|^{2+\delta} < \infty$, for some $\delta > 0$.
- (e4) $E|Y - \mu(X)|^4 < \infty$ and $\tau_{\theta_0}(x) := E[(Y - m_{\theta_0}(X))^4|X = x]$ is continuous on $\mathbb{R}^d$.
- (g1) The d.f. G of the design variable X has a uniformly continuous Lebesgue density g that is bounded from below on $\mathcal{I}$.
- (g2) The density g of the d.f. G is twice continuously differentiable.
- (p) Φ has a continuous Lebesgue density ϕ .

About the kernel functions K and K^* , we shall assume the following:

- (k) The kernels K, K^* are positive symmetric density functions on $[-1, 1]^d$ with finite variances. In addition, K is a bounded kernel and K^* satisfies a Lipschitz condition.

About the parametric family of functions to be fitted we need to assume the following:

(m1) For each θ , $m_\theta(x)$ is a.e. continuous in x , w.r.t. the integrating measure Φ on $\mathbb{R}^d$.

(m2) The function $m_\theta(x)$ is identifiable w.r.t. θ . i.e., if $m_{\theta_1}(x) = m_{\theta_2}(x)$ for almost all $x(\Phi)$, then $\theta_1 = \theta_2$.

(m3) For some positive continuous function ℓ on $\mathbb{R}^d$ with $E\ell(X) < \infty$ and some $\beta > 0$,

$$|m_{\theta_2}(x) - m_{\theta_1}(x)| \leq \|\theta_2 - \theta_1\|^\beta \ell(x), \quad \forall \theta_1, \theta_2 \in \Theta, x \in \mathcal{I}.$$

(m4) For every x , $m_\theta(x)$ is differentiable in θ in a neighborhood of θ_0 with the vector of derivatives $\dot{m}_\theta(x)$, which is continuous on $\mathbb{R}^d$ such that for every $0 < k < \infty$,

$$\sup_{\mathcal{C}} \frac{|m_\theta(X_i) - m_{\theta_0}(X_i) - (\theta - \theta_0)' \dot{m}_{\theta_0}(X_i)|}{\|\theta - \theta_0\|} = o_p(1),$$

where $\mathcal{C} := \{1 \leq i \leq n, \sqrt{nh_n^d} \|\theta - \theta_0\| \leq k\}$.

(m5) For every $0 < k < \infty$,

$$\sup_{\mathcal{C}} h_n^{-d/2} \|\dot{m}_\theta(X_i) - \dot{m}_{\theta_0}(X_i)\| = o_p(1), \quad \forall n > N_\varepsilon.$$

(m6) $\int \|\dot{m}_{\theta_0}\|^2 d\Phi < \infty$, and $\Sigma_0 := \int \dot{m}_{\theta_0} \dot{m}_{\theta_0}' d\Phi$ is positive definite.

About the bandwidth h_n we shall make the following assumptions:

(h1) $h_n \rightarrow 0$ as $n \rightarrow \infty$.

(h2) $nh_n^{2d} \rightarrow \infty$ as $n \rightarrow \infty$.

(h3) $h_n \sim n^{-a}$, where $a < \min(1/2d, 4/(d(d+4)))$.

Let $\hat{g}_h$ and g_w^* denote the kernel density estimators of g with bandwidth h and w , respectively. From Mack and Silverman (1982), we obtain that under (g1), (k), (h1) and (h2),

$$\begin{aligned} \sup_{x \in \mathcal{I}} |\hat{g}_h(x) - g(x)| &= o_p(1), & \sup_{x \in \mathcal{I}} |g_w^*(x) - g(x)| &= o_p(1), \\ \sup_{x \in \mathcal{I}} \left| \frac{g(x)}{g_w^*(x)} - 1 \right| &= o_p(1). \end{aligned} \quad (2.3)$$

These conclusions are often used in the proofs below.

In the sequel, $\varepsilon := Y - m_{\theta_0}(X)$. The integrals with respect to Φ and G measures are understood to be over the compact set $\mathcal{I}$. The inequality $(a+b)^2 \leq 2(a^2 + b^2)$, for any real numbers a, b , is often used without mention in the proofs below. The convergence in distribution is denoted by $\rightarrow_d$ and $\mathcal{N}_p(a, B)$ denotes the p -dimensional normal distribution with mean vector a and covariance matrix B , $p \geq 1$.

2.3 Consistency of θ_n^* and $\hat{\theta}_n$

This section proves the consistency of θ_n^* and $\hat{\theta}_n$. To state and prove these results we need some more notation. For a σ -finite measure α on d -dimensional Borel space $(\mathbb{R}^d, \mathcal{B}^d)$, let $L_2(\alpha)$ denote a class of square integrable real valued functions on $\mathbb{R}^d$ with respect to α . Define

$$\begin{aligned} \rho(\nu_1, \nu_2; \Phi) &:= \int_{\mathcal{I}} (\nu_1(x) - \nu_2(x))^2 d\Phi(x), \\ \rho(\nu_1, \nu_2) &:= \int_{\mathcal{I}} (\nu_1(x) - \nu_2(x))^2 dG(x), \end{aligned}$$

$$\begin{aligned}
\rho_n(\nu_1, \nu_2) &:= \int_{\mathcal{I}} (\nu_1(x) - \nu_2(x))^2 dG_n(x) \\
&= n^{-1} \sum_{j=1}^n (\nu_1(X_j) - \nu_2(X_j))^2 \mathcal{I}_j, \quad \nu_1, \nu_2 \in L_2(G),
\end{aligned}$$

and the maps

$$T(\nu, \Phi) := \operatorname{argmin}_{\theta \in \Theta} \rho(\nu, m_\theta; \Phi), \quad T(\nu) := \operatorname{argmin}_{\theta \in \Theta} \rho(\nu, m_\theta),$$

$$T_n(\nu) := \operatorname{argmin}_{\theta \in \Theta} \rho_n(\nu, m_\theta), \quad \nu \in L_2(G), n \geq 1.$$

The following lemma has its roots in Beran (1977) and is proved in Ni (2002).

Lemma 2.3.1 *Let m satisfy the conditions (m1), (m2), and (m3). Then the following hold.*

(a) $\forall \nu \in L_2(\Phi)$, $T(\nu; \Phi)$ always exists.

(b) If $T(\nu; \Phi)$ is unique, then $T(\nu; \Phi)$ is continuous at ν in the following sense:

For any sequence of $\{\nu_n\}, \nu \in L_2(\Phi)$, $\rho(\nu_n, \nu; \Phi) \rightarrow 0$ implies $T(\nu_n; \Phi) \rightarrow T(\nu; \Phi)$.

(c) $\forall \theta \in \Theta$, $T(m_\theta) = \theta$, uniquely.

We need an analog of this lemma for the random distance ρ_n and the corresponding T_n given as follows.

Lemma 2.3.2 *Let m satisfy the conditions (m1), (m2), and (m3) with Φ replaced by G . Then the following hold.*

(a) $\forall \nu \in L_2(G)$, $T(\nu)$ always exists, and $T_n(\nu)$ exists $\forall n \geq 1$, w.p.1.

(b) If $T(\nu)$ is unique, then the following holds. For any sequence of $\{\nu_n\}, \nu \in L_2(G)$,

$$\rho_n(\nu_n, \nu) \rightarrow_p 0, \quad \text{implies} \quad T_n(\nu_n) \rightarrow_p T(\nu), \quad \text{as } n \rightarrow \infty.$$

(c) $\forall \theta \in \Theta$, $T(m_\theta(\cdot)) = \theta$, uniquely, and $T_n(m_\theta(\cdot)) = \theta$ uniquely, for all $n \geq 1$, w.p.1.

Proof. The following proof is a suitable modification of the proof that appears in Ni.

Proof of Part (a). The existence of $T(\nu)$ follows from (a) of Lemma 2.3.1. We shall prove that the family of random functions $\theta \mapsto \rho_n(\nu, m_\theta)$, $n \geq 1$, is almost surely equi-continuous. Then the claim (a) pertaining to T_n follows from the compactness of Θ . By the Cauchy-Schwarz inequality, for any $\vartheta, \theta \in \Theta$,

$$|\rho_n(\nu, m_\vartheta) - \rho_n(\nu, m_\theta)| \leq \rho_n(m_\vartheta, m_\theta) + 2\rho_n^{1/2}(\nu, m_\theta)\rho_n^{1/2}(m_\vartheta, m_\theta).$$

But, by (m3),

$$\rho_n(m_\vartheta, m_\theta) = \int_{\mathcal{I}} (m_\vartheta(x) - m_\theta(x))^2 dG_n(x) \leq \|\vartheta - \theta\|^2 n^{-1} \sum_{i=1}^n \ell^2(X_i) \mathcal{I}_i.$$

Since ℓ is continuous on $\mathbb{R}^d$ and $\mathcal{I}$ is compact then ℓ is bounded on $\mathcal{I}$ and hence

$$\sup_{n \geq 1} \rho_n(m_\vartheta, m_\theta) \leq C \|\vartheta - \theta\|^2, \quad \text{w.p.1.}$$

Similarly, under (m3),

$$\rho(m_\vartheta, m_\theta) \leq C \|\vartheta - \theta\|^2 \quad \forall \vartheta, \theta \in \Theta.$$

Because of the SLLN's and because $m_\vartheta, \nu \in L_2(G)$, $\rho_n(\nu, m_\vartheta) \rightarrow \rho(\nu, m_\vartheta)$, a.s. for each $\vartheta \in \Theta$. Also in view of the above bounds, both functions $\vartheta \mapsto \rho_n(\cdot, m_\vartheta)$ and $\vartheta \mapsto \rho(\cdot, m_\vartheta)$ are Lipschitz(2) uniformly in n and with probability 1. These facts together with the compactness of Θ imply

$$\begin{aligned} \sup_{\theta \in \Theta} |\rho_n(\nu, m_\theta) - \rho(\nu, m_\theta)| &\rightarrow 0, \text{ a.s., as } n \rightarrow \infty, \\ \sup_n |\rho_n(\nu, m_\vartheta) - \rho_n(\nu, m_\theta)| &\rightarrow 0, \text{ a.s., as } \|\vartheta - \theta\| \rightarrow 0, \end{aligned} \tag{2.4}$$

thereby completing the proof of equi-continuity of the map $\theta \mapsto \rho_n(\nu, m_\theta)$, and of part (a).

Proof of part (b). Let $\{\nu_n\}$, ν in $L_2(G)$ be such that

$$\rho_n(\nu_n, \nu) = o_p(1). \quad (2.5)$$

Let $\theta = T(\nu)$, $\vartheta_n = T_n(\nu_n)$. For an $\epsilon > 0$, let

$$A_{n,\epsilon} := \{\rho_n(\nu_n, \nu) \leq \epsilon, |\rho_n(\nu, m_\theta) - \rho(\nu, m_\theta)| \leq \epsilon\}.$$

By (2.4) and (2.5), there is an N_ϵ such that

$$P(A_{n,\epsilon}) \geq 1 - \epsilon, \quad \forall n > N_\epsilon. \quad (2.6)$$

Now, by the definition of T_n ,

$$\rho_n(\nu_n, m_{\vartheta_n}) \leq \rho_n(\nu_n, m_\theta), \quad \forall n \geq 1, \text{ w.p.1.}$$

By subtracting and adding ν inside of the square of the integrand, expanding the quadratic and using the Cauchy-Schwarz inequality on the cross product term,

$$\rho_n(\nu_n, m_\theta) \leq \rho_n(\nu, m_\theta) + \rho_n(\nu_n, \nu) + 2\rho_n^{1/2}(\nu_n, \nu)\rho_n^{1/2}(\nu, m_\theta).$$

On $A_{n,\epsilon}$, we thus obtain

$$\rho_n(\nu_n, m_{\vartheta_n}) \leq \rho(\nu, m_\theta) + \epsilon + 2\epsilon^{1/2}(\epsilon + \rho(\nu, m_\theta))^{1/2}. \quad (2.7)$$

On the other hand, again by the definition of T , T_n , θ , and ϑ_n , $\rho_n(\nu, m_\theta) \leq \rho_n(\nu, m_{\vartheta_n})$, for all $n \geq 1$, a.s. This, together with an argument like the above,

implies

$$\begin{aligned}
& \rho_n(\nu_n, m_{\vartheta_n}) - \rho_n(\nu, m_\theta) \\
& \geq \rho_n(\nu_n, m_{\vartheta_n}) - \rho_n(\nu, m_{\vartheta_n}) \\
& \geq \rho_n(\nu_n, \nu) - 2\rho_n^{1/2}(\nu_n, \nu)\rho_n^{1/2}(\nu, m_{\vartheta_n}) \quad \forall n \geq 1, \text{ w.p.1.}
\end{aligned}$$

But,

$$\rho_n(\nu, m_{\vartheta_n}) \leq 6\rho_n(\nu_n, \nu) + 4\rho_n(\nu, m_\theta).$$

Hence, on $A_{n,\epsilon}$,

$$\begin{aligned}
& \rho_n(\nu_n, m_{\vartheta_n}) \\
& \geq \rho_n(\nu, m_\theta) + \rho_n(\nu_n, \nu) - 2\rho_n^{1/2}(\nu_n, \nu)\{6\rho_n(\nu_n, \nu) + 4\rho_n(\nu, m_\theta)\}^{1/2} \\
& \geq \rho_n(\nu, m_\theta) - \epsilon - 2\epsilon^{1/2}(2\epsilon + \rho(\nu, m_\theta))^{1/2} \\
& \geq \rho(\nu, m_\theta) - 2\epsilon - 2\epsilon^{1/2}(2\epsilon + \rho(\nu, m_\theta))^{1/2}.
\end{aligned}$$

Thus, in view of (2.6), (2.7), and the arbitrariness of ϵ , we obtain

$$\rho_n(\nu_n, m_{\vartheta_n}) = \rho(\nu, m_\theta) + o_p(1). \quad (2.8)$$

From these facts it follows that $\vartheta_n \rightarrow_p \theta$. For, suppose $\vartheta_n \not\rightarrow \theta$, in probability. Then, by the compactness of Θ , there are subsequences ϑ_{n_k} of $\{\vartheta_n\}$ such that $\vartheta_{n_k} \rightarrow \vartheta \neq \theta$, and by (2.8), $\rho_{n_k}(\nu_{n_k}, m_{\vartheta_{n_k}}) \rightarrow \rho(\nu, m_\vartheta)$, in probability. Hence, $\rho(\nu, m_\vartheta) = \rho(\nu, m_\theta)$, implying, in view of the uniqueness of $T(\nu)$, a contradiction, unless $\vartheta = \theta$.

Proof of part (c). The claim here follows from the identifiability condition (m2) with $\Phi = G$. □

As in K-N, for any average $L := n^{-1} \sum_{j=1}^n (\gamma(X_j)/g_w^*(X_j))$, the replacement of g_w^* by g is reflected by the notation $\tilde{L} := n^{-1} \sum_{j=1}^n (\gamma(X_j)/g(X_j))$. We also need to define for $x \in \mathbb{R}^d$, $\theta \in \mathbb{R}^q$,

$$\begin{aligned}
\mu_n(x, \theta) &:= \frac{1}{n} \sum_{i=1}^n K_h(x - X_i) m_\theta(X_i), \\
\dot{\mu}_n(x, \theta) &:= \frac{1}{n} \sum_{i=1}^n K_h(x - X_i) \dot{m}_\theta(X_i), \\
U_n(x, \theta) &:= \frac{1}{n} \sum_{i=1}^n K_h(x - X_i) Y_i - \mu_n(x, \theta) \\
&= \frac{1}{n} \sum_{i=1}^n K_h(x - X_i) [Y_i - m_\theta(X_i)], \quad U_n(x) := U_n(x, \theta_0) \\
\dot{\mu}_h(x) &:= E \dot{\mu}_n(x, \theta_0) = EK_h(x - X) \dot{m}_{\theta_0}(x), \\
Z_n(x, \theta) &:= \mu_n(x, \theta) - \mu_n(x, \theta_0) \\
&= \frac{1}{n} \sum_{i=1}^n K_h(x - X_i) [m_\theta(X_i) - m_{\theta_0}(X_i)], \\
\mathcal{S}_n &:= \int \frac{U_n(x) \dot{\mu}_h(x)}{g^2(x)} d\Phi(x), \quad C_{n1} := \frac{1}{n} \sum_{j=1}^n \left[\frac{U_n^2(X_j)}{g_w^{*2}(X_j)} \right] \mathcal{I}_j, \\
C_{n2}(\theta) &:= \frac{1}{n} \sum_{j=1}^n \left[\frac{\mu_n(X_j, \theta)}{g_w^*(X_j)} - m_\theta(X_j) \right]^2 \mathcal{I}_j, \\
\Sigma &:= \int \frac{\sigma_{\theta_0}^2(x) \dot{m}_{\theta_0}(x) \dot{m}_{\theta_0}(x)' \phi^2(x)}{g(x)} dx.
\end{aligned} \tag{2.9}$$

Many of these entities are the empirical analogues of the entities defined at (3.1) in K-N.

Now, we will summarize the results of K-N and K-S for the sake of completeness.

The next two results state the consistency of ϑ_n^* and $\hat{\vartheta}_n$.

Result 2-3-1. Suppose $\mathcal{H}_0$, (e1), (e2), (k), (g1), (h1), (h2), and (m1)-(m3) hold.

Then, $\vartheta_n^* \rightarrow \theta_0$, in probability.

Result 2-3-2. Suppose $\mathcal{H}_0$, (e1), (e2), (k), (g1), (h1), (h2), and (m1)-(m3) hold.

Then, $\hat{\vartheta}_n \rightarrow \theta_0$, in probability.

The following result gives asymptotic normality of $\hat{\vartheta}_n$.

Result 2-3-3. Suppose $\mathcal{H}_0$, (e1), (e2), (e3), (g1), (g2), (k), (p), (h3), and (m1)-(m5) hold. Then,

$$n^{1/2}(\hat{\vartheta}_n - \theta_0) = \Sigma_0^{-1} n^{1/2} \mathcal{S}_n + o_p(1),$$

Consequently, $n^{1/2}(\hat{\vartheta}_n - \theta_0) \rightarrow_d \mathcal{N}_q(0, \Sigma_0^{-1} \Sigma \Sigma_0^{-1})$, where $\mathcal{S}_n$ and Σ are as in (2.9) and Σ_0 is as in (m6).

The following result states asymptotic normality of the minimized distance $\mathcal{I}_n(\hat{\vartheta}_n)$.

Result 2-3-4. Suppose $\mathcal{H}_0$, (e1), (e2), (e4), (k), (p), (g1), (g2), (h3), and (m1)-(m5) hold. Then, $nh^{d/2}(\mathcal{I}_n(\hat{\vartheta}_n) - \hat{R}_n) \rightarrow_d N_1(0, \Gamma)$. Moreover, $|\hat{\mathcal{G}}_n \Gamma^{-1} - 1| = o_p(1)$, where $\hat{\mathcal{G}}_n$, $\hat{R}_n$, and Γ are as in (2.1) and (2.2).

The following result from K-S gives consistency of ϑ_n^* and $\hat{\vartheta}_n$ for $T(m)$, where m is a given regression function, different from the model being fitted.

Result 2-3-5. Suppose (k), (g1), (m3) hold, and m is a given regression function such that $m \notin \{m_\theta; \theta \in \Theta\}$, $m \in L_2(\Phi)$, and $T(m)$ is unique.

(a) In addition, suppose m is a.e. (Φ) continuous. Then $\vartheta_n^* = T(m) + o_p(1)$.

(b) In addition, suppose m is continuous on $\mathcal{I}$. Then $\hat{\vartheta}_n = T(m) + o_p(1)$.

Next, we shall prove consistency of empirical minimum distance estimates of the underlying parameter vectors under $\mathcal{H}_0$.

Corollary 2.3.1 Assume $\mathcal{H}_0$, (e1), (e2), (g1), (k), (m1)-(m3), (h1), and (h2) hold, with Φ replaced by G . Then $\theta_n^* \rightarrow \theta_0$, in probability.

Proof. Note that $M_n^*(\theta_0) = \rho_n(\hat{m}_n, m_{\theta_0})$, $\theta_n^* = T_n(\hat{m}_n)$, and by the identifiability condition (m2), $T(m_{\theta_0}) = \theta_0$. It thus suffices to prove

$$\rho_n(\hat{m}_n, m_{\theta_0}) = o_p(1). \quad (2.10)$$

To prove this, substitute $m_{\theta_0}(X_i) + \epsilon_i$ for Y_i inside the i^{th} summand of $M_n^*(\theta_0)$ and expand the quadratic summand to obtain that $\rho_n(\hat{m}_n, m_{\theta_0})$ is bounded above by the sum $2[C_{n1} + C_{n2}(\theta_0)]$, where C_{n1} , C_{n2} are as in (2.9). It thus suffices to show that both of these terms are $o_p(1)$.

Since ϵ_i is conditionally centered, given X_i , and by continuity of g and $\sigma_{\theta_0}^2$, assured by (e2), (g1), (k) and (h2), we obtain

$$\begin{aligned} & E \left(n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j)}{g(X_j)} \mathcal{I}_j \right]^2 \right) \\ &= n^{-3} \sum_{j=1}^n \sum_{i=1}^n E \left[\frac{K_h(X_j - X_i) \epsilon_i}{g(X_j)} \mathcal{I}_j \right]^2 \\ &= n^{-3} h^{-2d} K^2(0) \sum_{i=1}^n E \left[\frac{\epsilon_i}{g(X_i)} \mathcal{I}_i \right]^2 + n^{-3} \sum_{j \neq i} E \left[\frac{K_h(X_j - X_i) \epsilon_i}{g(X_j)} \mathcal{I}_j \right]^2 \\ &= \frac{K^2(0)}{(nh^d)^2} \int_{\mathcal{I}} \frac{\sigma_{\theta_0}^2(x)}{g(x)} dx + \frac{1}{nh^d} \int \int_{\mathcal{I}} \frac{\sigma_{\theta_0}^2(x - uh) g(x - uh) K^2(u)}{g(x)} dx du \\ &= O(1/nh^d). \end{aligned} \quad (2.11)$$

Hence

$$n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j)}{g(X_j)} \mathcal{I}_j \right]^2 = O_p((nh^d)^{-1}), \quad (2.12)$$

and, by (2.3), $C_{n1} = o_p(1)$. Next, we shall show

$$C_{n2}(\theta_0) = o_p(1). \quad (2.13)$$

By taking the summations for $i = j$ and $i \neq j$, and by using the inequality $(a + b)^2 \leq 2(a^2 + b^2)$, for any real numbers a, b ,

$$C_{n2}(\theta) \leq 2[\hat{C}_{n21}(\theta) + \hat{C}_{n22}(\theta)], \quad \theta \in \Theta, \quad (2.14)$$

where

$$\begin{aligned} \hat{C}_{n21}(\theta) &= n^{-3} \sum_{j=1}^n \left[\frac{(K_h(0) - K_w^*(0))m_\theta(X_j)}{g_w^*(X_j)} \mathcal{I}_j \right]^2, \\ \hat{C}_{n22}(\theta) &= n^{-3} \sum_{j=1}^n \left[\sum_{i \neq j} \frac{(K_h(X_j - X_i)m_\theta(X_i) - K_w^*(X_j - X_i)m_\theta(X_j))}{g_w^*(X_j)} \mathcal{I}_j \right]^2. \end{aligned} \quad (2.15)$$

By the compactness of Θ , every open cover of Θ has a finite subcover $\{\theta_j; 1 \leq j \leq k\}$, say. For any $\delta > 0$ such that $\|\theta - \theta_j\| \leq \delta$, and by (m3),

$$\begin{aligned} &\sup_{\theta \in \Theta} E(\tilde{C}_{n21}(\theta)) \\ &\leq \frac{2K^2(0)}{n^2} (h^{-2d} - w^{-2d}) \sup_{1 \leq j \leq k} \sup_{\|\theta - \theta_j\| \leq \delta} \\ &\quad \times \int_{\mathcal{I}} \left[\frac{(m_\theta(x) - m_{\theta_j}(x))^2}{g(x)} + \frac{(m_{\theta_j}(x))^2}{g(x)} \right] dx \\ &\leq \frac{2CK^2(0)}{n^2} (h^{-2d} - w^{-2d}) (\delta^{2\beta} + 1) \sup_{j=1}^k \int_{\mathcal{I}} \frac{\ell^2(x) + m_{\theta_j}^2(x)}{g(x)} dx. \end{aligned}$$

Thus by (g1), (k), and by continuity of $m_\theta, \forall \theta \in \Theta$, we obtain

$$\sup_{\theta \in \Theta} E(\tilde{C}_{n21}(\theta)) = O(nh^d)^{-1}. \quad (2.16)$$

Hence $\hat{C}_{n21}(\theta_0) = o_p(1)$ follows from (h2), and (2.3).

To deal with $\hat{C}_{n22}$, let, for $j \neq i$

$$e_h(x, \theta) = E[K_h(X_j - X_i)m_\theta(X_i)|X_j = x],$$

$$e_w^*(x, \theta) = E[K_w^*(X_j - X_i)m_\theta(X_j)|X_j = x].$$

By adding and subtracting $e_h(X_j, \theta)$ and $e_w^*(X_j, \theta)$ in the quadratic term of the summand of $\hat{C}_{n22}$, one obtains

$$\hat{C}_{n22}(\theta) \leq 3C_{n221}(\theta) + 3C_{n222}(\theta) + 3C_{n223}(\theta), \quad \theta \in \Theta, \quad (2.17)$$

where

$$\begin{aligned} C_{n221}(\theta) &= n^{-3} \sum_{j=1}^n \left[\frac{\sum_{i \neq j} [K_h(X_j - X_i)m_\theta(X_i) - e_h(X_j, \theta)]}{g_w^*(X_j)} \mathcal{I}_j \right]^2, \quad (2.18) \\ C_{n222}(\theta) &= n^{-3} \sum_{j=1}^n \left[\frac{\sum_{i \neq j} [K_w^*(X_j - X_i)m_\theta(X_j) - e_w^*(X_j, \theta)]}{g_w^*(X_j)} \mathcal{I}_j \right]^2, \\ C_{n223}(\theta) &= \frac{(n-1)^2}{n^3} \sum_{j=1}^n \left[\frac{e_h(X_j, \theta) - e_w^*(X_j, \theta)}{g_w^*(X_j)} \mathcal{I}_j \right]^2. \end{aligned}$$

By the fact that the variance is bounded above by the second moment, one obtains

$\forall \theta \in \Theta$,

$$E\tilde{C}_{n221}(\theta) \leq \frac{1}{n^3} \sum_{i \neq j} E \left[\frac{K_h(X_j - X_i)m_\theta(X_i)}{g(X_j)} \mathcal{I}_j \right]^2.$$

Again proceeding as for (2.16), for any $\delta > 0$ such that $\|\theta - \theta_j\| \leq \delta$, we obtain

$$\begin{aligned} &\sup_{\theta \in \Theta} E(\tilde{C}_{n221}(\theta)) \quad (2.19) \\ &\leq \sup_{1 \leq j \leq k} \sup_{\|\theta - \theta_j\| \leq \delta} \frac{2}{n^3} \sum_{i \neq j} E \left[\frac{K_h(X_j - X_i)(m_\theta(X_i) - m_{\theta_j}(X_i))}{g(X_j)} \mathcal{I}_j \right]^2 \\ &\quad + \sup_{1 \leq j \leq k} \frac{2}{n^3} \sum_{i \neq j} E \left[\frac{K_h(X_j - X_i)m_{\theta_j}(X_i)}{g(X_j)} \mathcal{I}_j \right]^2 \end{aligned}$$

$$\begin{aligned}
&\leq \sup_{1 \leq j \leq k} \sup_{\|\theta - \theta_j\| \leq \delta} \frac{2}{n^3} \sum_{i \neq j} \|\theta - \theta_j\|^{2\beta} E \left[\frac{K_h(X_j - X_i) \ell(X_i)}{g(X_j)} \mathcal{I}_j \right]^2 \\
&\quad + \sup_{1 \leq j \leq k} \frac{2}{n^3} \sum_{i \neq j} E \left[\frac{K_h(X_j - X_i) m_{\theta_j}(X_i)}{g(X_j)} \mathcal{I}_j \right]^2 \\
&\leq \frac{2\delta^{2\beta}}{nh^d} \int_{\mathcal{I}} \int \frac{K^2(u) \ell^2(y - uh) g(y - uh)}{g(y)} du dy \\
&\quad + \frac{2}{nh^d} \sup_{1 \leq j \leq k} \int_{\mathcal{I}} \int \frac{K^2(u) m_{\theta_j}^2(y - uh) g(y - uh)}{g(y)} dy du \\
&= O(nh^d)^{-1},
\end{aligned}$$

by (m3), (g1), (k), and by continuity of $m_\theta, \forall \theta \in \Theta$. Hence, $C_{n221}(\theta_0) = O_p((nh^d)^{-1})$ follows from (2.3). Similarly, we can obtain that $C_{n222}(\theta_0) = O_p((nh^d)^{-1})$.

Next, we shall show $C_{n223}(\theta_0) = o_p(1)$. By adding and subtracting $Ee_h(X_j, \theta)$ and $Ee_w^*(X_j, \theta)$ in the quadratic term of the summand of C_{n223} , one obtains

$$C_{n223}(\theta) \leq 3I_{n1}(\theta) + 3I_{n2}(\theta) + 3I_{n3}(\theta), \quad \theta \in \Theta, \quad (2.20)$$

where

$$\begin{aligned}
I_{n1}(\theta) &= \frac{1}{n} \sum_{j=1}^n \left[\frac{e_h(X_j, \theta) - Ee_h(X_j, \theta)}{g_w^*(X_j)} \mathcal{I}_j \right]^2, \\
I_{n2}(\theta) &= \frac{1}{n} \sum_{j=1}^n \left[\frac{e_w^*(X_j, \theta) - Ee_w^*(X_j, \theta)}{g_w^*(X_j)} \mathcal{I}_j \right]^2, \\
I_{n3}(\theta) &= \frac{1}{n} \sum_{j=1}^n \left[\frac{Ee_h(X_j, \theta) - Ee_w^*(X_j, \theta)}{g_w^*(X_j)} \mathcal{I}_j \right]^2.
\end{aligned} \quad (2.21)$$

But $\forall \theta \in \Theta$,

$$E\tilde{I}_{n1}(\theta) \leq \frac{1}{n} E \left[\int \frac{K^2(u) m_\theta^2(X - uh) g^2(X - uh) I(X \in \mathcal{I})}{g^2(X)} du \right].$$

Similar to the argument as in (2.19), by boundedness of Θ , (m3), (g1), (k), and by continuity of $m_\theta, \forall \theta \in \Theta$, we thus obtain

$$\sup_{\theta \in \Theta} E\tilde{I}_{n1}(\theta) = O(n)^{-1}. \quad (2.22)$$

Hence $I_{n1}(\theta_0) = o_p(1)$ follows from (2.3). Similarly, one can obtain that $I_{n2}(\theta_0) = o_p(1)$. Also, by continuity of m_θ and g , one readily sees that

$$E\tilde{I}_{n3}(\theta) = o(1), \quad \text{for each } \theta \in \Theta. \quad (2.23)$$

Hence, $I_{n3}(\theta_0) = o(1)$ follows from (2.3). This completes the proof of (2.13) and hence that of (2.10) and the Corollary 2.3.1. $\square$

Remark 2.3.1: The basic ideas of the above proof is the same as that of Corollary 3.1 in K-N. The only difference is in some details, like e.g. in the derivation of the bounds (2.11) and (2.19). This phenomena is true in many proofs that follow. So we shall be brief in these proofs whenever possible.

Before stating the next result we give a fact that is often used in the proofs below. Under (g1), (k), and (h2), independence of X_i 's, and for any continuous function α , one obtain

$$\begin{aligned} & n^{-1} \sum_{j=1}^n E \left[n^{-1} \sum_{i=1}^n K_h(X_j - X_i) \frac{\alpha(X_i)}{g(X_j)} \mathcal{I}_j \right]^2 \\ & \leq \frac{2K^2(0)}{n^3 h^{2d}} \sum_{j=1}^n E \left[\frac{\alpha(X_j)}{g(X_j)} \mathcal{I}_j \right]^2 + \frac{2}{n^3} \sum_{j=1}^n E \left[\sum_{i \neq j} \frac{K_h(X_j - X_i) \alpha(X_i)}{g(X_j)} \mathcal{I}_j \right]^2 \\ & = \frac{2K^2(0)}{(nh^d)^2} E \left[\frac{\alpha(X) I(X \in \mathcal{I})}{g(X)} \right]^2 + \frac{2}{n^3} \sum_{i \neq j} E \left[\frac{K_h(X_j - X_i) \alpha(X_i)}{g(X_j)} \mathcal{I}_j \right]^2 \\ & \quad + \frac{2}{n^3} \sum_{l \neq i \neq j} E \left[\frac{K_h(X_j - X_i) K_h(X_j - X_l) \alpha(X_i) \alpha(X_l)}{g^2(X_j)} \mathcal{I}_j \right] \end{aligned} \quad (2.24)$$

$$\begin{aligned}
&= \frac{2K^2(0)}{(nh^d)^2} \int_{\mathcal{I}} \frac{\alpha^2(x)}{g(x)} dx + \frac{2}{nh^d} \int \int_{\mathcal{I}} \frac{K^2(u)\alpha^2(x-uh)g(x-uh)}{g(x)} dx du \\
&\quad + 2 \int_{\mathcal{I}} \left[\int (K(u)\alpha(x-uh)g(x-uh) du \right]^2 \frac{1}{g(x)} dx \\
&= o(1) + o(1) + O(1) = O(1).
\end{aligned}$$

We now proceed to state and prove

Theorem 2.3.1 *Suppose $\mathcal{H}_0$, (e1), (e2), (g1), (k), (m1)-(m3), (h1), and (h2) hold with Φ replaced by G . Then,*

$$\hat{\theta}_n \rightarrow \theta_0, \quad \text{in probability.}$$

Proof. Arguing as in K-N, we shall again use part (b) of Lemma 2.3.2 with $\nu(x) \equiv m_{\theta_0}(x)$, $\nu_n(x) \equiv m_{\hat{\theta}_n}(x)$. Then by (m2), $\hat{\theta}_n = T_n(\nu_n)$, $\theta_0 = T(\nu)$, uniquely. It thus suffices to show that

$$\sup_{\theta} |M_n(\theta) - M_n^*(\theta)| = o_p(1). \quad (2.25)$$

For, (2.25) implies that

$$\begin{aligned}
M_n^*(\hat{\theta}_n) &= M_n(\hat{\theta}_n) + o_p(1), \quad M_n^*(\theta_n^*) = M_n(\theta_n^*) + o_p(1), \\
M_n^*(\hat{\theta}_n) - M_n^*(\theta_n^*) &= M_n(\hat{\theta}_n) - M_n(\theta_n^*) + o_p(1).
\end{aligned} \quad (2.26)$$

By the definitions of $\hat{\theta}_n$ and θ_n^* , for every n , the left hand side of (2.26) is non negative, while the first term on the right hand side is non-positive. Hence,

$$M_n^*(\hat{\theta}_n) - M_n^*(\theta_n^*) = o_p(1).$$

This together with the fact that $M_n^*(\theta_n^*) \leq M_n^*(\theta_0)$ and (2.10) then proves the required result.

Arguing as in the proof of Theorem 3.1 of K-N, it thus suffices to show that

$$\sup_{\theta} C_{n2}(\theta) = o_p(1), \quad \sup_{\theta} M_n(\theta) = O_p(1). \quad (2.27)$$

Using the same argument as in (2.14) - (2.23), one obtains, $\tilde{C}_{n2}(\theta) = o_p(1)$, for each $\theta \in \Theta$. This and (2.3) in turn imply that

$$C_{n2}(\theta) \leq \sup_{x \in \mathcal{I}} \frac{g^2(x)}{g_w^{*2}(x)} \tilde{C}_{n2}(\theta) = o_p(1), \quad \forall \theta \in \Theta. \quad (2.28)$$

By the Cauchy-Schwarz inequality, for any $\theta_1, \theta_2 \in \Theta$

$$|C_{n2}(\theta_2) - C_{n2}(\theta_1)| \leq 2(E_1 + E_2) + 4C_{n2}^{1/2}(\theta_1)[E_1 + E_2]^{1/2},$$

where, by (m3),

$$\begin{aligned} E_1 &:= n^{-3} \sum_{j=1}^n \left[\sum_{i=1}^n \frac{K_h(X_j - X_i)(m_{\theta_2}(X_i) - m_{\theta_1}(X_i))}{g_w^*(X_j)} \mathcal{I}_j \right]^2 \\ &\leq \|\theta_2 - \theta_1\|^{2\beta} \sup_{x \in \mathcal{I}} \frac{g^2(x)}{g_w^{*2}(x)} \left[n^{-3} \sum_{j=1}^n \left(\sum_{i=1}^n \frac{K_h(X_j - X_i)\ell(X_i)}{g(X_j)} \mathcal{I}_j \right)^2 \right], \end{aligned}$$

$$E_2 := n^{-1} \sum_{j=1}^n \left[[m_{\theta_2}(X_j) - m_{\theta_1}(X_j)] \mathcal{I}_j \right]^2 \leq \|\theta_2 - \theta_1\|^{2\beta} n^{-1} \sum_{j=1}^n [\ell(X_j) \mathcal{I}_j]^2.$$

Hence (2.24) applied with $\alpha \equiv \ell$, (2.3) together with the compactness of Θ and (2.28) completes the proof of the first part of (2.27).

To prove the second part of (2.27), note that by adding and subtracting $m_{\theta_0}(X_i)$ to the i^{th} summand in $M_n(\theta)$, we obtain

$$M_n(\theta) \leq 2 \sup_{x \in \mathcal{I}} \frac{g^2(x)}{g_w^{*2}(x)} \left(n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) \mathcal{I}_j}{g(X_j)} \right]^2 + n^{-1} \sum_{j=1}^n \left[\frac{Z_n(X_j, \theta) \mathcal{I}_j}{g(X_j)} \right]^2 \right).$$

But, by a similar argument as for (2.19), and by (2.24) applied with $\alpha \equiv \ell$,

$$\begin{aligned} & \sup_{\theta} \left[n^{-1} \sum_{j=1}^n \left[\frac{Z_n(X_j, \theta) \mathcal{I}_j}{g(X_j)} \right]^2 \right] \\ & \leq \|\theta - \theta_0\|^{2\beta} n^{-3} \sum_{j=1}^n \left[\sum_{i=1}^n \frac{K_h(X_j - X_i) \ell(X_i) \mathcal{I}_j}{g(X_j)} \right]^2 = O_p(1). \end{aligned} \quad (2.29)$$

This together with (2.12) complete the proof of the second part of (2.27), and hence that of Theorem 2.3.1. $\square$

2.4 Asymptotic distribution of $\hat{\theta}_n$

In this section we shall prove the asymptotic normality of $n^{1/2}(\hat{\theta}_n - \theta_0)$. Let

$$\begin{aligned} \dot{\mu}_h(x) &:= E \dot{\mu}_n(x, \theta_0) = EK_h(x - X) \dot{m}_{\theta_0}(X), \\ \hat{S}_n &:= n^{-1} \sum_{j=1}^n \frac{U_n(X_j) \dot{\mu}_h(X_j) \mathcal{I}_j}{g(X_j)}. \end{aligned} \quad (2.30)$$

We shall prove the following

Theorem 2.4.1 *Assume that (e1), (e2), (e3), (g1), (g2), (k), (m1)-(m5), and (h3) hold, and Φ is replaced by G . Then under $\mathcal{H}_0$,*

$$n^{1/2}(\hat{\theta}_n - \theta_0) = \Sigma_0^{-1} n^{1/2} \hat{S}_n + o_p(1), \quad (2.31)$$

Consequently, $n^{1/2}(\hat{\theta}_n - \theta_0) \rightarrow_d \mathcal{N}_q(0, \Sigma_0^{-1} \Sigma \Sigma_0^{-1})$, where Σ_0 and Σ are as in (m6) and (2.9), respectively.

Proof. The proof consists of several steps. The first is to show that

$$nh^d \|\hat{\theta}_n - \theta_0\|^2 = O_p(1). \quad (2.32)$$

Let

$$D_n(\theta) := n^{-1} \sum_{j=1}^n \left[\frac{Z_n(X_j, \theta) \mathcal{I}_j}{g(X_j)} \right]^2.$$

We claim

$$nh^d D_n(\hat{\theta}_n) = O_p(1). \quad (2.33)$$

To see this observe that

$$\begin{aligned} nh^d M_n(\theta_0) &= nh^d \left[n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) \mathcal{I}_j}{g_w^*(X_j)} \right]^2 \right] \\ &\leq nh^d \left[n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) \mathcal{I}_j}{g(X_j)} \right]^2 \right] \\ &\quad + nh^d \left[n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) \mathcal{I}_j}{g(X_j)} \right]^2 \right] \sup_{x \in \mathcal{I}} \left| \frac{g^2(x)}{g_w^{*2}(x)} - 1 \right| \\ &= O_p(1), \end{aligned}$$

by (2.12) and (2.3). But, by definition, $M_n(\hat{\theta}_n) \leq M_n(\theta_0)$, implying that $nh^d M_n(\hat{\theta}_n) = O_p(1)$. These facts together with the inequality $D_n(\theta) \leq 2[M_n(\theta_0) + M_n(\hat{\theta}_n)]$ proves (2.33).

To complete the proof of (2.32), arguing as in K-N, it suffices to show for any $0 < a < 1$, there exists an N_a such that

$$P \left(\frac{D_n(\hat{\theta}_n)}{\|\hat{\theta}_n - \theta_0\|^2} \geq a + \inf_{\|b\|=1} b^T \Sigma_0 b \right) > 1 - a, \quad \forall n > N_a, \quad (2.34)$$

where Σ_0 as in (m6). To prove (2.34), let

$$u_n := \hat{\theta}_n - \theta_0, \quad (2.35)$$

$$d_{ni} := m_{\hat{\theta}_n}(X_i) - m_{\theta_0}(X_i) - u_n' \dot{m}_{\theta_0}(X_i), \quad 1 \leq i \leq n,$$

$$\begin{aligned}
D_{n1} &:= n^{-1} \sum_{j=1}^n \left[\left(\frac{1}{n} \sum_{i=1}^n K_h(X_j - X_i) \left(\frac{d_{ni}}{\|u_n\|} \right) \mathcal{I}_j \right) / g(X_j) \right]^2, \\
D_{n2} &:= n^{-1} \sum_{j=1}^n \left[\frac{u_n' \dot{\mu}_n(X_j, \theta_0) \mathcal{I}_j}{\|u_n\| g(X_j)} \right]^2.
\end{aligned}$$

Then, we have $D_n(\hat{\theta}_n)/\|\hat{\theta}_n - \theta_0\|^2 \geq D_{n1} + D_{n2} - 2D_{n1}^{1/2}D_{n2}^{1/2}$.

By assumption (m4), consistency of $\hat{\theta}_n$, and by using (2.24) with $\alpha = 1$, one verifies that $D_{n1} = o_p(1)$. For the term D_{n2} , note that

$$D_{n2} \geq \inf_{\|b\|=1} \Sigma_n(b), \quad (2.36)$$

where

$$\Sigma_n(b) := n^{-1} \sum_{j=1}^n \left[\frac{b' \dot{\mu}_n(X_j, \theta_0) \mathcal{I}_j}{g(X_j)} \right]^2, \quad b \in \mathbb{R}^q.$$

Now, we will prove that for each $b \in \mathbb{R}^q$, $\|b\| = 1$, $\Sigma_n(b) \rightarrow b' \Sigma_0 b$, in probability.

For this it suffices to show that

$$E[\Sigma_n(b) - b' \Sigma_0 b]^2 = o(1), \quad \forall b \in \mathbb{R}^q. \quad (2.37)$$

Rewrite

$$\begin{aligned}
\Sigma_n(b) &= \frac{1}{n^3} \sum_{j=1}^n \sum_{i=1}^n \left(\frac{b' K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i) \mathcal{I}_j}{g(X_j)} \right)^2 \\
&\quad + \frac{1}{n^3} \sum_{j=1}^n \sum_{i \neq k} \frac{b' K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i) K_h(X_j - X_k) \dot{m}_{\theta_0}(X_k)' b \mathcal{I}_j}{g^2(X_j)} \\
&:= \Sigma_{n1}(b) + \Sigma_{n2}(b) + 2\Sigma_{n3}(b) + \Sigma_{n4}(b),
\end{aligned}$$

where

$$\Sigma_{n1}(b) = \frac{K^2(0)}{n^3 h^{2d}} \sum_{j=1}^n \left[\frac{b' \dot{m}_{\theta_0}(X_j)}{g(X_j)} \mathcal{I}_j \right]^2,$$

$$\begin{aligned}
\Sigma_{n2}(b) &= \frac{1}{n^3} \sum_i \sum_{j \neq i} \left[\frac{b' K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i)}{g(X_j)} \mathcal{I}_j \right]^2, \\
\Sigma_{n3}(b) &= \frac{K(0)}{n^3 h^d} \sum_i \sum_{j \neq i} \left[\frac{b' K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i) \dot{m}_{\theta_0}(X_j)' b}{g^2(X_j)} \mathcal{I}_j \right], \\
\Sigma_{n4}(b) &= \frac{1}{n^3} \sum_{i \neq j, k} \left[\frac{b' K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i) K_h(X_j - X_k) \dot{m}_{\theta_0}(X_k)' b}{g^2(X_j)} \mathcal{I}_j \right].
\end{aligned}$$

The left hand side of (2.37) is less than or equal to

$$6[E\Sigma_{n1}^2(b) + E\Sigma_{n2}^2(b) + E\Sigma_{n3}^2(b)] + 2E[\Sigma_{n4}(b) - b'\Sigma_0 b]^2.$$

Thus to prove (2.37), it is enough to show for each $b \in \mathbb{R}^q$,

$$E\Sigma_{n1}^2(b) = o(1), \quad E\Sigma_{n2}^2(b) = o(1), \quad (2.38)$$

$$E\Sigma_{n3}^2(b) = o(1), \quad E[\Sigma_{n4}(b) - b'\Sigma_0 b]^2 = o(1).$$

Now, we shall prove the first part of (2.38). By the Cauchy-Schwarz inequality,

$$\Sigma_{n1}^2(b) \leq \frac{K^4(0)}{n^5 h^{4d}} \sum_{j=1}^n \left[\frac{\|\dot{m}_{\theta_0}(X_j)\|^2}{g^2(X_j)} \mathcal{I}_j \right].$$

Therefore, by (g1), (k) and (h2), one obtains

$$\begin{aligned}
E \sup_b \Sigma_{n1}^2(b) &\leq \frac{K^4(0)}{n^4 h^{4d}} \int_{\mathcal{I}} \frac{\|\dot{m}_{\theta_0}(x)\|^2}{g(x)} dx \\
&= O(nh^d)^{-1} = o(1).
\end{aligned}$$

Similarly,

$$\begin{aligned}
E \sup_b \Sigma_{n3}^2(b) &\leq \frac{K^2(0)}{n^4 h^{4d}} E \sup_b \sum_i \sum_{j \neq i} \left[\frac{b' K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i) \dot{m}_{\theta_0}(X_j)' b}{g^2(X_j)} \mathcal{I}_j \right]^2 \\
&\leq \frac{K^2(0)}{n^2 h^{4d}} \int \int_{\mathcal{I}} \frac{K_h^4(x-y) \|\dot{m}_{\theta_0}(x)\|^2 \|\dot{m}_{\theta_0}(y)\|^2 g(y)}{g^3(x)} dx dy \\
&\leq \frac{K^2(0)}{n^2 h^{3d}} \int \int_{\mathcal{I}} \frac{K^4(u) \|\dot{m}_{\theta_0}(x)\|^2 \|\dot{m}_{\theta_0}(x-hu)\|^2 g(x-hu)}{g^3(x)} dx du \\
&= O(n^2 h^{3d})^{-1} = o(1).
\end{aligned}$$

by (g1), (k), (h2), and continuity of $\dot{m}_{\theta_0}$. This proves the third claim in (2.38).

Now we shall prove the fourth claim in (2.38). Since,

$$\begin{aligned}
E\Sigma_{n4}(b) &= E \left[\frac{b' K_h(X_2 - X_1) \dot{m}_{\theta_0}(X_1) K_h(X_2 - X_3) \dot{m}_{\theta_0}(X_3)' I_2 b}{g^2(X_2)} \right] \\
&= b' \int_{\mathcal{I}} \int \int K(u) K(v) \dot{m}_{\theta_0}(x - hu) \dot{m}_{\theta_0}(x - vh)' \\
&\quad \times \frac{g(x - hu) g(x - vh)}{g(x)} du dv dx b \\
&\rightarrow b' \Sigma_0 b.
\end{aligned}$$

Thus, to prove $E[\Sigma_{n4}(b) - b' \Sigma_0 b]^2 = o(1)$, it is enough to show $E\Sigma_{n4}^2(b) = (b' \Sigma_0 b)^2$.

Now,

$$\begin{aligned}
\Sigma_{n4}^2(b) &= \sum_{i \neq j \neq k} \sum_{l \neq m \neq n} z_{ijk} z_{lmn} \\
&\leq C \sum_{i \neq j \neq k \neq m \neq l} z_{ijk} (z_{lmi} + z_{lmj} + z_{ljm}) + \sum_{i \neq j \neq k \neq m \neq l \neq n} z_{ijk} z_{lmn} \\
&= \Sigma_{n41}(b) + \Sigma_{n42}(b) + \Sigma_{n43}(b) + \Sigma_{n44}(b), \quad \text{say,}
\end{aligned}$$

where

$$z_{ijk} := b' n^{-3} \left[\frac{K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i) K_h(X_j - X_k) \dot{m}_{\theta_0}(X_k)'}{g^2(X_j)} \mathcal{I}_j \right] b.$$

By independence of X_i 's, (k), (m1), (g1), and (h2), one obtain for each $b \in \mathbb{R}^q$,

$$\|b\| = 1,$$

$$\begin{aligned}
E\Sigma_{n41}(b) &\leq \frac{1}{nh^{2d}} \int \int \left[\int_{\mathcal{I}} K(u) K(v) b' \dot{m}_{\theta_0}(x - uh) \dot{m}_{\theta_0}(x - vh)' b \right. \\
&\quad \left. \times \frac{(g(x - uh) g(x - vh))^{1/2}}{g(x)} dx \right]^2 du dv \\
&= O(nh^{2d})^{-1} = o(1).
\end{aligned}$$

Similarly, one can obtain that $E\Sigma_{n42}(b) = o(1)$, for each $b \in \mathbb{R}^q$ such that $\|b\| = 1$.

Again by independence of X_i 's, (k), (m1), and (g1), one obtain for each $b \in \mathbb{R}^q$, $\|b\| = 1$,

$$\begin{aligned}
& E\Sigma_{n43}(b) \\
& \leq \frac{1}{n} \int \left[\int_{\mathcal{I}} \int_{\mathcal{I}} K(u)K(v) \|\dot{m}_{\theta_0}(x - uh)\| \|\dot{m}_{\theta_0}(x - vh)\| \right. \\
& \quad \left. \times g(x - hu)g(x - hv) du dv \right]^2 g^{-3}(x) dx \\
& = O(n^{-1}) = o(1).
\end{aligned}$$

Also by independence of X_i 's, $E\Sigma_{n44}(b) = (E\Sigma_{n4}(b))^2 \rightarrow (b'\Sigma_0 b)^2$, for each $b \in \mathbb{R}^q$, $\|b\| = 1$. This also completes the proof of (2.37).

Also note that for any $\lambda > 0$, and any two unit vectors $b_1, b_2 \in \mathbb{R}^q$, $\|b_1\| = 1 = \|b_2\|$, $\|b_2 - b_1\| \leq \lambda$, we have

$$\begin{aligned}
& |\Sigma_{n1}(b_2) - \Sigma_{n1}(b_1)| \\
& = \left| n^{-1} \sum_{j=1}^n \left[(b_2 - b_1)' \frac{1}{n} \sum_{i=1}^n \frac{K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i)}{g(X_j)} \mathcal{I}_j \right]^2 \right| \\
& \quad + 2 \left| n^{-1} \sum_{j=1}^n \left[(b_2 - b_1)' \frac{1}{n} \sum_{i=1}^n \frac{K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i)}{g(X_j)} \mathcal{I}_j \right] \right. \\
& \quad \left. \times \left[\frac{b_1'}{n} \sum_{i=1}^n \frac{K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i)}{g(X_j)} \mathcal{I}_j \right] \right| \\
& \leq \lambda(\lambda + 2)n^{-1} \sum_{j=1}^n \left\| \frac{1}{n} \sum_{i=1}^n \frac{K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i)}{g(X_j)} \mathcal{I}_j \right\|^2.
\end{aligned}$$

But similar to the argument as for Σ_n ,

$$\begin{aligned}
& E \left[\frac{1}{n} \sum_{j=1}^n \left\| n^{-1} \sum_{i=1}^n K_h(X_j - X_i) \dot{m}_{\theta_0}(X_i) / g(X_j) \mathcal{I}_j \right\|^2 \right] \\
& \rightarrow \int \|\dot{m}_{\theta_0}(x)\|^2 dG(x), \text{ in probability,}
\end{aligned}$$

Hence, the second factor is $O_p(1)$. From these observations and the compactness of $\{b \in \mathbb{R}^q : \|b\| = 1\}$, one obtains $\sup_{\|b\|=1} \|\Sigma_n(b) - b' \Sigma_0 b\| = o_p(1)$.

This fact, together with (2.36), implies (2.34), and also concludes the proof of (2.32).

Now, we shall prove asymptotic normality of $n^{1/2}(\hat{\theta}_n - \theta_0)$. Recall the definitions of (2.9) and (2.35), and let

$$\dot{M}_n(\theta) := -2n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j, \theta) \dot{\mu}_n(X_j, \theta)}{g_w^{*2}(X_j)} \mathcal{I}_j \right].$$

Since θ_0 is an interior point of Θ , by consistency, for sufficiently large n , $\hat{\theta}_n$ will be in the interior of Θ , and $\dot{M}_n(\hat{\theta}_n) = 0$, with arbitrarily large probability. But the equation $\dot{M}_n(\hat{\theta}_n) = 0$ is equivalent to

$$n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) \dot{\mu}_n(X_j, \hat{\theta}_n)}{g_w^{*2}(X_j)} \mathcal{I}_j \right] = n^{-1} \sum_{j=1}^n \left[\frac{\mathcal{Z}_n(X_j, \hat{\theta}_n) \dot{\mu}_n(X_j, \hat{\theta}_n)}{g_w^{*2}(X_j)} \mathcal{I}_j \right]. \quad (2.39)$$

In the final step of the proof we shall show that $\sqrt{n} \times$ the left hand side of this equation converges in distribution to a normal random variable, while the right hand side of this equation equals $Q_n(\hat{\theta}_n - \theta_0)$, for all $n \geq 1$, with $Q_n = \Sigma_0 + o_p(1)$. To establish the first of these two claims, rewrite this random variable as the sum of $S_n + S_{n1} + g_{n1} + g_{n2} + g_{n3} + g_{n4}$, where $\hat{S}_n$ is as in (2.30) and

$$\begin{aligned} S_{n1} &:= n^{-1} \sum_{j=1}^n [U_n(X_j) \dot{\mu}_h(X_j) (\hat{g}_w^{-2}(X_j) - g^{-2}(X_j)) \mathcal{I}_j], \\ g_{n1} &:= n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) [\dot{\mu}_n(X_j, \theta_0) - \dot{\mu}_h(X_j)]}{g^2(X_j)} \mathcal{I}_j \right], \\ g_{n2} &:= n^{-1} \sum_{j=1}^n [U_n(X_j) [\dot{\mu}_n(X_j, \theta_0) - \dot{\mu}_h(X_j)] (\hat{g}_w^{-2}(X_j) - g^{-2}(X_j)) \mathcal{I}_j], \end{aligned}$$

$$\begin{aligned}
g_{n3} &:= n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j)[\dot{\mu}_n(X_j, \hat{\theta}_n) - \dot{\mu}_n(X_j, \theta_0)]}{g^2(X_j)} \mathcal{I}_j \right], \\
g_{n4} &:= n^{-1} \sum_{j=1}^n [U_n(X_j)[\dot{\mu}_n(X_j, \hat{\theta}_n) - \dot{\mu}_n(X_j, \theta_0)](\hat{g}_w^{-2}(X_j) - g^{-2}(X_j)) \mathcal{I}_j].
\end{aligned}$$

Note that these g_{nj} 's are the empirical analogs of the similar entities in K-N. Analogous to the proof there, we need the following lemmas.

Lemma 2.4.1 *Suppose (e1), (e2), (g1), (k), (m1)-(m5), (h1), (h2) and $\mathcal{H}_0$ with Φ replaced by G hold.*

(i) *If, additionally, (e3) hold, then $\sqrt{n} \hat{S}_n \rightarrow_d N_q(0, \Sigma)$, where Σ is as in (2.9).*

(ii) *If, additionally, (g2) and (h3) holds, then $\forall \lambda \in \mathbb{R}^q$*

$$\sqrt{n} \|\lambda' S_{n1}\| = o_p(1). \quad (2.40)$$

Lemma 2.4.2 *Suppose (e1), (e2), (g1), (k), (m1), (m2), (m4), (m5), (h1), and (h2) with Φ replaced by G hold. Then, under $\mathcal{H}_0$, $\forall \lambda \in \mathbb{R}^q$*

$$n^{1/2} \|\lambda' g_{nk}\| = o_p(1), \quad k = 1, 2, 3, 4. \quad (2.41)$$

The proof of (2.40) is facilitated by the following lemma, which along with its proof appears as Theorem 2.2 part (2) in Bosq (1998).

Lemma 2.4.3 *Let $g_w^*(x)$, $x \in \mathbb{R}^d$, $d \geq 1$, be the kernel estimator associated with a kernel K^* which satisfies a Lipschitz condition. If (g2) holds and $w_n = a_n(\log n/n)^{1/(d+4)}$, where $a_n \rightarrow a_0 > 0$, then for any positive integer k ,*

$$(\log_k n)^{-1} (n/\log n)^{2/(d+4)} \sup_{x \in \mathcal{I}} |g_w^*(x) - g(x)| \rightarrow 0 \quad a.s.$$

Proof of Lemma 2.4.1. Let $\hat{\mathcal{S}}_n$ denote $\mathcal{S}_n$ of (2.9) with $\Phi = G_n$. To prove the first part of Lemma 2.4.1, by Slutsky's Theorem, it suffices to show that

$$\sqrt{n}\mathcal{S}_n \rightarrow_d N_q(0, \Sigma), \quad \sqrt{n}\|\hat{\mathcal{S}}_n - \mathcal{S}_n\| = o_p(1). \quad (2.42)$$

First part of (2.42) follows from Lemma 4.1 of K-N. To prove the second part of (2.42), it suffices to show that, $\forall \lambda \in \mathbb{R}^q$,

$$E[\sqrt{n}\lambda'(\hat{\mathcal{S}}_n - \mathcal{S}_n)]^2 = o(1). \quad (2.43)$$

Let $\dot{\nu}_h(x) = \lambda' \dot{\mu}_h(x)$ and $a_{ij} := n^{-3/2} \varepsilon_i c_{ij}$, where

$$c_{ij} = \left[\frac{K_h(X_j - X_i) \dot{\nu}_h(X_j)}{g^2(X_j)} \mathcal{I}_j - \int_{\mathcal{I}} \frac{K_h(x - X_i) \dot{\nu}_h(x)}{g^2(x)} dG(x) \right].$$

Now, the left hand side of equation (2.43) can be rewritten as the following sum:

$$6 \sum_i E(a_{ii}^2) + \sum_{j \neq i} E(a_{ij}^2) + 4 \sum_{j \neq i} E(a_{ij} a_{ii}) + \sum_{m \neq i \neq j} E(a_{ij} a_{im}). \quad (2.44)$$

To prove (2.43), it suffices to show that each of the four terms of (2.44) are $o(1)$. By continuity of $\dot{m}_{\theta_0}$, (k), and (g1), $\forall \lambda \in \mathbb{R}^q$ and $\forall x \in \mathcal{I}$, one obtains

$$\begin{aligned} \sup_n \|\dot{\nu}_h(x)\| &\leq EK_h(x - X) \|\lambda' \dot{m}_{\theta_0}(X)\| \\ &= \int K(u) \|\lambda' \dot{m}_{\theta_0}(x - hu)\| g(x - hu) du \\ &= O(1). \end{aligned} \quad (2.45)$$

Now, we shall show that $\sum_{j \neq i} E(a_{ij}^2) = o(1)$. By (2.45), (k), (g1), and continuity of $\sigma_{\theta_0}^2$, for $j \neq i$, and $\forall \lambda \in \mathbb{R}^q$,

$$Ea_{ij}^2 = n^{-3} E(\varepsilon_i^2 c_{ij}^2) = n^{-3} E(c_{ij}^2 \sigma_{\theta_0}^2(X_i))$$

$$\begin{aligned}
&= n^{-3} \int \sigma_{\theta_0}^2(y) \int_{\mathcal{I}} \frac{K_h^2(x-y) \dot{\nu}_h^2(x)}{g(x)} g(y) dx dy \\
&\leq C n^{-3} h^{-d} \int_{\mathcal{I}} \sigma_{\theta_0}^2(x-hu) \int \frac{K^2(u) g(x-hu)}{g(x)} du dx \\
&= O(n^{-3} h^{-d})
\end{aligned}$$

Hence, by (h2), $\sum_{j \neq i} E(a_{ij}^2) = O(nh^{-d})^{-1} = o(1)$, $\forall \lambda \in \mathbb{R}^q$. Similarly, we can show that $\sum_i E(a_{ii}^2) = o(1)$, $\forall \lambda \in \mathbb{R}^q$.

Note that, $\forall \lambda \in \mathbb{R}^q$, $\forall i \neq j$, $E[c_{ij}|X_i = y] = 0$. Thus, by the independence of X_i 's, $\sum_{j \neq i} E(a_{ij} a_{ii} | X_i) = 0 = \sum_{m \neq i \neq j} E(a_{ij} a_{im} | X_i)$, for all $n \geq 1$. This completes the proof of the second part of (2.42), and hence that of the first part of Lemma 2.4.1.

To prove (2.40), by the Cauchy-Schwarz inequality, (2.45), (2.12), by Lemma 2.4.3, and by (h3), we obtain $\forall \lambda \in \mathbb{R}^q$

$$\begin{aligned}
&n[\lambda' S_{n1}]^2 \\
&\leq n \left[\frac{1}{n} \sum_{j=1}^n \left[\frac{U_n(X_j) \dot{\nu}_h(X_j)}{g^2(X_j)} \mathcal{I}_j \right] \right]^2 \sup_{x \in \mathcal{I}} \left| \frac{g^2(x)}{g_w^{*2}(x)} - 1 \right|^2 \\
&\leq n \left[n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) \mathcal{I}_j}{g(X_j)} \right]^2 \right] \left[n^{-1} \sum_{j=1}^n \frac{\dot{\nu}_h(X_j)}{g(X_j)} \mathcal{I}_j \right]^2 \sup_{x \in \mathcal{I}} \left| \frac{g^2(x)}{g_w^{*2}(x)} - 1 \right|^2 \\
&\leq C n \left[n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) \mathcal{I}_j}{g(X_j)} \right]^2 \right] \left[n^{-1} \sum_{j=1}^n \frac{1}{g(X_j)} \mathcal{I}_j \right]^2 \sup_{x \in \mathcal{I}} \left| \frac{g^2(x)}{\hat{g}_w^2(x)} - 1 \right|^2 \\
&= n O_p((nh^d)^{-1}) O_p(1) O_p((\log_k n)^2 (\log n/n)^{\frac{4}{d+4}}) \\
&= O_p((\log_k n)^2 (\log n)^{\frac{4}{d+4}} n^{ad - \frac{4}{d+4}}) = o_p(1).
\end{aligned}$$

This completes the proof of Lemma 2.4.1. □

Proof of Lemma 2.4.2. Let $\dot{\nu}_n := \lambda' \dot{\mu}_n$, $\dot{\nu}_h := \lambda' \dot{\mu}_h$, and $\dot{\nu} := \lambda' \dot{m}$. By the Cauchy-Schwarz inequality, $\forall \lambda \in \mathbb{R}^q$

$$\begin{aligned} n^{1/2} \|\lambda' g_{n1}\| &\leq \left(n^{-1/2} \sum_{j=1}^n \left[\frac{U_n(X_j) \mathcal{I}_j}{g(X_j)} \right]^2 \right) \\ &\quad \times \left(n^{-1/2} \sum_{j=1}^n \left[\frac{\|\dot{\nu}_n(X_j, \theta_0) - \dot{\nu}_h(X_j)\|}{g(X_j)} \mathcal{I}_j \right]^2 \right). \end{aligned} \quad (2.46)$$

By (2.11) and (h2),

$$E n^{-1/2} \sum_{j=1}^n \left[\frac{U_n(X_j) \mathcal{I}_j}{g(X_j)} \right]^2 = O(n^{-1/2} h^{-d}) = o(1).$$

The second factor is bounded by $2b_{n1} + 2b_{n2}$, where

$$\begin{aligned} b_{n1} &= n^{-5/2} \sum_{j=1}^n \left[\frac{\|K_h(0) \dot{\nu}_{\theta_0}(X_j) - \int K_h(X_j - y) \dot{\nu}_{\theta_0}(y) dG(y)\|}{g(X_j)} \mathcal{I}_j \right]^2, \\ b_{n2} &= n^{-1/2} \sum_{j=1}^n \\ &\quad \times \left[n^{-1} \sum_{i \neq j} \frac{\|K_h(X_j - X_i) \dot{\nu}_{\theta_0}(X_i) - \int K_h(X_j - y) \dot{\nu}_{\theta_0}(y) dG(y)\|}{g(X_j)} \mathcal{I}_j \right]^2. \end{aligned}$$

Now, by using (g1), (k), (h2), and continuity of $\dot{m}_{\theta_0}$, we obtain that the expected value of b_{n1} is bounded above by

$$\begin{aligned} &\frac{2K^2(0)}{n^{3/2} h^{2d}} \int_{\mathcal{I}} \frac{\|\dot{\nu}_{\theta_0}(x)\|^2}{g(x)} dx \\ &\quad + \frac{2}{n^{3/2}} \int_{\mathcal{I}} \frac{(\int K(u) \|\dot{\nu}_{\theta_0}(x - hu)\| g(x - hu) du)^2}{g(x)} dx = o(1). \end{aligned}$$

To handle b_{n2} , first note that conditional on X_j , the inner term of b_{n2} is $(n-1)/n$ times the average of centered i.i.d. r.v.'s. Using the fact that the variance is bounded above by the second moment, we obtain that the expected value of b_{n2} is bounded

above by

$$\begin{aligned} & n^{-1/2} E \left[\frac{K_h(X_2 - X_1) \|\dot{\nu}_{\theta_0}(X_1)\|}{g(X_2)} \mathcal{I}_2 \right]^2 \\ &= \frac{1}{n^{1/2} h^d} \int \int \frac{K^2(u) \|\dot{\nu}_{\theta_0}(x - uh)\| g(x - uh)}{g(x)} du dx = o(1). \end{aligned}$$

This completes the proof of (2.41) for $k = 1$. This together with (2.1) implies (2.41)

for $k = 2$.

To prove (2.41) for $k = 3$, similarly by the Cauchy-Schwarz inequality, $\forall \lambda \in \mathbb{R}^q$

$$\begin{aligned} & \|n^{1/2} g_{n3}\|^2 \\ & \leq n \left(n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) \mathcal{I}_j}{g(X_j)} \right]^2 \right) \left(n^{-1} \sum_{j=1}^n \left[\frac{\|\dot{\nu}_n(X_j, \hat{\theta}_n) - \dot{\nu}_n(X_j, \theta_0)\|}{g(X_j)} \mathcal{I}_j \right]^2 \right). \end{aligned}$$

But the second summation is bounded above by

$$\begin{aligned} & \max_{1 \leq i \leq n} \|\dot{m}_{\hat{\theta}_n}(X_i) - \dot{m}_{\theta_0}(X_i)\|^2 n^{-3} \sum_{j=1}^n \left[\frac{\sum_{i=1}^n K_h(X_j - X_i) \mathcal{I}_j}{g(X_j)} \right]^2 \\ &= o_p(h^d) \times O_p(1), \end{aligned}$$

by (2.32) and the assumption (m5), and by (2.24) applied with $\alpha \equiv 1$. This together with (2.12) proves (2.41) for $k = 3$. The proof of (2.41) for $k = 4$ uses (2.41) for $k = 3$ and (2.3), thereby, completing the proof of Lemma 2.4.2.

Next, we shall show that the right hand side of (2.39) equals $Q_n(\hat{\theta}_n - \theta_0)$, where $Q_n = \Sigma_0 + o_p(1)$. Recall the notation at (2.9) and (2.35). Let

$$\begin{aligned} V_n &:= n^{-1} \sum_{j=1}^n \left[\dot{\nu}_n(X_j, \hat{\theta}_n) \cdot \frac{\hat{g}_h(X_j)}{g_w^{*2}(X_j)} \mathcal{I}_j \cdot \frac{d_{ni}}{\|u_n\|} \right], \\ L_n &:= n^{-1} \sum_{j=1}^n \left[\frac{\dot{\nu}_n(X_j, \theta_0) \dot{\nu}_n'(X_j, \hat{\theta}_n)}{g_w^{*2}(X_j)} \mathcal{I}_j \right], \quad \forall \lambda \in \mathbb{R}^q. \end{aligned}$$

So, the right hand side of (2.39) can be written as the sum $[V_n u'_n + L_n]u_n$. But,

$$\begin{aligned}\|V_n\| &\leq \max_{1 \leq i \leq n} \frac{|d_{ni}|}{\|u_n\|} \|V_{n1}\|, \\ \|V_{n1}\| &:= n^{-1} \sum_{j=1}^n \left[\frac{\hat{g}_h(X_j)}{g_w^{*2}(X_j)} \|\dot{\nu}_n(X_j, \hat{\theta}_n)\| \mathcal{I}_j \right] \leq \|V_{n11}\| + \|V_{n12}\|,\end{aligned}$$

where,

$$\begin{aligned}\|V_{n11}\| &:= \max_{1 \leq i \leq n} \|\dot{m}_{\hat{\theta}_n}(X_i) - \dot{m}_{\theta_0}(X_i)\| \\ &\quad \times \left[n^{-3} \sum_{j=1}^n \left(\frac{\sum_{i=1}^n K_h(X_j - X_i)}{g_w^*(X_j)} \mathcal{I}_j \right)^2 \right], \\ \|V_{n12}\| &:= n^{-1} \sum_{j=1}^n \left(\frac{n^{-1} \sum_{i=1}^n K_h(X_j - X_i) \|\dot{m}_{\theta_0}(X_i)\|}{g_w^*(X_j)} \mathcal{I}_j \right) \times \left(\frac{\hat{g}_h(X_j)}{g_w^*(X_j)} \mathcal{I}_j \right).\end{aligned}$$

By (2.3), (2.32), the assumption (m5), and by using (2.24) with $\alpha \equiv 1$, $\|V_{n11}\| = o_p(1)$. Also, by the Cauchy-Schwarz inequality, (2.3), and by using (2.24) with $\alpha \equiv \dot{m}_{\theta_0}$ and 1 on the first and second term, respectively, one obtains

$$\begin{aligned}\|V_{n12}\| &\leq \left[n^{-1} \sum_{j=1}^n \left(\frac{n^{-1} \sum_{i=1}^n K_h(X_j - X_i) \|\dot{m}_{\theta_0}(X_i)\|}{g_w^*(X_j)} \mathcal{I}_j \right)^2 \right]^{1/2} \\ &\quad \times \left[n^{-1} \sum_{j=1}^n \left(\frac{n^{-1} \sum_{i=1}^n K_h(X_j - X_i)}{g_w^*(X_j)} \mathcal{I}_j \right)^2 \right]^{1/2} \\ &= O_p(1) \times O_p(1) = O_p(1).\end{aligned}$$

Hence, $\|V_{n1}\| = O_p(1)$. This together with (m4) then implies that $\|V_n\| = o_p(1)$, and

by consistency of $\hat{\theta}_n$, we have $\|u'_n V_n u_n\| = o_p(1)$.

Next, we will show that $\|L_n u_n\| = o_p(1)$. For this write $L_n := L_{n1} + L_{n2}$, where

$$\begin{aligned}L_{n1} &:= n^{-1} \sum_{j=1}^n \left[\frac{\dot{\nu}_n(X_j, \theta_0) [\dot{\nu}_n(X_j, \hat{\theta}_n) - \dot{\nu}_n(X_j, \theta_0)]'}{g_w^{*2}(X_j)} \mathcal{I}_j \right], \\ L_{n2} &:= n^{-1} \sum_{j=1}^n \left[\frac{\dot{\nu}_n(X_j, \theta_0) \dot{\nu}_n(X_j, \theta_0)'}{g_w^{*2}(X_j)} \mathcal{I}_j \right]\end{aligned}$$

But, by (2.3), (m5), by the Cauchy-Schwarz inequality and (2.24), $\forall \lambda \in \mathbb{R}^q$, $\|L_{n1}\| = o_p(1)$, while

$$\begin{aligned} & \left\| L_{n2} - n^{-1} \sum_{j=1}^n \frac{\dot{\nu}_h(X_j, \theta_0) \dot{\nu}_h(X_j, \theta_0)' \mathcal{I}_j}{g_w^{*2}(X_j)} \right\| \\ & \leq n^{-1} \sum_{j=1}^n \left[\frac{\|(\dot{\nu}_n(X_j, \theta_0) - \dot{\nu}_h(X_j, \theta_0))\|^2}{g_w^{*2}(X_j)} \mathcal{I}_j \right] \\ & \quad + 2n^{-1} \sum_{j=1}^n \left[\frac{\|(\dot{\nu}_n(X_j, \theta_0) - \dot{\nu}_h(X_j, \theta_0))\| \|\dot{\nu}_h(X_j, \theta_0)\|}{g_w^{*2}(X_j)} \mathcal{I}_j \right] \end{aligned}$$

But, by using same argument as in the second factor of the right hand side of (2.46), and by (2.3), this upper bound is $o_p(1)$. Moreover, similar to the argument used for Σ_n in (2.37) and using (2.3), one obtains $\forall \lambda \in \mathbb{R}^q$,

$$n^{-1} \sum_{j=1}^n \frac{\|\dot{\nu}_h(X_j, \theta_0) \dot{\nu}_h(X_j, \theta_0)'\|}{g_w^{*2}(X_j)} \mathcal{I}_j = \Sigma_0 + o_p(1).$$

This proves $Q_n = \Sigma_0 + o_p(1)$, thereby also completing the proof of Theorem 2.4.1.

2.5 Asymptotic normality of $M_n(\hat{\theta}_n)$

This section contains a proof of the asymptotic normality of the minimized distance $M_n(\hat{\theta}_n)$. The replacement of g_w^* by g in M_n and $\mathcal{T}_n$ is reflected by notation $\tilde{M}_n$ and $\tilde{\mathcal{T}}_n$. The main result proved in this section is the following

Theorem 2.5.1 *Suppose (e1), (e2), (e4), (g1), (g2), (k), (m1)-(m5) and (h3) with Φ replaced by G hold. Then under $\mathcal{H}_0$, $nh^{d/2}(M_n(\hat{\theta}_n) - \hat{C}_n) \rightarrow_d \mathcal{N}_1(0, \Gamma)$. Moreover, $|\hat{\Gamma}_n \Gamma^{-1} - 1| = o_p(1)$.*

Consequently, the test that rejects H_0 whenever $nh^{d/2}\hat{\Gamma}_n^{-1/2}|M_n(\hat{\theta}_n) - \hat{C}_n| > z_{\alpha/2}$ is of the asymptotic size α , where z_{α} is the $100(1 - \alpha)\%$ percentile of the standard normal distribution.

Analogous to the proof in K-N, the proof of this theorem is facilitated by the following five lemmas.

Lemma 2.5.1 *Suppose (e1), (e2), (e4), (g1), (k), (h1), and (h2) with Φ replaced by G hold. Then, under $\mathcal{H}_0$,*

$$nh^{d/2}(\tilde{M}_n(\theta_0) - \tilde{C}_n) \rightarrow_d N_1(0, \Gamma).$$

Lemma 2.5.2 *Suppose (e1), (e2), (g1), (k), (m3)-(m5), (h1), and (h2) with Φ replaced by G hold. Then under $\mathcal{H}_0$,*

$$nh^{d/2}|M_n(\hat{\theta}_n) - M_n(\theta_0)| = o_p(1).$$

Lemma 2.5.3 *Suppose (e1), (e2), (g1), (g2), (k), (m3)-(m5), and (h3) hold with Φ replaced by G . Then under $\mathcal{H}_0$,*

$$nh^{d/2}|M_n(\theta_0) - \tilde{M}_n(\theta_0)| = o_p(1).$$

Lemma 2.5.4 *Under the same conditions as in Lemma 2.5.3,*

$$nh^{d/2}|\hat{C}_n - \tilde{C}_n| = o_p(1).$$

Lemma 2.5.5 *Under the same conditions as in Lemma 2.5.2, $\hat{\Gamma}_n - \Gamma = o_p(1)$,
Consequently, the positive definiteness of Γ implies $|\hat{\Gamma}_n\Gamma^{-1} - 1| = o_p(1)$.*

The proof of Lemma 2.5.1 is facilitated by Theorem 1 of Hall (1984) which is reproduced here for the sake of completeness.

Theorem 2.5.2 *Let $\tilde{X}_i, 1 \leq i \leq n$, be i.i.d. random vectors, and let*

$$U_n := \sum_{1 \leq i < j \leq n} H_n(\tilde{X}_i, \tilde{X}_j), \quad G_n(x, y) := EH_n(\tilde{X}_1, x)H_n(\tilde{X}_1, y),$$

where H_n is a sequence of measurable functions symmetric under permutation with

$$E[H_n(\tilde{X}_1, \tilde{X}_2)|\tilde{X}_1] = 0, \quad EH_n^2(\tilde{X}_1, \tilde{X}_2) < \infty \quad \forall n \geq 1.$$

If, additionally,

$$\frac{EG_n^2(\tilde{X}_1, \tilde{X}_2) + n^{-1} EH_n^4(\tilde{X}_1, \tilde{X}_2)}{[EH_n^2(\tilde{X}_1, \tilde{X}_2)]^2} \rightarrow 0, \quad \text{as } n \rightarrow \infty,$$

then U_n is asymptotically normally distributed with the mean 0 and the variance $n^2 EH_n^2(\tilde{X}_1, \tilde{X}_2)/2$.

Proof of Lemma 2.5.1. Let

$$\begin{aligned} \tilde{T}_n(\theta) &:= \int_{\mathcal{I}} \left[\frac{1}{n} \sum_{i=1}^n K_h(x - X_i)(Y_i - m_\theta(X_i)) \right]^2 \frac{d\Phi(x)}{g^2(x)}, \\ \tilde{R}_n &:= \frac{1}{n^2} \sum_{i=1}^n \int_{\mathcal{I}} \frac{K_h^2(x - X_i) \hat{\epsilon}_i^2}{g^2(x)} d\Phi(x). \end{aligned}$$

To prove Lemma 2.5.1, by Slutsky's Theorem, it suffices to show that

$$nh^{d/2}(\tilde{T}_n(\theta_0) - \tilde{R}_n) \rightarrow_d N_1(0, \Gamma), \quad (2.47)$$

$$nh^{d/2}|\tilde{M}_n(\theta_0) - \tilde{T}_n(\theta_0)| = o_p(1),$$

$$nh^{d/2}(\tilde{R}_n - \tilde{C}_n) = o_p(1).$$

The first claim in (2.47) is proved in Lemma 5.1 of K-N.

For the second claim, it suffices to show that $E[nh^{d/2}|\tilde{M}_n(\theta_0) - \tilde{T}_n(\theta_0)|]^2 = o(1)$.

Let

$$f_{ijk} := \left[\frac{K_h(X_j - X_i)K_h(X_j - X_k)}{g^2(X_j)} \mathcal{I}_j - \int_{\mathcal{I}} \frac{K_h(x - X_i)K_h(x - X_k)}{g^2(x)} dG(x) \right],$$

and $e_{ijk} = \frac{h^{d/2} \epsilon_i \epsilon_k}{n^2} f_{ijk}$. Hence,

$$[nh^{d/2}(\tilde{M}_n(\theta_0) - \tilde{T}_n(\theta_0))] = \sum_i \sum_j \sum_k e_{ijk}.$$

Expanding the quadratic and using the fact $E[f_{ijk}|X_i, X_k] = 0, \forall j \neq i, k$, we obtain

$$\begin{aligned} & E[nh^{d/2}(\tilde{M}_n(\theta_0) - \tilde{T}_n(\theta_0))]^2 \\ & \leq C E \left[\sum_i e_{iii}^2 + \sum_{j \neq i} [e_{iii}e_{jjj} + e_{iji}e_{jij} + e_{iji}^2 + e_{ijj}e_{iij} + e_{ijj}e_{jii} + e_{ijj}^2 \right. \\ & \quad \left. + e_{ijj}e_{jji} + e_{iij}e_{jji} + e_{iij}^2 + e_{iij}e_{jii} + e_{iji}e_{jjj} + e_{jij}e_{iii}] \right]. \end{aligned} \quad (2.48)$$

To prove the required claim, it suffices to show that all terms on the right hand side of inequality (2.48) are $o(1)$.

By (e4), (g1), (k), and (h2), one obtains

$$\begin{aligned} & E \left[\sum_i e_{iii}^2 \right] \\ & = \frac{h^d}{n^3} E \left[\tau_{\theta_0}(X_1) \left[\frac{K_h^2(0)I_1}{g^2(X_1)} - \int_{\mathcal{I}} \frac{K_h^2(x - X_1)}{g(x)} dx \right]^2 \right] \\ & \leq \frac{2K^4(0)}{(nh^d)^3} \int_{\mathcal{I}} \frac{\tau_{\theta_0}(x)}{g^3(x)} dx + \frac{2}{n^3 h^{2d}} \int_{\mathcal{I}} \int \frac{K^4(u)g(x - hu)\tau_{\theta_0}(x - hu)}{g^2(x)} du dx \\ & = O(nh^d)^{-1} = o(1). \end{aligned}$$

Similarly, using the independence of X_i 's, by (e4), (g1), (k), and (h2), one shows

$$E[\sum_{j \neq i} [e_{iii}e_{jjj}]] = o(1).$$

Next,

$$E \left[\sum_{j \neq i} e_{iji}e_{jij} \right] = \frac{h^d}{n^2} E[\sigma_{\theta_0}^2(X_1)\sigma_{\theta_0}^2(X_2)f_{121}f_{212}].$$

By the independence of X_i 's, (g1), and (k) and continuity of $\sigma_{\theta_0}^2$,

$$\begin{aligned}
& E[\sigma_{\theta_0}^2(X_1)\sigma_{\theta_0}^2(X_2)f_{121}f_{212}] \\
&= \frac{1}{h^{3d}} \int_{\mathcal{I}} \int \frac{K^4(u)\sigma_{\theta_0}^2(x)\sigma_{\theta_0}^2(x-hu)}{g(x)g(x-hu)} du dx \\
&\quad - 2 \frac{1}{h^{2d}} \int \int \int_{\mathcal{I}} K^2(u)K^2(v)\sigma_{\theta_0}^2(x-hu)\sigma_{\theta_0}^2(x-hu-hv) \\
&\quad \quad \quad \times \frac{g(x-hu-hv)}{g(x)g(x-hu)} dx du dv \\
&\quad + \frac{1}{h^{2d}} \left[\int \int_{\mathcal{I}} \frac{K^2(u)g(x-hu)\sigma_{\theta_0}^2(x-hu)}{g(x)} dx du \right]^2 = O(h^d)^{-3}.
\end{aligned}$$

Hence, $E[\sum_{j \neq i} e_{iji}e_{jjj}] = O(nh^{2d})^{-1} = o(1)$. Similarly, one can show that other terms of (2.48) are $o(1)$, thereby completing the proof of the second claim in (2.47).

To prove the third claim in (2.47), it suffices to show that

$$E[nh^{d/2}(\tilde{R}_n - \tilde{C}_n)]^2 = o(1).$$

Let

$$d_{ij} := \left[\frac{K_h^2(X_j - X_i)}{g^2(X_j)} \mathcal{I}_j - \int_{\mathcal{I}} \frac{K_h^2(x - X_i)}{g^2(x)} dG(x) \right],$$

and $b_{ij} = \frac{h^{d/2}\epsilon_i^2}{n^2} d_{ij}$. Then, $[nh^{d/2}(\tilde{R}_n - \tilde{C}_n)] = \sum_i \sum_j b_{ij}$. Expanding the quadratic and using the fact that $E[d_{ij}|X_i] = 0, \forall j \neq i$, we obtain

$$\begin{aligned}
& E[nh^{d/2}(\tilde{R}_n - \tilde{C}_n)]^2 \\
& \leq C E \left[\sum_i b_{ii}^2 + \sum_{m \neq i} [b_{ii}b_{mm} + b_{im}(b_{mm} + b_{im} + b_{mi})] \right].
\end{aligned}$$

To prove the required claim, it suffices to show that all the terms on the right hand side of above inequality are $o(1)$.

Thus by (e4), (g1), (k), and (h2) and continuity of τ_{θ_0} , one obtains

$$\begin{aligned} E\left[\sum_i b_{ii}^2\right] &\leq \frac{2K^4(0)}{(nh^d)^3} \int_{\mathcal{I}} \frac{\tau_{\theta_0}(x)}{g^3(x)} dx \\ &\quad + \frac{2}{n^3 h^{2d}} \int \left[\int_{\mathcal{I}} \frac{K^2(u) g^{1/2}(x-hu) \tau_{\theta_0}^{1/2}((x-hu))}{g(x)} dx \right]^2 du \\ &= O(nh^d)^{-1} = o(1). \end{aligned}$$

For the fourth term, since

$$E\left[\sum_{m \neq i} b_{im} b_{mi}\right] = \frac{h^d}{n^2} E[\sigma_{\theta_0}^2(X_1) \sigma_{\theta_0}^2(X_2) d_{12} d_{21}],$$

thus, by independence of X_i 's, (g1), and (k) and continuity of $\sigma_{\theta_0}^2$, one obtains

$$\begin{aligned} &E[\sigma_{\theta_0}^2(X_1) \sigma_{\theta_0}^2(X_2) d_{12} d_{21}] \\ &= \frac{1}{h^{3d}} \int_{\mathcal{I}} \int \frac{K^4(u) \sigma_{\theta_0}^2(x) \sigma_{\theta_0}^2(x-hu)}{g(x) g(x-hu)} du dx \\ &\quad - 2 \frac{1}{h^{2d}} \int \int \int_{\mathcal{I}} K^2(u) K^2(v) \sigma_{\theta_0}^2(x-uh) \sigma_{\theta_0}^2(x-vh-uh) \\ &\quad \quad \quad \times \frac{g(x-vh-hu)}{g(x) g(x-hu)} dx du dv \\ &\quad + \frac{1}{h^{2d}} \left[\int \int_{\mathcal{I}} \frac{K^2(u) g(x-hu) \sigma_{\theta_0}^2(x-uh)}{g(x)} dx du \right]^2 = O(h^d)^{-3}. \end{aligned}$$

Hence, $E[\sum_{m \neq i} b_{im} b_{mi}] = O(nh^{2d})^{-1} = o(1)$. By similar arguments, one can show that other terms are also $o(1)$. Hence we are done with the third part of (2.47) and also with the proof of Lemma 2.5.1. $\square$

Proof of Lemma 2.5.2. Recall the definitions of U_n and Z_n from (2.9). Add and subtract $m_{\theta_0}(X_i)$ to the i^{th} summand inside the squared integrand of $M_n(\hat{\theta}_n)$,

to obtain that

$$\begin{aligned}
& M_n(\theta_0) - M_n(\hat{\theta}_n) \\
&= 2n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) Z_n(X_j, \hat{\theta}_n) \mathcal{I}_j}{g_w^{*2}(X_j)} \right] \\
&\quad n^{-1} \sum_{j=1}^n \left[\frac{Z_n^2(X_j, \hat{\theta}_n) \mathcal{I}_j}{g_w^{*2}(X_j)} \right] \\
&= 2Q_1 - Q_2, \quad \text{say.}
\end{aligned}$$

It thus suffices to show that

$$nh^{d/2}Q_1 = o_p(1), \quad nh^{d/2}Q_2 = o_p(1) \quad (2.49)$$

Add and subtract $(\hat{\theta}_n - \theta_0)' \dot{m}_{\theta_0}(X_i)$ to the i^{th} summand of $Z_n(X_j, \hat{\theta}_n)$, we can rewrite

$$\begin{aligned}
Q_1 &= n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) \hat{g}_h(X_j) d_{ni}}{g_w^{*2}(X_j)} \mathcal{I}_j \right] \\
&\quad + (\hat{\theta}_n - \theta_0)' n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) \dot{m}_n(X_j, \theta_0)}{g_w^{*2}(X_j)} \mathcal{I}_j \right] \\
&= Q_{11} + Q_{12}, \quad \text{say,}
\end{aligned}$$

where d_{ni} are as in (2.35). By using (2.3) and (2.24) with $\alpha \equiv 1$, one obtains

$$n^{-1} \sum_{j=1}^n \left[\frac{\hat{g}_h(X_j)}{g_w^{*2}(X_j)} \mathcal{I}_j \right]^2 = O_p(1) \quad (2.50)$$

By the Cauchy-Schwarz inequality, (2.3), (2.12) and (2.50), one obtain that $(nh^{d/2}|Q_{11}|)$ is bounded above by

$$n^{1/2} \|\hat{\theta}_n - \theta_0\| (nh^d)^{1/2} \cdot O_p((nh^d)^{-1/2}) \cdot \max_i \frac{|d_{ni}|}{\|\hat{\theta}_n - \theta_0\|}.$$

But, by (m4) and (2.32), this entire bound in turn is $o_p(1)$. Hence, to prove the first part of (2.49), it remains to prove that $nh^{d/2}|Q_{12}| = o_p(1)$. But Q_{12} can be rewritten as

$$\begin{aligned} & (\hat{\theta}_n - \theta_0)' n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) \dot{\mu}_n(X_j, \hat{\theta}_n)}{g_w^{*2}(X_j)} \mathcal{I}_j \right] \\ & - (\hat{\theta}_n - \theta_0)' n^{-1} \sum_{j=1}^n \left[\frac{U_n(X_j) [\dot{\mu}_n(X_j, \hat{\theta}_n) - \dot{\mu}_n(X_j, \theta_0)]}{g_w^{*2}(X_j)} \mathcal{I}_j \right] \\ & = Q_{121} - Q_{122}, \quad \text{say.} \end{aligned}$$

But, by the Cauchy-Schwarz inequality, (2.3), (2.12), (2.50), one obtains $(nh^{d/2} \|Q_{122}\|)^2$ is bounded above by

$$n^2 h^d \|\hat{\theta}_n - \theta_0\|^2 O_p(nh^d)^{-1} \max_i \|\dot{m}_{\hat{\theta}_n}(X_i) - \dot{m}_{\theta_0}(X_i)\|.$$

By assumptions (m5) and (h2), and consistency of $\hat{\theta}_n$ for θ_0 , this entire bound is $o_p(1)$.

Next, note that the average in Q_{121} is the same as the expression in the left hand side of (2.39). Thus it is equal to

$$\begin{aligned} & (\hat{\theta}_n - \theta_0)' n^{-1} \sum_{j=1}^n \left[\frac{Z_n(X_j, \hat{\theta}_n) \dot{\mu}_n(X_j, \hat{\theta}_n)}{g_w^{*2}(X_j)} \mathcal{I}_j \right] \tag{2.51} \\ & = (\hat{\theta}_n - \theta_0)' n^{-1} \sum_{j=1}^n \left[\frac{Z_n(X_j, \hat{\theta}_n) \dot{\mu}_n(X_j, \theta_0)}{g_w^{*2}(X_j)} \mathcal{I}_j \right] \\ & + (\hat{\theta}_n - \theta_0)' n^{-1} \sum_{j=1}^n \left[\frac{Z_n(X_j, \hat{\theta}_n) [\dot{\mu}_n(X_j, \hat{\theta}_n) - \dot{\mu}_n(X_j, \theta_0)]}{g_w^{*2}(X_j)} \mathcal{I}_j \right] \\ & = D_1 + D_2, \quad \text{say,} \end{aligned}$$

But, by the Cauchy-Schwarz inequality, (2.3), (2.29) and (2.24) with $\alpha \equiv \dot{m}_{\theta_0}$, $\forall \lambda \in \mathbb{R}^q$,

$$\begin{aligned} & E(nh^{d/2}\|D_1\|) \\ & \leq nh^{d/2}\|(\hat{\theta}_n - \theta_0)\|^2 n^{-1} \sum_{j=1}^n \left[\frac{Z_n(X_j, \hat{\theta}_n)}{g_w^*(X_j)} \mathcal{I}_j \right]^2 n^{-1} \sum_{j=1}^n \left[\frac{\|\dot{\nu}_n(X_j, \hat{\theta}_n)\|}{g_w^*(X_j)} \mathcal{I}_j \right]^2 \\ & = nh^{d/2}\|(\hat{\theta}_n - \theta_0)\|^2 O(1) = o(1), \end{aligned}$$

by Theorem 2.4.1 and the assumption (m5) and (h2). Hence, $nh^{d/2}\|D_1\| = o_p(1)$. Similarly, one obtains $(nh^{d/2}\|D_2\|)$ is bounded above by $nh^{d/2}\|(\hat{\theta}_n - \theta_0)\|^2 o_p(1) = o_p(1)$.

This completes the proof of the first part of (2.49). The proof of the second part of (2.49) is similar. $\square$

Proof of Lemma 2.5.3. By (2.11) and Lemma 2.4.3.

$$\begin{aligned} & nh^{d/2}|M_n(\theta_0) - \tilde{M}_n(\theta_0)| \\ & \leq nh^{d/2} \left[n^{-1} \sum_{j=1}^n \frac{U_n^2(X_j)}{g^2(X_j)} \mathcal{I}_j \right] \sup_{x \in \mathcal{I}} \left| \frac{g^2(x)}{\hat{g}_w^2(x)} - 1 \right| \\ & = nh^{d/2} O_p((nh^d)^{-1}) \cdot O_p((\log_k n)(\log n/n)^{1/d+4}) = o_p(1). \end{aligned} \quad \square$$

Proof of Lemma 2.5.4. Let

$$t_i = m_{\hat{\theta}_n}(X_i) - m_{\theta_0}(X_i), \quad \Delta_w(x) := g^2(x) \left(g_w^{*-2}(x) - g^{-2}(x) \right).$$

Rewrite $\hat{C}_n$ as

$$\hat{C}_n = n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i)(\epsilon_i - t_i)^2}{g_w^{*2}(X_j)} \mathcal{I}_j \right]$$

$$\begin{aligned}
&= n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i)(\epsilon_i - t_i)^2}{g^2(X_j)} \mathcal{I}_j \right] \\
&\quad + n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i)(\epsilon_i - t_i)^2 \Delta_w(X_j)}{g_w^{*2}(X_j)} \mathcal{I}_j \right] \\
&= A_{n1} + A_{n2}, \quad \text{say.}
\end{aligned}$$

Now, it suffices to prove that

$$nh^{d/2}(A_{n1} - \tilde{C}_n) = o_p(1), \quad nh^{d/2}A_{n2} = o_p(1). \quad (2.52)$$

By expanding the quadratic term in the summand, A_{n1} can be written as the sum of $\tilde{C}_n$, A_{n12} , and A_{n13} , where

$$\begin{aligned}
A_{n12} &= n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i)t_i^2}{g^2(X_j)} \mathcal{I}_j \right], \\
A_{n13} &= -2n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i)\epsilon_i t_i}{g^2(X_j)} \mathcal{I}_j \right]
\end{aligned}$$

By (m4) and (2.32), one obtains that $\max_{1 \leq i \leq n} |t_i|^2 = O_p((nh^d)^{-1})$. Moreover,

by (g1), (k), and (h2), one obtains that

$$\begin{aligned}
&E \left[n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i)}{g^2(X_j)} \mathcal{I}_j \right] \right] \\
&= \frac{K^2(0)}{n^2 h^{2d}} \sum_{i=1}^n E \left[\frac{1}{g^2(X_i)} \mathcal{I}_i \right] + \frac{1}{n^3} \sum_{i \neq j} E \left[\frac{K_h^2(X_j - X_i)}{g^2(X_j)} \mathcal{I}_j \right] \\
&= \frac{K^2(0)}{n^2 h^{2d}} \int_{\mathcal{I}} \frac{1}{g(x)} dx + \frac{1}{n h^d} \int_{\mathcal{I}} \int \frac{K^2(u)g(x - uh)}{g(x)} du dx \\
&= O((nh^d)^{-1}).
\end{aligned}$$

Hence,

$$\begin{aligned}
|A_{n12}| &\leq \max_{1 \leq i \leq n} |t_i|^2 n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i)}{g^2(X_j)} \mathcal{I}_j \right] \\
&= O_p((nh^d)^{-1}) \times O_p((nh^d)^{-1}) = O_p((nh^d)^{-2}).
\end{aligned} \quad (2.53)$$

Next, by (g1), (k), and (h2),

$$\begin{aligned}
& E \left[n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i) |\epsilon_i|}{g^2(X_j)} \mathcal{I}_j \right] \right] \\
&= \frac{K^2(0)}{n^3 h^{2d}} \sum_{i=1}^n E \left[\frac{|\epsilon_i|}{g^2(X_i)} \mathcal{I}_i \right] + \frac{1}{n^3} \sum_{i \neq j} E \left[\frac{K_h^2(X_j - X_i) |\epsilon_i|}{g^2(X_j)} \mathcal{I}_j \right] \\
&\leq \frac{K^2(0)}{(nh^d)^2} \int_{\mathcal{I}} \frac{\sigma_{\theta_0}(x)}{g(x)} dx \\
&\quad + \frac{1}{(nh^d)} \int \int_{\mathcal{I}} \frac{K^2(u) \sigma_{\theta_0}(x - uh) g(x - uh)}{g(x)} dx du \\
&= O((nh^d)^{-1}).
\end{aligned}$$

Hence,

$$\begin{aligned}
|A_{n13}| &\leq 2 \max_{1 \leq i \leq n} |t_i| n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i) |\epsilon_i|}{g^2(X_j)} \mathcal{I}_j \right] \\
&= O_p((nh^d)^{-1/2}) \times O_p((nh^d)^{-1}) = O_p((nh^d)^{-3/2}).
\end{aligned}$$

Consequently,

$$\begin{aligned}
|nh^{d/2}(A_{n1} - \tilde{C}_n)| &= nh^{d/2}(O_p((nh^d)^{-2}) + O_p((nh^d)^{-3/2})) \\
&= O_p((nh^{-3d/2})^{-1}) + O_p((nh^{2d})^{-1/2}) = o_p(1).
\end{aligned}$$

To prove the second claim in (2.52), rewrite A_{n2} as the sum of A_{n21} , A_{n22} and

A_{n23} , where

$$\begin{aligned}
A_{n21} &= n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i) \epsilon_i^2 \Delta w(X_j)}{g^2(X_j)} \mathcal{I}_j \right], \\
A_{n22} &= n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i) t_i^2 \Delta w(X_j)}{g^2(X_j)} \mathcal{I}_j \right], \\
A_{n23} &= n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i) \epsilon_i t_i \Delta w(X_j)}{g^2(X_j)} \mathcal{I}_j \right].
\end{aligned}$$

By (e2), (g1), (k), and (h2), one obtains that

$$\begin{aligned}
& E \left[n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i) \epsilon_i^2}{g^2(X_j)} \mathcal{I}_j \right] \right] \\
&= \frac{K^2(0)}{n^3 h^{2d}} \sum_{i=1}^n E \left[\frac{\epsilon_i^2}{g^2(X_i)} \mathcal{I}_i \right] + \frac{1}{n^3} \sum_{i \neq j} E \left[\frac{K_h^2(X_j - X_i) \epsilon_i^2}{g^2(X_j)} \mathcal{I}_j \right] \\
&= \frac{K^2(0)}{n^2 h^{2d}} \int_{\mathcal{I}} \frac{\sigma_{\theta_0}^2(x)}{g(x)} dx + \frac{1}{(nh^d)} \int \int_{\mathcal{I}} \frac{K^2(u) \sigma_{\theta_0}^2(x - uh) g(x - uh)}{g(x)} dx du \\
&= O((nh^d)^{-1}).
\end{aligned}$$

Hence,

$$\begin{aligned}
|nh^{d/2} A_{n21}| &\leq nh^{d/2} \sup_{x \in \mathcal{I}} |\Delta_w(x)| n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i) \epsilon_i^2}{g^2(X_j)} \mathcal{I}_j \right] \\
&= nh^{d/2} O_p(\log_k n (\log n/n)^{2/d+4}) O_p((nh^d)^{-1}) \\
&= O_p(h^{-d/2} \log_k n (\log n/n)^{2/d+4}) = o_p(1),
\end{aligned}$$

by Lemma 2.4.3 and (2.3). Similarly, one obtains that

$$\begin{aligned}
|nh^{d/2} A_{n22}| &\leq nh^{d/2} \sup_{x \in \mathcal{I}} |\Delta_w(x)| \max_{1 \leq i \leq n} |t_i|^2 n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i)}{g^2(X_j)} \mathcal{I}_j \right] \\
&= nh^{d/2} O_p(\log_k n (\log n/n)^{2/d+4}) O_p((nh^d)^{-1}) O_p((nh^d)^{-1}) \\
&= o_p((nh^{3d/2})^{-1}) = o_p(1),
\end{aligned}$$

and

$$\begin{aligned}
|nh^{d/2} A_{n23}| &\leq 2nh^{d/2} \sup_{x \in \mathcal{I}} |\Delta_w(x)| \max_{1 \leq i \leq n} |t_i| n^{-3} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{K_h^2(X_j - X_i) |\epsilon_i|}{g^2(X_j)} \mathcal{I}_j \right]
\end{aligned}$$

$$\begin{aligned}
&= nh^{d/2} O_p(\log_k n (\log n/n)^{2/d+4}) O_p((nh^d)^{-1/2}) O_p((nh^d)^{-1}) \\
&= o_p((nh^{2d})^{-1/2}) = o_p(1),
\end{aligned}$$

thereby completing the proof of the claim in (2.52), and hence that of the Lemma 2.5.4. $\square$

Proof of Lemma 2.5.5. Recall the notation from (2.1), (2.2). Let

$$\mathcal{G}_n := 2h^d \int \int \left[\frac{EK_h(x-X)K_h(y-X)\sigma_{\theta_0}^2(X)}{g(x)g(y)} \right]^2 dG(x) dG(y)$$

To prove the first part of Lemma 2.5.5, we need to prove the following steps:

$$|\hat{\Gamma}_n - \tilde{\Gamma}_n| = o_p(1), \quad |\tilde{\Gamma}_n - \tilde{\mathcal{G}}_n| = o_p(1) \quad (2.54)$$

$$|\tilde{\mathcal{G}}_n - \mathcal{G}_n| = o_p(1), \quad \mathcal{G}_n \rightarrow \Gamma.$$

Now, we shall prove the first part of (2.54). For the sake of convenience, write $K_h(X_j - X_i)$ by $K_i(X_j)$ and $\Delta_h(x) := g^2(x)(\hat{g}_h^{-2}(x) - g^{-2}(x))$. Now, rewrite $\hat{\Gamma}_n$ as the sum of the following three terms:

$$\begin{aligned}
B_1 &:= 2h^d n^{-2} \sum_{i \neq j} \left[n^{-1} \sum_l \frac{K_i(X_l)K_j(X_l)(\epsilon_i - t_i)(\epsilon_j - t_j)}{g^2(X_l)} \mathcal{I}_l \right]^2, \\
B_2 &:= 2h^d n^{-2} \sum_{i \neq j} \left[n^{-1} \sum_l \frac{K_i(X_l)K_j(X_l)(\epsilon_i - t_i)(\epsilon_j - t_j)\Delta_h(X_l)}{g^2(X_l)} \mathcal{I}_l \right]^2, \\
B_3 &:= 2h^d n^{-2} \sum_{i \neq j} \left[\left(n^{-1} \sum_l \frac{K_i(X_l)K_j(X_l)(\epsilon_i - t_i)(\epsilon_j - t_j)}{g^2(X_l)} \mathcal{I}_l \right) \right. \\
&\quad \left. \times \left(n^{-1} \sum_l \frac{K_i(X_l)K_j(X_l)(\epsilon_i - t_i)(\epsilon_j - t_j)\Delta_h(X_l)}{g^2(X_l)} \mathcal{I}_l \right) \right].
\end{aligned}$$

In order to prove the first part of (2.54), it suffices to prove that

$$B_1 - \tilde{\Gamma}_n = o_p(1), \quad B_2 = o_p(1), \quad \text{and} \quad B_3 = o_p(1). \quad (2.55)$$

For this, we shall show that

$$B := h^d n^{-2} \sum_{i \neq j} \left[n^{-1} \sum_l \frac{K_i(X_l) K_j(X_l) |\epsilon_i| |\epsilon_j|}{g^2(X_l)} \mathcal{I}_l \right]^2 = O_p(1). \quad (2.56)$$

This expression is bounded by the sum of the following two terms:

$$\begin{aligned} H_1 &:= \frac{2K^2(0)}{n^4 h^d} \sum_{i \neq j} \epsilon_i^2 \epsilon_j^2 \left[\frac{K_j(X_i)}{g^2(X_i)} \mathcal{I}_i + \frac{K_i(X_j)}{g^2(X_j)} \mathcal{I}_j \right]^2, \\ H_2 &:= 2h^d n^{-4} \sum_{i \neq j} \epsilon_i^2 \epsilon_j^2 \left[\sum_{l \neq i, j} \frac{K_i(X_l) K_j(X_l)}{g^2(X_l)} \mathcal{I}_l \right]^2. \end{aligned}$$

By using (g1), (k), and continuity of $\sigma_{\theta_0}^2$, we obtain

$$\begin{aligned} &E \left[\frac{K_1^2(X_2) \epsilon_1^2 \epsilon_2^2}{g^4(X_2)} \mathcal{I}_2 \right] \\ &= h^{-d} \int_{\mathcal{I}} \int \frac{K^2(u) \sigma_{\theta_0}^2(x - uh) \sigma_{\theta_0}^2(x) g(x - uh)}{g^3(x)} du dx \\ &= O(h^{-d}). \end{aligned}$$

Hence, $EH_1 = O(nh^d)^{-2} = o(1)$, by (h2), and $H_1 = o_p(1)$.

Next, rewrite H_2 as the sum of the following two terms:

$$\begin{aligned} H_{21} &= h^d n^{-4} \sum_{l \neq i \neq j} \frac{K_i^2(X_l) K_j^2(X_l) \epsilon_i^2 \epsilon_j^2}{g^4(X_l)} \mathcal{I}_l, \\ H_{22} &= h^d n^{-4} \sum_{m \neq i \neq j \neq l} \frac{K_i(X_l) K_j(X_l) K_i(X_m) K_j(X_m) \epsilon_i^2 \epsilon_j^2}{g^2(X_l) g^2(X_m)} \mathcal{I}_l \mathcal{I}_m. \end{aligned}$$

By (e2), (g1), (k), (h2), and independence of X_i 's,

$$\begin{aligned} EH_{21} &= (nh^d)^{-1} E \left[\frac{K_1^2(X_3) K_2^2(X_3) \sigma_{\theta_0}^2(X_1) \sigma_{\theta_0}^2(X_2)}{g^4(X_3)} \mathcal{I}_3 \right] \\ &= (nh^d)^{-1} \int_{\mathcal{I}} \left[\int K^2(u) \sigma_{\theta_0}^2(x - uh) g(x - uh) du \right]^2 g^{-3}(x) dx \\ &= o(1). \end{aligned}$$

Similarly, by (e2), (g1), and (k),

EH_{22}

$$\begin{aligned}
&= \int_{\mathcal{I}} \int \int \int \left[K(u)K(w)K(v+u)K(w+v)\sigma_{\theta_0}^2(x-uh-vh)\sigma_{\theta_0}^2(x-vh-wh) \right. \\
&\quad \left. \times \frac{g(x-vh-wh)g(x-vh-uh)}{g(x-vh)g(x)} dw du dv dx \right] \\
&= O(1).
\end{aligned}$$

Hence, $H_2 = O_p(1)$, and (2.56) is proved.

By a similar argument, under (2.56), (e2), (g1), (k), and (h2) one obtains

$$h^{d_n-2} \sum_{i \neq j} \left[n^{-1} \sum_l \frac{K_i(X_l)K_j(X_l)|\epsilon_i|}{g^2(X_l)} \mathcal{I}_l \right]^2 = O_p(1), \quad (2.57)$$

$$h^{d_n-2} \sum_{i \neq j} \left[n^{-1} \sum_l \frac{K_i(X_l)K_j(X_l)}{g^2(X_l)} \right]^2 = O_p(1). \quad (2.58)$$

Furthermore,

$$\sup_{x \in \mathcal{I}} |\Delta_h(x)| = o_p(1), \quad \text{by (2.3),} \quad (2.59)$$

$$\max_{1 \leq i \leq n} |t_i| = o_p(1), \quad \text{by (m4) and (2.32).} \quad (2.60)$$

Note that by expanding $(\epsilon_i - t_i)(\epsilon_j - t_j)$ and the quadratic terms, $|B_1 - \tilde{\Gamma}_n|$ is

bounded above by the sum of B_{12} and B_{13} , where

$$\begin{aligned}
B_{12} &:= 2h^{d_n-2} \sum_{i \neq j} \left[n^{-1} \sum_l \frac{K_i(X_l)K_j(X_l)(|t_i t_j| + |\epsilon_i t_i| + |t_i \epsilon_j|)}{g^2(X_l)} \mathcal{I}_l \right]^2, \\
B_{13} &:= 4h^{d_n-2} \sum_{i \neq j} \left(n^{-1} \sum_l \frac{K_i(X_l)K_j(X_l)|\epsilon_i \epsilon_j|}{g^2(X_l)} \mathcal{I}_l \right) \\
&\quad \times \left(n^{-1} \sum_l \frac{K_i(X_l)K_j(X_l)(|t_i t_j| + |\epsilon_i t_i| + |t_i \epsilon_j|)}{g^2(X_l)} \mathcal{I}_l \right).
\end{aligned}$$

But $B_{12} = o_p(1)$ by (2.57), (2.58), (2.60), and the fact that t_i 's are free of X_l . Similarly, by applying the Cauchy-Schwarz inequality to the double sum and by (2.56) $B_{13} = o_p(1)$. Hence $|B_1 - \tilde{\Gamma}_n| = o_p(1)$.

Next, consider B_2 . By using the inequality

$$|\epsilon_i - t_i||\epsilon_j - t_j| \leq |\epsilon_i \epsilon_j| + |t_i t_j| + |\epsilon_i t_i| + |t_i \epsilon_j|,$$

and by (2.59),

$$B_2 \leq 2 \sup_{x \in \mathcal{I}} |\Delta_h(x)| [B_{12} + B] = o_p(1).$$

Similarly, an application of the Cauchy-Schwarz inequality to the double sum yields $B_3 = o_p(1)$. This completes the proof of (2.55), and hence that of the first part of (2.54).

To prove the second part of (2.54), it suffices to show that $E[\tilde{\Gamma}_n - \tilde{\mathcal{G}}_n]^2 = o(1)$.

Let

$$s_{ijkl} := \left[\frac{K_i(X_k)K_j(X_k)K_i(X_l)K_j(X_l)\mathcal{I}_k\mathcal{I}_l}{g^2(X_k)g^2(X_l)} - \int \int \frac{K_i(x)K_j(x)K_i(y)K_j(y)}{g(x)g(y)} dx dy \right],$$

and $u_{ijkl} = \frac{h^d}{n^4} \epsilon_i^2 \epsilon_j^2 s_{ijkl}$. Hence, $[\tilde{\Gamma}_n - \tilde{\mathcal{G}}_n] = \sum_{j \neq i} \sum_k \sum_l u_{ijkl}$.

Expanding the quadratic, we obtain

$$[\tilde{\Gamma}_n - \tilde{\mathcal{G}}_n]^2 = \sum_i \sum_{j \neq i} \sum_k \sum_l \sum_m \sum_{n \neq m} \sum_p \sum_q u_{ijkl} u_{mnpq}$$

Since we have four kernel terms in each s_{ijkl} term, thus by using (g1) and (k) maximum order of $E s_{ijkl} s_{mnpq} = O(h^{-8d})$, and hence that of $E[u_{ijkl} u_{mnpq}] =$

$\frac{h^{2d}}{n^8} \times O(h^{-8d}) = O(n^{-8}h^{-6d})$. Also, we have eight summations, if there are at most five different subscripts in the summations, then the expected value of those terms which has at most five different subscripts is at most $n^5 \times O(n^{-8}h^{-6d}) = O(n^{-3}h^{-6d})$, and hence by (h2), it is $o(1)$.

So, we will only discuss those terms, which are involved with more than five different subscripts in the summations. According to that criteria, $[\bar{\Gamma}_n - \tilde{\mathcal{G}}_n]^2$ is bounded by the following terms:

$$\begin{aligned}
U_1 &= \sum_{j \neq i} \sum_k \sum_l \sum_{n \neq m} \sum_{p \neq i, j, k, l, m, n} \sum_{q \neq i, j, k, l, m, n, p} u_{ijkl} u_{mnpq}, \quad (2.61) \\
U_2 &= \sum_{j \neq i} \sum_{n \neq m} \sum_{k \neq i, j, m, n} \sum_{l \neq i, j, k, m, n} \sum_{p \neq i, j, k, l, m, n} u_{ijkl} \\
&\quad \times [u_{mnp i} + u_{mnp j} + u_{mnp m} + u_{mnp n}], \\
U_3 &= \sum_{j \neq i} \sum_{n \neq m} \sum_{k \neq i, j, m, n} \sum_{p \neq i, j, k, m, n} u_{mnpk} \\
&\quad \times [u_{ijii} + u_{ijij} + u_{ijim} + u_{ijin} + u_{ijji} + u_{ijjj} + u_{ijjm} + u_{ijjn} \\
&\quad + u_{ijmi} + u_{ijmj} + u_{ijmm} + u_{ijmn} + u_{ijn i} + u_{ijn j} + u_{ijn m} + u_{ijn n}], \\
U_4 &= \sum_{p \neq i \neq j \neq k \neq m \neq n \neq l} u_{ijkl} u_{mnp l}, \\
U_5 &= \sum_{p \neq i \neq j \neq k \neq m \neq n} \{u_{ijk n} (u_{mip k} + u_{mj p k} + u_{mnp k}) \\
&\quad + u_{mnp k} (u_{ijk i} + u_{ijk j} + u_{ijk k}) + u_{ijk m} (u_{inp k} + u_{jn p k} + u_{mnp k}) \\
&\quad + (u_{mnp i} + u_{mnp j} + u_{mnp m} + u_{mnp n}) (u_{ijk i} + u_{ijk j} + u_{ijk k} + u_{ijk m} \\
&\quad + u_{ijk n}) + (u_{ijk i} + u_{ijk j} + u_{ijk k}) u_{mnp k} + (u_{inp m} + u_{jn p m} + u_{kn p m}) u_{ijk m}\}.
\end{aligned}$$

To prove the claim, it suffices to show that the expected value of all of these terms is $o(1)$.

Note that, for all $k, l \neq i, j$ and $k \neq l$, $E[s_{ijkl}|X_i, X_j] = 0$. Now, by using this fact, $\forall p, q \neq i, j, k, m, n, l$, and $p \neq q$,

$$E(u_{ijkl} u_{mnpq}|X_i, X_j, X_k, X_m, X_n, X_l) = u_{ijkl} E(u_{mnpq}|X_m, X_n) = 0.$$

Hence, $E(U_1) = 0$.

Similarly, for all $k, l \neq i, j, m, n, p$ and $k \neq l$,

$$E(u_{ijkl} u_{mnpq}|X_i, X_j, X_m, X_n, X_p) = 0$$

and, expected value of other terms of U_2 is zero. Hence, $EU_2 = 0$. Similarly, for all $k, p \neq i, j, m, n$ and $k \neq p$, $E(U_3) = 0$.

Again by the above fact,

$$\begin{aligned} E(U_4) &= \frac{h^{2d}}{n} E[\epsilon_1^2 \epsilon_2^2 \epsilon_5^2 \epsilon_6^2 s_{1234} s_{5674}] \\ &= \frac{h^{2d}}{n} E[\{E[\sigma_{\theta_0}^2(X_1) \sigma_{\theta_0}^2(X_2) s_{1234}|X_3, X_4]\}^2] \\ &= \frac{h^{2d}}{n} E[\sigma_{\theta_0}^2(X_1) \sigma_{\theta_0}^2(X_2) \{E[s_{1234}|X_3, X_4]\}^2] = 0. \end{aligned}$$

By independence of X_i 's, expected value of the first term of U_5 is equal to

$$n^6 E[u_{ijkn} u_{mipk}] = \frac{h^{2d}}{n^2} E[\tau_{\theta_0}(X_1) \sigma_{\theta_0}^2(X_2) \sigma_{\theta_0}^2(X_5) s_{1234} s_{5163}].$$

By independence of X_i 's, (g1), (k) and continuity of σ_{θ_0} and τ_{θ_0} , one obtains

$$\begin{aligned} &E[\tau_{\theta_0}(X_1) \sigma_{\theta_0}^2(X_2) \sigma_{\theta_0}^2(X_5) s_{1234} s_{5123}] \\ &= E[\tau_{\theta_0}(X_1) \{E[\sigma_{\theta_0}^2(X_2) s_{1234}|X_1, X_3, X_4]\}^2] \\ &\leq \frac{2}{h^{4d}} \left[\int_{\mathcal{I}} \int \int \left[\int \tau_{\theta_0}^{1/2}(x - hv - hw) \sigma_{\theta_0}^2(x - hu - hw) K(u) K(v) \right. \right. \\ &\quad \left. \left. \times K(w + v) K(w + u) g(x - hu - hw) du \right]^2 \frac{g(x - hv - hw)}{g^3(x - hw) g^3(x)} dv dw dx \right] \end{aligned}$$

$$\begin{aligned}
& + \int \left[\int_{\mathcal{I}} \int \int \tau_{\theta_0}^{1/2}(x - hv - hw) \sigma_{\theta_0}^2(x - hu - hw) K(u) K(v) \right. \\
& \quad \times K(w + v) K(w + u) \frac{g(x - hu - hw) g^{1/2}(x - hv - hw)}{g(x) g(x - wh)} du dv dw dx \left. \right]^2 dv \\
& = O(h^{-4d}).
\end{aligned}$$

Hence, the expected value of the first term of U_5 is equal to $O(nh^d)^{-2} = o(1)$, by (h2). Similarly, by using (g1), (k) and (h2), expected value of the other terms of U_5 is $o(1)$. Hence, the second part of (2.54) is proved.

The proof of the third and fourth part of (2.54) is given in Lemma 5.5 of K-N.

Hence (2.54) is proved, and so is the Lemma 2.5.5. $\square$

CHAPTER 3

Minimum Distance

Goodness-Of-Fit Tests

For Current Status Data

3.1 Introduction

This chapter discusses a minimum distance method for fitting a parametric model to the distribution function of the event occurrence time in the one sample set up with current status data. Let X and T denote the event occurrence and inspection times, respectively. Let $F(G)$ denote the d.f. of $X(T)$. Assume X and T are independent. In the current status data set up, one observes $\delta = I[X \leq T]$ and T , where $I[A]$ denotes the indicator function of the event A . Let $\mathcal{A} := \{F_\theta(t) : t \in \mathbb{R}^+, \theta \in \Theta \subset \mathbb{R}^q, q \geq 1\}$ be a given parametric family of d.f.'s. Let $\mathcal{I}$ be a compact sub-interval of $[0, \infty)$. The

problem of interest here is to test the hypothesis

$$H_{01} : F(t) = F_{\theta_0}(t), \text{ for all } t \in \mathcal{I}, \text{ for some } \theta_0 \in \Theta,$$

against the alternative

$$H_{11} : H_{01} \text{ is not true.}$$

based on the random sample $\{(T_i, \delta_i) : 1 \leq i \leq n\}$ from the distribution of (T, δ) .

In this chapter, we adapt the inference procedures discussed in chapter 2 to the current status data. More precisely, let $\sigma_\theta^2(T_i) = F_\theta(T_i)(1 - F_\theta(T_i))$, and consider the regression model

$$\delta_i = F_\theta(T_i) + \sigma_\theta(T_i)\zeta_i, \quad 1 \leq i \leq n.$$

Here $\{\zeta_i\}$ are i.i.d. r.v's such that $E(\zeta_i|T_i) = 0$ and $E(\zeta_i^2|T_i) = 1$, $1 \leq i \leq n$. We shall be using the notation of chapter 2 with $X, Y, \mu(x)$ and m_θ replaced by $T, \delta, F(t)$ and F_θ , respectively, where now $d = 1$. Thus e.g., now

$$\mathcal{T}_n(\theta) := \int_{\mathcal{I}} \left[\frac{\frac{1}{n} \sum_{i=1}^n K_h(t - T_i)(\delta_i - F_\theta(T_i))}{g_w^*(t)} \right]^2 d\Phi(t), \quad (3.1)$$

$$M_n(\theta) := \frac{1}{n} \sum_{j=1}^n \left[\frac{1}{n} \sum_{i=1}^n \frac{K_h(T_j - T_i)(\delta_i - F_\theta(T_i))}{g_w^*(t)} \mathcal{I}_j \right]^2,$$

$$\mathcal{D}_n := nh^{1/2} \hat{\mathcal{G}}_n^{-1/2} (\mathcal{T}_n(\hat{\vartheta}_n) - \hat{R}_n),$$

$$\hat{\vartheta}_n := \operatorname{argmin}_{\theta \in \Theta} \mathcal{T}_n(\theta), \quad \hat{\theta}_n := \operatorname{argmin}_{\theta \in \Theta} M_n(\theta),$$

$$\hat{R}_n := n^{-2} \sum_{i=1}^n \int_{\mathcal{I}} \frac{K_h^2(t - T_i) \hat{\varepsilon}_i^2}{g_w^{*2}(t)} d\Phi(t), \quad \hat{\varepsilon}_i := \delta_i - F_{\hat{\theta}_n}(T_i), \quad 1 \leq i \leq n,$$

$$\hat{\mathcal{G}}_n := 2n^{-2} h^d \sum_{i \neq j} \left(\int_{\mathcal{I}} \frac{K_h(t - T_i) K_h(t - T_j) \hat{\varepsilon}_i \hat{\varepsilon}_j}{\hat{g}_h^2(t)} d\Phi(t) \right)^2,$$

$$\hat{C}_n := n^{-3} \sum_{j=1}^n \sum_{i=1}^n \frac{K_h^2(T_j - T_i) \hat{\varepsilon}_i^2}{g_w^{*2}(T_j)} \mathcal{I}_j,$$

$$\hat{\Gamma}_n := 2h^{d_n-4} \sum_{i \neq j} \left(\sum_l \frac{K_h(T_l - T_i) K_h(T_l - T_j) \hat{\varepsilon}_i \hat{\varepsilon}_j}{\hat{g}_h^2(T_l)} \mathcal{I}_l \right)^2.$$

This chapter is organized as follows. Section 2 adapts the results discussed in chapter 2 based on minimum distance statistic to the current status data. First, we discuss consistency of ϑ_n^* and $\hat{\vartheta}_n$ for $T(F^*)$, where $F^* \in L_2(\Phi)$ is a d.f., different from the null model $\mathcal{A}$. Then, we discuss the consistency of ϑ_n^* , $\hat{\vartheta}_n$ and asymptotic normality of $\hat{\vartheta}_n$ and $\mathcal{D}_n$, under H_{01} . Similar to Koul and Song (2006) (K-S), we also obtain consistency of $\mathcal{D}_n$ against a fixed alternative, under some regularity conditions.

Additionally, we obtain asymptotic power of the proposed minimum distance tests under a class of local alternatives $H_{1n} : F(t) = F_{\theta_0}(t) + \psi(t)/nh^{1/2}$, where ψ is a continuously differentiable function such that $\int \psi^2 d\Phi < \infty$ and $\int F_{\theta} \psi d\Phi = 0$, for all $\theta \in \Theta$.

Section 4 adapts the results of chapter 2 based on empirical minimum distance statistic to the current status data and discusses consistency of θ_n^* and $\hat{\theta}_n$ and asymptotic normality of $\hat{\theta}_n$ and $nh^{1/2} \hat{\Gamma}_n^{-1/2} (M_n(\hat{\theta}_n) - \hat{C}_n)$, under H_{01} .

Section 5 reports results of the three simulation studies. The first simulation study investigates Monte Carlo size and power of empirical minimum distance test. The finite sample level approximates the nominal level well for large sample sizes. Simulation results also show little bias in the estimator $\hat{\theta}_n$, for all the chosen sample sizes.

The second simulation study investigates the empirical size and power behavior

of the Cramér-Von-Mises test CV_1 , where CV_1 is defined in chapter 1. Since the asymptotic distribution of CV_1 is not known, so in order to find the Monte Carlo levels and powers of this test, we need to estimate its cut off points. Estimated cut off points are obtained by first getting 10,000 values of CV_1 and then by finding percentiles from the distribution of 10,000 values. The finite sample level approximates the nominal level well for all the chosen sample sizes. In our simulations, $\hat{F}$ is computed by the one step procedure for calculating the nonparametric maximum likelihood estimator, based on isotonic regression, cf. Groeneboom and Wellner (1992).

The third simulation study investigates Monte Carlo size comparison of the empirical minimum distance test with the tests of Koul and Yi (2006) (KY) and CV_1 . Simulation results show that the empirical levels of CV_1 and KY tests are better than $M_n(\hat{\theta}_n)$, when sample size is less than 200. But when the sample size is 200, the significance levels of all the three tests are comparable to each other.

3.2 Minimum Distance Statistics and Tests

In this section we adapt the results discussed in section 2 based on a class of minimum distance statistics to the current status data. Here we shall be using the same assumptions discussed in chapter 2 with X , Y , $\mu(x)$ and m_θ replaced by T , δ , $F(t)$ and F_θ , respectively, where now $d = 1$ and $\mathcal{I}$ is a bounded interval in $[0, \infty)$. Also under the current status data set up, assumptions (e1), (e2), (e3), and (e4) are automatically satisfied.

First, we discuss the consistency of ϑ_n^* and $\hat{\vartheta}_n$ for $T(F^*)$, where $F^* \in L_2(\Phi)$ is a

d.f., different from the null model $\mathcal{A}$. Let $H_{11} : F(t) = F^*(t)$, $t \in \mathcal{I}$.

Lemma 3.2.1 *Suppose assumptions (k), (g1), and (m3) of chapter 2 hold with m_θ replaced by F_θ , $d = 1$ and $\mathcal{I}$ is a bounded sub-interval in $[0, \infty)$. Let F^* be a given d.f. such that $F^* \notin \mathcal{A}$, $F^* \in L_2(\Phi)$, and $T(F^*)$ is unique.*

(a) *In addition, suppose F^* is a.e. (Φ) continuous. Then, under H_{11} , $\vartheta_n^* = T(F^*) + o_p(1)$.*

(b) *In addition, suppose F^* is continuous on $\mathcal{I}$. Then, under H_{11} , $\hat{\vartheta}_n = T(F^*) + o_p(1)$.*

Upon taking $F^* = F_{\theta_0}$ in the above result one immediately obtains the following:

Corollary 3.2.1 *Suppose assumptions (g1), (k), (m1)-(m3), (h1), and (h2) of chapter 2 hold with m_θ replaced by F_θ , where now $d = 1$ and $\mathcal{I}$ is a bounded sub-interval in $[0, \infty)$. Then, under H_{01} , $\vartheta_n^* \rightarrow \theta_0$, and $\hat{\vartheta}_n \rightarrow \theta_0$ in probability.*

Next result gives the asymptotic normality of $n^{1/2}(\hat{\vartheta}_n - \theta_0)$ under H_{01} . Let

$$\begin{aligned}
F_n(t, \theta) &:= n^{-1} \sum_{i=1}^n K_h(t - T_i) F_\theta(T_i), \\
U_n(t, \theta) &:= \frac{1}{n} \sum_{i=1}^n K_h(t - T_i) [\delta_i - F_\theta(T_i)], \quad U_n(t) := U_n(t, \theta_0), \\
\dot{F}_n(t, \theta) &:= n^{-1} \sum_{i=1}^n K_h(t - T_i) \dot{F}_\theta(T_i), \\
\dot{F}_h(t) &:= E \dot{F}_n(t, \theta_0) = E K_h(t - T) \dot{F}_{\theta_0}(T), \\
S_n &:= \int_{\mathcal{I}} \frac{U_n(t) \dot{F}_h(t)}{g^2(t)} d\Phi(t), \\
\Sigma &:= \int_{\mathcal{I}} \frac{F_{\theta_0}(t)(1 - F_{\theta_0}(t)) \dot{F}_{\theta_0}(t) \dot{F}_{\theta_0}(t)' \phi^2(t)}{g(t)} dt.
\end{aligned} \tag{3.2}$$

Corollary 3.2.2 *Suppose assumptions (g1), (g2), (p), (k), (m1)-(m5), and (h3) of chapter 2 hold with m_θ replaced by F_θ , where now $d = 1$ and $\mathcal{I}$ is a bounded sub-interval in $[0, \infty)$. Then, under H_{01} , $n^{1/2}(\hat{\vartheta}_n - \theta_0) = \Sigma_0^{-1}n^{1/2}\mathcal{S}_n + o_p(1)$. Consequently,*

$$n^{1/2}(\hat{\vartheta}_n - \theta_0) \rightarrow_d N_q(0, \Sigma_0^{-1}\Sigma\Sigma_0^{-1}).$$

Next, we state the asymptotic normality result about $\mathcal{D}_n$ under H_{01} . Let

$$\Gamma := 2 \int_{\mathcal{I}} \left\{ \frac{F_{\theta_0}(t)(1 - F_{\theta_0}(t))\phi(t)}{g(t)} \right\}^2 dt \int \left(\int K(u)K(v+u)du \right)^2 dv.$$

Corollary 3.2.3 *Suppose assumptions (g1), (g2), (p), (k), (m1)-(m5) and (h3) of chapter 2 hold with m_θ replaced by F_θ , where now $d = 1$ and $\mathcal{I}$ is a bounded sub-interval of $[0, \infty)$. Then under H_{01} , $\mathcal{D}_n \rightarrow_d N_1(0, \Gamma)$ and $|\hat{\mathcal{G}}_n\Gamma^{-1} - 1| = o_p(1)$, where $\hat{\mathcal{G}}_n$ is as in (3.1).*

Consequently, the test that rejects H_{01} whenever $|\mathcal{D}_n| > z_\alpha/2$ is of the asymptotic size α , where z_α is the $100(1 - \alpha)\%$ percentile of the standard normal distribution.

The following corollary provides a set of sufficient conditions under which $|\mathcal{D}_n| \rightarrow \infty$, in probability, for any sequence of consistent estimators $\hat{\vartheta}_n$ of $T(F^*)$ under the fixed alternative H_{11} .

Corollary 3.2.4 *Suppose assumptions (g1), (g2), (p), (k), (m3), (h3) of chapter 2 hold with m_θ replaced by F_θ , where now $d = 1$ and $\mathcal{I}$ is a bounded sub-interval of $[0, \infty)$. Assume the alternative hypothesis H_{11} hold with the additional assumption that $\inf_\theta \rho(F^*, F_\theta) > 0$. Then, for any sequence of consistent estimator $\hat{\vartheta}_n$ of $T(F^*)$, $|\mathcal{D}_n| \rightarrow \infty$, in probability. Consequently, the test that rejects whenever $|\mathcal{D}_n| > z_\alpha$ is consistent against the fixed alternative H_{11} .*

Its proof is similar to that of Theorem 5.1 in K-S adapted to the current status data.

Next, let ψ be a known continuously differentiable real valued function. In addition, assume $\psi \in L_2(\Phi)$ and

$$\int F_\theta \psi d\Phi = 0, \quad \forall \theta \in \Theta. \quad (3.3)$$

Consider the sequence of local alternatives

$$H_{1n} : F(t) = F_{\theta_0}(t) + \gamma_n \psi(t), \quad \gamma_n = 1/(nh^{1/2})^{1/2}. \quad (3.4)$$

The following corollary gives asymptotic power of the minimum distance test against the local alternative H_{1n} . Its proof is similar to that of Theorem 5.3 in K-S adapted to the current status data.

Corollary 3.2.5 *Suppose assumptions (g1), (g2), (p), (k), (m4), and (h3) of chapter 2 hold with m_θ replaced by F_θ , where now $d = 1$ and $\mathcal{I}$ is a bounded subinterval of $[0, \infty)$, then under the local alternative hypothesis (3.3) and (3.4), $\mathcal{D}_n \rightarrow_d N(\Gamma^{-1/2} \int \psi^2 d\Phi, 1)$.*

The following corollary gives the asymptotic distribution of $\hat{\vartheta}_n$ under H_{1n} . Its proof is similar to that of Theorem 5.2 in K-S adapted to the current status data.

Corollary 3.2.6 *Suppose assumptions (g1), (g2), (p), (k), (m1)-(m6), (h3) of chapter 2 hold with m_θ replaced by F_θ , where now $d = 1$ and $\mathcal{I}$ is a bounded subinterval of $[0, \infty)$, then under the local alternative (3.3) and (3.4), $n^{1/2}(\hat{\vartheta}_n - \theta_0) \rightarrow_d N_q(0, \Sigma_0^{-1} \Sigma \Sigma_0^{-1})$.*

3.3 Empirical Minimum Distance Statistic

In this section, we adapt the results of chapter 2 based on empirical minimum distance statistic to the current status data. First, we discuss the consistency of θ_n^* and $\hat{\theta}_n$. The consistency of these estimators for θ_0 under H_{01} follows from Lemma 2.3.2 of chapter 2. Applying Corollary 2.3.1 to the current status set up, we have

Corollary 3.3.1 *Suppose assumptions (g1), (k), (m1)-(m3), (h1), and (h2) of chapter 2 hold with m_θ and Φ replaced by F_θ and G , respectively, where now $d = 1$ and $\mathcal{I}$ is a bounded sub-interval of $[0, \infty)$. Then under H_{01} , $\theta_n^* \rightarrow \theta_0$, in probability.*

Applying Theorem 2.3.1 to the current status set up, we have

Corollary 3.3.2 *Suppose assumptions (g1), (k), (m1)-(m3), (h1), and (h2) of chapter 2 hold with m_θ and Φ replaced by F_θ and G , respectively, where now $d = 1$ and $\mathcal{I}$ is a bounded sub-interval in $[0, \infty)$. Then, under H_{01} ,*

$$\hat{\theta}_n \rightarrow \theta_0, \quad \text{in probability.}$$

Now, we discuss asymptotic normality of $n^{1/2}(\hat{\theta}_n - \theta_0)$. Let

$$\hat{S}_n := n^{-1} \sum_{j=1}^n \frac{U_n(T_j) \dot{F}_h(T_j) \mathcal{I}_j}{g(T_j)}.$$

Applying Theorem 2.4.1 to the current status set up, we obtain

Corollary 3.3.3 *Suppose (e1), (g1), (g2), (k), (m1)-(m5) and (h3) of chapter 2 hold with m_θ and Φ replaced by F_θ and G , respectively, where now $d = 1$ and $\mathcal{I}$ is a bounded sub-interval in $[0, \infty)$. Then under H_{01} ,*

$$n^{1/2}(\hat{\theta}_n - \theta_0) = \Sigma_0^{-1} n^{1/2} \hat{S}_n + o_p(1),$$

Consequently, $n^{1/2}(\hat{\theta}_n - \theta_0) \rightarrow_d \mathcal{N}_q(0, \Sigma_0^{-1} \Sigma \Sigma_0^{-1})$, where Σ_0 and Σ are as in (m6) and (3.2), respectively.

Next, we discuss asymptotic distribution of the empirical minimized distance $M_n(\hat{\theta}_n)$. It follows from Theorem 2.5.1 adapted to current status data.

Corollary 3.3.4 *Suppose (e1), (g1), (g2), (k), (m1)-(m5) and (h3) of chapter 2 hold with m_ϑ and Φ replaced by F_ϑ and G , respectively, where now $d = 1$ and $\mathcal{I}$ is a bounded sub-interval in $[0, \infty)$. Then under H_{01} , $nh^{1/2}(M_n(\hat{\theta}_n) - \hat{C}_n) \rightarrow_d \mathcal{N}_1(0, \Gamma)$. Moreover, $|\hat{\Gamma}_n \Gamma^{-1} - 1| = o_p(1)$, where $\hat{\Gamma}_n$ and $\hat{C}_n$ are as in (3.1).*

Consequently, the test that rejects H_0 whenever $nh^{1/2}\hat{\Gamma}_n^{-1/2}|M_n(\hat{\theta}_n) - \hat{C}_n| > z_{\alpha/2}$ is of the asymptotic size α .

3.4 Simulations

This section contains the results of three simulation studies. The first one assesses finite sample level and power behavior of the empirical minimum distance test statistic $M_n(\theta)$. The second simulation study investigates finite sample level behavior of the Cramér-Von-Mises test CV_1 . The third simulation study investigates a Monte Carlo size comparison of $M_n(\theta)$, CV_1 , and KY test. The simulations are done using Matlab. The kernel functions and the bandwidths used in the simulations are

$$\begin{aligned} K(x) &= K^*(x) = \frac{3}{4}(1 - x^2)I(|x| \leq 1) \\ h &= c_1 n^{-1/3}, \quad w = c_2 n^{-1/5}(\log n)^{1/5}, \end{aligned}$$

with some choices for c_1 and c_2 . In the tables below, $\exp(\theta)$, $\theta > 0$, denotes the exponential distribution with parameter θ under the null hypothesis $H_{01} : F = \exp(\theta)$, for some $\theta > 0$. The Weibull distribution with density $w(t) := ba^{-b}t^{b-1}\exp(-t/a)^b$ is denoted by $W(a, b)$ and $G(a, b)$ represents the Gamma distribution with density $g(t) := \frac{1}{b^a\Gamma(a)}t^{a-1}\exp(-t/b)$, $a > 0$, $b > 0$. The asymptotic level is taken to be 0.05 in all the cases. The sample sizes chosen are 50, 100, 200, each repeated 1,000 times.

Table 3.1 reports the Monte Carlo mean and the $\text{MSE}(\hat{\theta}_n)$ under H_{01} which are obtained by minimizing $M_n(\theta)$ and employing the Newton-Raphson algorithm. The sample sizes chosen are 50, 100, 200, 500, each repeated 1,000 times. One can see there appears to be little bias in $\hat{\theta}_n$ for all the chosen sample sizes and MSE decreases as the sample size increases.

To assess the effect of the choice of (c_1, c_2) that appears in the bandwidths on the level and power, we ran the simulations for various choices of (c_1, c_2) , ranging from 0.1 to 1. Table 3.2 reports the simulation results for those (c_1, c_2) which gave the best results. The entries in the tables for $M_n(\hat{\theta}_n)$ are obtained by computing the number of times $(|nh^{1/2}\hat{\Gamma}_n^{-1/2}(M_n(\hat{\theta}_n) - \hat{C}_n)| \geq 1.96)/1,000$. Table 3.2 summarizes the empirical levels for test statistic $M_n(\hat{\theta}_n)$. It shows that as the sample size increases the simulated levels are getting closer to the asymptotic level 0.05.

Table 3.3 represents the power for test statistic $M_n(\hat{\theta}_n)$ for four different alternatives, when $(c_1, c_2) = (.9, 1)$. It shows that the power is getting better as the sample size increases.

The second simulation study investigates the behavior of the Cramér-Von-Mises

test CV_1 . Since the asymptotic distribution of CV_1 is not known, so in order to find the Monte Carlo levels and powers of this test, we need to estimate its cut off points. Estimated cut off points are obtained by first generating 10,000 values of CV_1 and then by finding percentiles from these 10,000 values. After that, for CV_1 , the empirical level and power are obtained by computing the number of $(CV_1 \geq \text{estimated cut off point})/1,000$. In our simulations, $\hat{F}$ is obtained by the one step procedure for the calculation of the nonparametric maximum likelihood estimator, based on isotonic regression, cf. Groeneboom and Wellner (1992).

Table 3.4 contains the simulated 90th, 95th, 97.5th, 99th, and 99.5th percentiles of CV_1 when distributions of X, T are $\exp(1)$. Table 3.5 represents simulated significance levels by using the corresponding simulated percentiles given in Table 3.4 when testing $F = \exp(1)$ and the distribution of T is $\exp(1)$. It shows that the simulated significance levels of CV_1 for different chosen sample sizes are very close to the true nominal sizes.

Let

$$\hat{\zeta}_n := \operatorname{argmin}_{\zeta \in \Theta} CV_1(\zeta).$$

Table 3.6 reports the Monte Carlo mean and the MSE of $(\hat{\zeta}_n)$ under $F = \exp(1)$ which are obtained by minimizing CV_1 and employing the Newton-Raphson algorithm. One can see there appears to be little bias in $\hat{\theta}_n$ for all the chosen sample sizes and MSE decreases as the sample size increases.

Table 3.7 represents the power of CV_1 for five different alternatives, when distribution of T is $\exp(1)$. It shows that the power is getting better as the sample size

increases.

In the third simulation study, we make a comparison of the Monte Carlo level of the proposed empirical minimum distance test $M_n(\hat{\theta}_n)$ with the other two tests CV_1 and KY.

KY test. Let $\hat{\theta}_{MLE}$ denote the maximum likelihood estimator, obtained by using the the following score statistics $S_n(\theta)$ given in (3.3) of KY:

$$S_n(\theta) := \sum_{i=1}^n \left[\frac{\delta_i}{1 - e^{-\theta T_i}} - 1 \right] T_i.$$

Let

$$\hat{\mathcal{U}}_n(t) := n^{-1/2} \sum_{i=1}^n \left[I(T_i \leq t) - n^{-1} \sum_{j=1}^n \frac{T_j T_i e^{-\hat{\theta}(T_i+T_j)/2} (1 - e^{-\hat{\theta} T_i})^{-1/2} (1 - e^{-\hat{\theta} T_j})^{-1/2} I(T_j \leq t \wedge T_i)}{n^{-1} \sum_{k=1}^n T_k^2 e^{-\hat{\theta} T_k} (1 - e^{-\hat{\theta} T_k})^{-1} I(T_k \geq T_j)} \right] \hat{\varepsilon}_i.$$

Let G_n denote the empirical distribution of the design variable T_i , $1 \leq i \leq n$ and $t_0 = 99^{th}$ percentile of G_n . The KY test statistic is

$$\hat{K}_n := \sup_{0 \leq t \leq t_0} \frac{|\hat{\mathcal{U}}_n(t)|}{\sqrt{G_n(t_0)}}.$$

As shown in KY, the limiting null distribution of $\hat{K}_n$ is the same as that of $\sup_{0 \leq t \leq 1} |B(t)|$, where B is the standard Brownian motion. The 95th percentile of this distribution is approximately equal to 2.24241, which is obtained from the fact

$$P\left(\sup_{0 \leq t \leq 1} |B(t)| < b\right) = P(B(1) < b) + 2 \sum_{i=1}^{\infty} (-1)^i P((2i-1)b < B(1) < (2i+1)b).$$

The empirical size is computed by using $\#\{\hat{K}_n > 2.24241\}/1000$.

Table 3.8 shows comparison of simulated significance levels for $M_n(\hat{\theta}_n)$, CV_1 and $\hat{K}_n$. For the simulated significance levels of CV_1 we used the percentiles given in

Table 3.4. It shows that the empirical levels of statistics CV_1 and $\hat{K}_n$ tests are better than $M_n(\hat{\theta}_n)$, when sample size is less than 200. But when the sample size is 200, the significance levels of all the three tests are comparable to each other.

Table 3.1: Mean and MSE of $\hat{\theta}_n$, X , $T \sim \exp(1)$, $\theta_0 = 1$

Sample Size	50	100	200	500
Mean	1.0442	1.0172	1.0025	1.0016
MSE	0.3324	0.1780	0.1216	0.0748

Table 3.2: Empirical sizes of $M_n(\hat{\theta}_n)$, X , $T \sim \exp(1)$

$c_1, c_2 \backslash n$	50	100	200
0.5, 0.2	0.024	0.044	0.049
0.8, 0.6	0.061	0.058	0.046
0.8, 0.7	0.088	0.062	0.056
0.9, 0.8	0.07	0.061	0.05
0.9, .9	0.07	0.059	0.049
0.9, 1	0.094	0.058	0.045

Table 3.3: Power of $M_n(\hat{\theta}_n)$, $T \sim \exp(1)$, $(c_1, c_2) = (.9, 1)$

X \ n	50	100	200
G(2,1)	0.975	1	1
G(1,3)	0.927	0.999	1
W(1,5)	0.211	0.404	0.628
W(1,2)	0.452	0.693	0.929

Table 3.4: Simulated percentiles of CV_1 , $X, T \sim \exp(1)$

Percentile \ n	50	100	200
99.5	0.0412	0.0245	0.0152
99	0.0359	0.0216	0.0131
97.5	0.0298	0.0177	0.0112
95	0.0236	0.0156	0.0095
90	0.0183	0.0125	0.0077

Table 3.5: Empirical sizes of CV_1 , $X, T \sim \exp(1)$

True level \ n	50	100	200
0.005	0.004	0.003	0.004
0.01	0.011	0.009	0.008
0.025	0.025	0.029	0.026
0.05	0.048	0.056	0.049
0.1	0.101	0.106	0.09

Table 3.6: Mean and MSE of $\hat{\zeta}_n$, $X, T \sim \exp(1)$, $\theta_0 = 1$

Sample Size	50	100	200
Mean	1.6815	1.4594	1.3088
MSE	0.5558	0.3152	0.1878

Table 3.7: Power of CV_1 , $T = \exp(1)$.

Dist. of $X \setminus n$	50	100	200
G(1,3)	0.95	0.999	1
G(2,1)	0.975	1	1
W(1,.5)	0.413	0.624	0.891
W(1,1.5)	0.394	0.45	0.575
W(1,2)	0.619	0.779	0.957

Table 3.8: Empirical sizes, $X, T \sim \exp(1)$, $(c_1, c_2) = (.9, .8)$

Tests \ n	50	100	200
M_n	.074	0.07	0.055
$\hat{K}_n$	.04	0.049	0.052
CV_1	0.049	0.048	0.051

CHAPTER 4

Testing the equality of two distributions with Current Status Data

4.1 Introduction

This chapter discusses the problem of testing the equality of two distribution functions based on current status data. Accordingly, let X, S (Y, T) denote the event occurrence and inspection times, respectively, from the first (second) population. Let F_1 (F_2) denote d.f. of X (Y) and G_1 (G_2) denote the d.f. of S (T). Let $X_1, \dots, X_{n_1}$ ($Y_1, \dots, Y_{n_2}$) be i.i.d. F_1 (F_2) and $S_1, \dots, S_{n_1}$ ($T_1, \dots, T_{n_2}$) be i.i.d. G_1 (G_2) random variables. Assume all random variables are mutually independent. In the two sample current status data set up, one observes (δ_i, S_i) , $1 \leq i \leq n_1$, and (η_j, T_j) , $1 \leq j \leq n_2$, where $\delta = I[X \leq S]$, $\eta = I[Y \leq T]$.

The problem of interest here is to test the null hypothesis that the two event occurrence distributions are the same, i.e.

$$H_{02} : F_1(x) = F_2(x), \quad \text{for all } x \in \mathcal{I},$$

against the alternative

$$H_{12} : F_1(x) \neq F_2(x), \quad \text{for some } x \in \mathcal{I},$$

where $\mathcal{I}$ is a compact sub-interval of $[0, \infty)$.

In this chapter we adapt the test proposed by Koul and Schick (2003)(K-Sh) to the two sample current status data. More precisely, let $\sigma_1^2(S_i) = F_1(S_i)(1 - F_1(S_i))$, $\sigma_2^2(T_j) = F_2(T_j)(1 - F_2(T_j))$ and consider the regression models

$$\delta_i := F_1(S_i) + \sigma_1(S_i)\zeta_{1i}, \quad 1 \leq i \leq n_1,$$

$$\eta_j := F_2(T_j) + \sigma_2(T_j)\zeta_{2j}, \quad 1 \leq j \leq n_2.$$

Here ζ_{1i} , ζ_{2j} are i.i.d. r.v.'s such that $E(\zeta_{1i}|S_i) = 0 = E(\zeta_{2j}|T_j)$ and $E(\zeta_{1i}^2|S_i) = 1 = E(\zeta_{2j}^2|T_j)$, $1 \leq i \leq n_1$, $1 \leq j \leq n_2$. Assume also that G_1 , G_2 have positive densities g_1 and g_2 on $[0, \infty)$, respectively, and that F_1 , F_2 have bounded densities.

Let $\mathcal{U}$ denote the set of all nonnegative functions that vanish off $\mathcal{I}$ and whose restrictions to $\mathcal{I}$ are continuous. Consider the integral

$$\Gamma = \int u(x)[F_1(x) - F_2(x)]dx, \quad u \in \mathcal{U}.$$

A possible choice for u is the indicator $1_{\mathcal{I}}$ of the interval $\mathcal{I}$. The integral Γ is 0 if the null hypothesis holds, and is non-zero under the alternative H_{12} , for all $u \in \mathcal{U}$.

Let K be a symmetric density with compact support $[-1,1]$ and $a = a_n$ be a bandwidth sequence, and let

$$\mathcal{T} = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \frac{\sqrt{u(S_i)} \sqrt{u(T_j)}}{g_1(S_i) g_2(T_j)} (\delta_i - \eta_j) K_a(S_i - T_j).$$

Observe that

$$E(\mathcal{T}) = \int \int \sqrt{u(s)} \sqrt{u(t)} [F_1(s) - F_2(t)] K_a(s - t) ds dt,$$

which is close to Γ for small a . Thus $\mathcal{T}$ provides an estimate of Γ if g_1 and g_2 are known, which is rarely the case.

This suggests to replace the densities in $\mathcal{T}$ by their estimates. Accordingly, let

$$\hat{\mathcal{T}} = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \hat{\nu}_1(S_i) \hat{\nu}_2(T_j) (\delta_i - \eta_j) K_a(S_i - T_j), \quad (4.1)$$

where $\hat{\nu}_k$ is an estimate of $\nu_k = \sqrt{u}/g_k$, $k = 1, 2$, constructed from the pooled sample such that $\hat{\nu}_k(x) = 0$ for $x \notin \mathcal{I}$. Similar to K-Sh, the estimators of ν_k can be obtained when u is known and when $u = u_\gamma$ as described in Remark 4.2.5 below. So that the asymptotic normality of $\hat{\mathcal{T}}$ both under the null hypothesis and under local alternatives

$$F_1 = F_2 + N^{-\frac{1}{2}} \gamma, \quad N := \frac{n_1 n_2}{n_1 + n_2}, \quad (4.2)$$

can be obtained, where γ is a non-negative continuous function such that $\gamma(0) = 0$, $\gamma(\infty) = 0$ and $0 < \int u(x) \gamma(x) dx < \infty$.

The rest of the chapter is organized as follows. Section 2 discusses asymptotic normality of $\hat{\mathcal{T}}$ under a general set of assumptions on the estimates $\hat{\nu}_1$ and $\hat{\nu}_2$. Section 3 reports the numerical results of the two simulation studies. The first one assesses

the finite sample level and power behavior of $\hat{T}$ test. The simulation results of the test statistic $\hat{T}$ are consistent with asymptotic theory.

In the second study, the finite sample comparison of $\hat{T}$ and CV_2 tests is made, where CV_2 is defined in chapter 1. Since the asymptotic distribution of CV_2 is not known, so in order to find Monte Carlo levels and powers of this test, we need to estimate its cut off points. Estimated cut off points are obtained by first getting 10,000 values of CV_2 and then by finding percentiles from the distribution of these 10,000 values. Simulation results show that for all the chosen alternatives and bandwidths, significance level of CV_2 is better than $\hat{T}$, and power of $\hat{T}$ is better than CV_2 , when sample sizes are 50 and 100. But when sample size is 200, significance level and power of $\hat{T}$ and CV_2 tests are comparable. In our simulations, $\hat{F}_1$ and $\hat{F}_2$ are computed by the one step procedure for calculating the nonparametric maximum likelihood estimator, based on isotonic regression, cf. Groeneboom and Wellner (1992).

4.2 Asymptotic behavior under the null hypothesis and local alternatives

This section discusses the behavior of the test statistic $\hat{T}$ given in (4.1) under the null hypothesis and under the alternatives (4.2). Note that the choice $\gamma = 0$ in (4.2) corresponds to the null hypothesis. To stress the dependence of local alternative on the parameter γ we write P_γ for the underlying probability measure and E_γ for the corresponding expectation.

Arguing as in K-Sh, we shall describe the asymptotic behavior of $\hat{T}$ as the sample sizes n_1 and n_2 tend to ∞ . For this we need the following assumptions.

(A.1) The function $u \in \mathcal{U}$, the set of all non-negative functions that vanish off $\mathcal{I}$ and whose restrictions to $\mathcal{I}$ are continuous.

(A.2) The densities g_1 and g_2 are bounded and their restrictions to $\mathcal{I}$ are positive and continuous.

(A.3) For any pairs of d.f.'s (F_1, G_1) , and (F_2, G_2) , $P(0 < F_1(S) < 1) = 1$ and $P(0 < F_2(T) < 1) = 1$.

(A.4) The weight function K is a symmetric Lipschitz-continuous density with compact support $[-1, 1]$.

(A.5) The bandwidth a is chosen such that $a^2 N \rightarrow 0$ and $a N^c \rightarrow \infty$, for some $c < 1$.

Note that $\sigma_1(S) = 0$ a.s., implies either $F_1(S) = 0$ or $F_1(S) = 1$ a.s. In the former case $E(\delta|S) = 0$ implies $\delta = 0$ a.s. Hence $\delta - F_1(S) = 0$ a.s.. Similarly $F_1(S) = 1$ a.s. implies $\delta = 1$ a.s. and hence $\delta - F_1(S) = 0$ a.s. Thus $\sigma_1(S) = 0$ a.s., implies $\delta - F_1(S) = 0$ a.s., and, under (A.3), $P(\sigma_1(S) > 0) = P(0 < F_1(S) < 1) = 1$. Similarly, $P(0 < F_2(T) < 1) = P(\sigma_2(T) > 0) = 1$.

The condition (A.3) is a joint condition on the supports of (F_1, G_1) and (F_2, G_2) . For example, if distributions of X and S are exponential with the scale parameter $\theta_1 > 0$, then $P(0 < F_1(S) < 1) = P(0 < e^{-\theta_1 S} < 1) = 1$. But if the distributions of X and S are $U(0, 1)$ and exponential with the scale parameter $\theta > 0$, respectively, then $P(0 < F_1(S) < 1) = P(0 < S < 1) = 1 - e^{-\theta}$, and hence in this case the

first part of (A.3) does not hold. A sufficient condition for (A.3) is that F_1 (F_2) be strictly increasing on the support of G_1 (G_2).

Also note that under (A.2) and (A.3) the functions g_1 , g_2 and σ_1^2 , σ_2^2 are bounded and bounded away from zero on interval $\mathcal{I}$ and so are the functions σ_1^2/g_1 and σ_2^2/g_2 .

To establish the asymptotic normality of $\hat{T}$, rewrite this statistic as

$$\frac{1}{n_1} \sum_{i=1}^{n_1} \hat{r}_1(S_i) \sigma_1(S_i) \zeta_{1i} - \frac{1}{n_2} \sum_{j=1}^{n_2} \hat{r}_2(T_j) \sigma_2(T_j) \zeta_{2j} + \frac{N^{-\frac{1}{2}}}{n_1} \sum_{i=1}^{n_1} \hat{r}_1(S_i) \delta(S_i) + \hat{T}_4,$$

where

$$\begin{aligned} \hat{r}_1(x) &= \hat{\nu}_1(x) \frac{1}{n_2} \sum_{j=1}^{n_2} \hat{\nu}_2(T_j) K_a(x - T_j), \\ \hat{r}_2(x) &= \hat{\nu}_2(x) \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{\nu}_1(S_i) K_a(x - S_i), \quad x \in [0, \infty), \end{aligned}$$

and

$$\hat{T}_4 = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \hat{\nu}_1(S_i) \hat{\nu}_2(T_j) (F_2(S_i) - F_2(T_j)) K_a(S_i - T_j).$$

As in K-Sh, the following additional definitions and assumptions are made to analyze the asymptotic behavior of $\hat{T}$. Let $\mathbf{S} = (S_1, \dots, S_{n_1})$, $\mathbf{T} = (T_1, \dots, T_{n_2})$, $\boldsymbol{\delta} = (\delta_1, \dots, \delta_{n_1})$, $\boldsymbol{\eta} = (\eta_1, \dots, \eta_{n_2})$, and $\boldsymbol{\delta}_i$ ($\boldsymbol{\eta}_j$) be the vector obtained from $\boldsymbol{\delta}$ ($\boldsymbol{\eta}$) by removing δ_i (η_j).

Definition 4.2.1 We say the estimator $\hat{r}_k$ is consistent and cross-validated (CCV) on $\mathcal{I}$ for the function r_k if the following conditions hold:

$$\begin{aligned} \frac{N}{n_1^2} \sum_{i=1}^{n_1} I_{\mathcal{I}}(S_i) E_{\gamma}[(\hat{r}_1(S_i) - r_1(S_i))^2 | \mathbf{S}] &= o_{p_{\gamma}}(1), \\ \frac{N}{n_2^2} \sum_{j=1}^{n_2} I_{\mathcal{I}}(T_j) E_{\gamma}[(\hat{r}_2(T_j) - r_2(T_j))^2 | \mathbf{T}] &= o_{p_{\gamma}}(1), \end{aligned}$$

$$N \max_{1 \leq i \leq n_1} \sup_{x \in \mathcal{I}} E_\gamma[(\hat{r}_1(x) - E_\gamma[\hat{r}_1(x)|\mathbf{S}, \boldsymbol{\delta}_i])^2|\mathbf{S}] = o_{p_\gamma}(1),$$

$$N \max_{1 \leq j \leq n_2} \sup_{x \in \mathcal{I}} E_\gamma[(\hat{r}_2(x) - E_\gamma[\hat{r}_2(x)|\mathbf{T}, \boldsymbol{\eta}_j])^2|\mathbf{T}] = o_{p_\gamma}(1).$$

We say $\tilde{r}_k$ is a modification of $\hat{r}_k$ if $P_\gamma(\sup_{x \in \mathcal{I}} |\tilde{r}_k(x) - \hat{r}_k(x)| > 0) \rightarrow 0$. We say $\hat{r}_k$ is essentially CCV on $\mathcal{I}$ for r_k if there exists a modification of $\hat{r}_k$ which is CCV on $\mathcal{I}$ for r_k .

Assumption 4.2.2 The estimate $\hat{r}_k$ is essentially CCV on $\mathcal{I}$ for $r_k = u/g_k$ for $k = 1, 2$.

The following result gives a sufficient condition for Assumption 4.2.2. Its proof is similar to that of Lemma 2.4 in K-Sh and hence no details are given.

Lemma 4.2.1 Suppose there are modifications $\tilde{\nu}_k$ of $\hat{\nu}_k$ such that, for $k=1,2$

$$0 \leq \tilde{\nu}_k(x) \leq K, \quad x \in \mathcal{I}, \quad (4.3)$$

for some finite constant K ,

$$\frac{1}{n_1} \sum_{i=1}^{n_1} E_\gamma[(\tilde{\nu}_1(S_i) - \nu_1(S_i))^2|\mathbf{S}] = o_{p_\gamma}(1), \quad (4.4)$$

$$\frac{1}{n_2} \sum_{j=1}^{n_2} E_\gamma[(\tilde{\nu}_2(T_j) - \nu_2(T_j))^2|\mathbf{T}] = o_{p_\gamma}(1), \quad (4.5)$$

$$N \max_{1 \leq i \leq n_1} \sup_{x \in \mathcal{I}} E_\gamma[(\tilde{\nu}_1(x) - E_\gamma[\tilde{\nu}_1(x)|\mathbf{S}, \boldsymbol{\delta}_i])^2|\mathbf{S}] = o_{p_\gamma}(1), \quad (4.6)$$

$$N \max_{1 \leq j \leq n_2} \sup_{x \in \mathcal{I}} E_\gamma[(\tilde{\nu}_2(x) - E_\gamma[\tilde{\nu}_2(x)|\mathbf{T}, \boldsymbol{\eta}_j])^2|\mathbf{T}] = o_{p_\gamma}(1). \quad (4.7)$$

Then, Assumption 4.2.2 holds.

The next result gives the asymptotic distribution of $\hat{\mathcal{T}}$ under the alternative (4.2) for any γ including the case $\gamma = 0$.

Proposition 4.2.1 *Suppose the conditions (A.1)-(A.5) and Assumption 4.2.2 hold.*

Then, under P_γ , $N^{1/2}(\hat{T} - \Gamma)/\tau$ converges in distribution to a $N(0, 1)$ r.v., where

$$\tau^2 = \int u^2(x)[q_1\psi_1(x) + q_2\psi_2(x)]dx,$$

$$\psi_1 = F_1(1 - F_1)/g_1, \psi_2 = F_2(1 - F_2)/g_2, q_1 = N/n_1 \text{ and } q_2 = N/n_2.$$

Details of the proof of this result are similar to those appearing in K-Sh and left out for the sake of brevity.

Remark 4.2.3 The above result suggests a test which rejects H_{02} for large values of $|\hat{T}|$. To implement such a test we need a consistent estimate $\hat{\tau}^2$ of τ^2 . Given such an estimator $\hat{\tau}^2$, we have under the above assumptions that $N^{1/2}(\hat{T} - \Gamma)/\hat{\tau}$ is asymptotically standard normal under P_γ , where $\Gamma = 0$ under H_{02} . Let now Φ denote the standard normal distribution function and $z_{\alpha/2}$ be its $(1 - \alpha/2)$ -quantile. Then a test that rejects H_{02} if $|(N^{1/2}\hat{T})/\hat{\tau}| \geq z_{\alpha/2}$, has the asymptotic level α . Moreover, from the above result, the asymptotic power of this test, under P_γ , is $1 - \Phi(z_{\alpha/2} - \kappa) + \Phi(-z_{\alpha/2} - \kappa)$, where

$$\kappa = \frac{\int u(x)\gamma(x)dx}{\tau}.$$

Note that the value of κ does not change if we replace u by cu , with c a positive constant.

Remark 4.2.4 Optimal u . Similar to K-Sh, the optimal u can be achieved such that it maximizes the asymptotic power, or equivalently the function κ , under (4.2) for a specific function γ . An application of the Cauchy-Schwarz inequality shows that

κ is maximized by the choice

$$u = u_\gamma = \frac{\gamma I_{\mathcal{I}}}{q_1 \psi_1(x) + q_2 \psi_2(x)}, \quad (4.8)$$

and the maximal value of κ is

$$\kappa_\gamma = \left(\int \frac{\gamma^2(x) I_{\mathcal{I}}(x)}{q_1 \psi_1(x) + q_2 \psi_2(x)} dx \right)^{1/2}.$$

The optimal u_γ depends on the sample sizes, the density functions g_1 and g_2 , and the distribution functions F_1 and F_2 . Next, we shall present the estimates of g_k , σ_k^2 , τ^2 and ν_k .

Estimates of g_k , σ_k^2 , τ^2 and ν_k . Similar to K-Sh, estimates of ν_k , $k = 1, 2$, can be found for fixed given u and for the (unknown) optimal $u = u_\gamma$. For this we need estimates of the inspection time densities and variance function.

The inspection time densities g_1 and g_2 can be estimated by the kernel density estimates

$$\begin{aligned} \hat{g}_1(t) &= \frac{1}{n_1} \sum_{i=1}^{n_1} K_{h_1}(t - S_i), \\ \hat{g}_2(t) &= \frac{1}{n_2} \sum_{j=1}^{n_2} K_{h_2}(t - T_j), \quad t \in \mathbb{R}^+, \end{aligned}$$

with bandwidth h_k , $k = 1, 2$. Its expected value is

$$\bar{g}_k(t) = \int g_k(t + h_k y) K(y) d(y), \quad t \in \mathbb{R}^+.$$

Lemma 4.2.2 *Suppose (A.2), (A.4) hold and the bandwidth h_k is such that $h_k \rightarrow 0$ and $h_k n_k^c \rightarrow \infty$ for some $c < 1$. Then the following hold:*

$$\sup_{t \in \mathcal{I}} |\hat{g}_k(t) - \bar{g}_k(t)| = o_{p_\gamma}(1), \quad (4.9)$$

$$\int (\bar{g}_k(t) - g_k(t))^2 dt \rightarrow 0. \quad (4.10)$$

Details of the proof of this result are similar to those appearing in K-Sh and left out for the sake of brevity.

Next, consider the following estimator of σ_k^2 , $k = 1, 2$:

$$\begin{aligned}\hat{\sigma}_1^2(t) &= \frac{\sum_{i=1}^{n_1} (\delta_i - \hat{\mu}_1(S_i))^2 K_{c_1}(t - S_i)}{\sum_{i=1}^{n_1} K_{c_1}(t - S_i)}, \\ \hat{\sigma}_2^2(t) &= \frac{\sum_{j=1}^{n_2} (\eta_j - \hat{\mu}_2(T_j))^2 K_{c_2}(t - T_j)}{\sum_{j=1}^{n_2} K_{c_2}(t - T_j)}, \quad t \in \mathbb{R}^+, \end{aligned}$$

where $\hat{\mu}_k$, $k = 1, 2$ is the kernel regression estimate

$$\begin{aligned}\hat{\mu}_1^2(t) &= \frac{\sum_{i=1}^{n_1} \delta_i K_{b_1}(t - S_i)}{\sum_{i=1}^{n_1} K_{b_1}(t - S_i)}, \\ \hat{\mu}_2^2(t) &= \frac{\sum_{j=1}^{n_2} \eta_j K_{b_2}(t - T_j)}{\sum_{j=1}^{n_2} K_{b_2}(t - T_j)}, \quad t \in \mathbb{R}^+. \end{aligned}$$

Here the bandwidths b_k and c_k are chosen such that $b_k \rightarrow 0$, $c_k \rightarrow 0$, $n_k^c(b_k + c_k) \rightarrow \infty$, for some $c < \frac{1}{2}$. The following lemma gives the needed properties of this estimator.

It follows from Lemma 3.3 of K-Sh.

Lemma 4.2.3 *Suppose (A.2)-(A.5) hold, and the conditional fourth moment v_k is bounded on an open interval $\mathcal{I}$. Then $\hat{\sigma}_k^2$ is essentially CCV on $\mathcal{I}$ for σ_k^2 and*

$$\sup_{x \in \mathcal{I}} |\hat{\sigma}_k^2 - \sigma_k^2| = o_{p_\gamma}(1), \quad k = 1, 2. \quad (4.11)$$

Now, consider the following estimator of variance function τ^2 :

$$\hat{\tau}^2 = q_1 \frac{1}{n_1} \sum_{i=1}^{n_1} \frac{\hat{u}^2(S_i)}{\hat{g}_1^2(S_i)} \hat{\sigma}_1^2(S_i) + q_2 \frac{1}{n_2} \sum_{j=1}^{n_2} \frac{\hat{u}^2(T_j)}{\hat{g}_2^2(T_j)} \hat{\sigma}_2^2(T_j). \quad (4.12)$$

The following Lemma proves the consistency of this estimator.

Lemma 4.2.4 *Suppose the assumptions of Proposition 4.2.1 and Lemma 4.2.3 hold and $\hat{u}$ be a uniformly consistent estimator of u on $\mathcal{I}$. Then, $\hat{\tau}^2 = \tau^2 + o_{p_\gamma}(1)$, where*

$$\begin{aligned}\tau^2 &= \tau_1^2 + \tau_2^2 \\ &= q_1 \int \frac{u^2(x)F_1(x)(1-F_1(x))}{g_1^2(x)} dG_1(x) + q_2 \int \frac{u^2(x)F_2(x)(1-F_2(x))}{g_2^2(x)} dG_2(x).\end{aligned}$$

Proof. Note that (4.12) can be written as $\hat{\tau}^2 = \hat{\tau}_1^2 + \hat{\tau}_2^2$, where

$$\hat{\tau}_1^2 = q_1 \frac{1}{n_1} \sum_{i=1}^{n_1} \frac{\hat{u}^2(S_i)}{\hat{g}_1^2(S_i)} \hat{\sigma}_1^2(S_i), \quad \hat{\tau}_2^2 = q_2 \frac{1}{n_2} \sum_{j=1}^{n_2} \frac{\hat{u}^2(T_j)}{\hat{g}_2^2(T_j)} \hat{\sigma}_2^2(T_j).$$

Let $\tilde{\tau}^2 = \tilde{\tau}_1^2 + \tilde{\tau}_2^2$, where

$$\tilde{\tau}_1^2 = q_1 \frac{1}{n_1} \sum_{i=1}^{n_1} \frac{u^2(S_i)}{\hat{g}_1^2(S_i)} \hat{\sigma}_1^2(S_i), \quad \tilde{\tau}_2^2 = q_2 \frac{1}{n_2} \sum_{j=1}^{n_2} \frac{u^2(T_j)}{\hat{g}_2^2(T_j)} \hat{\sigma}_2^2(T_j).$$

In order to prove $\hat{\tau}^2 = \tau^2 + o_{p_\gamma}(1)$, it suffices to prove that

$$\hat{\tau}^2 = \tilde{\tau}^2 + o_{p_\gamma}(1), \quad \tilde{\tau}^2 = \tau^2 + o_{p_\gamma}(1). \quad (4.13)$$

For the first claim in (4.13), it suffices to show that

$$\hat{\tau}_1^2 = \tilde{\tau}_1^2 + o_{p_\gamma}(1), \quad \hat{\tau}_2^2 = \tilde{\tau}_2^2 + o_{p_\gamma}(1). \quad (4.14)$$

By the choice of n_1 and n_2 , $0 < q_1 < 1$. Thus, for the first claim in (4.14), it suffices to show that

$$\left| \frac{1}{n_1} \sum_{i=1}^{n_1} \left[\frac{[\hat{u}^2(S_i) - u^2(S_i)] \hat{\sigma}_1^2(S_i)}{\hat{g}_1^2(S_i)} \right] \right| = o_{p_\gamma}(1). \quad (4.15)$$

Now, the left hand side of (4.15) is bounded above by

$$\sup_{x \in \mathcal{I}} \left[|\hat{u}^2(x) - u^2(x)| \frac{\hat{\sigma}_1^2(x)}{\hat{g}_1^2(x)} \right] = o_{p_\gamma}(1),$$

by uniform consistency of $\hat{u}$, $\hat{\sigma}_1$, $\hat{g}_1$, (A.2) and (A.3). This completes the proof of the first claim in (4.14). The proof of the second claim in (4.14) is similar, thereby completing the proof of the first part of (4.13).

To prove the second claim in (4.13), it suffices to show that

$$\tilde{\tau}_1^2 = \tau_1^2 + o_{p_\gamma}(1), \quad \tilde{\tau}_2^2 = \tau_2^2 + o_{p_\gamma}(1). \quad (4.16)$$

By the choice of n_1 and n_2 , $0 < q_1 < 1$. Thus, for the first claim in (4.16), it suffices to show that

$$\left| \frac{1}{n_1} \sum_{i=1}^{n_1} \left[\frac{u^2(S_i) \hat{\sigma}_1^2(S_i)}{\hat{g}_1^2(S_i)} \right] - \int \frac{u^2(x) \sigma_1^2(x)}{g_1^2(x)} dG_1(x) \right| = o_{p_\gamma}(1). \quad (4.17)$$

By the Law of Large Numbers,

$$\frac{1}{n_1} \sum_{i=1}^{n_1} \left[\frac{u^2(S_i) \sigma_1^2(S_i)}{g_1^2(S_i)} \right] \rightarrow \int \frac{u^2(x) \sigma_1^2(x)}{g_1^2(x)} dG_1(x), \quad \text{in probability.}$$

Thus, to prove (4.17), it remains to prove that

$$\left| \frac{1}{n_1} \sum_{i=1}^{n_1} u^2(S_i) \left[\frac{\hat{\sigma}_1^2(S_i)}{\hat{g}_1^2(S_i)} - \frac{\sigma_1^2(S_i)}{g_1^2(S_i)} \right] \right| = o_{p_\gamma}(1). \quad (4.18)$$

By the triangle inequality, the left hand side of (4.18) is bounded above by the sum of the following two terms:

$$\begin{aligned} A_1 &:= \left| \frac{1}{n_1} \sum_{i=1}^{n_1} \frac{u^2(S_i)}{\hat{g}_1^2(S_i)} \left[\hat{\sigma}_1^2(S_i) - \sigma_1^2(S_i) \right] \right|, \\ A_2 &:= \left| \frac{1}{n_1} \sum_{i=1}^{n_1} u^2(S_i) \sigma_1^2(S_i) \left[\frac{1}{\hat{g}_1^2(S_i)} - \frac{1}{g_1^2(S_i)} \right] \right|. \end{aligned}$$

Note that, for u known and $u = u_\gamma$, $\frac{1}{n_1} \sum_{i=1}^{n_1} u^2(S_i) = O_{p_\gamma}(1)$, and by (A.2), $\inf_{x \in \mathcal{I}} g_1(x) > 0$. Hence

$$\begin{aligned} A_1 &\leq \sup_{x \in \mathcal{I}} \frac{|\hat{\sigma}_1^2(x) - \sigma_1^2(x)|}{\hat{g}_1^2(x)} \cdot \frac{1}{n_1} \sum_{i=1}^{n_1} u^2(S_i) \\ &= o_{p_\gamma}(1) O_{p_\gamma}(1) = o_{p_\gamma}(1), \end{aligned}$$

by uniform consistency of $\hat{g}_1$ and (4.11). Similarly, because $\sigma_1^2(S) < 1, \forall S$,

$$\begin{aligned} A_2 &\leq \sup_{x \in \mathcal{I}} \frac{|\hat{g}_1^2(x) - g_1^2(x)|}{\hat{g}_1^2(x)g_1^2(x)} \cdot \frac{1}{n_1} \sum_{i=1}^{n_1} u^2(S_i) \\ &= o_{p_\gamma}(1)O_{p_\gamma}(1) = o_{p_\gamma}(1), \end{aligned}$$

by uniform consistency of $\hat{g}_1$ and (A.2). This completes the proof of the first claim in (4.16). The proof of the second claim in (4.16) is similar, thereby completing the proof of Lemma 4.2.4.

Remark 4.2.5 Estimation of ν_k . *Estimation of ν_k when u is known.* Assume (A.1), (A.2) and (A.4) hold. Then ν_k can be estimated by $\sqrt{u}/\hat{g}_k$ with h_k as mentioned in Lemma 4.2.2. We shall now show that these estimates satisfy the assumptions of Lemma 4.2.1 and hence Assumption 4.2.2. It follows from (A.2) that $g_k(t) > 4\beta$ for all $t \in \mathcal{I}$ and for some $\beta > 0$. Thus, by (A.4), $\bar{g}_k(t) > 2\beta$ for all $t \in \mathcal{I}$. In view of (4.9), $\tilde{\nu}_k = \sqrt{u}/(\hat{g}_k \vee \beta)$ is a modification of $\hat{\nu}_k$. We then obtain (4.3) from the boundedness of u , while (4.4) and (4.5) follows from (4.9) and (4.10). Of course, (4.6) and (4.7) holds as $\tilde{\nu}_k$ does not depend on δ and η .

Estimation of ν_k when $u = u_\gamma$. Here we shall discuss the estimation of $\nu_k = \sqrt{u_\gamma}/g_k$, where γ is a known non negative continuous function. In view of (4.8), an obvious estimate of u_γ is

$$\hat{u}_\gamma = \frac{\gamma I_{\mathcal{I}}}{q_1 \hat{\psi}_1 + q_2 \hat{\psi}_2}, \text{ where } \hat{\psi}_k = \frac{\hat{\sigma}_k^2}{\hat{g}_k}, k = 1, 2.$$

Similar to K-Sh, we can easily verify the assumptions of Lemma 4.2.1 for $\nu_k = \sqrt{u_\gamma}/g_k$ by using Lemma 4.2.2 and 4.2.3.

4.3 Simulations

This section examines the Monte Carlo comparison of the test statistics $\hat{T}$ and CV_2 based on 10,000 replications. For simplicity we took $\mathcal{I} = (0, 5)$ and $u(x) = 1_{\mathcal{I}}$. The simulations are done using Matlab. The kernel function used for w , g_1 and g_2 in the simulations is $\frac{3}{4}(1 - x^2)I(|x| \leq 1)$. Let c_1 be the bandwidth used for w . Similar to K-Sh, the values chosen for c_1 are 0.2 and 0.25. Also the bandwidths used for densities g_1 , g_2 in the simulations are $h_1 = h_2 = c_2(\log(n)/n)^{1/5}$. In the tables below, $exp(\lambda)$ denote the exponential distribution with parameter λ and $wei(a, b)$ represents the weibull distribution with density $w(t) := ba^{-b}t^{b-1}exp(-t/a)^b$. The asymptotic level is taken to be 0.05 in all the cases. We used $\hat{\tau}^2$ of (4.12) to compute $\hat{T}$. The entries in the table for $\hat{T}$ test statistic are obtained by computing the number of $(|\hat{T}| \geq 1.96)/10,000$.

Since the asymptotic distribution of CV_2 is not known, so in order to find the Monte Carlo levels and the Monte Carlo powers of this test, we need to estimate its cut off points. Estimated cut off points are obtained by first getting 10,000 values of CV_2 and then by finding percentiles from the distribution of these 10,000 values. After that, for CV_2 , the significance levels and powers are obtained by computing the number of $(CV_2 \geq \text{estimated cut off point})/10,000$.

Table 4.1 summarizes the empirical levels for test statistic $\hat{T}$ when sample sizes for both the populations are the same with chosen values of c_1 and c_2 . The sample sizes chosen here are 50, 100 and 200. It shows that as the sample size increases the simulated levels are getting closer to the asymptotic level 0.05.

Table 4.2 represents the empirical levels for test statistic $\hat{T}$ when sample sizes for the two populations are not the same for all the chosen values of c_1 and c_2 and chosen inspection time densities. It shows that the simulated levels are consistent with the asymptotic theory when sample sizes are not the same for the two populations.

Table 4.3 shows the simulated power of $\hat{T}$ for six different alternatives and chosen values of c_1 and c_2 when sample size for both the populations is 50. It shows that the power is getting better as the parameter of exponential distribution increases.

Table 4.4 represents the simulated 95th, 97.5th, 99th, 99.5th and 90th percentiles of CV_2 for sample sizes 40, 80, 100, 200 when distribution of X, Y is $\exp(1)$ and distribution of S, T is $\exp(1.5)$. Table 4.5 represents the simulated significance level by using the corresponding simulated percentiles given in table 4.4 for sample sizes 40, 80, 100, 200 when distribution of X, Y is $\exp(1)$ and distribution of S, T is $\exp(1.5)$. It shows that the simulated significance levels of CV_2 for different chosen sample sizes are very close to the true nominal size.

Table 4.6 represents the simulated 95th percentile of CV_2 for sample sizes 50, 100, 200 and for all the chosen inspection time densities.

Table 4.7 shows comparison of simulated significance levels for $\hat{T}$ and CV_2 for different inspection time densities and different sample sizes. For the simulated significance levels of CV_2 we used the percentiles given in Table 4.6. It shows that the empirical levels of statistics CV_2 is better than $\hat{T}$, when sample size is small. But when sample size is large, then the results of $\hat{T}$ and CV_2 are close to each other.

Table 4.8 represents the comparison of power between $\hat{T}$ and CV_2 for different chosen alternatives and sample sizes. For the power of CV_2 we used the percentiles

given in Table 4.6. It shows that the power of statistics $\hat{T}$ is better than CV_2 , when sample size is small. But when sample size is large, then power of statistics $\hat{T}$ and CV_2 is comparable to each other.

Table 4.1: Empirical sizes of $\hat{T}$, $X, Y \sim \exp(1)$, $S, T \sim \exp(1)$

c_1, c_2	$n_1=n_2=50$	$n_1=n_2=100$	$n_1=n_2=200$
0.2, 0.6	0.1151	0.0852	0.062
0.2, 0.9	0.1112	0.0812	0.0564
0.25, 0.8	0.0763	0.0698	0.0585
0.25, 0.9	0.0843	0.0655	0.0595
0.25, 1	0.1016	0.086	0.052

Table 4.2: Empirical sizes of $\hat{T}$, $X, Y \sim \exp(1)$, $(n_1, n_2) = (180, 200)$

c_1, c_2	$S, T := \exp(1.5)$	$S, T := \exp(1)$	$S := \exp(1), T := \exp(1.5)$
0.2, 0.6	0.061	0.0589	0.062
0.2, 0.9	0.056	0.0551	0.0573
0.25, 0.8	0.0573	0.058	0.0549
0.25, 0.9	0.0598	0.0586	0.059
0.25, 1	0.0535	0.0522	0.0560

Table 4.3: Power of $\hat{T}$, $S, T \sim \exp(1)$, $X \sim \exp(1)$, $n_1 = n_2 = 50$

$c_1, c_2 \backslash Y$	exp(.5)	exp(1.5)	exp(2)	exp(3)	exp(4)	exp(5)
0.2, 0.6	0.5381	0.2779	0.5762	0.9095	0.9634	0.9916
0.2, 0.9	0.4875	0.2767	0.5454	0.8891	0.9688	0.9940
0.25, 0.6	0.5253	0.2684	0.5960	0.9197	0.9772	0.9965
0.25, 0.8	0.5213	0.2777	0.5687	0.8980	0.9800	0.9913
0.25, 0.9	0.5227	0.2739	0.5900	0.8732	0.9761	0.9947
0.25, 1	0.5040	0.2645	0.5459	0.8943	0.9676	0.9848

Table 4.4: Simulated percentiles of CV_2 , $X, Y \sim \exp(1)$, $S, T \sim \exp(1.5)$

Percentile $\backslash n_1 = n_2$	40	80	100	200
99.5	0.1896	0.108	0.0866	0.0466
99	0.1433	0.0934	0.0755	0.041
97.5	0.1413	0.0787	0.0649	0.0318
95	0.1189	0.0671	0.0563	0.0321
90	0.0974	0.0571	0.0471	0.0275

Table 4.5: Empirical sizes of CV_2 , $X, Y \sim \exp(1)$, $S, T \sim \exp(1.5)$

True level $\backslash n_1 = n_2$	40	80	100	200
0.005	0.00498	0.0053	0.0051	0.0049
0.01	0.0099	0.0105	0.011	0.0121
0.025	0.02456	0.0254	0.0255	0.0249
0.05	0.05	0.0502	0.0501	0.0510
0.1	0.1022	0.1015	0.1014	0.0998

Table 4.6: Simulated 95th percentile of CV_2 , $X, Y \sim \exp(1)$.

Dist. of S, T	$n_1=n_2=50$	$n_1=n_2=100$	$n_1=n_2=200$
$\exp(1)$	0.0999	0.0551	0.031
$\exp(1.5)$	0.1008	0.0563	0.0321
$\exp(1), \exp(1.5)$	0.1011	0.0556	0.0311

Table 4.7: Empirical sizes, $X, Y \sim \exp(1)$, $(c_1, c_2) = (.25, 1)$.

S, T	$\exp(1.5)$		$\exp(1)$		$\exp(1), \exp(1.5)$	
$n_1 = n_2$	$\hat{T}$	CV	$\hat{T}$	CV	$\hat{T}$	CV
50	0.1013	0.0510	0.0965	0.0486	0.1115	0.0480
100	0.853	0.0497	0.0729	0.0494	0.0876	0.0504
200	0.0521	0.0482	0.0559	0.0466	0.058	0.0465

Table 4.8: Power, $S, T \sim \exp(1)$, $X \sim \exp(1)$, $(c_1, c_2) = (.2, .9)$.

$n_1 = n_2$	50		100		200	
Dist. of Y	$\hat{T}$	CV	$\hat{T}$	CV	$\hat{T}$	CV
$\exp(0.5)$	0.5016	0.2845	0.7067	0.5328	0.8941	0.8465
$\exp(1.5)$	0.2677	0.1478	0.4136	0.2365	0.6511	0.4085
$\exp(2)$	0.5624	0.3606	0.8223	0.6200	0.9756	0.8699
$\exp(3)$	0.7976	0.7209	0.9812	0.9429	1	1
$w(.2, 1)$	0.9468	0.9488	0.9997	0.9993	1	1
$w(.5, 1)$	0.5487	0.3876	0.8281	0.6586	0.9804	0.8999
$w(1.5, 1)$	0.2138	0.1298	0.3438	0.2129	0.5120	0.3922
$w(2, 1)$	0.4445	0.3086	0.6568	0.5418	0.8673	0.8488

BIBLIOGRAPHY

- [1] Ayer, M.; Brunk, H.D.; Ewing, G.M.; Reid, W.T.; Silverman, E. (1955). An empirical distribution function for sampling with incomplete information. *Ann. Math. Statist.* **26**, 641-647.
- [2] Beran, R.J. (1977). Minimum Hellinger distance estimates for parametric models. *Ann. Statist.* **5**, 445-463.
- [3] Carroll, R.J.; Hall, P. (1992). Semiparametric comparison of regression curves via normal likelihoods. *Austral. J. Statist.* **34**, 471-487.
- [4] Delgado, M.A. (1993). Testing the equality of nonparametric regression curves. *Statist. Probab. Lett.* **17**, 199-204.
- [5] Diamond, I.D.; Mcdonald, J. W.; Shah, I. H. (1986). Proportional hazards models for current status data: application to the study of differentials in age at weaning in Pakistan *Demography*, **23**, 607-620.
- [6] Diamond, I.D.; Mcdonald, J. W. (1991). Analysis of current status data, In *Demographic Applications of Event History Analysis* (J.Trussell, R. Hankinson and J.Tilton, Eds.), 231-252. Oxford Univ. Press.
- [7] Finkelstein, D.M.; Wolfe, R.A. (1985). A semiparametric model for regression analysis of interval-censored failure time data. *Biometrics*, **41**, 933-945.
- [8] Finkelstein, D.M. (1986). A proportional hazards model for interval-censored failure time data. *Biometrics*, **42**, 845-854.
- [9] Groeneboom, P.; Wellner, J. A. (1992). *Information bounds and nonparametric maximum likelihood estimation*. DMV Seminar, 19, Birkhauser Verlag , Basel.
- [10] Hall, P. (1984). Central limit theorem for integrated square error of multivariate nonparametric density estimators. *J. Mult. Analysis.* **14**, 1-16.

- [11] Hall, P.; Hart, J. D. (1990). Bootstrap test for difference between means in nonparametric regression. *J. Amer. Statist. Assoc.* **85**, 1039-1049.
- [12] Hart, J.D. (1997). *Nonparametric smoothing and lack-of-fit tests*. Springer-Verlag, New York, Inc.
- [13] Hoel, D. G.; Walburg; H. E. (1972). Statistical analysis of survival experiment. *J. National Cancer Institute*, **49**, 361-372.
- [14] Jewell, N. P.; Van der Laan, M. (2004). Current status data : review, recent developments and open problems. *Advances in survival analysis*, 625-642, Handbook of Statist., 23, Elsevier, Amsterdam.
- [15] Keiding, N. (1991). Age-specific incidence and prevalence: A statistical perspective (with discussion) *J.Roy. Statist. Soc. Ser. A.*, **154**, 371-412.
- [16] Khmaladze, E .V.(1981) Martingale approach in the theory of goodness-of-fit tests. *Theor. Probability Appl.* **26**, 240-257.
- [17] King, E; Hart, J.D.; Wehrly, T.E. (1991). Testing the equality of two regression curves using linear smoothers. *Statist. Probab. Lett.* **12**, 239-247.
- [18] Koul, H. L.; Ni, P. (2004). Minimum distance regression model checking, *J. Stat. Plann. Inference*, **119**, No.1, 109-141.
- [19] Koul, H. L.; Schick, A. (1997). Testing for the equality of two nonparametric regression curves. *J. Statist. Plann.Inference.* **65**, 293-314.
- [20] Koul, H. L.; Schick, A. (2003). Testing for superiority among two regression curves. *J. Statist. Plann. Inference.* **117**, 15-33.
- [21] Koul, H. L.; Song, W. (2008) Minimum distance regression model checking with Berkson measurement errors. To appear in *Ann. Math. Statist.*
- [22] Koul, H. L.; Yi, T. (2006). Goodness-of-fit testing in interval censoring case 1. *Statist. Probab. Letters.* **76**, 709-718.
- [23] Kulasekera, K. B. (1995). Comparison of regression curves using quasi-residuals. *J. Amer. Statist. Assoc.* **431**, 1085-1093.
- [24] Mack, Y.P.; Silverman, B.W. (1982). Weak and strong uniform consistency of kernel regression estimates, *Z. Wahrsch. Gebiete*, **61**, 405-415.

- [25] Neumeyer, N.; Dette, H. (2003). Nonparametric comparison of regression curves: an empirical process approach. *Ann. Statist.* **31**, 880-920.
- [26] Ni, P. (2002). Minimum distance regression and autoregressive model fitting. In *Thesis*. The Department of Statistics and Probability, Michigan State University.
- [27] Shen, X. (2000). Linear regression with current status data. *J. Amer. Statist. Assoc.* **451**, 842-852.
- [28] Stute, W.; Thies, S.; Zhu, L.X. (1998). Model checks for regression: an innovation process approach. *Ann. Statist.* **26**, 1916-1934.