SADDLE POINT APPROXIMATION

BY JIXIN LI

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Statistics

Written under the direction of

John Kolassa

and approved by

New Brunswick, New Jersey

January, 2009

ABSTRACT OF THE DISSERTATION

Saddle Point Approximation

by Jixin Li

Dissertation Director: John Kolassa

We extend known saddlepoint tail probability approximations to multivariate cases, including multivariate conditional cases. Our first approximation applies to both continuous and lattice variables, and requires the existence of a cumulant generating function. The method is applied to some examples, including a real data set from a case-control study of endometrial cancer. The method contains less terms, is easier to implement than existing methods, and shows an accuracy comparable to that of existing methods. The drawback of the first method is that the coefficient for the main term is not 1, and therefore it may be hard to show the reflexivity property which in general does not hold, because the route of path of the integral used in the saddlepoint method has to have positive real part. Our second method uses a different approach for the main term. We show that in the bivariate case, the reflexivity property holds. We applies the method to a three dimensional example, and our method demonstrates better accuracy than the normal approximation.

Acknowledgements

I would like to express my deepest gratitude to my advisor, Professor John Kolassa. The thesis could not have been written without his enduring patience, enlightening advice and warm encouragement. He not only provides me with helpful insights into our research subject, but also teaches me the spirit to overcome difficulties and how to face challenges arising from life. I would like to express my cordial thanks to Professor Donald Hoover, who had great influence on my choice of statistics as my PhD area, and help me gain hands-on experience through external projects. I thank Professor Minge Xie and Professor Pamela Ohman Strickland, together with Professor Hoover and Professor Kolassa for serving as my committee members.

I would also like to thank my wife, Minya, for her wholehearted support and care, in research and in life.

Table of Contents

Chapter 1

Introduction

Let $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n}$ be independent and identically distributed random vectors from a density $f_{\mathbf{X}}(\cdot)$ on $\mathbf{R}^d$. We construct two accurate multivariate saddlepoint approximation to the tail probability of the mean random vector $\bar{\mathbf{X}} = (\mathbf{X_1} + \mathbf{X_2} + \cdots + \mathbf{X_n})/n$. We also develop similar approximations to conditional tail probabilities. The first approximation has a relative error of $O(n^{-1})$ uniformly over a compact set of $\bar{\mathbf{x}}$, a realization of $\bar{\mathbf{X}}$, under some general conditions. The second approximation removes the leading coefficient of the main term and therefore has the reflexive property. Our methods utilizes the likelihood ratio statistic, routinely calculated by standard software, which makes the approximation easy to implement. I will also investigate applications of this approximation.

The Edgeworth expansion is a natural competitor to the saddlepoint approximation. This expansion has a uniformly bounded absolute error and works well in the center of the distribution being approximated. However, the approximation deteriorates at the far tail of the distribution, where it can sometimes even attain negative values. [Daniels 1954] first applied saddlepoint techniques to the approximation of a probability density function. Saddlepoint approximation addresses the problem of degradation outside a region of radius $O(n^{-\frac{1}{2}})$ about $E(\mathbf{X}_i)$, by bounding the relative error, rather than the absolute error of the approximation over the admissible range.

Daniels ([Daniels 1954]) discussed approximating the density of $\bar{X}$ when the dimension $d = 1$, i.e. the univariate case. The approximation achieved a relative error of $O(n^{-1})$ uniformly over the whole admissible range of the variable, under some conditions. The method uses the Fourier inversion formula, which involves moment generating or characteristic functions, and complex integration. In this approach, the path of integration is shifted, so that it passes through the saddlepoint of the integrand and follows the steepest descent curve at the neighborhood of the saddlepoint. The asymptotic property is proved by the lemma due to [Watson 1948].

Extensions of univariate saddlepoint approximation to tail probabilities $P(\bar{X} > \bar{x})$ for the mean of independent random variables have also been studied. This calculation is more difficult in that, unlike the density function case, the integrand of the Fourier inversion integral for tail probabilities has a pole at zero.

Robinson ([Robinson 1982]) presented a general saddlepoint approximation technique that can be applied to tail probability approximation, based on Laplace approximation to the integrated saddlepoint density, with an error of $O(n^{-1})$. Robinson used an argument involving an conjugate exponentially shifted distribution family and the Edgeworth expansion. The terms of the expansion then can be integrated termwise. There is no direct explicit formula for the integration of each term, but the terms may be computed recursively. This method applies when $\bar{x} \geq E(X)$. When $\bar{x} < E(X)$, Boole's law and reflection of the distribution must be used.

Lugannani and Rice ([Lugannani and Rice 1980]) provided an alternative approximation. Daniels ([Daniels 1987]) derived this technique, using a transformation of variables to directly address the local quadratic behavior of the numerator. The integral then is split into two parts, one that contains a pole but can be integrated exactly and explicitly, and the other one that only has removable singularities and can be expanded and approximated accurately. The virtue of this method is that the approximation is compact and can be computed without recursion, and the formula is valid over the whole range of admissible $\bar{x}$.

Reid ([Reid 1988]) thoroughly discussed the usefulness of saddlepoint method, in a review of the saddlepoint method focusing on a variety of applications to statistical inference.

Kolassa ([Kolassa 2003]) generalized the univariate Robinson approach under the Daniels framework and achieved an error of size $O(n^{-1})$. The method uses integral expressions for the tail probability in the multivariate case and presents a multivariate expansion of the numerator of the integrand and a termwise multivariate integration using recursion. This approach shares the drawback of Robinson's approach in that it required a positivity constraint on the ordinate.

Wang ([Wang 1991]) generalized Lugannani & Rice's method to the case of bivariate probability distribution function using variable transformations. As summarized in [Kolassa 2003], he used a different method of proof, and showed that the error term is of order $O(n^{-1})$. His method is limited to $d = 2$. Furthermore, Wang's development involves an inversion integral in which the pole of one variable depends on the values of other variables, and in general the problem can not be solved by a simple linear transformation.

Wang's proof of the error rate in the neighborhood of the pole is incomplete. In this paper, a way of effectively extending the Lugannani & Rice's method to multivariate case, which uses a different transformation formula from Wang's and can be used in the case $d > 2$, is proposed. The method uses fewer terms and is extended to multivariate conditional cases.

My saddlepoint approximation may be used to test null and alternative hypothesis concerning a multivariate parameter, when the hypotheses are specified by systems of linear inequalities. Kolassa ([Kolassa 2004]) applied the method of [Kolassa 2003], in conjunction with the adjusted profile likelihood, in such a case. For instance, Kolassa ([Kolassa 2004]) refers to data presented by [Stokes *et al.* 1995] on 63 case-control pairs of women with endometrial cancer. The occurrence of endometrial cancer is influenced by explanatory variables including gall bladder disease, hypertension and non-estrogen drug use. The test of whether hypertension or non-estrogen drug use is associated with an increase in endometrial cancer will be performed conditional on the sufficient statistic value associated with gall bladder disease. This type of inferential problem will be discussed in this thesis.

The remainder of the paper is organized as follows. Section 2 provides the unified framework, under which both unconditional and conditional tail probability approximations are considered. Section 3 derives formulas for unconditional bivariate distributions. Section 4 focuses on conditional distributions. Section 5 presents an alternative multivariate saddlepoint approximation.

Chapter 2

Representations of tail probabilities as multiple complex integrals, and integral decompositions

The unconditional and conditional tail probability approximation share some common characteristics. I derive them in a unified way. Applying the Fourier inversion theorem and Fubini's theorem as in [Kolassa 2003], we find that both the unconditional and conditional tail probability approximations require the evaluation of an integral of form

$$\frac{n^{d-d_0}}{(2\pi i)^d}\int_{\mathbf{c}-i\mathbf{K}}^{\mathbf{c}+i\mathbf{K}} \frac{\exp(n[K(\boldsymbol{\tau})-\boldsymbol{\tau}^T\mathbf{t}^*])}{\prod_{j=1}^{d_0}\rho(\tau_j)}\, d\boldsymbol{\tau}, \tag{2.0.1}$$

where K is the cumulant generating function, which is the natural logarithm of the moment generating function, and $\mathbf{c}$ is any positive d dimensional vector. In the unconditional case, for continuous variables, $\mathbf{K}$ is a vector of length d, with every entry infinity, $\mathbf{t}^* = \mathbf{t}$, $\rho(\tau) = \tau$, and for variables confined to unit lattice, $\mathbf{K}$ is a vector of length d, with every entry π, $\mathbf{t}^*$ is $\mathbf{t}$ corrected for continuity, $\rho(\tau) = 2\sinh(\tau/2)$, and $d = d_0$. In the conditional case, the setting is same, except that d_0 equals d minus the dimension of the conditioning variables.

Daniels ([Daniels 1987]) recast a great deal of the saddlepoint literature in terms of inversion integrals of form (2.0.1), rescaled so that the exponent is exactly quadratic. This rescaling includes the multiplier for the linear term in the exponent; this linear term is the signed root of the likelihood ratio statistic. Kolassa ([Kolassa 1997]) defines a multivariate version of this reparameterization, and also defines the multiplier for the linear terms; again these are signed roots of likelihood ratio statistics, but this time for a sequence of nested models.

$$-\frac{1}{2}\hat{\mathbf{w}}^T\hat{\mathbf{w}} = \min_{\boldsymbol{\gamma}}(K(\boldsymbol{\gamma}) - \boldsymbol{\gamma}^T\mathbf{t}^*)$$

and

$$-\frac{1}{2}(\mathbf{w}-\hat{\mathbf{w}})^T(\mathbf{w}-\hat{\mathbf{w}}) = K(\boldsymbol{\tau}) - \boldsymbol{\tau}^T\mathbf{t}^* - \min_{\boldsymbol{\gamma}}(K(\boldsymbol{\gamma}) - \boldsymbol{\gamma}^T\mathbf{t}^*).$$

Further specification of $\hat{\mathbf{w}}$ and $\mathbf{w}$ is needed. For any vector $\mathbf{v}$ of length d, let $\mathbf{v}_j$ be the vector consisting of the first j elements, i.e., $(v_1, v_2, \cdots, v_d)^T$. For instance, $\boldsymbol{\gamma}_j = (\gamma_1, \gamma_2, \cdots, \gamma_j)^T$, $\boldsymbol{\tau}_j = (\tau_1, \tau_2, \cdots, \tau_j)^T$ and $\mathbf{0}_j$ is the zero vector $(0, 0, \cdots, 0)^T$ with dimension j. Let $\mathbf{v}_{-j}$ be the vector consisting all but the first j elements of $\mathbf{v}$, i.e., $(v_{j+1}, v_{j+2}, \cdots, v_d)^T$. [Kolassa 1997], Chapter 6 defines $\hat{\mathbf{w}}$ and $\mathbf{w}$ using:

$$-\frac{1}{2}\hat{w}_j^2 = \min_{\boldsymbol{\gamma},\boldsymbol{\gamma}_{j-1}=\mathbf{0}_{j-1}}(K(\boldsymbol{\gamma}) - \boldsymbol{\gamma}^T\mathbf{t}^*) - \min_{\boldsymbol{\gamma},\boldsymbol{\gamma}_j=\mathbf{0}_j}(K(\boldsymbol{\gamma}) - \boldsymbol{\gamma}^T\mathbf{t}^*) \tag{2.0.2a}$$

$$-\frac{1}{2}(w_j - \hat{w}_j)^2 = \min_{\boldsymbol{\gamma},\boldsymbol{\gamma}_{j-1}=\boldsymbol{\tau}_{j-1}}(K(\boldsymbol{\gamma}) - \boldsymbol{\gamma}^T\mathbf{t}^*) - \min_{\boldsymbol{\gamma},\boldsymbol{\gamma}_j=\boldsymbol{\tau}_j}(K(\boldsymbol{\gamma}) - \boldsymbol{\gamma}^T\mathbf{t}^*). \tag{2.0.2b}$$

This definition is not symmetric with regard to the order of the coordinates. Also note that w_j is a function of only $\boldsymbol{\tau}_j$, but not of any element of $\boldsymbol{\tau}_{-j}$, $\forall j$. The same holds true for τ_j as a function of $\mathbf{w}$.

We now construct more explicit formulas for $\hat{\mathbf{w}}$ and $\mathbf{w}$. Let

$$\tilde{\boldsymbol{\tau}}_j(\boldsymbol{\gamma}_j) = (\gamma_1, \gamma_2, \cdots, \gamma_j, \tilde{\tau}_{j+1}(\boldsymbol{\gamma}_j), \tilde{\tau}_{j+2}(\boldsymbol{\gamma}_j), \cdots, \tilde{\tau}_d(\boldsymbol{\gamma}_j))$$

be the minimizer of $(K(\boldsymbol{\gamma}) - \boldsymbol{\gamma}^T\mathbf{t}^*)$ when the first j variables are fixed. The function $\tilde{\tau}_k(\boldsymbol{\gamma}_j)$ above is the minimizer for variable k when the first j variables are fixed, for $k > j$.

Using the notation above, the definition of $\hat{\mathbf{w}}$ and $\mathbf{w}$ can be rewritten as

$$-\frac{1}{2}\hat{w}_j^2 = K(\tilde{\boldsymbol{\tau}}_{j-1}(\mathbf{0}_{j-1})) - \tilde{\boldsymbol{\tau}}_{j-1}(\mathbf{0}_{j-1})^T\mathbf{t}^* - (K(\tilde{\boldsymbol{\tau}}_j(\mathbf{0}_j)) - \tilde{\boldsymbol{\tau}}_j(\mathbf{0}_j)^T\mathbf{t}^*) \tag{2.0.3a}$$

$$-\frac{1}{2}(w_j - \hat{w}_j)^2 = K(\tilde{\boldsymbol{\tau}}_{j-1}(\boldsymbol{\tau}_{j-1})) - \tilde{\boldsymbol{\tau}}_{j-1}(\boldsymbol{\tau}_{j-1})^T\mathbf{t}^* - (K(\tilde{\boldsymbol{\tau}}_j(\boldsymbol{\tau}_j)) - \tilde{\boldsymbol{\tau}}_j(\boldsymbol{\tau}_j)^T\mathbf{t}^*), \tag{2.0.3b}$$

where $\tilde{\boldsymbol{\tau}}_{j-1}(\cdot)$ is set to $\hat{\boldsymbol{\tau}}$ when $j = 1$ for expression succinctness.

By choosing a sign to make $\hat{\mathbf{w}}$ and $\mathbf{w}$ increasing functions of $\hat{\boldsymbol{\tau}}$ and $\boldsymbol{\tau}$ respectively, we can further specify them as below:

$$\hat{w}_j = \text{sign}(\tilde{\tau}_j(\mathbf{0}_{j-1}))\cdot \sqrt{-2[K(\tilde{\boldsymbol{\tau}}_{j-1}(\mathbf{0}_{j-1})) - \tilde{\boldsymbol{\tau}}_{j-1}(\mathbf{0}_{j-1})^T\mathbf{t}^* - (K(\tilde{\boldsymbol{\tau}}_j(\mathbf{0}_j)) - \tilde{\boldsymbol{\tau}}_j(\mathbf{0}_j)^T\mathbf{t}^*)]} \tag{2.0.4a}$$

$$w_j = \hat{w}_j + \text{sign}(\tau_j - \tilde{\tau}_j(\boldsymbol{\tau}_{j-1}))\cdot \sqrt{-2[K(\tilde{\boldsymbol{\tau}}_{j-1}(\boldsymbol{\tau}_{j-1})) - \tilde{\boldsymbol{\tau}}_{j-1}(\boldsymbol{\tau}_{j-1})^T\mathbf{t}^* - (K(\tilde{\boldsymbol{\tau}}_j(\boldsymbol{\tau}_j)) - \tilde{\boldsymbol{\tau}}_j(\boldsymbol{\tau}_j)^T\mathbf{t}^*)]}. \tag{2.0.4b}$$

The derivation of the [Lugannani and Rice 1980] approximation provided by Daniels ([Daniels 1987]) requires identification of the simple pole in the inversion integrand. We need to match zeros in the denominator of the multivariate integrand with functions of the variables in the new parameterization; the points at which this matching occurs will be denoted by a tilde. The quantities above, such as $\hat{\boldsymbol{\tau}}$, $\hat{\mathbf{w}}$, $\tilde{\tau}_j(\boldsymbol{\tau}_{j-1})$ and functional relationships between $\boldsymbol{\tau}$ and $\mathbf{w}$, etc., can be solved numerically by Newton-Raphson methods, or even analytically in some cases. Finally, we define a function $\tilde{w}_j(\mathbf{w}_{j-1})$, such that $\tau_j(w_1, w_2, \cdots, \tilde{w}_j(\mathbf{w}_{j-1})) = 0$, for $j > 1$.

It can be verified that the following properties hold:

$$\boldsymbol{\tau}_j = 0 \text{ if and only if } \mathbf{w}_j = 0, \tag{2.0.5a}$$

$$\tilde{w}_j(\mathbf{0}_{j-1}) = 0, \text{ for } j > 1, \tag{2.0.5b}$$

$$\tau_j = \tilde{\tau}_j(\boldsymbol{\tau}_{j-1}) \text{ if and only if } w_j = \hat{w}_j, \text{ for } j > 1, \tag{2.0.5c}$$

$$\boldsymbol{\tau}_j = \hat{\boldsymbol{\tau}}_j \text{ if and only if } \mathbf{w}_j = \hat{\mathbf{w}}_j. \tag{2.0.5d}$$

Below a superscript of a function denotes differentiation with respect to the corresponding argument of the function. T denotes transpose of matrix. In other cases, we will follow the same use of superscripts in the subsequent text of the paper except that when the superscript is a set, it denotes difference as defined at the end of this section. Furthermore, let $\breve{w}_j = \tilde{w}_j(\hat{\mathbf{w}}_{j-1})$ and $\breve{w}_j^k = \tilde{w}_j^k(\hat{\mathbf{w}}_{j-1})$. Substitute $w_j = \breve{w}_j$, $\tau_j = 0$, $\boldsymbol{\tau}_{j-1} = \hat{\boldsymbol{\tau}}_{j-1}$ and $\boldsymbol{\tau}_j = (\hat{\tau}_1, \hat{\tau}_2, \cdots, \hat{\tau}_{j-1}, 0)^T = (\hat{\boldsymbol{\tau}}_{j-1}, 0)^T$ into (2.0.4b), to obtain

$$\breve{w}_j = \hat{w}_j + \text{sign}(0 - \hat{\tau}_j)\sqrt{-2[K(\hat{\boldsymbol{\tau}}) - \hat{\boldsymbol{\tau}}^T\mathbf{t}^* - (K(\tilde{\boldsymbol{\tau}}_j(\hat{\boldsymbol{\tau}}_{j-1}, 0)) - (\hat{\boldsymbol{\tau}}_{j-1}, 0)^T\mathbf{t}^*)]}. \tag{2.0.6}$$

Differentiate (2.0.3b) with respect to w_k, and rearrange terms, to obtain

$$\breve{w}_j^k = \frac{\sum_{l=k}^{j-1} (K^l(\tilde{\boldsymbol{\tau}}_j(\hat{\boldsymbol{\tau}}_{j-1},0)) \cdot \frac{d\tau_l}{dw_k}\Big|_{\hat{\mathbf{w}}_l} - t_l^*)}{\breve{w}_j - \hat{w}_j}, \tag{2.0.7}$$

for $k < j$. The derivatives $\frac{d\tau_l}{dw_k}$ evaluated at point $\mathbf{w}_l$ can be obtained by differentiating (2.0.3b) with respect to w_k once or twice depending on whether $w_j = \hat{w}_j$ or not, and solving the resulting equation system. In particular,

$$\frac{d\tau_j}{dw_j}\Big|_{\mathbf{w}_j} = \begin{cases} \sqrt{\frac{1}{\sum_{l=j}^d K^{jl}(\tilde{\boldsymbol{\tau}}_{j-1}(\boldsymbol{\tau}_{j-1}))\tau_l^j([\tilde{\boldsymbol{\tau}}_{j-1}(\boldsymbol{\tau}_{j-1})]_j)}} & \text{if } w_j = \hat{w}_j \\ \frac{w_j - \hat{w}_j}{K^j(\tilde{\boldsymbol{\tau}}_j(\tau_j)) - t_j^*} & \text{if } w_j \neq \hat{w}_j, \end{cases} \tag{2.0.8}$$

for $j \leq d_0$, where $[\cdot]_j$ denotes the first j elements, $\tau_l^j([\tilde{\boldsymbol{\tau}}_{j-1}(\boldsymbol{\tau}_{j-1})]_j)$ is the partial derivative with respect to the jth argument of $\tau_l(\cdot)$, and

$$\prod_{j=d_0+1}^{d} \frac{d\tau_j}{dw_j}\Bigg|_{(\mathbf{w}_{d_0}, \hat{\mathbf{w}}_{-d_0})} = \prod_{j=d_0+1}^{d} \sqrt{\frac{1}{\sum_{l=j}^d K^{jl}(\tilde{\boldsymbol{\tau}}_{d_0}(\boldsymbol{\tau}_{d_0}))\tau_l^j([\tilde{\boldsymbol{\tau}}_{d_0}(\boldsymbol{\tau}_{d_0})]_j))}}, \tag{2.0.9}$$

where for expression succinctness, we define $\tau_l^j(\cdot)$ to be one when $l = j$. For $l > j$, we obtain $\tau_l^j(\cdot)$ by differentiating both sides of the definition of $\tau_l^j(\cdot)$, i.e. $K^l(\cdot) = t_l^*$ with respect to τ_j $\forall l > j$, and solving the equation system.

Under this variable transformation from $\boldsymbol{\tau}$ to $\mathbf{w}$, the Jacobian is just the product of the diagonal terms of the Jacobian matrix, and (2.0.1) can be expressed as

$$\begin{aligned} &\frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0} \rho(\tau_j(\mathbf{w}_j))} \prod_{j=1}^{d} \frac{d\tau_j}{dw_j}\, d\mathbf{w} \\ =&\frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0} (w_j - \tilde{w}_j(\mathbf{w}_{j-1}))} \cdot \prod_{j=1}^{d} \frac{d\tau_j}{dw_j} \frac{\prod_{j=1}^{d_0} (w_j - \tilde{w}_j(\mathbf{w}_{j-1}))}{\prod_{j=1}^{d_0} \rho(\tau_j(\mathbf{w}_j))}\, d\mathbf{w} \\ \sim&\frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0} (w_j - \tilde{w}_j(\mathbf{w}_{j-1}))} G(\boldsymbol{\tau})\, d\mathbf{w}, \end{aligned} \tag{2.0.10}$$

where $G(\boldsymbol{\tau}) = \prod_{j=1}^{d_0} \left(\frac{d\tau_j}{dw_j} \frac{w_j - \tilde{w}_j(\mathbf{w}_{j-1})}{\rho(\tau_j(\mathbf{w}_j))}\right) \cdot \prod_{j=d_0+1}^{d} \frac{d\tau_j}{dw_j}\Big|_{(\mathbf{w}_{d_0}, \hat{\mathbf{w}}_{-d_0})}$, and for notational succinctness, set $\tilde{w}_j(\mathbf{w}_{j-1})$ to zero for $j = 1$. The product $\prod_{j=d_0+1}^{d} \frac{d\tau_j}{dw_j}\Big|_{(\mathbf{w}_{d_0}, \hat{\mathbf{w}}_{-d_0})}$ can be ignored for unconditional case, where $d = d_0$. For convenience later, we write $G(\boldsymbol{\tau})$ as a function $\boldsymbol{\tau}$ instead of $\mathbf{w}$. The relation $\sim$ in the last step indicates exact equality in unconditional case, where $d = d_0$, but holds with a relative error of $O(n^{-1})$ in conditional case, which we will discuss in Section 4. Hereafter, we use $\sim$ to denote approximation with a relative error of $O(n^{-\frac{1}{2}})$ to both the left hand side and the tail probability, and we use $\dot{\sim}$ ($\sim$ with a dot on it) in the case that the right hand side is an approximation with a relative error of $O(n^{-\frac{1}{2}})$ to the left hand side.

The last integral in (2.0.10) will be evaluated by splitting it into rather simple terms involving poles, and more complicated terms involving analytic functions. We can decompose (2.0.10) into 2^n terms. Let $U = \{1, 2, \cdots, d_0\}$ be the index set from integer 1 to d_0. For set $s \subset U$, define $G^s(\boldsymbol{\tau}) = G(\boldsymbol{\tau}^s)$, where the vector $\boldsymbol{\tau}^s$ is defined by

$$\tau_j^s = \begin{cases} \tau_j & \text{if } j \in s \\ 0 & \text{if } j \notin s. \end{cases}$$

For example, suppose $d_0 = 3$. Then $G^{\{1,2\}}(\boldsymbol{\tau}) = G(\tau_1, \tau_2, 0)$. Now for $t \subset U$, define $H^t = \sum_{s\subset t}(-1)^{|t-s|}G^s(\boldsymbol{\tau})$, where $|\cdot|$ denotes the cardinality, i.e. the number of elements of a set. For example, $H^{\{1,2\}} = G^{\{1,2\}}(\boldsymbol{\tau}) - G^{\{1\}}(\boldsymbol{\tau}) - G^{\{2\}}(\boldsymbol{\tau}) + G^{\emptyset}(\boldsymbol{\tau}) = G(\tau_1, \tau_2, 0) - G(\tau_1, 0, 0) - G(0, \tau_2, 0) + G(0, 0, 0)$, where $\emptyset$ denotes empty set. We conclude that $G(\boldsymbol{\tau}) = \sum_{t\subset U} H^t$. This decomposition holds by induction on d_0. Noting that $\forall s \subset U$ and $a \in s$, $H^s(\boldsymbol{\tau}^{\{a\}}) = 0$, we see that

$$\frac{H^t(\boldsymbol{\tau})}{\prod_{j\in t}(w_j - \tilde{w}_j(\mathbf{w}_{j-1}))}$$

is analytic. In other words, $|t|$ product terms in the denominator of the integrand in (2.0.10) are "absorbed" by $H^t(\boldsymbol{\tau})$, leaving the rest $(d_0 - |t|)$ product terms unabsorbed. As explained in [Kolassa 2003], each term that is absorbed contributes a relative error of $O(n^{\frac{1}{2}})$. Therefore, if we let I^t be the integral corresponding to H^t, we obtain

$$\frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0}(w_j - \tilde{w}_j(\mathbf{w}_{j-1}))} G(\boldsymbol{\tau})\, d\mathbf{w} \sim \sum_{|t|\leq 1, t\subset U} I^t. \tag{2.0.11}$$

Now we have an approximation as the sum of $d_0 + 1$ integrals as shown above. In the next chapter, we will examine each of the integrals in detail for bivariate distribution approximation and provide two examples.

Chapter 3

Bivariate distribution approximations

In the bivariate case, we consider three terms $I^{\emptyset}$, $I^{\{1\}}$ and $I^{\{2\}}$, where $I^{\emptyset}$ is the main term and $I^{\{1\}}$ and $I^{\{2\}}$ are terms of relative error $O(n^{-\frac{1}{2}})$. These terms in general can not be computed exactly, and Watson's lemma can not be applied directly. We use some techniques to circumvent the problem. We will start with continuous distributions, and then the derivation for variables confined to unit lattice are similar. We give two examples, one for continuous case, and the other for unit lattice case. In both examples, our approximation shows superior results than normal approximation and the approximation presented in [Kolassa 2003].

3.1 Continuous distributions

In the continuous bivariate case, i.e. the approximation of $P(\bar{\mathbf{X}} \geq \bar{\mathbf{x}})$, we let $\mathbf{T} = \bar{\mathbf{X}}$, $\mathbf{t} = \bar{\mathbf{x}}$ and let $\mathbf{K}$ be the vector with d components, all of them infinity, $\mathbf{t}^* = \mathbf{t}$, $\rho(\tau) = \tau$ and $d = d_0 = 2$ in (2.0.1). Then (2.0.1) becomes

$$\frac{1}{(2\pi i)^2} \int_{\mathbf{c}-i\infty}^{\mathbf{c}+i\infty} \frac{\exp(n[K(\tau_1, \tau_2) - \tau_1 \bar{x}_1 - \tau_2 \bar{x}_2])}{\tau_1 \tau_2} d\boldsymbol{\tau}. \tag{3.1.1}$$

The definition of $\hat{\mathbf{w}}$ and $\mathbf{w}$ becomes

$$-\frac{1}{2}\hat{w}_1^2 = \min_{\gamma_1, \gamma_2}(K(\gamma_1, \gamma_2) - \gamma_1 \bar{x}_1 - \gamma_2 \bar{x}_2) - \min_{\gamma_2}(K(0, \gamma_2) - \gamma_2 \bar{x}_2), \tag{3.1.2a}$$

$$-\frac{1}{2}\hat{w}_2^2 = \min_{\gamma_2}(K(0, \gamma_2) - \gamma_2 \bar{x}_2), \tag{3.1.2b}$$

$$-\frac{1}{2}(w_1 - \hat{w}_1)^2 = \min_{\gamma_1, \gamma_2}(K(\gamma_1, \gamma_2) - \gamma_1 \bar{x}_1 - \gamma_2 \bar{x}_2) - \min_{\gamma_2}(K(\tau_1, \gamma_2) - \tau_1 \bar{x}_1 - \gamma_2 \bar{x}_2), \tag{3.1.2c}$$

$$-\frac{1}{2}(w_2 - \hat{w}_2)^2 = \min_{\gamma_2}(K(\tau_1, \gamma_2) - \tau_1 \bar{x}_1 - \gamma_2 \bar{x}_2) - (K(\tau_1, \tau_2) - \tau_1 \bar{x}_1 - \tau_2 \bar{x}_2), \tag{3.1.2d}$$

which, as in (2.0.3a) and (2.0.3b), can be rewritten as

$$-\frac{1}{2}\hat{w}_1^2 = K(\hat{\tau}_1, \hat{\tau}_2) - \hat{\tau}_1\bar{x}_1 - \hat{\tau}_2\bar{x}_2 - (K(0, \tilde{\tau}_2(0)) - \tilde{\tau}_2(0)\bar{x}_2), \tag{3.1.3a}$$

$$-\frac{1}{2}\hat{w}_2^2 = K(0, \tilde{\tau}_2(0)) - \tilde{\tau}_2(0)\bar{x}_2, \tag{3.1.3b}$$

$$-\frac{1}{2}(w_1 - \hat{w}_1)^2 = K(\hat{\tau}_1, \hat{\tau}_2) - \hat{\tau}_1\bar{x}_1 - \hat{\tau}_2\bar{x}_2 - (K(\tau_1, \tilde{\tau}_2(\tau_1)) - \tau_1\bar{x}_1 - \tilde{\tau}_2(\tau_1)\bar{x}_2), \tag{3.1.3c}$$

$$-\frac{1}{2}(w_2 - \hat{w}_2)^2 = K(\tau_1, \tilde{\tau}_2(\tau_1)) - \tau_1\bar{x}_1 - \tilde{\tau}_2(\tau_1)\bar{x}_2 - (K(\tau_1, \tau_2) - \tau_1\bar{x}_1 - \tau_2\bar{x}_2), \tag{3.1.3d}$$

or more specifically, as in (2.0.4a) and (2.0.4b)

$$\hat{w}_1 = \text{sign}(\hat{\tau}_1)\sqrt{-2[K(\hat{\tau}_1, \hat{\tau}_2) - \hat{\tau}_1\bar{x}_1 - \hat{\tau}_2\bar{x}_2 - (K(0, \tilde{\tau}_2(0)) - \tilde{\tau}_2(0)\bar{x}_2)]}, \tag{3.1.4a}$$

$$\hat{w}_2 = \text{sign}(\hat{\tau}_2)\sqrt{-2[K(0, \tilde{\tau}_2(0)) - \tilde{\tau}_2(0)\bar{x}_2]}, \tag{3.1.4b}$$

$$w_1 = \hat{w}_1 + \text{sign}(\tau_1 - \hat{\tau}_1)\sqrt{-2[K(\hat{\tau}_1, \hat{\tau}_2) - \hat{\tau}_1\bar{x}_1 - \hat{\tau}_2\bar{x}_2 - (K(\tau_1, \tilde{\tau}_2(\tau_1)) - \tau_1\bar{x}_1 - \tilde{\tau}_2(\tau_1)\bar{x}_2)]}, \tag{3.1.4c}$$

$$w_2 = \hat{w}_2 + \text{sign}(\tau_2 - \tilde{\tau}_2(\tau_1))\sqrt{-2[K(\tau_1, \tilde{\tau}_2(\tau_1)) - \tau_1\bar{x}_1 - \tilde{\tau}_2(\tau_1)\bar{x}_2 - (K(\tau_1, \tau_2) - \tau_1\bar{x}_1 - \tau_2\bar{x}_2)]}. \tag{3.1.4d}$$

Properties of (2.0.5a)–(2.0.5d) in the two dimensional case as listed below hold:

$$\tau_1 = 0 \text{ if and only if } w_1 = 0 \tag{3.1.5a}$$

$$\tilde{w}_2(0) = 0 \tag{3.1.5b}$$

$$\tau_2 = \tilde{\tau}_2(\tau_1) \text{ if and only if } w_2 = \hat{w}_2 \tag{3.1.5c}$$

$$\tau_1 = \hat{\tau}_1 \text{ if and only if } w_1 = \hat{w}_1. \tag{3.1.5d}$$

We also have

$$G(\boldsymbol{\tau}) = \left(\frac{w_1}{\tau_1}\frac{d\tau_1}{dw_1}\right)\left(\frac{w_2 - \tilde{w}_2(w_1)}{\tau_2}\frac{d\tau_2}{dw_2}\right). \tag{3.1.6}$$

First of all, $G(0,0) = \lim_{\tau_1 \to 0, \tau_2 \to 0} \left(\frac{w_1}{\tau_1}\frac{d\tau_1}{dw_1}\right)\left(\frac{w_2 - \tilde{w}_2(w_1)}{\tau_2}\frac{d\tau_2}{dw_2}\right) = 1$, and

$$I^{\emptyset} = \int_{\hat{\mathbf{w}} - i\infty}^{\hat{\mathbf{w}} + i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2])}{(2\pi i)^2 w_1(w_2 - \tilde{w}_2(w_1))}\, d\mathbf{w}. \tag{3.1.7}$$

Because of the presence of $\tilde{w}_2(w_1)$ in the denominator, $I^{\emptyset}$ does not have a closed-form expression. Let $u_1 = w_1$ and $u_2 = w_2 - \tilde{w}_2(w_1)$. By changing variables, with Jacobian equal to 1, we have

$$\begin{aligned} I^{\emptyset} &= \int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{\exp(n[\frac{1}{2}u_1^2 + \frac{1}{2}(u_2 + \tilde{w}_2(u_1))^2 - \hat{w}_1 u_1 - \hat{w}_2(u_2 + \tilde{w}_2(u_1))])}{(2\pi i)^2 u_1 u_2}\, d\mathbf{u} \\ &= \int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{\exp(n[g(u_1, u_2)])}{(2\pi i)^2 u_1 u_2}\, d\mathbf{u}, \end{aligned} \tag{3.1.8}$$

where $g(u_1, u_2) = \frac{1}{2}u_1^2 + \frac{1}{2}(u_2 + \tilde{w}_2(u_1))^2 - \hat{w}_1 u_1 - \hat{w}_2(u_2 + \tilde{w}_2(u_1))$.

The integration in (3.1.7) can not be performed exactly in general; however, using the same argument as in [Kolassa 2003], we approximate it by expanding $g(u_1, u_2)$ about $(\hat{u}_1, \hat{u}_2)$ up to the third degree; after termwise integration, the resulting approximation to $I^{\emptyset}$ has relative error $O(n^{-1})$. So $I^{\emptyset}$ can be approximated by

$$\begin{aligned} I^{\emptyset} = &\int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{1}{(2\pi i)^2 u_1 u_2} \exp(n[\hat{g} + \hat{g}^1(u_1 - \hat{u}_1) + \hat{g}^2(u_2 - \hat{u}_2) + \\ &\frac{1}{2}\hat{g}^{11}(u_1 - \hat{u}_1)^2 + \frac{1}{2}\hat{g}^{22}(u_2 - \hat{u}_2)^2 + \hat{g}^{12}(u_1 - \hat{u}_1)(u_2 - \hat{u}_2)]) \cdot \\ &\left(1 + \frac{n}{6} \sum_{i,j,k \in \{1,2\}} \hat{g}^{ijk}(u_i - \hat{u}_i)(u_j - \hat{u}_j)(u_k - \hat{u}_k)\right) d\mathbf{u} \\ = &\exp(n[-\frac{1}{2}\hat{w}_1^2 - \frac{1}{2}\hat{w}_2^2]) \int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{1}{(2\pi i)^2 u_1 u_2} \exp(n[\frac{1}{2}(1 + (\breve{w}_2')^2)(u_1 - \hat{u}_1)^2 + \\ &\frac{1}{2}(u_2 - \hat{u}_2)^2 + \breve{w}_2'(u_1 - \hat{u}_1)(u_2 - \hat{u}_2)])\, d\mathbf{u} + \\ &\int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{1}{(2\pi i)^2 u_1 u_2} \exp(n[\hat{g} + \hat{g}^1(u_1 - \hat{u}_1) + \hat{g}^2(u_2 - \hat{u}_2) + \\ &\frac{1}{2}\hat{g}^{11}(u_1 - \hat{u}_1)^2 + \frac{1}{2}\hat{g}^{22}(u_2 - \hat{u}_2)^2 + \hat{g}^{12}(u_1 - \hat{u}_1)(u_2 - \hat{u}_2)]) \cdot \\ &\frac{n}{6} \sum_{i,j,k \in \{1,2\}} \hat{g}^{ijk}(u_i - \hat{u}_i)(u_j - \hat{u}_j)(u_k - \hat{u}_k)\, d\mathbf{u}, \end{aligned} \tag{3.1.9}$$

where, for brevity, we write $\hat{g}^r$ for $g^r(\hat{u}_1, \hat{u}_2)$. The computation of the second integral is addressed in [Kolassa 2003]. The details involve partial derivatives of some functions up to the second or third degree, which are algebraically complicated and therefore omitted here. For the first integral, rearrange the terms in the numerator in the order of degree of $\mathbf{u}$. Expansion (3.1.9) can be written as

$$\begin{aligned} & C\int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{1}{(2\pi i)^2 u_1 u_2}\exp(n[\frac{1}{2}(1+(\breve{w}_2')^2)u_1^2+\frac{1}{2}u_2^2+\breve{w}_2' u_1 u_2 - \\ & \quad ((1+(\breve{w}_2')^2)\hat{u}_1+\breve{w}_2'\hat{u}_2)u_1-(\breve{w}_2'\hat{u}_1+\hat{u}_2)u_2])\,d\mathbf{u} \\ & = C\bar{\Phi}(\frac{\sqrt{n}[(1+(\breve{w}_2')^2)\hat{w}_1+\breve{w}_2'(\hat{w}_2-\breve{w}_2)]}{\sqrt{1+(\breve{w}_2')^2}}, \sqrt{n}[\breve{w}_2'\hat{w}_1+\hat{w}_2-\breve{w}_2], \frac{\breve{w}_2'}{1+(\breve{w}_2')^2}), \end{aligned} \tag{3.1.10}$$

where

$$C=\exp(n[(\breve{w}_2-\breve{w}_2'\hat{w}_1)(\frac{1}{2}\breve{w}_2-\frac{1}{2}\breve{w}_2'\hat{w}_1-\hat{w}_2)]), \tag{3.1.11}$$

and $\bar{\Phi}(\cdot,\cdot,\rho)$ is the tail probability of a bivariate normal distribution with means 0, variances 1 and correlation coefficient ρ. The quantity $\hat{\mathbf{w}}$, $\breve{w}_2$, $\breve{w}_2'$ can be computed using (2.0.4a), (2.0.4b), (2.0.6), (2.0.7) and (2.0.8).

From (2.0.11) we have

$$I^{\{2\}}=\int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty}\frac{\exp(n[\frac{1}{2}w_1^2+\frac{1}{2}w_2^2-w_1\hat{w}_1-w_2\hat{w}_2])}{(2\pi i)^2 w_1}\left(\frac{G(0,\tau_2)-G(0,0)}{w_2-\tilde{w}_2(w_1)}\right)d\mathbf{w}. \tag{3.1.12}$$

The function $H(w_1,w_2)=\frac{G(0,\tau_2)-G(0,0)}{w_2-\tilde{w}_2(w_1)}$ is analytic, since $H(w_1,w_2)=\frac{G(0,\tau_2)-G(0,0)}{\tau_2}\times\frac{\tau_2}{w_2-\tilde{w}_2(w_1)}$ and $\tau_2\to 0 \Leftrightarrow w_2-\tilde{w}_2(w_1)\to 0$ by definition. However, $\frac{H(w_1,w_2)}{w_1}$ is not analytic, and we can not use Watson's Lemma directly. We decompose $H(w_1,w_2)/w_1$ as following,

$$\frac{H(w_1,w_2)}{w_1}=\frac{H(0,w_2)}{w_1}+\frac{H(w_1,w_2)-H(0,w_2)}{w_1} \tag{3.1.13}$$

The second term in the equation is now analytic, and

$$\begin{aligned} I^{\{2\}} & \dot{\sim}\int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty}\frac{\exp(n[\frac{1}{2}w_1^2+\frac{1}{2}w_2^2-\hat{w}_1 w_1-\hat{w}_2 w_2])}{(2\pi i)^2 w_1}H(0,w_2)\,d\mathbf{w} \\ & =\int_{\hat{w}_1-i\infty}^{\hat{w}_1+i\infty}\frac{\exp(n[\frac{1}{2}w_1^2-\hat{w}_1 w_1])}{(2\pi i)w_1}\,dw_1\times \\ & \quad \int_{\hat{w}_2-i\infty}^{\hat{w}_2+i\infty}\frac{\exp(n[\frac{1}{2}w_2^2-\hat{w}_2 w_2])}{2\pi i}H(0,w_2)\,dw_2 \\ & \dot{\sim}\frac{1}{\sqrt{n}}H(0,\hat{w}_2)\bar{\Phi}(\sqrt{n}\hat{w}_1)\phi(\sqrt{n}\hat{w}_2), \end{aligned} \tag{3.1.14}$$

where, by the definition of $G(\tau_1,\tau_2)$, (2.0.5b), (2.0.5c) and (2.0.8),

$$\begin{aligned} H(0,\hat{w}_2) & =\frac{G(0,\tau_2(0,\hat{w}_2))-G(0,0)}{\hat{w}_2-\tilde{w}_2(0)}=\frac{[\frac{w_1}{\tau_1}\frac{d\tau_1}{dw_1}]\Big|_0\cdot\frac{\hat{w}_2}{\tau_2(0,\hat{w}_2)}\cdot\frac{d\tau_2}{dw_2}\Big|_{(0,\hat{w}_2)}-1}{\hat{w}_2} \\ & =\frac{1}{\tilde{\tau}_2(0)\sqrt{K^{22}(0,\tilde{\tau}_2(0))}}-\frac{1}{\hat{w}_2} \end{aligned}$$

If $EX_2 = \bar{x}_2$, then both $\tilde{\tau}_2(0) = 0$ and $\hat{w}_2 = 0$. The treatment of this special case follows the discussion in [Yang and Kolassa 2002] and [Zhang and Kolassa 2008]. Since the handling of this special case is not theoretically difficult, but algebraically messy, we omit it here. In all other cases, we have

$$I^{\{2\}} \dot{\sim} \frac{1}{\sqrt{n}} \left(\frac{1}{\tilde{\tau}_2(0)\sqrt{K^{22}(0,\tilde{\tau}_2(0)}} - \frac{1}{\hat{w}_2} \right) \bar{\Phi}(\sqrt{n}\hat{w}_1)\phi(\sqrt{n}\hat{w}_2). \tag{3.1.15}$$

Now we evaluate $I^{\{1\}}$. From (2.0.11), we have

$$\begin{aligned} I^{\{1\}} &= \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2])}{(2\pi i)^2 (w_2 - \tilde{w}_2(w_1))} \left(\frac{G(\tau_1,0) - G(0,0)}{w_1} \right) d\mathbf{w} \\ &= \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2])}{(2\pi i)^2 (w_2 - \tilde{w}_2(w_1))} \left(\frac{1}{\tau_1(w_1)} \frac{d\tau_1}{dw_1} - \frac{1}{w_1} \right) d\mathbf{w}. \end{aligned} \tag{3.1.16}$$

Let $u_1 = w_1$ and $u_2 = w_2 - \tilde{w}_2(w_1)$. By change of variables, (3.1.16) transforms to

$$\begin{aligned} I^{\{1\}} = \int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} & \frac{1}{(2\pi i)^2 u_2} \exp(n[\frac{1}{2}u_1^2 + \frac{1}{2}(u_2 + \tilde{w}_2(u_1))^2 - \hat{w}_1 u_1 - \\ & \hat{w}_2(u_2 + \tilde{w}_2(u_1))]) \left(\frac{1}{\tau_1(u_1)} \frac{d\tau_1}{du_1} - \frac{1}{u_1} \right) d\mathbf{u} \end{aligned} \tag{3.1.17}$$

Noting that

$$h(u_1) = \left(\frac{1}{\tau_1(u_1)} \frac{d\tau_1}{du_1} - \frac{1}{u_1} \right) \tag{3.1.18}$$

is analytic since $\frac{G(\tau_1,0)-G(0,0)}{w_1} = \frac{G(\tau_1,0)-G(0,0)}{\tau_1} \frac{\tau_1}{w_1}$ is analytic, again we can use the same technique and reasoning presented in [Kolassa 2003] to obtain

$$\begin{aligned} I^{\{1\}} \dot{\sim} C \int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} & \frac{1}{(2\pi i)^2 u_2} \exp(n[\frac{1}{2}(1 + (\breve{w}_2')^2)u_1^2 + \frac{1}{2}u_2^2 + \breve{w}_2' u_1 u_2 - \\ & ((1 + (\breve{w}_2')^2)\hat{u}_1 + \tilde{w}_2'(\hat{u}_1)\hat{u}_2)u_1 - (\breve{w}_2'\hat{u}_1 + \hat{u}_2)u_2])h(u_1)\, d\mathbf{u}, \end{aligned} \tag{3.1.19}$$

where C is defined in (3.1.11). Here the third degree terms contribute an error of $O(n^{-\frac{1}{2}})$ to $I^{\{1\}}$, which is itself $I^{\emptyset}O(n^{-\frac{1}{2}})$, and therefore can be omitted.

Integrals of the general form

$$\int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{\exp(n[\frac{a_1}{2}u_1^2 + \frac{a_2}{2}u_2^2 + cu_1u_2 - b_1u_1 - b_2u_2])}{(2\pi i)^2 u_2} h(u_1)\, d\mathbf{u} \tag{3.1.20}$$

can be computed. Using the transformation, $v_1 = \sqrt{a_1}(u_1 + \frac{c}{a_1}u_2)$ and $v_2 = \sqrt{a_2 - \frac{c^2}{a_1}}u_2$, we have:

$$\int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(n[\frac{1}{2}v_1^2 + \frac{1}{2}v_2^2 - \frac{b_1}{\sqrt{a_1}}v_1 - \frac{b_2 - \frac{b_1 c}{a_1}}{\sqrt{a_2 - \frac{c^2}{a_1}}}v_2])}{(2\pi i)^2 v_2 \sqrt{a_1}} h(\frac{v_1}{\sqrt{a_1}} - \frac{cv_2}{a_1\sqrt{a_2 - \frac{c^2}{a_1}}})\, d\mathbf{v} \tag{3.1.21}$$

Applying the same trick as in (3.1.13) to $h(\frac{v_1}{\sqrt{a_1}} - \frac{cv_2}{a_1\sqrt{a_2-\frac{c^2}{a_1}}})$, the integral (3.1.20) may be approximated, with relative error $O(n^{-\frac{1}{2}})$, by

$$\frac{1}{\sqrt{a_1}}\phi(\frac{b_1}{\sqrt{a_1}})\bar{\Phi}(\frac{b_2 - \frac{b_1 c}{a_1}}{\sqrt{a_2 - \frac{c^2}{a_1}}})h(\hat{u}_1). \quad (3.1.22)$$

Here again, we omit the special case that $\hat{w}_1 = 0$.

Now come back to $I^{\{1\}}$. Compare the coefficients of (3.1.19) and (3.1.20), and substitute the corresponding quantities into the equivalence of (3.1.20) and (3.1.22), to obtain

$$I^{\{1\}}\dot{\sim}C\frac{h(\hat{w}_1)}{\sqrt{n}\sqrt{1+(\check{w}_2')^2}}\phi(\frac{\sqrt{n}((1+(\check{w}_2')^2)\hat{w}_1+\check{w}_2'(\hat{w}_2-\check{w}_2))}{\sqrt{1+(\check{w}_2')^2}})\bar{\Phi}(\frac{\sqrt{n}(\hat{w}_2-\check{w}_2)}{\sqrt{1+(\check{w}_2')^2}}). \quad (3.1.23)$$

3.2 An example for continuous case

We consider the bivariate random vector (Y_1, Y_2), with $Y_1 = X_1+X_2$ and $Y_2 = X_2+X_3$, where X_1, X_2 and X_3 are independent and identically distributed random variables following the exponential distribution, which has a density function $f(x) = e^{-x}$ for $x > 0$. The moment generating function of (Y_1, Y_2) is

$$\begin{aligned} M_{(Y_1,Y_2)}(\tau_1,\tau_2) =& Ee^{\tau_1 Y_1+\tau_2 Y_2} \\ =& Ee^{\tau_1(X_1+X_2)+\tau_2(X_2+X_3)} \\ =& Ee^{\tau_1 X_1}Ee^{(\tau_1+\tau_2)X_2}Ee^{\tau_2 X_3} \\ =& M_{X_1}(\tau_1)M_{X_2}(\tau_1+\tau_2)M_{X_3}(\tau_2) \\ =& \frac{1}{(1-\tau_1)(1-\tau_1-\tau_2)(1-\tau_2)}, \end{aligned} \quad (3.2.1)$$

for $\tau_1 < 1$, $\tau_2 < 1$ and $\tau_1 + \tau_2 < 1$. The cumulative generating function is, therefore, $K(\tau_1, \tau_2) = \log(M_{(Y_1,Y_2)}(\tau_1, \tau_2)) = -\log(1 - \tau_1) - \log(1 - \tau_1 - \tau_2) - \log(1 - \tau_2)$. The global minimum of $K(\tau_1, \tau_2) - \tau_1\bar{x} - \tau_2\bar{y}$ can be solve with the equation system

$$\begin{cases} K^1(\tau_1,\tau_2) = \frac{1}{1-\tau_1} + \frac{1}{1-\tau_2-\tau_2} = \bar{y}_1 \\ K^2(\tau_1,\tau_2) = \frac{1}{1-\tau_2} + \frac{1}{1-\tau_2-\tau_2} = \bar{y}_2 \end{cases}. \quad (3.2.2)$$

From the first equation, we have

$$\tau_2 = (1-\tau_1) - \frac{1}{\bar{y}_1 - \frac{1}{1-\tau_1}}$$

Substitute it into the second equation and simplify to obtain

$$(\bar{y}_2-\bar{y}_1)\bar{y}_1\tau_1^3+[(\bar{y}_2-\bar{y}_1)(2-2\bar{y}_1)-2\bar{y}_1]\tau_1^2+[(\bar{y}_2-\bar{y}_1)(\bar{y}_1-3)+3\bar{y}_1-3]\tau_1+(\bar{y}_2-2\bar{y}_1+2) = 0.$$

The polynomial equation with a degree of 3, and can be solved numerically using the Newton-Raphson method.

Given τ_1, the $\tilde{\tau}_2(\tau_1)$ that minimize $K(\cdot,\cdot)$ is obtained by solving $K^2(\tau_1,\tau_2) = \frac{1}{1-\tau_2} + \frac{1}{1-\tau_1-\tau_2} = \bar{y}_2$ for τ_2. Combining the constraint that $\tau_1 < 1$, $\tau_2 < 1$ and $\tau_1 + \tau_2 < 1$, we have

$$\tilde{\tau}_2(\tau_1) = 1 - \frac{\bar{y}_2\tau_1 + 2 + \sqrt{\bar{y}_2^2\tau_1^2 + 4}}{2\bar{y}_2},$$

and in particular, $\tilde{\tau}_2(0) = 1 - \frac{2}{y_2}$. With $\hat{\tau}_1$ and $\hat{\tau}_2$, we can use (3.1.4a) and (3.1.4b) to obtain $\hat{w}_1$ and $\hat{w}_2$.

By (2.0.6), we have

$$\breve{w}_2 = \hat{w}_2 + \text{sign}(0-\hat{\tau}_2)\sqrt{-2[K(\hat{\tau}_1,\hat{\tau}_2) - \hat{\tau}_2\bar{y}_2 - K(\hat{\tau}_1,0)]}. \tag{3.2.3}$$

To obtain $\breve{w}_2'$, by (2.0.7) we have

$$\breve{w}_2' = \frac{[K^1(\hat{\tau}_1,0) - \bar{y}_1] \left.\frac{d\tau_1}{dw_1}\right|_{\hat{w}_1}}{\breve{w}_2 - \hat{w}_2}, \tag{3.2.4}$$

where $\left.\frac{d\tau_1}{dw_1}\right|_{\hat{w}_1}$ can be computed by (2.0.8), i.e.,

$$\left.\frac{d\tau_1}{dw_1}\right|_{\hat{w}_1} = \sqrt{\frac{1}{K^{11}(\hat{\tau}_1,\hat{\tau}_2) + K^{12}(\hat{\tau}_1,\hat{\tau}_2)\tilde{\tau}_2'(\hat{\tau}_1)}}, \tag{3.2.5}$$

where the second derivatives of $K(\cdot,\cdot)$ can be calculated by the following formula:

$$\begin{cases} K^{11}(\tau_1,\tau_2) = \frac{1}{(1-\tau_1)^2} + \frac{1}{(1-\tau_1-\tau_2)^2} \\ K^{22}(\tau_1,\tau_2) = \frac{1}{(1-\tau_2)^2} + \frac{1}{(1-\tau_1-\tau_2)^2} \\ K^{12}(\tau_1,\tau_2) = \frac{1}{(1-\tau_1-\tau_2)^2} \end{cases}. \tag{3.2.6}$$

The last thing that we need to compute is $h(\hat{w}_1) = \frac{1}{\hat{\tau}_1}\left.\frac{d\tau_1}{dw_1}\right|_{\hat{w}_1} - \frac{1}{\hat{w}_1}$, which is readily available.

The results for approximating $P(\bar{Y}_1 \geq \bar{y}_1, \bar{Y}_2 \geq \bar{y}_2)$ when $n = 5$ are listed in Table 3.1 below, where "P. approx." stands for saddlepoint approximation proposed in this paper, "K. approx." stands for saddlepoint approximation presented at [Kolassa 2003], and "N. approx" stands for bivariate normal approximation. The "Exact" column shows the exact tail probability values computed by [Mathematica 5.0 2005]. "Relative Error" column shows the relative error of "P. approx". The case that $(\bar{y}_1 = 2.5, \bar{y}_2 = 3.0)$ and $(\bar{y}_1 = 3.0, \bar{y}_2 = 4.0)$ is the special case that $\hat{w}_1 = 0$ and is omitted here. The normal approximation deteriorates at the far tail, while both saddlepoint approximations show much better and more stable relative errors. In almost all cases, the new method shows smaller relative errors than that in [Kolassa 2003].

Table 3.1: Results of saddlepoint approximation compared with other approximations in the continuous case.

$\bar{y}_1$	$\bar{y}_2$	P. approx.	K. approx	N. approx.	Exact	Relative Error
2.5	2.5	9.12×10^{-2}	8.98×10^{-2}	9.65×10^{-2}	9.22×10^{-2}	−1.08%
2.5	3.5	1.41×10^{-2}	1.41×10^{-2}	6.54×10^{-3}	1.41×10^{-2}	0.00%
2.5	4.0	3.90×10^{-3}	3.99×10^{-3}	6.69×10^{-3}	3.93×10^{-3}	−0.76%
3.0	3.0	2.20×10^{-2}	2.14×10^{-2}	1.46×10^{-2}	2.22×10^{-2}	−0.90%
3.0	3.5	8.96×10^{-3}	8.73×10^{-3}	3.52×10^{-3}	8.96×10^{-3}	0.00%
3.5	3.5	4.40×10^{-3}	4.25×10^{-3}	1.09×10^{-3}	4.40×10^{-3}	0.00%
3.5	4.0	1.67×10^{-3}	1.61×10^{-3}	1.78×10^{-4}	1.66×10^{-3}	0.60%
4.0	4.0	7.67×10^{-4}	7.34×10^{-4}	3.88×10^{-5}	7.58×10^{-4}	1.19%

3.3 Unit lattice distributions

Bivariate tail probability approximations for unit lattice variables follow the same route. In unit lattice case, we consider the inversion integral for $P(\bar{\mathbf{x}} \leq \bar{\mathbf{X}} < \bar{\mathbf{x}}_0)$. We deform the path of integration to run through $\mathbf{c}$ for some $\mathbf{c} > 0$, then one can pass the limit as $\bar{\mathbf{x}}_0 \to \infty$. The integral (3.1.1) becomes

$$\begin{aligned} &\frac{1}{(2\pi i)^2} \int_{\mathbf{c}-i\pi}^{\mathbf{c}+i\pi} \frac{\exp(n[K(\tau_1,\tau_2) - \tau_1(\bar{x} - \frac{1}{2n}) - \tau_2(\bar{y} - \frac{1}{2n})])}{2\sinh(\frac{\tau_1}{2}) 2\sinh(\frac{\tau_2}{2})} d\boldsymbol{\tau} \\ =&\frac{1}{(2\pi i)^2} \int_{\mathbf{c}-i\pi}^{\mathbf{c}+i\pi} \frac{\exp(n[K(\tau_1,\tau_2) - \tau_1\bar{x}^* - \tau_2\bar{y}^*])}{2\sinh(\frac{\tau_1}{2}) 2\sinh(\frac{\tau_2}{2})} d\boldsymbol{\tau}, \end{aligned} \quad (3.3.1)$$

where $\bar{x}^* = \bar{x} - \frac{1}{2n}$, $\bar{y}^* = \bar{y} - \frac{1}{2n}$, and $\sinh(x)$ is the hyperbolic sin function defined as $\frac{e^x - e^{-x}}{2}$. The definitions (3.1.3a)-(3.1.3d) becomes

$$-\frac{1}{2}\hat{w}_1^2 = K(\hat{\tau}_1, \hat{\tau}_2) - \hat{\tau}_1\bar{x}_1^* - \hat{\tau}_2\bar{x}_2^* - (K(0, \tilde{\tau}_2(0)) - \tilde{\tau}_2(0)\bar{x}_2^*), \tag{3.3.2a}$$

$$-\frac{1}{2}\hat{w}_2^2 = K(0, \tilde{\tau}_2(0)) - \tilde{\tau}_2(0)\bar{x}_2^*, \tag{3.3.2b}$$

$$-\frac{1}{2}(w_1 - \hat{w}_1)^2 = K(\hat{\tau}_1, \hat{\tau}_2) - \hat{\tau}_1\bar{x}_1^* - \hat{\tau}_2\bar{x}_2^* - (K(\tau_1, \tilde{\tau}_2(\tau_1)) - \tau_1\bar{x}_1^* - \tilde{\tau}_2(\tau_1)\bar{x}_2^*), \tag{3.3.2c}$$

$$-\frac{1}{2}(w_2 - \hat{w}_2)^2 = K(\tau_1, \tilde{\tau}_2(\tau_1)) - \tau_1\bar{x}_1^* - \tilde{\tau}_2(\tau_1)\bar{x}_2^* - (K(\tau_1, \tau_2) - \tau_1\bar{x}_1^* - \tau_2\bar{x}_2^*), \tag{3.3.2d}$$

The definition (3.1.4a)-(3.1.4d) becomes

$$\hat{w}_1 = \text{sign}(\hat{\tau}_1)\sqrt{-2[K(\hat{\tau}_1, \hat{\tau}_2) - \hat{\tau}_1\bar{x}_1^* - \hat{\tau}_2\bar{x}_2^* - (K(0, \tilde{\tau}_2(0)) - \tilde{\tau}_2(0)\bar{x}_2^*)]}, \tag{3.3.3a}$$

$$\hat{w}_2 = \text{sign}(\hat{\tau}_2)\sqrt{-2[K(0, \tilde{\tau}_2(0)) - \tilde{\tau}_2(0)\bar{x}_2^*]}, \tag{3.3.3b}$$

$$\begin{aligned} w_1 =& \hat{w}_1 + \\ & \text{sign}(\tau_1 - \hat{\tau}_1)\sqrt{-2[K(\hat{\tau}_1, \hat{\tau}_2) - \hat{\tau}_1\bar{x}_1^* - \hat{\tau}_2\bar{x}_2^* - (K(\tau_1, \tilde{\tau}_2(\tau_1)) - \tau_1\bar{x}_1^* - \tilde{\tau}_2(\tau_1)\bar{x}_2^*)]}, \end{aligned} \tag{3.3.3c}$$

$$\begin{aligned} w_2 =& \hat{w}_2 + \\ & \text{sign}(\tau_2 - \tilde{\tau}_2(\tau_1))\sqrt{-2[K(\tau_1, \tilde{\tau}_2(\tau_1)) - \tau_1\bar{x}_1^* - \tilde{\tau}_2(\tau_1)\bar{x}_2^* - (K(\tau_1, \tau_2) - \tau_1\bar{x}_1^* - \tau_2\bar{x}_2^*)]}. \end{aligned} \tag{3.3.3d}$$

And (3.1.6) becomes

$$G(\boldsymbol{\tau}) = \left(\frac{w_1}{2\sinh(\tau_1/2)}\frac{d\tau_1}{dw_1}\right)\left(\frac{w_2 - \tilde{w}_2(w_1)}{2\sinh(\tau_2/2)}\frac{d\tau_2}{dw_2}\right). \tag{3.3.4}$$

Since $\lim_{x\to 0}(2\sinh(x/2)/x) = 1$, any analytic property in continuous case still holds in lattice case. With this in mind, we obtain the exactly same formula for $I^{\emptyset}$ as in (3.1.9) and (3.1.10),

$$I^{\{2\}} \dot{\sim} \frac{1}{\sqrt{n}}\left(\frac{1}{2\sinh(\tilde{\tau}_2(0)/2)\sqrt{K^{22}(0, \tilde{\tau}_2(0)}} - \frac{1}{\hat{w}_2}\right)\bar{\Phi}(\sqrt{n}\hat{w}_1)\phi(\sqrt{n}\hat{w}_2) \tag{3.3.5}$$

and

$$\begin{aligned} I^{\{1\}} \dot{\sim} & C\frac{h(\hat{w}_1)}{\sqrt{n}\sqrt{1 + (\breve{w}_2')^2}} \cdot \\ & \phi\left(\frac{\sqrt{n}((1 + (\breve{w}_2')^2)\hat{w}_1 + \breve{w}_2'(\hat{w}_2 - \breve{w}_2))}{\sqrt{1 + (\breve{w}_2')^2}}\right)\bar{\Phi}\left(\frac{\sqrt{n}(\hat{w}_2 - \breve{w}_2)}{\sqrt{1 + (\breve{w}_2')^2}}\right), \end{aligned} \tag{3.3.6}$$

where here $h(z) = \frac{1}{2\sinh(\tau_1(z)/2)} \frac{d\tau_1}{dw_1}\Big|_z - \frac{1}{z}$.

In summary, we have

Theorem 3.3.1.

$$P(T \geq t) \sim I^{\emptyset} + I^{\{2\}} + I^{\{1\}},$$

which can be computed as in (3.1.9), (3.1.10), (3.1.15) and (3.1.23) for continuous variables and (3.1.9), (3.1.10), (3.3.5) and (3.3.6) for unit lattice variables.

3.4 An example for lattice case

In the second example, we consider the bivariate random vector (Y_1, Y_2), with $Y_1 = X_1 + X_2$ and $Y_2 = X_2 + X_3$, where X_1, X_2 and X_3 are independent and identically distributed random variables following binomial distribution, which has a mass function $\binom{N}{x} p^x (1-p)^{N-x}$ for $0 \leq x \leq N$ and a moment generating function $(1 - p + pe^{\tau})^N$. Using the same technique, the moment generating function of (Y_1, Y_2) is:

$$M_{X_1}(\tau_1) M_{X_2}(\tau_1 + \tau_2) M_{X_3}(\tau_2) = [(1 - p + pe^{\tau_1})(1 - p + pe^{\tau_1+\tau_2})(1 - p + pe^{\tau_2})]^N.$$

The cumulative generating function is, therefore, $K(\tau_1, \tau_2) = N[\log(1 - p + pe^{\tau_1}) + \log(1 - p + pe^{\tau_1+\tau_2}) + \log(1 - p + pe^{\tau_2})]$. The global minimum of $K(\tau_1, \tau_2) - \tau_1 \bar{y}_1^* - \tau_2 \bar{y}_2^*$ can be solve with the equation system

$$\begin{cases} K^1(\tau_1, \tau_2) = N[\frac{pe^{\tau_1}}{1-p+pe^{\tau_1}} + \frac{pe^{\tau_1+\tau_2}}{1-p+pe^{\tau_1+\tau_2}}] = \bar{y}_1^* \\ K^2(\tau_1, \tau_2) = N[\frac{pe^{\tau_2}}{1-p+pe^{\tau_2}} + \frac{pe^{\tau_1+\tau_2}}{1-p+pe^{\tau_1+\tau_2}}] = \bar{y}_2^* \end{cases} \tag{3.4.1}$$

To simplify notation, let $a = \bar{y}_1^*/N$, $b = \bar{y}_2^*/N$, $z_1 = e^{\tau_1}$ and $z_2 = e^{\tau_2}$. From the first equation, we have

$$z_2 = \frac{(1-p)(a - \frac{pz_1}{1-p+pz_1})}{pz_1(1 - a + \frac{pz_1}{1-p+pz_1})}$$

Substitute it into the second equation and simplify, we have

$$
\begin{aligned}
&p^2(2-a)(-a+b+1)z_1^3 \\
&+p[(1-p)(1-a)-(1-p)(a-b)(2-a)-(1-p)(a-b)(1-a)+p(a-b)(1-a)]z_1^2 \\
&+(1-p)[p(1-a)-(1-p)(a-b)(1-a)+p(a-b)(1-a)-p(a-b)a]z_1 \\
&-a(1-p)^2(a-b+1)=0.
\end{aligned}
\tag{3.4.2}
$$

It is a polynomial equation with a degree of 3, and can be solved numerically using Newton-Raphson method.

Given τ_1, the $\tilde{\tau}_2(\tau_1)$ that minimize $K(\cdot,\cdot)$ is obtained by solving:

$$K^2(\tau_1,\tau_2)=N[\frac{pe^{\tau_2}}{1-p+pe^{\tau_2}}+\frac{pe^{\tau_1+\tau_2}}{1-p+pe^{\tau_1+\tau_2}}]=\bar{y}_2^*$$

for τ_2. The solution is

$$\tilde{\tau}_2(\tau_1)=\log\left(\frac{1-p}{p}\frac{\sqrt{(z_1+1)^2(1-b)^2+4(2-b)bz_1}-(z_1+1)(1-b)}{2(2-b)z_1}\right), \tag{3.4.3}$$

and in particular,

$$\tilde{\tau}_2(0)=\log\left(\frac{1-p}{p}\frac{b}{2-b}\right).$$

Similar to the continuous case, with $\hat{\tau}_1$ and $\hat{\tau}_2$, we can use (3.3.3a) and (3.3.3b) to obtain $\hat{w}_1$ and $\hat{w}_2$.

By (2.0.6), we have

$$\check{w}_2=\hat{w}_2+\text{sign}(0-\hat{\tau}_2)\sqrt{-2[K(\hat{\tau}_1,\hat{\tau}_2)-\hat{\tau}_2\bar{y}_2^*-K(\hat{\tau}_1,0)]}.$$

To obtain $\check{w}_2'$, by (2.0.7) we have

$$\check{w}_2'=\frac{[K^1(\hat{\tau}_1,0)-\bar{y}_1^*]\left.\frac{d\tau_1}{dw_1}\right|_{\hat{w}_1}}{\check{w}_2-\hat{w}_2},$$

where again $\left.\frac{d\tau_1}{dw_1}\right|_{\hat{w}_1}$ can be computed by (2.0.8), i.e.,

$$\left.\frac{d\tau_1}{dw_1}\right|_{\hat{w}_1}=\sqrt{\frac{1}{K^{11}(\hat{\tau}_1,\tilde{\tau}_2(\hat{\tau}_1))+K^{12}(\hat{\tau}_1,\tilde{\tau}_2(\hat{\tau}_1))\tilde{\tau}_2'(\hat{\tau}_1)}},$$

where the second derivatives of $K(\cdot,\cdot)$ can be calculated by the following formula:

$$\begin{cases} K^{11}(\tau_1,\tau_2) = Np(1-p)(\frac{z_1}{1-p+pz_1} + \frac{z_1z_2}{1-p+pz_1z_2}) \\ K^{12}(\tau_1,\tau_2) = Np(1-p)(\frac{z_1z_2}{1-p+pz_1z_2}) \\ K^{22}(\tau_1,\tau_2) = Np(1-p)(\frac{z_2}{1-p+pz_2} + \frac{z_1z_2}{1-p+pz_1z_2}) \end{cases} \tag{3.4.4}$$

The last thing that we need to compute is $h(\hat{w}_1) = \frac{1}{2\sinh(\frac{\hat{\tau}_1}{2})} \frac{d\tau_1}{dw_1}\Big|_{\hat{w}_1} - \frac{1}{\hat{w}_1}$,

The results for approximating $P(\bar{Y}_1 \geq \bar{y}_1, \bar{Y}_2 \geq \bar{y}_2)$ when $N = 10$, $p = 0.2$ and $n = 8$ are listed in Table 3.2 below, We can again see from the table that the normal approximation (with adjustment for continuity) deteriorates at the far tail, while the saddlepoint approximations show much better and more stable relative errors. In most cases, the new approximation shows better accuracy than that of [Kolassa 2003].

Table 3.2: Results of saddlepoint approximation compared with other approximations in the unit lattice case.

$\bar{y}_1$	$\bar{y}_2$	P. approx.	K. approx.	N. approx.	Exact	Relative Error
4.5	4.5	1.15×10^{-1}	1.16×10^{-1}	1.16×10^{-1}	1.15×10^{-1}	0.00%
4.5	5.0	4.43×10^{-2}	4.51×10^{-2}	4.28×10^{-2}	4.44×10^{-2}	−0.23%
4.5	5.5	1.04×10^{-2}	1.05×10^{-2}	8.73×10^{-3}	1.04×10^{-2}	0.00%
4.5	6.0	1.46×10^{-3}	1.45×10^{-3}	9.50×10^{-4}	1.46×10^{-3}	0.00%
5.0	5.0	2.07×10^{-2}	2.12×10^{-2}	1.92×10^{-2}	2.08×10^{-2}	−0.48%
5.0	5.5	5.89×10^{-3}	6.04×10^{-3}	4.85×10^{-3}	5.91×10^{-3}	−0.34%
5.0	6.0	9.91×10^{-4}	1.01×10^{-3}	6.40×10^{-4}	9.94×10^{-4}	−0.30%
5.5	5.5	2.11×10^{-3}	2.16×10^{-3}	1.57×10^{-3}	2.11×10^{-3}	0.00%
5.5	6.0	4.45×10^{-4}	4.56×10^{-4}	2.69×10^{-4}	4.47×10^{-4}	-0.45%
6.0	6.0	1.21×10^{-4}	1.24×10^{-4}	6.14×10^{-5}	1.21×10^{-4}	0.00%

Chapter 4

Multivariate conditional distribution approximations

4.1 Conditional continuous distributions

Consider a multivariate canonical exponential family. In practice, we are often interested in only a subset of the parameters in a given statistical model, with the other model parameters usually treated as nuisance parameters. The distribution of the sufficient statistics associated with parameters of interest, conditional on the sufficient statistics associated with the nuisance parameters, depends on the parameters of interest and not the nuisance parameters. We can therefore use the conditional distributions instead of the original distributions for inference. For instance, in testing equality of proportions for a 2×2 contingency table, we condition on the row or column margins; another example is logistic regression, where inference on some regression parameters is often performed conditionally on sufficient statistics associated with nuisance parameters.

Certain hypotheses involving parameters of interest, particularly order-restricted hypotheses, may be tested by computing the tail probabilities for the conditional distribution $P(\mathbf{T}_{d_0} \geq \mathbf{t}_{d_0} | \mathbf{T}_{-d_0} = \mathbf{t}_{-d_0})$. Skovgaard ([Skovgaard 1987]) applies double saddlepoint approximation to the problem in the case that $d_0 = 1$, $d > 1$ and $\mathbf{T}$ is the mean of independent and identically distributed random vectors. Here we propose a method that extends the results to $d_0 > 1$ and $d > d_0$, using the idea in the previous sections.

First, consider $\mathbf{T}$, the mean of independent and identically distributed continuous random vectors. Then

$$P(\mathbf{T}_{d_0} \geq \mathbf{t}_{d_0} | \mathbf{T}_{-d_0} = \mathbf{t}_{-d_0}) = \frac{\int_{\mathbf{t}_{d_0}}^{\infty} f_{\mathbf{T}}(y_1, \cdots, y_{d_0}, t_{d_0+1}, \cdots, t_d) \, d\mathbf{y}_{d_0}}{f_{\mathbf{T}_{-d_0}}(\mathbf{t}_{-d_0})},$$

where $f_{\mathbf{T}}(\cdot)$ is the joint density and $f_{\mathbf{T}_{-d_0}}(\cdot)$ is the marginal density of $\mathbf{T}_{-d_0}$. Again, we use the Fourier inversion formula to obtain:

$$P(\mathbf{T}_{d_0} \geq \mathbf{t}_{d_0} | \mathbf{T}_{-d_0} = \mathbf{t}_{-d_0}) = \frac{\frac{n^{d-d_0}}{(2\pi i)^d} \int_{\mathbf{c}-i\infty}^{\mathbf{c}+i\infty} \frac{\exp(n[K(\boldsymbol{\tau}) - \boldsymbol{\tau}^T \mathbf{t}])}{\prod_{j=1}^{d_0} \tau_j} d\boldsymbol{\tau}}{f_{\mathbf{T}_{-d_0}}(\mathbf{t}_{-d_0})}, \tag{4.1.1}$$

where $K(\boldsymbol{\tau})$ is the cumulant generating function of the random vector $\mathbf{T}$. The numerator is just a special case of (2.0.1).

Approximation (2.0.10) holds because of the following lemma, which will allow us to apply previous unconditional results, by substituting components of $\hat{\mathbf{w}}$ for components of $\mathbf{w}$, when the components correspond to variables in the conditioning event.

Lemma 4.1.1.

$$\begin{aligned} &\frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0} (w_j - \tilde{w}_j(\mathbf{w}_{j-1}))} \cdot \prod_{j=1}^{d} \frac{d\tau_j}{dw_j} \frac{\prod_{j=1}^{d_0} (w_j - \tilde{w}_j(\mathbf{w}_{j-1}))}{\prod_{j=1}^{d_0} \rho(\tau_j(\mathbf{w}_j))} d\mathbf{w} \\ =& \frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0} (w_j - \tilde{w}_j(\mathbf{w}_{j-1}))} G(\boldsymbol{\tau})\, d\mathbf{w}(1 + O(n^{-1})), \end{aligned} \tag{4.1.2}$$

where $G(\boldsymbol{\tau}) = \frac{\prod_{j=1}^{d_0} (w_j - \tilde{w}_j(\mathbf{w}_{j-1}))}{\prod_{j=1}^{d_0} \rho(\tau_j(\mathbf{w}_j))} \cdot \prod_{j=1}^{d} \frac{d\tau_j}{dw_j}\Big|_{(\mathbf{w}_{d_0}, \hat{\mathbf{w}}_{-d_0})}$.

Proof. By Watson's lemma, given fixed $\mathbf{w}_{d_0}$, we have

$$\begin{aligned} &\int_{\hat{\mathbf{w}}_{-d_0}-i\mathbf{K}}^{\hat{\mathbf{w}}_{-d_0}+i\mathbf{K}} \exp(n[\frac{1}{2}\mathbf{w}_{-d_0}^T \mathbf{w}_{-d_0} - \hat{\mathbf{w}}_{-d_0}^T \mathbf{w}_{-d_0}]) \prod_{j=d_0+1}^{d} \frac{d\tau_j}{dw_j}\, d\mathbf{w}_{-d_0} \\ =& \int_{\hat{\mathbf{w}}_{-d_0}-i\mathbf{K}}^{\hat{\mathbf{w}}_{-d_0}+i\mathbf{K}} \exp(n[\frac{1}{2}\mathbf{w}_{-d_0}^T \mathbf{w}_{-d_0} - \hat{\mathbf{w}}_{-d_0}^T \mathbf{w}_{-d_0}]) \prod_{j=d_0+1}^{d} \frac{d\tau_j}{dw_j}\Big|_{(\mathbf{w}_{d_0}, \hat{\mathbf{w}}_{-d_0})} \cdot \\ &\left(1 + \frac{E_n(\mathbf{w}_{d_0})}{n}\right) d\mathbf{w}_{-d_0}, \end{aligned}$$

for some analytic function $E_n(\mathbf{w}_{d_0})$ of $O(1)$. Therefore,

$$\begin{aligned} LHS =& \frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{w}}_{d_0}-i\mathbf{K}}^{\hat{\mathbf{w}}_{d_0}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}_{d_0}^T \mathbf{w}_{d_0} - \hat{\mathbf{w}}_{d_0}^T \mathbf{w}_{d_0}])}{\prod_{j=1}^{d_0} \rho(\tau_j(\mathbf{w}_j))} \prod_{j=1}^{d_0} \frac{d\tau_j}{dw_j} \cdot \\ & \int_{\hat{\mathbf{w}}_{-d_0}-i\mathbf{K}}^{\hat{\mathbf{w}}_{-d_0}+i\mathbf{K}} \exp(n[\frac{1}{2}\mathbf{w}_{-d_0}^T \mathbf{w}_{-d_0} - \hat{\mathbf{w}}_{-d_0}^T \mathbf{w}_{-d_0}]) \prod_{j=d_0+1}^{d} \frac{d\tau_j}{dw_j}\Big|_{(\mathbf{w}_{d_0}, \hat{\mathbf{w}}_{-d_0})} \cdot \\ & \left(1 + \frac{E(\mathbf{w}_{d_0})}{n}\right) d\mathbf{w}_{-d_0} \cdot d\mathbf{w}_{d_0} \\ =& A(1 + \frac{1}{n}\frac{B}{A}), \end{aligned}$$

where

$$A = \frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0}(w_j - \tilde{w}_j(\mathbf{w}_{j-1}))} G(\boldsymbol{\tau})\, d\mathbf{w}$$

and

$$B = \frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0}(w_j - \tilde{w}_j(\mathbf{w}_{j-1}))} G(\boldsymbol{\tau}) E(\mathbf{w}_{d_0}) d\mathbf{w}$$

If A and B are expanded according to [Kolassa 2003], each integral is approximated by a tilting term times a normal multivariate tail probability, up to relative order $O(1/\sqrt{n})$. The expression for B is also multiplied by the leading term of E. Hence $A/B = O(1)$, and therefore, $LHS = A(1 + O(n^{-1}))$. □

To deal with the denominator in (4.1.1), we have the following lemma:

Lemma 4.1.2.

$$\left(\frac{n}{2\pi i}\right)^{d-d_0} \int_{\hat{\mathbf{w}}_{-d_0}-i\infty}^{\hat{\mathbf{w}}_{-d_0}+i\infty} \exp(n[\frac{1}{2}\mathbf{w}_{-d_0}^T\mathbf{w}_{-d_0} - \hat{\mathbf{w}}_{-d_0}^T\mathbf{w}_{-d_0}]) \prod_{j=d_0+1}^{d} \frac{d\tau_j}{dw_j}\Bigg|_{(\mathbf{0}_{d_0},\hat{\mathbf{w}}_{-d_0})} d\mathbf{w}_{-d_0}$$

$$= f_{\mathbf{T}_{-d_0}}(\mathbf{t}_{-d_0})(1 + O(n^{-1})) \tag{4.1.3}$$

Proof. This development is similar to that of [Kolassa 1997], p. 147. Substitute $\mathbf{w}_d = \mathbf{0}_d$, and by property (2.0.5a), $\boldsymbol{\tau}_d = \mathbf{0}_d$, into (2.0.2a) and (2.0.2b) to obtain

$$-\frac{1}{2}\hat{w}_j^2 = \min_{\boldsymbol{\gamma}_{-d_0},\boldsymbol{\gamma}_{j-1}=\mathbf{0}_{j-1}} (K_{-d_0}(\boldsymbol{\gamma}_{-d_0}) - \boldsymbol{\gamma}_{-d_0}^T\mathbf{t}_{-d_0}^*) - \min_{\boldsymbol{\gamma}_{-d_0},\boldsymbol{\gamma}_j=\mathbf{0}_j} (K_{-d_0}(\boldsymbol{\gamma}_{-d_0}) - \boldsymbol{\gamma}_{-d_0}^T\mathbf{t}_{-d_0}^*) \tag{4.1.4a}$$

$$-\frac{1}{2}(w_j - \hat{w}_j)^2 = \min_{\boldsymbol{\gamma}_{-d_0},\boldsymbol{\gamma}_{j-1}=\boldsymbol{\tau}_{j-1}} (K_{-d_0}(\boldsymbol{\gamma}_{-d_0}) - \boldsymbol{\gamma}_{-d_0}^T\mathbf{t}_{-d_0}^*) - \min_{\boldsymbol{\gamma}_{-d_0},\boldsymbol{\gamma}_j=\boldsymbol{\tau}_j} (K_{-d_0}(\boldsymbol{\gamma}_{-d_0}) - \boldsymbol{\gamma}_{-d_0}^T\mathbf{t}_{-d_0}^*), \tag{4.1.4b}$$

where $K_{-d_0}(\cdot)$ is the cumulant generating function of the random variable $\mathbf{T}_{-d_0}$. Change variables from $\mathbf{w}_{-d_0}$ to $\boldsymbol{\tau}_{-d_0}$ to obtain

$$\left(\frac{n}{2\pi i}\right)^{d-d_0} \int_{\hat{\mathbf{w}}_{-d_0}-i\infty}^{\hat{\mathbf{w}}_{-d_0}+i\infty} \exp(n[\frac{1}{2}\mathbf{w}_{-d_0}^T\mathbf{w}_{-d_0} - \hat{\mathbf{w}}_{-d_0}^T\mathbf{w}_{-d_0}]) \prod_{j=d_0+1}^{d} \frac{d\tau_j}{dw_j}\Bigg|_{(\mathbf{0}_{d_0},\mathbf{w}_{-d_0})} d\mathbf{w}_{-d_0}$$

$$= f_{\mathbf{T}_{-d_0}}(\mathbf{t}_{-d_0}).$$

By Watson's lemma, the left hand side is just the density approximation of $f_{\mathbf{T}_{-d_0}}(\mathbf{t}_{-d_0})$ up to $O(n^{-1})$. $\square$

With a continuous distribution, we can decompose A according to (2.0.11), with

$$G(\boldsymbol{\tau}) = \prod_{j=1}^{d_0}\left(\frac{w_j - \tilde{w}_j(\mathbf{w}_{j-1})}{\tau_j}\frac{d\tau_j}{dw_j}\right)\prod_{j=d_0+1}^{d}\frac{d\tau_j}{dw_j}\Bigg|_{(\mathbf{w}_{d_0},\hat{\mathbf{w}}_{-d_0})}. \tag{4.1.5}$$

Denote the left hand side of (4.1.3) as J_{-d_0}. Note that $G(\mathbf{0}) = \prod_{j=d_0+1}^{d}\frac{d\tau_j}{dw_j}\Big|_{(\mathbf{0}_{d_0},\hat{\mathbf{w}}_{-d_0})}$. Now, consider the case with $d_0 = 2$. Then the main term is

$$\begin{aligned} I^{\emptyset} =& \frac{n^{d-2}}{(2\pi i)^d}\int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty}\frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\hat{\mathbf{w}}])}{(2\pi i)^2 w_1(w_2 - \tilde{w}_2(w_1))}\prod_{j=3}^{d}\frac{d\tau_j}{dw_j}\Bigg|_{(0,0,\hat{\mathbf{w}}_{-2})} d\mathbf{w} \\ =& \int_{\hat{\mathbf{w}}_2-i\infty}^{\hat{\mathbf{w}}_2+i\infty}\frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2])}{(2\pi i)^2 w_1(w_2 - \tilde{w}_2(w_1))}\, d\mathbf{w}_2 \cdot J_{-2} \\ \sim& \int_{\hat{\mathbf{w}}_2-i\infty}^{\hat{\mathbf{w}}_2+i\infty}\frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2])}{(2\pi i)^2 w_1(w_2 - \tilde{w}_2(w_1))}\, d\mathbf{w}_2 \cdot f_{\mathbf{T}_{-2}}(\mathbf{t}_{-2}), \end{aligned} \tag{4.1.6}$$

where

$$\int_{\hat{\mathbf{w}}_2-i\infty}^{\hat{\mathbf{w}}_2+i\infty}\frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2])}{(2\pi i)^2 w_1(w_2 - \tilde{w}_2(w_1))}\, d\mathbf{w}_2 \tag{4.1.7}$$

can be obtained by formula (3.1.9) and (3.1.10).

Using the same technique as in (3.1.12)-(3.1.15), we have

$$\begin{aligned} I^{\{2\}}\dot{\sim}& \frac{n^{d-2}}{(2\pi d)^d}\int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty}\frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\hat{\mathbf{w}}])}{w_1}H(0, w_2)\, d\mathbf{w} \\ =& \frac{1}{n\prod_{j=3}^{d}\frac{d\tau_j}{dw_j}\Big|_{(0,0,\hat{\mathbf{w}}_{-2})}}\int_{\hat{w}_1-i\infty}^{\hat{w}_1+i\infty}\frac{\exp(n[\frac{1}{2}w_1^2 - \hat{w}_1 w_1])}{(2\pi i)w_1}\, dw_1 \cdot \\ & \frac{n}{2\pi i}\int_{\hat{w}_2-i\infty}^{\hat{w}_2+i\infty}\exp(n[\frac{1}{2}w_2^2 - \hat{w}_2 w_2])H(0, w_2)\, dw_2 \cdot J_{-2} \\ \dot{\sim}& \frac{H(0, \hat{w}_2)}{\sqrt{n}\prod_{j=3}^{d}\frac{d\tau_j}{dw_j}\Big|_{(0,0,\hat{\mathbf{w}}_{-2})}}\bar{\Phi}(\sqrt{n}\hat{w}_1)\phi(\sqrt{n}\hat{w}_2)f_{\mathbf{T}_{-2}}(\mathbf{t}_{-2}), \end{aligned} \tag{4.1.8}$$

at $O(n^{-1})$, where

$$\begin{aligned} H(0, \hat{w}_2) =& \frac{G(0, \tau_2(0, \hat{w}_2)) - G(0, 0)}{\hat{w}_2 - \tilde{w}_2(0)} \\ =& \frac{1}{\tilde{\tau}_2(0)}\left.\frac{d\tau_2}{dw_2}\right|_{0,\hat{w}_2}\prod_{j=3}^{d}\frac{d\tau_j}{dw_j}\Bigg|_{(0,\hat{w}_2,\hat{\mathbf{w}}_{-2})} - \frac{1}{\hat{w}_2}\prod_{j=3}^{d}\frac{d\tau_j}{dw_j}\Bigg|_{(0,0,\hat{\mathbf{w}}_{-2})} \end{aligned} \tag{4.1.9}$$

and $\prod_{j=3}^{d} \frac{d\tau_j}{dw_j}\Big|_{(w_1,w_2,\hat{\mathbf{w}}_{-2})}$ can be obtained using (2.0.9).

Using the argument as in (3.1.16)-(3.1.23), we have

$$\begin{aligned}
I^{\{1\}} &\dot{\sim} \frac{n^{d-2}}{(2\pi d)^d} \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\hat{\mathbf{w}}])}{w_2 - \tilde{w}_2(w_1)} \left(\frac{G(\tau_1, 0) - G(0,0)}{w_1} \right) d\mathbf{w} \\
&= \frac{1}{n \prod_{j=3}^{d} \frac{d\tau_j}{dw_j}\Big|_{(0,0,\hat{\mathbf{w}}_{-2})}} \cdot \\
&\quad \frac{n}{(2\pi i)^2} \int_{\hat{\mathbf{w}}_2-i\infty}^{\hat{\mathbf{w}}_2+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2])}{(2\pi i) w_1} \cdot h(w_1)\, d\mathbf{w}_2 \cdot J_{-2} \qquad (4.1.10) \\
&\dot{\sim} \frac{C}{\prod_{j=3}^{d} \frac{d\tau_j}{dw_j}\Big|_{0,0,\hat{\mathbf{w}}_{-2}}} \frac{h(\hat{w}_1)}{\sqrt{n}\sqrt{1+(\breve{w}_2')^2}} \cdot \\
&\quad \phi\Big(\frac{\sqrt{n}((1+(\breve{w}_2')^2)\hat{w}_1 + \breve{w}_2'(\hat{w}_2 - \breve{w}_2))}{\sqrt{1+(\breve{w}_2')^2}}\Big) \bar{\Phi}\Big(\frac{\sqrt{n}(\hat{w}_2 - \breve{w}_2)}{\sqrt{1+(\breve{w}_2')^2}}\Big) \cdot f_{\mathbf{T}_{-2}}(\mathbf{t}_{-2}),
\end{aligned}$$

where

$$h(z) = \frac{1}{\tau_1(z)} \frac{d\tau_1}{dw_1}\Big|_z \prod_{j=3}^{d} \frac{d\tau_j}{dw_j}\Big|_{(z,\tilde{w}_2(z),\hat{\mathbf{w}}_{-2})} - \frac{1}{z} \prod_{j=3}^{d} \frac{d\tau_j}{dw_j}\Big|_{(0,0,\hat{\mathbf{w}}_{-2})}. \qquad (4.1.11)$$

4.2 An example of arising from continuous distributions

In the case of $d = 3$ and $d_0 = 2$, we rewrite (2.0.2a)-(2.0.2b) as

$$-\frac{1}{2}\hat{w}_1^2 = \min_{\gamma_1,\gamma_2,\gamma_3}(K(\gamma_1,\gamma_2,\gamma_3) - \gamma_1\bar{x}_1 - \gamma_2\bar{x}_2 - \gamma_3\bar{x}_3) - \min_{\gamma_2,\gamma_3}(K(0,\gamma_2,\gamma_3) - \gamma_2\bar{x}_2 - \gamma_3\bar{x}_3) \qquad (4.2.1a)$$

$$-\frac{1}{2}\hat{w}_2^2 = \min_{\gamma_2,\gamma_3}(K(0,\gamma_2,\gamma_3) - \gamma_2\bar{x}_2 - \gamma_3\bar{x}_3) - \min_{\gamma_3}(K(0,0,\gamma_3) - \gamma_3\bar{x}_3) \qquad (4.2.1b)$$

$$-\frac{1}{2}\hat{w}_3^2 = \min_{\gamma_3}(K(0,0,\gamma_3) - \gamma_3\bar{x}_3) \qquad (4.2.1c)$$

$$-\frac{1}{2}(w_1-\hat{w}_1)^2 = \min_{\gamma_1,\gamma_2,\gamma_3}(K(\gamma_1,\gamma_2,\gamma_3) - \gamma_1\bar{x}_1 - \gamma_2\bar{x}_2 - \gamma_3\bar{x}_3) - \min_{\gamma_2,\gamma_3}(K(\tau_1,\gamma_2,\gamma_3) - \tau_1\bar{x}_1 - \gamma_2\bar{x}_2 - \gamma_3\bar{x}_3) \qquad (4.2.1d)$$

$$-\frac{1}{2}(w_2-\hat{w}_2)^2 = \min_{\gamma_2,\gamma_3}(K(\tau_1,\gamma_2,\gamma_3) - \tau_1\bar{x}_1 - \gamma_2\bar{x}_2 - \gamma_3\bar{x}_3) - \min_{\gamma_3}(K(\tau_1,\tau_2,\gamma_3) - \tau_1\bar{x}_1 - \tau_2\bar{x}_2 - \gamma_3\bar{x}_3) \qquad (4.2.1e)$$

$$-\frac{1}{2}(w_3-\hat{w}_3)^2 = \min_{\gamma_3}(K(\tau_1,\tau_2,\gamma_3) - \tau_1\bar{x}_1 - \tau_2\bar{x}_2 - \gamma_3\bar{x}_3) - (K(\tau_1,\tau_2,\tau_3) - \tau_1\bar{x}_1 - \tau_2\bar{x}_2 - \tau_3\bar{x}_3) \qquad (4.2.1f)$$

We can also rewrite (2.0.3a),(2.0.3b) and (2.0.4a), (2.0.4b) as

$$-\frac{1}{2}\hat{w}_1^2 = K(\hat{\tau}_1,\hat{\tau}_2,\hat{\tau}_3) - \hat{\tau}_1\bar{x}_1 - \hat{\tau}_2\bar{x}_2 - \hat{\tau}_3\bar{x}_3 - (K(0,\tilde{\tau}_2(0),\tilde{\tau}_3(0)) - \tilde{\tau}_2(0)\bar{x}_2 - \tilde{\tau}_3(0)\bar{x}_3) \tag{4.2.2a}$$

$$-\frac{1}{2}\hat{w}_2^2 = K(0,\tilde{\tau}_2(0),\tilde{\tau}_3(0)) - \tilde{\tau}_2(0)\bar{x}_2 - \tilde{\tau}_3(0)\bar{x}_3 - (K(0,0,\tilde{\tau}_3(0,0)) - \tilde{\tau}_3(0,0)\bar{x}_3) \tag{4.2.2b}$$

$$-\frac{1}{2}\hat{w}_3^2 = K(0,0,\tilde{\tau}_3(0,0)) - \tilde{\tau}_3(0,0)\bar{x}_3 \tag{4.2.2c}$$

$$-\frac{1}{2}(w_1-\hat{w}_1)^2 = K(\hat{\tau}_1,\hat{\tau}_2,\hat{\tau}_3) - \hat{\tau}_1\bar{x}_1 - \hat{\tau}_2\bar{x}_2 - \hat{\tau}_3\bar{x}_3 - (K(\tau_1,\tilde{\tau}_2(\tau_1),\tilde{\tau}_3(\tau_1)) - \tau_1\bar{x}_1 - \tilde{\tau}_2(\tau_1)\bar{x}_2 - \tilde{\tau}_3(\tau_1)\bar{x}_3) \tag{4.2.2d}$$

$$-\frac{1}{2}(w_2-\hat{w}_2)^2 = K(\tau_1,\tilde{\tau}_2(\tau_1),\tilde{\tau}_3(\tau_1)) - \tilde{\tau}_2(\tau_1)\bar{x}_2 - \tilde{\tau}_3(\tau_1)\bar{x}_3 - (K(\tau_1,\tau_2,\tilde{\tau}_3(\tau_1,\tau_2)) - \tau_2\bar{x}_2 - \tilde{\tau}_3(\tau_1,\tau_2)\bar{x}_3) \tag{4.2.2e}$$

$$-\frac{1}{2}(w_3-\hat{w}_3)^2 = K(\tau_1,\tau_2,\tilde{\tau}_3(\tau_1,\tau_2)) - \tilde{\tau}_3(\tau_1,\tau_2)\bar{x}_3 - (K(\tau_1,\tau_2,\tau_3) - \tau_3\bar{x}_3), \tag{4.2.2f}$$

where $\tilde{\tau}_3(0)$ is the minimizer of $K(0,\tau_2,\tau_3) - \tau_2\bar{x}_2 - \tau_3\bar{x}_3$ and $\tilde{\tau}_3(0,0)$ is the minimizer of $K(0,0,\tau_3) - \tau_3\bar{x}_3$, and

$$\hat{w}_1 = \text{sign}(\hat{\tau}_1)\cdot \sqrt{2[K(\hat{\tau}_1,\hat{\tau}_2,\hat{\tau}_3) - \hat{\tau}_1\bar{x}_1 - \hat{\tau}_2\bar{x}_2 - \hat{\tau}_3\bar{x}_3 - (K(0,\tilde{\tau}_2(0),\tilde{\tau}_3(0)) - \tilde{\tau}_2(0)\bar{x}_2 - \tilde{\tau}_3(0)\bar{x}_3)]} \tag{4.2.3a}$$

$$\hat{w}_2 = \text{sign}(\tilde{\tau}_2(0))\cdot \sqrt{2[K(0,\tilde{\tau}_2(0),\tilde{\tau}_3(0)) - \tilde{\tau}_2(0)\bar{x}_2 - \tilde{\tau}_3(0)\bar{x}_3 - (K(0,0,\tilde{\tau}_3(0,0)) - \tilde{\tau}_3(0,0)\bar{x}_3)]} \tag{4.2.3b}$$

$$\hat{w}_3 = \text{sign}(\tilde{\tau}_3(0,0))\sqrt{2[K(0,0,\tilde{\tau}_3(0,0)) - \tilde{\tau}_3(0,0)\bar{x}_3]} \tag{4.2.3c}$$

$$w_1 = \hat{w}_1 + \text{sign}(\tau_1 - \hat{\tau}_1)\cdot \sqrt{2[K(\hat{\tau}_1,\hat{\tau}_2,\hat{\tau}_3) - \hat{\tau}_1\bar{x}_1 - \hat{\tau}_2\bar{x}_2 - \hat{\tau}_3\bar{x}_3 - (K(\tau_1,\tilde{\tau}_2(\tau_1),\tilde{\tau}_3(\tau_1)) - \tau_1\bar{x}_1 - \tilde{\tau}_2(\tau_1)\bar{x}_2 - \tilde{\tau}_3(\tau_1)\bar{x}_3)]} \tag{4.2.3d}$$

$$w_2 = \hat{w}_2 + \text{sign}(\tau_2 - \tilde{\tau}_2(\tau_1))\cdot \sqrt{2[K(\tau_1,\tilde{\tau}_2(\tau_1),\tilde{\tau}_3(\tau_1)) - \tilde{\tau}_2(\tau_1)\bar{x}_2 - \tilde{\tau}_3(\tau_1)\bar{x}_3 - (K(\tau_1,\tau_2,\tilde{\tau}_3(\tau_1,\tau_2)) - \tau_2\bar{x}_2 - \tilde{\tau}_3(\tau_1,\tau_2)\bar{x}_3)]} \tag{4.2.3e}$$

$$w_3 = \hat{w}_3 + \text{sign}(\tau_3 - \tilde{\tau}_3(\tau_1,\tau_2))\cdot \sqrt{2[K(\tau_1,\tau_2,\tilde{\tau}_3(\tau_1,\tau_2)) - \tilde{\tau}_3(\tau_1,\tau_2)\bar{x}_3 - (K(\tau_1,\tau_2,\tau_3) - \tau_3\bar{x}_3)]} \tag{4.2.3f}$$

respectively.

Let $X_i,\ i = 1,2,3$ be independent and identically distributed random variables following the exponential distribution, as in the first example. Consider the random vector (Y_1, Y_2, Y_3) with $Y_1 = X_2$, $Y_2 = X_3$ and $Y_3 = X_1 + X_2 + X_3$. The moment generating function of (Y_1, Y_2, Y_3) is

$$\begin{aligned} M_{(Y_1,Y_2,Y_3)}(\tau_1, \tau_2, \tau_3) =& Ee^{\tau_1 Y_1 + \tau_2 Y_2 + \tau_3 Y_3} \\ =& Ee^{\tau_3 X_1 + (\tau_1+\tau_3) X_2 + (\tau_2+\tau_3) X_3} \\ =& Ee^{\tau_3 X_1} Ee^{(\tau_1+\tau_3) X_2} Ee^{(\tau_2+\tau_3) X_3} \\ =& M_{X_1}(\tau_3) M_{X_2}(\tau_1 + \tau_3) M_{X_3}(\tau_2 + \tau_3) \\ =& \frac{1}{(1-\tau_3)(1-\tau_1-\tau_3)(1-\tau_2-\tau_3)}, \end{aligned}$$

for $\tau_1 + \tau_3 < 1$, $\tau_2 + \tau_3 < 1$ and $\tau_3 < 1$. The cumulative generating function is, therefore, $K(\tau_1, \tau_2, \tau_3) = -\log(1-\tau_3) - \log(1-\tau_1-\tau_3) - \log(1-\tau_2-\tau_3)$.

The global minimum of $K(\tau_1, \tau_2, \tau_3) - \tau_1 \bar{y}_1 - \tau_2 \bar{y}_2 - \tau_3 \bar{y}_3$ can be obtained by solving

$$\begin{cases} K^1(\tau_1, \tau_2, \tau_3) = \frac{1}{1-\tau_1-\tau_3} = \bar{y}_1 \\ K^2(\tau_1, \tau_2, \tau_3) = \frac{1}{1-\tau_2-\tau_3} = \bar{y}_2 \\ K^3(\tau_1, \tau_2, \tau_3) = \frac{1}{1-\tau_3} + \frac{1}{1-\tau_1-\tau_3} + \frac{1}{1-\tau_2-\tau_3} = \bar{y}_3 \end{cases} \tag{4.2.4}$$

The solution is

$$\begin{cases} \hat{\tau}_1 = \frac{1}{\bar{y}_3-\bar{y}_1-\bar{y}_2} - \frac{1}{\bar{y}_1} \\ \hat{\tau}_2 = \frac{1}{\bar{y}_3-\bar{y}_1-\bar{y}_2} - \frac{1}{\bar{y}_2} \\ \hat{\tau}_3 = 1 - \frac{1}{\bar{y}_3-\bar{y}_1-\bar{y}_2} \end{cases} \tag{4.2.5}$$

Given τ_1 and τ_2, the $\tilde{\tau}_3(\tau_1, \tau_2)$ is obtained by solving $K^3(\tau_1, \tau_2, \tau_3) = \frac{1}{1-\tau_3} + \frac{1}{1-\tau_1-\tau_3} + \frac{1}{1-\tau_2-\tau_3} = \bar{y}_3$ for τ_3, or equivalently

$$\bar{y}_3(1-\tau_3)^3 - (\bar{y}_3(\tau_1+\tau_2)+3)(1-\tau_3)^2 + (\bar{y}_3\tau_1\tau_2 + 2(\tau_1+\tau_2))(1-\tau_3) - \tau_1\tau_2 = 0 \tag{4.2.6}$$

This is a polynomial equation with degree 3, but when $\tau_2 = 0$ it reduces to

$$\bar{y}_3(1-\tau_3)^2 - (\bar{y}_3\tau_1 + 3)(1-\tau_3) + 2\tau_1 = 0$$

Noting the constraint that $\tau_3 < 1$, we have

$$\tilde{\tau}_3(\tau_1, 0) = 1 - \frac{\bar{y}_3\tau_1 + 3 + \sqrt{(\bar{y}_3\tau_1+3)^2 - 8\bar{y}_3\tau_1}}{2\bar{y}_3}$$

The other root $1 - \frac{\bar{y}_3\tau_1+3-\sqrt{(\bar{y}_3\tau_1+3)^2-8\bar{y}_3\tau_1}}{2\bar{y}_3}$ either does not satisfy $\tau_3 < 1$ in the case that $\tau_1 \le 0$ or does not satisfy $\tau_1 + \tau_3 < 1$ in the case that $\tau_1 > 0$. By symmetry, we also have

$$\tilde{\tau}_3(0,\tau_2) = 1 - \frac{\bar{y}_3\tau_2 + 3 + \sqrt{(\bar{y}_3\tau_2+3)^2 - 8\bar{y}_3\tau_2}}{2\bar{y}_3}.$$

We also need $\tilde{\tau}_3^2(0,\tau_2)$. Take derivative of (4.2.6) with respect to τ_2, to obtain

$$\begin{aligned}&-3\bar{y}_3(1-\tilde{\tau}_3(\tau_1,\tau_2))^2\cdot\tilde{\tau}_3^2(\tau_1,\tau_2) - \bar{y}_3(1-\tilde{\tau}_3(\tau_1,\tau_2))^2 + \\ &2(\bar{y}_3(\tau_1+\tau_2)+3)(1-\tilde{\tau}_3(\tau_1,\tau_2))\cdot\tilde{\tau}_3^2(\tau_1,\tau_2) + (\bar{y}_3\tau_1+2)(1-\tilde{\tau}_3(\tau_1,\tau_2)) - \\ &(\bar{y}_3\tau_1\tau_2 + 2(\tau_1+\tau_2))\cdot\tilde{\tau}_3^2(\tau_1,\tau_2) - \tau_1 = 0\end{aligned}$$

Substitute $\tau_1 = 0$ and reorganize the terms to obtain

$$\tilde{\tau}_3^2(0,\tau_2) = \frac{-\bar{y}_3(1-\tilde{\tau}_3(0,\tau_2))^2 + 2(1-\tilde{\tau}_3(0,\tau_2))}{3\bar{y}_3(1-\tilde{\tau}_3(0,\tau_2))^2 - 2(\bar{y}_3\tau_2+3)(1-\tilde{\tau}_3(0,\tau_2)) + 2\tau_2}$$

We also need to compute $\tilde{\tau}_2(\tau_1)$ and $\tilde{\tau}_3(\tau_1)$. They are obtained by solving the following equation system in terms of τ_1

$$\begin{cases} K^2(\tau_1,\tau_2,\tau_3) = \frac{1}{1-\tau_2-\tau_3} = \bar{y}_2 \\ K^3(\tau_1,\tau_2,\tau_3) = \frac{1}{1-\tau_3} + \frac{1}{1-\tau_1-\tau_3} + \frac{1}{1-\tau_2-\tau_3} = \bar{y}_3 \end{cases}$$

Substitute the first equation into the second and simplify, to obtain

$$(\bar{y}_3-\bar{y}_2)(1-\tau_3)^2 - (\tau_1(\bar{y}_3-\bar{y}_2)+2)(1-\tau_3) + \tau_1 = 0$$

With the restriction that $\tau_3 < 1$, $\tau_1+\tau_3 < 1$ and $\tau_2+\tau_3 < 1$, we can solve the equation to get $\tilde{\tau}_2(\tau_1)$. The solution is

$$\begin{cases} \tilde{\tau}_2(\tau_1) = \frac{(\tau_1(\bar{y}_3-\bar{y}_2)+2)+\sqrt{\tau_1^2(\bar{y}_3-\bar{y}_2)^2+4}}{2(\bar{y}_3-\bar{y}_2)} - \frac{1}{\bar{y}_2} \\ \tilde{\tau}_3(\tau_1) = 1 - \frac{(\tau_1(\bar{y}_3-\bar{y}_2)+2)+\sqrt{\tau_1^2(\bar{y}_3-\bar{y}_2)^2+4}}{2(\bar{y}_3-\bar{y}_2)} \end{cases} \tag{4.2.7}$$

From the above equations we can obtain $\tilde{\tau}_2'(\tau_1)$ and $\tilde{\tau}_3'(\tau_1)$ shown below

$$\begin{cases} \tilde{\tau}_2'(\tau_1) = \frac{1}{2}(1 + \frac{\tau_1(\bar{y}_3-\bar{y}_2)}{\sqrt{\tau_1^2(\bar{y}_3-\bar{y}_2)^2+4}}) \\ \tilde{\tau}_3'(\tau_1) = -\frac{1}{2}(1 + \frac{\tau_1(\bar{y}_3-\bar{y}_2)}{\sqrt{\tau_1^2(\bar{y}_3-\bar{y}_2)^2+4}}) \end{cases} \tag{4.2.8}$$

We will need the second derivatives of $K(\cdot,\cdot,\cdot)$, which can be easily calculated by the following formula:

$$\begin{cases} K^{11}(\tau_1,\tau_2,\tau_3) = \frac{1}{(1-\tau_1-\tau_3)^2} \\ K^{12}(\tau_1,\tau_2,\tau_3) = 0 \\ K^{13}(\tau_1,\tau_2,\tau_3) = \frac{1}{(1-\tau_1-\tau_3)^2} \\ K^{22}(\tau_1,\tau_2,\tau_3) = \frac{1}{(1-\tau_2-\tau_3)^2} \\ K^{23}(\tau_1,\tau_2,\tau_3) = \frac{1}{(1-\tau_2-\tau_3)^2} \\ K^{33}(\tau_1,\tau_2,\tau_3) = \frac{1}{(1-\tau_3)^2} + \frac{1}{(1-\tau_1-\tau_3)^2} + \frac{1}{(1-\tau_2-\tau_3)^2} \end{cases} \tag{4.2.9}$$

Again, we need to compute $\tau_1(w_1)$. Substitute the previous results into (4.2.2d) and rearrange the terms. The resulting equation can then be solved by Newton-Raphson method.

The quantities $\hat{\mathbf{w}}$ can be obtained by (4.2.3a)-(4.2.3c). Using (2.0.9), we have

$$\left.\frac{d\tau_3}{dw_3}\right|_{w_3=\hat{w}_3} = \sqrt{\frac{1}{K^{33}(\tau_1,\tau_2,\tilde{\tau}_3(\tau_1,\tau_2))}}.$$

In particular, we have

$$\left.\frac{d\tau_3}{dw_3}\right|_{0,0,\hat{w}_3} = \sqrt{\frac{1}{K^{33}(0,0,\tilde{\tau}_3(0,0))}},$$

$$\left.\frac{d\tau_3}{dw_3}\right|_{0,\hat{w}_2,\hat{w}_3} = \sqrt{\frac{1}{K^{33}(0,\tilde{\tau}_2(0),\tilde{\tau}_3(0))}},$$

$$\text{and} \quad \left.\frac{d\tau_3}{dw_3}\right|_{w_1,\tilde{w}_2(w_1),\hat{w}_3} = \sqrt{\frac{1}{K^{33}(\tau_1,0,\tilde{\tau}_3(\tau_1,0))}}.$$

Obtain $\check{w}_2$ using (4.2.3e). By (2.0.7), we have

$$\check{w}_2' = \frac{[K^1(\hat{\tau}_1,0,\tilde{\tau}_3(\tau_1,0)) - \bar{x}]\left.\frac{d\tau_1}{dw_1}\right|_{\hat{w}_1}}{\check{w}_2 - \hat{w}_2},$$

where

$$\left.\frac{d\tau_1}{dw_1}\right|_{\hat{w}_1} = \frac{1}{\sqrt{K^{11}(\hat{\tau}_1,\hat{\tau}_2,\hat{\tau}_3) + K^{12}(\hat{\tau}_1,\hat{\tau}_2,\hat{\tau}_3)\tilde{\tau}_2'(\hat{\tau}_1) + K^{13}(\hat{\tau}_1,\hat{\tau}_2,\hat{\tau}_3)\tilde{\tau}_3'(\hat{\tau}_1)}},$$

by (2.0.8). Similarly,

$$\left.\frac{d\tau_2}{dw_2}\right|_{0,\hat{w}_2} = \sqrt{\frac{1}{K^{22}(0,\tilde{\tau}_2(0),\tilde{\tau}_3(0)) + K^{23}(0,\tilde{\tau}_2(0),\tilde{\tau}_3(0))\tilde{\tau}_3^2(0,\tilde{\tau}_2(0))}}.$$

We want to compare the approximation results with normal approximation. The mean vector $(E[\bar{Y}_1|\bar{Y}_3 = \bar{y}_3], E[\bar{Y}_2|\bar{Y}_3 = \bar{y}_3])$ for multivariate normal distribution is $(\frac{\bar{y}_3}{3}, \frac{\bar{y}_3}{3})$ by symmetry. The covariance matrix of $(\bar{Y}_1, \bar{Y}_2, \bar{Y}_3)$ can be calculated and the result is

$$V = \begin{pmatrix} \frac{1}{n} & 0 & \frac{1}{n} \\ 0 & \frac{1}{n} & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \frac{3}{n} \end{pmatrix}$$

It can be verified that for normal approximation, the covariance matrix for $\bar{Y}_1|\bar{Y}_3 = \bar{y}_3$ and $\bar{Y}_2|\bar{Y}_3 = \bar{y}_3$ is

$$(V^{-1}_{[(1,2),(1,2)]})^{-1} = \begin{pmatrix} 2n & n & -n \\ n & 2n & -n \\ -n & -n & n \end{pmatrix}^{-1}_{[(1,2),(1,2)]} = \begin{pmatrix} 2n & n \\ n & 2n \end{pmatrix}^{-1} = \begin{pmatrix} \frac{2}{3n} & -\frac{1}{3n} \\ -\frac{1}{3n} & \frac{2}{3n} \end{pmatrix},$$

where $[(1,2),(1,2)]$ denotes the submatrix containing the first two rows and first two columns.

The results for approximating $P(\bar{Y}_1 \geq \bar{y}_1, \bar{Y}_2 \geq \bar{y}_2|\bar{Y}_3 = \bar{y}_3)$, when $n = 10$, are shown below in Table 3. The case that $\bar{y}_1 = 2.0$, $\bar{y}_2 = 2.5$ and $\bar{y}_3 = 7.0$ is the special case that $\tilde{\tau}_2(0) = 0$, and hence $\hat{w}_2 = 0$, as we discussed in Section 3, and is omitted here. The cases that $\bar{y}_1 = 2.0$, $\bar{y}_2 = 3.0$ and $\bar{y}_3 = 7.0$, and $\bar{y}_1 = 2.0$, $\bar{y}_2 = 2.5$ and $\bar{y}_3 = 6.5$, are the cases that $\hat{w}_1 = 0$ and are also omitted. The exact values are computed by [Mathematica 5.0 2005].

Table 4.1: Results of saddlepoint approximation compared with bivariate normal approximation in the conditional continuous case.

$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	P. approx.	N. approx.	Exact	Relative Error
2.0	2.0	7.0	4.42×10^{-1}	8.04×10^{-2}	4.38×10^{-1}	0.91%
2.5	2.5	7.0	6.25×10^{-2}	2.04×10^{-2}	6.32×10^{-2}	−1.11%
2.5	3.0	7.0	8.00×10^{-3}	4.14×10^{-5}	8.54×10^{-3}	−6.32%
3.0	3.0	7.0	3.02×10^{-4}	1.00×10^{-8}	3.46×10^{-4}	−12.7%
2.0	2.0	6.5	2.93×10^{-1}	1.16×10^{-1}	2.91×10^{-1}	0.69%
2.0	3.0	6.5	1.09×10^{-2}	6.48×10^{-5}	1.14×10^{-2}	−4.39%
2.5	2.5	6.5	1.49×10^{-2}	6.96×10^{-4}	1.56×10^{-2}	−4.49%
2.5	3.0	6.5	5.25×10^{-4}	1.57×10^{-7}	6.09×10^{-4}	−13.8%
3.0	3.0	6.5	9.63×10^{-7}	3.67×10^{-12}	1.10×10^{-6}	12.5%

4.3 Unit lattice distributions

As in the unconditional case, when (X_1, X_2, X_3) is integer lattice random variable, using the Fourier inversion formula and using summation instead of integration, we have the following:

$$\frac{n}{(2\pi i)^3}\int_{\mathbf{c}-i\infty}^{\mathbf{c}+i\infty}\frac{\exp(n[K(\tau_1,\tau_2,\tau_3)-\tau_1\bar{x}_1^*-\tau_2\bar{x}_2^*-\tau_3\bar{x}_3^*])}{2\sinh(\frac{\tau_1}{2})2\sinh(\frac{\tau_2}{2})}d\boldsymbol{\tau}/f_{\bar{X}_3}(\bar{x}_3), \tag{4.3.1}$$

where $\bar{x}_1^* = \bar{x}_1 - \frac{1}{2n}$, $\bar{x}_2^* = \bar{x}_2 - \frac{1}{2n}$, $\bar{x}_3^* = \bar{x}_3 - \frac{1}{2n}$. Again, the definitions of $\mathbf{w}$ and τ are same as defined in Section 4.2, except that $\bar{x}_1$, $\bar{x}_2$ and $\bar{x}_3$ should be replaced by $\bar{x}_1^*$, $\bar{x}_2^*$ and $\bar{x}_3^*$.

A counterpart of Lemma 4.1.1 for discrete case also exists.

Lemma 4.3.1.

$$\begin{aligned}&\frac{n}{(2\pi i)^3}\int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty}\frac{\exp(n[\frac{1}{2}w_1^2+\frac{1}{2}w_2^2+\frac{1}{2}w_3^2-\hat{w}_1w_2-\hat{w}_2w_2-\hat{w}_3w_3])}{2\sinh(\frac{\tau_1(w_1)}{2})2\sinh(\frac{\tau_2(w_1,w_2)}{2})}\frac{d\tau_1}{dw_1}\frac{d\tau_2}{dw_2}\frac{d\tau_3}{dw_3}d\mathbf{w}\\ =&A\cdot\left(1+O(\frac{1}{n})\right),\end{aligned} \tag{4.3.2}$$

where

$$\begin{aligned}A=&\frac{n}{(2\pi i)^3}\int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty}\frac{\exp(n[\frac{1}{2}w_1^2+\frac{1}{2}w_2^2+\frac{1}{2}w_3^2-\hat{w}_1w_2-\hat{w}_2w_2-\hat{w}_3w_3])}{2\sinh(\frac{\tau_1(w_1)}{2})2\sinh(\frac{\tau_2(w_1,w_2)}{2})}\cdot\\ &\left.\frac{d\tau_1}{dw_1}\frac{d\tau_2}{dw_2}\frac{d\tau_3}{dw_3}\right|_{w_1,w_2,\hat{w}_3}d\mathbf{w}\end{aligned} \tag{4.3.3}$$

Again, we can decompose A into 4 parts of the same form as in Section 4.1 with the only difference that

$$G(\boldsymbol{\tau}) = \left.\left(\frac{w_1}{2\sinh(\tau_1/2)}\frac{d\tau_1}{dw_1}\right)\left(\frac{w_2-\tilde{w}_2(w_1)}{2\sinh(\tau_2/2)}\frac{d\tau_2}{dw_2}\right)\prod_{j=3}^{d}\frac{d\tau_j}{dw_j}\right|_{(w_1,w_2,\hat{\mathbf{w}}_{-2})}. \tag{4.3.4}$$

Formula (4.1.6), (4.1.8) and (4.1.10) still hold, but here

$$H(0,\hat{w}_2) = \frac{1}{2\sinh(\tilde{\tau}_2(0)/2)}\left.\frac{d\tau_2}{dw_2}\right|_{0,\hat{w}_2}\left.\prod_{j=3}^{d}\frac{d\tau_j}{dw_j}\right|_{(0,\hat{w}_2,\hat{\mathbf{w}}_{-2})} - \frac{1}{\hat{w}_2}\left.\prod_{j=3}^{d}\frac{d\tau_j}{dw_j}\right|_{(0,0,\hat{\mathbf{w}}_{-2})}, \tag{4.3.5}$$

and

$$h(z) = \frac{1}{2\sinh(\tau_1(z)/2)}\left.\frac{d\tau_1}{dw_1}\right|_{z}\left.\prod_{j=3}^{d}\frac{d\tau_j}{dw_j}\right|_{(w_1,\tilde{w}_2(w_1),\hat{\mathbf{w}}_{-2})} - \frac{1}{z}\left.\prod_{j=3}^{d}\frac{d\tau_j}{dw_j}\right|_{(0,0,\hat{\mathbf{w}}_{-2})} \tag{4.3.6}$$

From (4.1.1), Lemma 4.1.1 and Lemma 4.3.1, we conclude that

Theorem 4.3.1.

$$P(\mathbf{T}_2 \geq \mathbf{t}_2 | \mathbf{T}_{-2} = \mathbf{t}_{-2}) \sim (I^{\emptyset} + I^{\{2\}} + I^{\{1\}})/f_{\mathbf{T}_{-2}}(\mathbf{t}_{-2}), \qquad (4.3.7)$$

where $I^{\emptyset}$, $I^{\{2\}}$ *and* $I^{\{1\}}$ *can be obtained by (4.1.6), (4.1.8), (4.1.9), (4.1.10) and (4.1.11) for continuous variables and (4.1.6), (4.1.8), (4.1.10), (4.3.5)and (4.3.6) for unit lattice variables.*

4.4 An example arising from variables confined to unit lattice

This example was used in [Kolassa 2003] and [Kolassa 2004], which refers to data presented by [Stokes *et al.* 1995]. The data consist of 63 case-control pairs of women with endometrial cancer. The relationship between the occurrence of endometrial cancer and explanatory variables including gall bladder disease, hypertension, and non-estrogen drug use, is modeled with logistic regression. Stokes ([Stokes *et al.* 1995]) noted that the likelihood for these data is equivalent to that of a logistic regression in which the units of observation are the matched pairs, the explanatory variables are those of the case member minus those of the control member, and the response variable is 1.

The number of pairs with each configuration of differences of the three variables are shown in Table 4. Let $\mathbf{z}_j, j = 1, 2, \ldots, 63$ denote the differences of covariates between

Table 4.2: Differences between cases and controls for endometrial cancer data.

Gall bladder disease	-1	-1	-1	0	0	0	0	0
Hypertension	-1	0	1	-1	-1	0	0	1
Non-estrogen drug use	0	-1	0	-1	0	0	1	0
Number of pairs	1	1	1	2	6	14	10	12
Gall bladder disease	0	1	1	1	1	1	1	1
Hypertension	1	-1	-1	0	0	0	1	1
Non-estrogen drug use	1	0	1	-1	0	1	0	1
Number of pairs	4	3	1	1	4	1	1	1

cases and controls as listed in Table 4. Consider the situation under null hypothesis, where the linear coefficients are zero. Let $\mathbf{Z}_j, j = 1, 2, \cdots, 63$ be the random vectors that take value $\mathbf{z}_j$ with a probability of $\frac{1}{2}$ and $\mathbf{0}$ with a probability of $\frac{1}{2}$. Let $\mathbf{Z}$ be matrix whose rows are $\mathbf{Z}_j$ and $\mathbf{T} = \mathbf{Z}'\mathbf{1}$, for $\mathbf{1}$ a column vector with dimension 63. Then

$K(\boldsymbol{\tau}) = \sum_j m_j[\log(\frac{1+\exp(\mathbf{z}_j\boldsymbol{\tau})}{2})]$. Kolassa ([Kolassa 2004]) tested association of hypertension or non-estrogen drug use with an increase in endometrial cancer, conditional on the sufficient statistic value associated with gall bladder disease, or more formally

$$H_0 : \beta_2 = \beta_3 = 0$$
$$H_a : \beta_2 > 0 \text{ or } \beta_3 > 0,$$

where $\boldsymbol{\beta}$ are regression coefficients corresponding to the three covariates. The test statistic was $S = \min(P(T_2 \geq t_2|T_1 = t_1, T_3 = t_3), P(T_3 \geq t_3|T_1 = t_1, T_2 = t_2))$. We could then compute the level α, and the corresponding rectangular critical region. After that, we could compute the multivariate probability of the region, which required evaluating the quantity $P(T_2 \geq 10 \text{ or } T_3 \geq 13|T_1 = 9)$ for $\mathbf{T} = (T_1, T_2, T_3)$, By Boole's law, this probability can be computed by

$$P(T_2 \geq 10|T_1 = 9) + P(T_3 \geq 13|T_1 = 9) - P(T_2 \geq 10, T_3 \geq 13|T_1 = 9).$$

The global minimum of $K(\tau_1, \tau_2, \tau_3) - \tau_1\bar{t}_1^* - \tau_2\bar{t}_2^* - \tau_3\bar{t}_3^*$ can be obtained by solving

$$\begin{cases} K^1(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j1}\frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{1+\exp(\mathbf{z_j}\boldsymbol{\tau})} = \bar{t}_1^* \\ K^2(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j2}\frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{1+\exp(\mathbf{z_j}\boldsymbol{\tau})} = \bar{t}_2^* \\ K^3(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j3}\frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{1+\exp(\mathbf{z_j}\boldsymbol{\tau})} = \bar{t}_3^* \end{cases}, \tag{4.4.1}$$

which has to be solved using the multivariate Newton-Raphson method. Given τ_1 and τ_2, $\tilde{\tau}_3(\tau_1, \tau_2)$ is obtained by solving $K^3(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j3}\frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{1+\exp(\mathbf{z_j}\boldsymbol{\tau})} = \bar{t}_3^*$ for τ_3, which can be solved using the Newton-Raphson method. We also need $\tilde{\tau}_3^2(\tau_1, \tau_2)$. Taking the derivative of the above equation with respect to τ_2, we have

$$K^{23}(\tau_1, \tau_2, \tilde{\tau}_3(\tau_1, \tau_2)) + K^{33}(\tau_1, \tau_2, \tilde{\tau}_3(\tau_1, \tau_2))\tau_3^2(\tau_1, \tau_2) = 0,$$

by which we know that $\tau_3^2(\tau_1, \tau_2) = -K^{23}(\tau_1, \tau_2, \tilde{\tau}_3(\tau_1, \tau_2))/K^{33}(\tau_1, \tau_2, \tilde{\tau}_3(\tau_1, \tau_2))$. We also need to compute $\tilde{\tau}_2(\tau_1)$ and $\tilde{\tau}_3(\tau_1)$. They are obtained by solving the following equation system in terms of τ_1

$$\begin{cases} K^2(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j2}\frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{1+\exp(\mathbf{z_j}\boldsymbol{\tau})} = \bar{t}_2^* \\ K^3(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j3}\frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{1+\exp(\mathbf{z_j}\boldsymbol{\tau})} = \bar{t}_3^*, \end{cases}$$

which can be calculated by the Newton-Raphson method. Differentiate the above equations with respect to τ_1 and use the results $\tilde{\tau}_2 = \tilde{\tau}_2(\tau_1)$, $\tilde{\tau}_3 = \tilde{\tau}_3(\tau_1)$ we computed above, to obtain

$$\begin{cases} K^{12}(\tau_1, \tilde{\tau}_2, \tilde{\tau}_3) + K^{22}(\tau_1, \tilde{\tau}_2, \tilde{\tau}_3)\tilde{\tau}_2'(\tau_1) + K^{23}(\tau_1, \tilde{\tau}_2, \tilde{\tau}_3)\tilde{\tau}_3'(\tau_1) = 0 \\ K^{13}(\tau_1, \tilde{\tau}_2, \tilde{\tau}_3) + K^{23}(\tau_1, \tilde{\tau}_2, \tilde{\tau}_3)\tilde{\tau}_2'(\tau_1) + K^{33}(\tau_1, \tilde{\tau}_2, \tilde{\tau}_3)\tilde{\tau}_3'(\tau_1) = 0 \end{cases} . \qquad (4.4.2)$$

This is a linear equation system and can be solved easily.

The second derivatives of $K(\cdot, \cdot, \cdot)$ can be calculated by the following formula:

$$\begin{cases} K^{11}(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j1}^2 \frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{(1+\exp(\mathbf{z_j}\boldsymbol{\tau}))^2} \\ K^{12}(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j1} z_{j2} \frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{(1+\exp(\mathbf{z_j}\boldsymbol{\tau}))^2} \\ K^{13}(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j1} z_{j3} \frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{(1+\exp(\mathbf{z_j}\boldsymbol{\tau}))^2} \\ K^{22}(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j2}^2 \frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{(1+\exp(\mathbf{z_j}\boldsymbol{\tau}))^2} \\ K^{23}(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j2} z_{j3} \frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{(1+\exp(\mathbf{z_j}\boldsymbol{\tau}))^2} \\ K^{33}(\tau_1, \tau_2, \tau_3) = \sum_j m_j z_{j3}^2 \frac{\exp(\mathbf{z_j}\boldsymbol{\tau})}{(1+\exp(\mathbf{z_j}\boldsymbol{\tau}))^2} \end{cases} . \qquad (4.4.3)$$

Again, we need to compute $\tau_1(w_1)$. Substitute the previous results into

$$-\frac{1}{2}(w_1 - \hat{w}_1)^2 = K(\hat{\tau}_1, \hat{\tau}_2, \hat{\tau}_3) - \hat{\tau}_1\bar{t}_1^* - \hat{\tau}_2\bar{t}_2^* - \hat{\tau}_3\bar{t}_3^* - [K(\tau_1, \tilde{\tau}_2(\tau_1), \tilde{\tau}_3(\tau_1)) - \tau_1\bar{t}_1^* - \tilde{\tau}_2(\tau_1)\bar{t}_2^* - \tilde{\tau}_3(\tau_1)\bar{t}_3^*]$$

and rearrange the terms, the resulting equation can then be solved by the Newton-Raphson method.

The results for approximating $P(T_2 \geq 10, T_3 \geq 13 | T_1 = 9)$ compared to those listed in [Kolassa 2004] are shown in Table 5, where "N. app." stands for normal approximation, "E. app." stands for Edgeworth approximation, "K. app" stands for the approximation presented in [Kolassa 2004] and "P. app" is the proposed approximation. Approximation results of $P(T_2 \geq t_2, T_3 \geq t_3 | T_1 = 9)$ for other values of t_2 and t_3 are also listed in the table. We can see that the proposed method achieves better results than other methods, except for the [Kolassa 2004] method, which is far more complicated computationally.

Table 4.3: Endometrial cancer results for some (t_2, t_3) instances

Method	$(10, 13)$	$(9, 12)$	$(8, 11)$	$(7, 10)$	$(6, 9)$
N. app.	3.50×10^{-4}	1.78×10^{-3}	7.26×10^{-3}	2.39×10^{-2}	6.39×10^{-2}
E. app.	3.31×10^{-4}	1.72×10^{-3}	7.13×10^{-3}	2.37×10^{-2}	6.37×10^{-2}
K. app.	1.51×10^{-4}	1.07×10^{-3}	5.37×10^{-3}	2.01×10^{-2}	5.84×10^{-2}
P. app.	1.62×10^{-4}	1.13×10^{-3}	5.60×10^{-3}	2.08×10^{-2}	6.00×10^{-2}
Exact	1.52×10^{-4}	1.09×10^{-3}	5.48×10^{-3}	2.05×10^{-2}	5.95×10^{-2}

Chapter 5

An alternative multivariate saddlepoint approximation

5.1 Theoretical development

The drawback of the method introduced in Section 3is that the coefficient C of the leading term $I^{\emptyset}$ is not equal to 1. We propose another method that solves the problem.

From (2.0.5b), we know that $\tilde{w}_2(\hat{w}_1)/w_1$ is analytic, and we denote it by $a(w_1)$. Equation (3.1.7) can then be rewritten as

$$I^{\emptyset} = \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - w_1\hat{w}_1 - w_2\hat{w}_2])}{(2\pi i)^2 w_1(w_2 - a(w_1)w_1)} \, d\mathbf{w} \tag{5.1.1}$$

We can approximate $a(w_1)$ by $a(w_1) \sim b_0 + b_1(w_1 - \hat{w}_1)$, where $b_0 = a(\hat{w}_1) - \frac{1}{2}a''(\hat{w}_1)\hat{w}_1^2$ and $b_1 = a'(\hat{w}_1) + \frac{1}{2}a''(\hat{w}_1)\hat{w}_1$. The approximation is justified by the following theorem

Theorem 5.1.1.

$$I^{\emptyset} = \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - w_1\hat{w}_1 - w_2\hat{w}_2])}{(2\pi i)^2 w_1(w_2 - b_0 w_1 - b_1 w_1(w_1 - \hat{w}_1))} \, d\mathbf{w}(1 + O(\frac{1}{n})) \tag{5.1.2}$$

Proof. First we do the same change of variables as in (2.12). We have

$$I^{\emptyset} = \int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{\exp(ng_1(u_1,u_2))}{(2\pi i)^2 u_1 u_2} \, d\mathbf{u},$$

where $g_1(u_1,u_2) = \frac{1}{2}u_1^2 + \frac{1}{2}(u_2 + \tilde{w}_2(u_1))^2 - \hat{w}_1 u_1 - \hat{w}_2(u_2 + \tilde{w}_2(u_1))$. Expand $g_1(u_1,u_2)$ around $(\hat{u}_1,\hat{u}_2)$ up to the third degree, where $\hat{u}_1 = \hat{w}_1$ and $\hat{u}_2 = \hat{w}_2 - \tilde{w}_2(\hat{w}_1)$. We have

$$\begin{aligned} g_1(u_1,u_2) =& \hat{g}_1 + \frac{1}{2}\hat{g}_1^{11}(u_1-\hat{u}_1)^2 + \frac{1}{2}\hat{g}_1^{22}(u_2-\hat{u}_2)^2 + \hat{g}_1^{12}(u_1-\hat{u}_1)(u_2-\hat{u}_2) + \\ & \frac{1}{6}\sum_{i,j,k}\hat{g}_1^{ijk}(u_i-\hat{u}_i)(u_j-\hat{u}_j)(u_k-\hat{u}_k) + \\ & \frac{1}{24}\sum_{i,j,k,l} g_1^{ijkl}(\boldsymbol{\xi})((u_i-\hat{u}_i)(u_j-\hat{u}_j)(u_k-\hat{u}_k)(u_l-\hat{u}_l), \end{aligned}$$

where $\boldsymbol{\xi}$ lies between $\mathbf{u}$ and $\hat{\mathbf{u}}$.

We use the same technique to expand the integrals in (5.1.1) and (5.1.2), and show that they yield the same result to relative order $O(n^{-1})$. Expression (5.1.1) represents the more general case, in which $a(w_1)$ is restricted only to be analytic; (5.1.2) represents the special case in which $a(w_1)$ is linear: $b_0 + b_1(w_1 - \hat{w}_1) = a(\hat{w}_1) - a'(\hat{w}_1)\hat{w}_1 + \frac{1}{2}a''(\hat{w}_1)\hat{w}_1^2 + (a'(\hat{w}_1) + \frac{1}{2}a''(\hat{w}_1))w_1$. Let the corresponding quadratic terms in the exponent be $g_2(v_1, v_2)$, where $v_1 = w_1$ and $v_2 = w_2 - (b_0 + b_1(w_1 - \hat{w}_1)w_1$. The proof then follows the argument of [Kolassa 2003]. We only need to prove that the coefficients $\hat{g}_1$, $\hat{g}_1^{ij}$ and $\hat{g}_1^{ijk}$ coincide with $\hat{g}_2$, $\hat{g}_2^{ij}$ and $\hat{g}_2^{ijk}$. This is true, since $\hat{g}_1 = \hat{g}_2 = -\frac{1}{2}\hat{w}_1^2 - \frac{1}{2}\hat{w}_2^2$, $\hat{g}_1^{11} = \hat{g}_2^{11} = 1 + (\check{w}_2')^2$, $\hat{g}_1^{12} = \hat{g}_2^{12} = \check{w}_2'$, $\hat{g}_1^{22} = \hat{g}_2^{22} = 1$, $\hat{g}_1^{111} = \hat{g}_2^{111} = 3\check{w}_2'\check{w}_2''$, $\hat{g}_1^{112} = \hat{g}_2^{112} = \check{w}_2''$ and $\hat{g}_1^{122} = \hat{g}_2^{122} = \hat{g}_1^{222} = \hat{g}_2^{222} = 0$. □

Let

$$f(b_1) = \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - w_1\hat{w}_1 - w_2\hat{w}_2])}{(2\pi i)^2 w_1(w_2 - b_0w_1 - b_1w_1(w_1 - \hat{w}_1))}\, d\mathbf{w}.$$

Expand $f(b_1)$ using Taylor's theorem, for two terms, i.e.,

$$\begin{aligned} f(b_1) =& f(0) + f'(0)b_1 + \frac{f''(b_1^*)}{2}b_1^2 \\ =& \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - w_1\hat{w}_1 - w_2\hat{w}_2])}{(2\pi i)^2 w_1(w_2 - b_0w_1)}\, d\mathbf{w} \\ &+ b_1 \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - w_1\hat{w}_1 - w_2\hat{w}_2])(w_1 - \hat{w}_1)}{(2\pi i)^2 (w_2 - b_0w_1)^2}\, d\mathbf{w} \\ &+ \frac{b_1^2}{2} \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - w_1\hat{w}_1 - w_2\hat{w}_2])w_1(w_1 - \hat{w}_1)^2}{(2\pi i)^2 (w_2 - b_0w_1 - b_1^*w_1(w_1 - \hat{w}_1))^3}\, d\mathbf{w}, \end{aligned}$$

where $b_1^* \in (0, b_1)$. Change variables to $u_1 = \sqrt{n}w_1$, $u_2 = \sqrt{n}w_2$ and $\hat{u}_1 = \sqrt{n}\hat{w}_1$, $\hat{u}_2 = \sqrt{n}\hat{w}_2$ to obtain

$$\begin{aligned} f(b_1) =& \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - w_1\hat{w}_1 - w_2\hat{w}_2])}{(2\pi i)^2 w_1(w_2 - b_0w_1)}\, d\mathbf{w} \\ &+ b_1 \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - w_1\hat{w}_1 - w_2\hat{w}_2])(w_1 - \hat{w}_1)}{(2\pi i)^2 (w_2 - b_0w_1)^2}\, d\mathbf{w} \\ &+ \frac{b_1^2}{2n} \int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{\exp(\frac{1}{2}u_1^2 + \frac{1}{2}u_2^2 - u_1\hat{u}_1 - u_2\hat{u}_2)(u_1 - \hat{u}_1)^3}{(2\pi i)^2 (u_2 - b_0u_1 - b_1^*u_1(u_1 - \hat{u}_1)/\sqrt{n})^3}\, d\mathbf{u} \\ &+ \frac{b_1^2}{2n} \int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{\exp(\frac{1}{2}u_1^2 + \frac{1}{2}u_2^2 - u_1\hat{u}_1 - u_2\hat{u}_2)\hat{u}_1(u_1 - \hat{u}_1)^2}{(2\pi i)^2 (u_2 - b_0u_1 - b_1^*u_1(u_1 - \hat{u}_1)/\sqrt{n})^3}\, d\mathbf{u} \\ =& I_0 + I_1 + I_2 + I_3. \end{aligned}$$

To obtain I_0, we can now do a change of variables. Let $s_1 = w_1$ and $s_2 = w_2 - b_0 w_1$, then $w_1 = s_1$ and $w_2 = s_2 + b_0 s_1$, and the Jacobian is 1. So we obtain

$$\begin{aligned} I_0 &= \int_{\hat{\mathbf{s}}-i\infty}^{\hat{\mathbf{s}}+i\infty} \frac{\exp(n[\frac{1}{2}s_1^2 + \frac{1}{2}(s_2 + b_0 s_1)^2 - s_1\hat{w}_1 - (s_2 + b_0 s_1)\hat{w}_2])}{(2\pi i)^2 s_1 s_2}\, d\mathbf{s} \\ &= \int_{\hat{\mathbf{s}}-i\infty}^{\hat{\mathbf{s}}+i\infty} \frac{\exp(n[\frac{1}{2}(1+a_0^2)s_1^2 + \frac{1}{2}s_2^2 + b_0 s_1 s_2 - (\hat{w}_1 + b_0\hat{w}_2)s_1 - \hat{w}_2 s_2])}{(2\pi i)^2 s_1 s_2}\, d\mathbf{s}. \end{aligned} \tag{5.1.3}$$

To further simplify the formula, let $t_1 = \sqrt{n}\sqrt{1+b_0^2}s_1$ and $t_2 = \sqrt{n}s_2$. Then $s_1 = \frac{1}{\sqrt{1+b_0^2}\sqrt{n}}t_1$ and $s_2 = t_2/\sqrt{n}$. Then

$$\begin{aligned} I_0 &= \int_{\hat{\mathbf{t}}-i\infty}^{\hat{\mathbf{t}}+i\infty} \frac{\exp(\frac{1}{2}t_1^2 + \frac{1}{2}t_2^2 + \frac{b_0}{\sqrt{1+b_0^2}}t_1 t_2 - \sqrt{n}\frac{\hat{w}_1+b_0\hat{w}_2}{\sqrt{1+b_0^2}}t_1 - \sqrt{n}\hat{w}_2 t_2)}{(2\pi i)^2 t_1 t_2}\, d\mathbf{t} \\ &= \bar{\Phi}(\sqrt{n}\frac{\hat{w}_1 + b_0\hat{w}_2}{\sqrt{1+b_0^2}}, \sqrt{n}\hat{w}_2, \frac{b_0}{\sqrt{1+b_0^2}}). \end{aligned} \tag{5.1.4}$$

Do the same change of variable to I_1, to obtain

$$\begin{aligned} I_1 =& \frac{b_1}{\sqrt{n}(1+b_0^2)} \int_{\hat{\mathbf{t}}-i\infty}^{\hat{\mathbf{t}}+i\infty} \frac{\exp(\frac{1}{2}t_1^2 + \frac{1}{2}t_2^2 + \frac{b_0}{\sqrt{1+b_0^2}}t_1 t_2 - \sqrt{n}\frac{\hat{w}_1+b_0\hat{w}_2}{\sqrt{1+b_0^2}}t_1 - \sqrt{n}\hat{w}_2 t_2)}{(2\pi i)^2 t_2^2} \cdot \\ & (t_1 - \sqrt{n}\sqrt{1+b_0^2}\hat{w}_1)\, d\mathbf{t} \\ =& \frac{b_1}{\sqrt{n}(1+b_0^2)} \int_{\hat{\mathbf{t}}-i\infty}^{\hat{\mathbf{t}}+i\infty} \frac{\exp(\frac{1}{2}t_1^2 + \frac{1}{2}t_2^2 + \rho t_1 t_2 - x t_1 - y t_2)(t_1 - \hat{t}_1)}{(2\pi i)^2 t_2^2}\, d\mathbf{t} \\ =& \frac{b_1}{\sqrt{n}(1+b_0^2)} \int_{\hat{\mathbf{t}}-i\infty}^{\hat{\mathbf{t}}+i\infty} \frac{\exp(\frac{1}{2}t_1^2 + \frac{1}{2}t_2^2 + \rho t_1 t_2 - x t_1 - y t_2)t_1}{(2\pi i)^2 t_2^2}\, d\mathbf{t} \\ & - \frac{\hat{t}_1 b_1}{\sqrt{n}(1+b_0^2)} \int_{\hat{\mathbf{t}}-i\infty}^{\hat{\mathbf{t}}+i\infty} \frac{\exp(\frac{1}{2}t_1^2 + \frac{1}{2}t_2^2 + \rho t_1 t_2 - x t_1 - y t_2)}{(2\pi i)^2 t_2^2}\, d\mathbf{t} \\ =& I_{10} - I_{11}, \end{aligned} \tag{5.1.5}$$

where $x = \sqrt{n}\frac{\hat{w}_1+b_0\hat{w}_2}{\sqrt{1+b_0^2}}$, $y = \sqrt{n}\hat{w}_2$ $\rho = \frac{b_0}{\sqrt{1+b_0^2}}$ and $\hat{t}_1 = \sqrt{n}\sqrt{1+b_0^2}\hat{w}_1$. To calculate I_{10} and I_{11}, we use the following technique. Consider $I_{10} = I_{10}(x, y, \rho)$ and $I_{11} = I_{11}(x, y, \rho)$ as a function of x, y and ρ. We then have

$$I_{10}^{22} = -\frac{b_1}{\sqrt{n}(1+b_0^2)}\phi^1(x, y, \rho)$$

and

$$I_{11}^{22} = \frac{\hat{t}_1 b_1}{\sqrt{n}(1+b_0^2)}\phi(x, y, \rho).$$

Integrate twice, to obtain

$$\begin{aligned} I_{11} =& \frac{\hat{t}_1 b_1}{\sqrt{n}(1+b_0^2)}\phi(x)[\sqrt{1-\rho^2}\phi\left(\frac{y-\rho x}{\sqrt{1-\rho^2}}\right) \\ & -(y-\rho x)\bar{\Phi}(\frac{y-\rho x}{\sqrt{1-\rho^2}})] \end{aligned} \tag{5.1.6}$$

and

$$\begin{aligned} I_{10} =& \frac{b_1}{\sqrt{n}(1+b_0^2)}\phi(x)[\sqrt{1-\rho^2}\phi\left(\frac{y-\rho x}{\sqrt{1-\rho^2}}\right)x \\ & -(\rho+xy-\rho x^2)\bar{\Phi}(\frac{y-\rho x}{\sqrt{1-\rho^2}}))]. \end{aligned} \tag{5.1.7}$$

Therefore,

$$\begin{aligned} I_1 =& I_{10} - I_{11} \\ =& \frac{b_1\phi(x)}{\sqrt{n}(1+b_0^2)}[\sqrt{1-\rho^2}\phi\left(\frac{y-\rho x}{\sqrt{1-\rho^2}}\right)(x-\hat{t}_1) \\ & -(\rho-\hat{t}_1 y-\rho x^2+\hat{t}_1\rho x+xy)\bar{\Phi}(\frac{y-\rho x}{\sqrt{1-\rho^2}}))]. \end{aligned} \tag{5.1.8}$$

To compute I_0 and I_1, we need x, y, ρ, b_0, b_1 and $\hat{t}$. We have

$$\begin{aligned} b_0 &= a(\hat{w}_1) - \frac{1}{2}a''(\hat{w}_1)\hat{w}_1^2 = \breve{w}_2' - \frac{1}{2}\breve{w}_2''\hat{w}_1 \\ b_1 &= a'(\hat{w}_1) + \frac{1}{2}a''(\hat{w}_1)\hat{w}_1 = \frac{1}{2}\breve{w}_2'', \end{aligned}$$

where we can use (3.1.3d) to obtain

$$\begin{aligned} \breve{w}_2'' =& [(K^{11}(\hat{\tau}_1,0) - K^{11}(\hat{\tau}_1,\hat{\tau}_2) - K^{12}(\hat{\tau}_1,\hat{\tau}_2)\tilde{\tau}_2'(\hat{\tau}_1))\left(\left.\frac{d\tau_1}{dw_1}\right|_{\hat{w}_1}\right)^2 + \\ & (K^1(\hat{\tau}_1,0)-\bar{x}_1)\left.\frac{d^2\tau_1}{dw_1^2}\right|_{\hat{w}_1} - (\breve{w}_2')^2]/(\breve{w}_2-\hat{w}_2), \end{aligned} \tag{5.1.9}$$

use (3.1.3c) to obtain

$$\begin{aligned} \left.\frac{d^2\tau_1}{dw_1^2}\right|_{\hat{w}_1} =& -[(K^{111}(\hat{\tau}_1,\hat{\tau}_2)+2K^{112}(\hat{\tau}_1,\hat{\tau}_2)\tilde{\tau}_2'(\hat{\tau}_1)+K^{122}(\hat{\tau}_1,\hat{\tau}_2)\tilde{\tau}_2'(\hat{\tau}_1)^2 + \\ & K^{12}(\hat{\tau}_1,\hat{\tau}_2)\tilde{\tau}_2''(\hat{\tau}_1))(\left.\frac{d\tau_1}{dw_1}\right|_{\hat{w}_1})^2]/(3(K^{11}(\hat{\tau}_1,\hat{\tau}_2)+K^{12}(\hat{\tau}_1,\hat{\tau}_2)\tilde{\tau}_2'(\hat{\tau}_1))), \end{aligned} \tag{5.1.10}$$

and use equation $K^2(\tau_1,\tilde{\tau}_2(\tau_1)) = \bar{x}_2$ to obtain

$$\tilde{\tau}_2''(\hat{\tau}_1) = -[K^{112}(\hat{\tau}_1,\hat{\tau}_2)+2K^{122}(\hat{\tau}_1,\hat{\tau}_2)\tilde{\tau}_2'(\hat{\tau}_1)+K^{222}(\hat{\tau}_1,\hat{\tau}_2)\tilde{\tau}_2'(\hat{\tau}_1)^2]/K^{22}(\hat{\tau}_1,\hat{\tau}_2). \tag{5.1.11}$$

The special case that $\hat{w}_1 = 0$ also involves partial derivatives of some functions up to the second or third degree, which are algebraically complicated and therefore omitted here. Other quantities can be obtained accordingly.

The remaining terms I_2 and I_3 are of $O(n^{-1})$ as shown and can be omitted. This error is relative, times a linear term in $\hat{w}_1$, as can be seen by the same reparameterization and use of Kolassa (2003), as before.

5.2 Reflexivity

In (2.0.1), one of the restrictions is that $\mathbf{c}$ in the route of path should be greater than zero. In general this restriction require us to have $\hat{\boldsymbol{\tau}} > 0$. Suppose in a bivariate setting, $\hat{\tau}_1 < 0$ and $\hat{\tau}_2 > 0$. Noting that, for continuous distribution,

$$P(\bar{\mathbf{X}}_1 \geq \bar{\mathbf{x}}_1, \bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2) = P(\bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2) - P((-\bar{\mathbf{X}}_1) \geq (-\bar{\mathbf{x}}_1), \bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2),$$

we can circumvent the problem with $P(\bar{\mathbf{X}}_1 \geq \bar{\mathbf{x}}_1, \bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2)$ by approximating $P(\bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2)$ and $P((-\bar{\mathbf{X}}_1) \geq (-\bar{\mathbf{x}}_1), \bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2)$. It can be easily checked that the latter bivariate approximation satisfies the positivity restriction. Similarly, in the case of both $\hat{\tau}_1 < 0$ and $\hat{\tau}_2 < 0$, we have to apply the previous detour twice. We now show that we can ignore this restriction for this approximation method and apply the method directly by the following theorem.

Theorem 5.2.1. *Let $SA1(\bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2)$ be the approximation of Lugannani and Rice. Let $SA2(\bar{\mathbf{X}}_1 \geq \bar{\mathbf{x}}_1, \bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2)$ be the saddlepoint approximation in this chapter. Then*

$$SA2(\bar{\mathbf{X}}_1 \geq \bar{\mathbf{x}}_1, \bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2) = SA1(\bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2) - SA2((-\bar{\mathbf{X}}_1) \geq (-\bar{\mathbf{x}}_1), \bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2),$$

Proof. Consider two random vectors, i.e., $(\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2)$ and $(-\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2)$, respectively. To distinguish them, we append a subscript $_{(1)}$ for the former and $_{(2)}$ for the latter to other notation, when necessary. For instance, we use $\bar{\mathbf{X}}_{1(1)}$ and $\bar{\mathbf{x}}_{1(1)}$ for $\bar{\mathbf{X}}_1$ and $\bar{\mathbf{x}}_1$, and

$\bar{\mathbf{X}}_{1(2)}$ and $\bar{\mathbf{x}}_{1(2)}$ for $-\bar{\mathbf{X}}_1$ and $-\bar{\mathbf{x}}_1$. Then, directly,

$$K_{(2)}(\tau_1,\tau_2) = K_{(1)}(-\tau_1,\tau_2), \tag{5.2.1a}$$

$$K^1_{(2)}(\tau_1,\tau_2) = -K^1_{(1)}(-\tau_1,\tau_2), K^2_{(2)}(\tau_1,\tau_2) = K^2_{(1)}(-\tau_1,\tau_2)$$

$$K^{11}_{(2)}(\tau_1,\tau_2) = K^{11}_{(1)}(-\tau_1,\tau_2), K^{12}_{(2)}(\tau_1,\tau_2) = -K^{12}_{(1)}(-\tau_1,\tau_2), K^{22}_{(2)}(\tau_1,\tau_2) = K^{22}_{(1)}(-\tau_1,\tau_2)$$

$$K^{ijk}_{(2)}(\tau_1,\tau_2) = -K^{ijk}_{(1)}(-\tau_1,\tau_2), \text{ if the number of 1's in index } ijk \text{ is odd}$$

$$K^{ijk}_{(2)}(\tau_1,\tau_2) = K^{ijk}_{(1)}(-\tau_1,\tau_2), \text{ if the number of 1's in index } ijk \text{ is even}$$

$$\hat{\tau}_{1(2)} = -\hat{\tau}_{1(1)}, \hat{\tau}_{2(2)} = \hat{\tau}_{2(1)}, \text{ so we let } \hat{\tau}_2 \text{ be the common value,} \tag{5.2.1b}$$

$$\tilde{\tau}_{2(2)}(\tau_1) = \tilde{\tau}_{2(1)}(-\tau_1), \tag{5.2.1c}$$

$$\hat{w}_{1(2)} = -\hat{w}_{1(1)}, \hat{w}_{2(2)} = \hat{w}_{2(1)}, \text{ and we } \hat{w}_2 \text{ denote the common value,} \tag{5.2.1d}$$

$$\check{w}_{2(2)} = \check{w}_{2(1)}, \check{w}'_{2(2)} = -\check{w}'_{2(1)}, \check{w}''_{2(2)} = \check{w}''_{2(1)}, \tag{5.2.1e}$$

$$b_{0(2)} = -b_{0(1)}, b_{1(2)} = b_{1(1)}, \tag{5.2.1f}$$

$$x_{(2)} = -x_{(1)}, y_{(2)} = y_{(1)}, \rho_{(2)} = -\rho_{(1)} \text{ and we still use the notation } y, \tag{5.2.1g}$$

$$\hat{t}_{1(2)} = -\hat{t}_{2(1)}, \tag{5.2.1h}$$

$$\tau_{1(2)}(w_1) = -\tau_{1(2)}(-w_1). \tag{5.2.1i}$$

Relation (5.2.1a) is true, since

$$K_{(2)}(\tau_1,\tau_2) = E\exp(\tau_1 X_{1(2)} + \tau_2 X_2) = E\exp((-\tau_1)X_{1(1)} + \tau_2 X_2) = K_{(1)}(-\tau_1,\tau_2),$$

and the rest follows easily. Relation (5.2.1b) is true because of (5.2.1a),

$$\begin{cases} K^1_{(1)}(\hat{\tau}_{1(1)}, \hat{\tau}_{2(2)}) = \bar{x}_{1(1)} \\ K^2_{(1)}(\hat{\tau}_{1(1)}, \hat{\tau}_{2(2)}) = \bar{x}_2, \end{cases}$$

and

$$\begin{cases} K^1_{(2)}(\hat{-\tau}_{1(1)}, \hat{\tau}_{2(2)}) = \bar{x}_{1(2)} \\ K^2_{(2)}(\hat{-\tau}_{1(1)}, \hat{\tau}_{2(2)}) = \bar{x}_2. \end{cases}$$

By definition and (5.2.1a),

$$\bar{x}_2 = K^2_{(2)}(\tau_1, \tilde{\tau}_{2(2)}(\tau_1)) = K^2_{(1)}(-\tau_1, \tilde{\tau}_{2(2)}(\tau_1)),$$

and therefore (5.2.1c) holds. Relation (5.2.1d) follows from definition (3.1.4a), (3.1.4b) and (5.2.1a), (5.2.1b). The first equation of (5.2.1e) follows from (3.2.3), (5.2.1a), (5.2.1b) and (5.2.1d). The second comes from (3.2.4), (3.2.5) and (5.2.1a). The third comes from (5.1.9)-(5.1.11). (5.2.1f) follows directly from the definition, (5.2.1d) and (5.2.1e). Using this result, relation (5.2.1g) and (5.2.1h) are obvious. To show that (5.2.1i) is true, first use (3.1.3c) but instead of w_1, we use $-w_1$ to obtain

$$-\frac{1}{2}(-w_1 - \hat{w}_{1(1)})^2 = K_{(1)}(\hat{\tau}_{1(1)}, \hat{\tau}_2) - \hat{\tau}_{1(1)}\bar{x}_{1(1)} - \hat{\tau}_2\bar{x}_2 - \\ (K_{(1)}(\tau_{1(1)}(-w_1), \tilde{\tau}_{2(1)}(\tau_{1(1)}(-w_1))) - \tau_{1(1)}(-w_1)\bar{x}_{1(1)} - \tilde{\tau}_{2(1)}(\tau_{1(1)}(-w_1))\bar{x}_2).$$

Substitute facts (5.2.1a)–(5.2.1d) into the above equation to obtain

$$-\frac{1}{2}(w_1 - \hat{w}_{1(2)})^2 = K_{(2)}(\hat{\tau}_{1(2)}, \hat{\tau}_2) - \hat{\tau}_{1(2)}\bar{x}_{1(2)} - \hat{\tau}_2\bar{x}_2 - \\ (K_{(2)}(-\tau_{1(1)}(-w_1), \tilde{\tau}_{2(2)}(-\tau_{1(1)}(-w_1))) - (-\tau_{1(1)}(-w_1)\bar{x}_{1(2)}) - \tilde{\tau}_{2(1)}(-\tau_{1(2)}(-w_1))\bar{x}_2),$$

which shows that (5.2.1i) is true.

To prove the theorem, we will show that

$$\bar{\Phi}(\sqrt{n}\hat{w}) + \frac{\phi(\sqrt{n}\hat{w})}{\sqrt{n}}\left(\frac{1}{\hat{\tau}\sqrt{K''(\hat{\tau})}} - \frac{1}{\hat{w}}\right) = (I^{\emptyset}_{(1)} + I^{\{1\}}_{(1)} + I^{\{2\}}_{(1)}) + (I^{\emptyset}_{(2)} + I^{\{1\}}_{(2)} + I^{\{2\}}_{(2)}),$$

where the left hand side is $SA1(\bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2)$, and the right hand side is the sum of the two bivariate saddlepoint approximations. Here, $\hat{\tau}$ and $\hat{w}$ are quantities from the univariate saddlepoint approximation $SA1(\bar{\mathbf{X}}_2 \geq \bar{\mathbf{x}}_2)$. We can see that the cumulant generating function for X_2 is just $K_{(1)}(0, \tau_2)$. Then, by definition of $\tilde{\tau}(\cdot)$, we have $\hat{\tau} = \tilde{\tau}_{(1)}(0)$, and therefore by (3.1.3b), $\hat{w} = \hat{w}_2$. Thus, we can rewrite the above equation as

$$\bar{\Phi}(\sqrt{n}\hat{w}_2) + \frac{\phi(\sqrt{n}\hat{w}_2)}{\sqrt{n}}\left(\frac{1}{\tilde{\tau}_{(1)}(0)\sqrt{K^{22}_{(1)}(0, \tilde{\tau}_{(1)}(0))}} - \frac{1}{\hat{w}_2}\right) = \\ (I^{\emptyset}_{(1)} + I^{\emptyset}_{(2)}) + (I^{\{1\}}_{(1)} + I^{\{1\}}_{(2)}) + (I^{\{2\}}_{(1)} + I^{\{2\}}_{(2)}).$$

First, we have $\bar{\Phi}(\sqrt{n}\hat{w}_2) = I^{\emptyset}_{(1)} + I^{\emptyset}_{(2)}$. As shown in the last section, $I^{\emptyset}_{(1)} + I^{\emptyset}_{(2)} = (I_{0(1)} + I_{0(2)}) + (I_{10(1)} + I_{10(2)}) - (I_{11(1)} + I_{11(2)})$. Using the facts (5.2.1f)–(5.2.1h) and the simple fact $\phi(x)=\phi(-x)$, we have $I_{10(1)}+I_{10(2)}=I_{11(1)}+I_{11(2)}=0$. And

$$
\begin{aligned}
&I_{0(1)}+I_{0(2)}\\
=&\bar{\Phi}(x_{(1)},y,\rho_{(1)})+\bar{\Phi}(x_{(2)},y,\rho_{(2)})=\bar{\Phi}(x_{(1)},y,\rho_{(1)})+\bar{\Phi}(-x_{(1)},y,-\rho_{(1)})\\
=&\int_y^\infty \frac{\phi(u_2)}{\sqrt{1-\rho^2}}\left(\int_{x_{(1)}}^\infty \phi(\frac{u_1-\rho u_2}{\sqrt{1-\rho}})\,du_1+\int_{-x_{(1)}}^\infty \phi(\frac{u_1+\rho u_2}{\sqrt{1-\rho}})\,du_1\right)du_2\\
=&\int_y^\infty \frac{\phi(u_2)}{\sqrt{1-\rho^2}}\left(\int_{x_{(1)}}^\infty \phi(\frac{u_1-\rho u_2}{\sqrt{1-\rho}})\,du_1+\int_{-\infty}^{x_{(1)}} \phi(\frac{u_1-\rho u_2}{\sqrt{1-\rho}})\,du_1\right)du_2\\
=&\int_y^\infty \phi(u_2)\int_{-\infty}^\infty \phi(\frac{u_1-\rho u_2}{\sqrt{1-\rho}})/\sqrt{1-\rho^2}\,du_1\,du_2\\
=&\int_y^\infty \phi(u_2)\,du_2=\bar{\Phi}(y),
\end{aligned}
$$

and $y=\sqrt{n}\hat{w}_2$ as defined. Next we show that

$$
\frac{\phi(\sqrt{n}\hat{w}_2)}{\sqrt{n}}\left(\frac{1}{\tilde{\tau}_{(1)}(0)\sqrt{K^{22}_{(1)}(0,\tilde{\tau}_{(1)}(0))}}-\frac{1}{\hat{w}_2}\right)=I^{\{2\}}_{(1)}+I^{\{2\}}_{(2)}.
$$

By (3.1.15), we have

$$
\begin{aligned}
I^{\{2\}}_{(1)}+I^{\{2\}}_{(2)}=&\frac{\phi(\sqrt{n}\hat{w}_2)}{\sqrt{n}}\left(\frac{1}{\tilde{\tau}_{(1)}(0)\sqrt{K^{22}_{(1)}(0,\tilde{\tau}_{(1)}(0))}}-\frac{1}{\hat{w}_2}\right)\bar{\Phi}(\sqrt{n}\hat{w}_{1(1)})+\\
&\frac{\phi(\sqrt{n}\hat{w}_2)}{\sqrt{n}}\left(\frac{1}{\tilde{\tau}_{(2)}(0)\sqrt{K^{22}_{(2)}(0,\tilde{\tau}_{(2)}(0))}}-\frac{1}{\hat{w}_2}\right)\bar{\Phi}(\sqrt{n}\hat{w}_{1(2)})\\
=&\frac{\phi(\sqrt{n}\hat{w}_2)}{\sqrt{n}}\left(\frac{1}{\tilde{\tau}_{(1)}(0)\sqrt{K^{22}_{(1)}(0,\tilde{\tau}_{(1)}(0))}}-\frac{1}{\hat{w}_2}\right)\bar{\Phi}(\sqrt{n}\hat{w}_{1(1)})+\\
&\frac{\phi(\sqrt{n}\hat{w}_2)}{\sqrt{n}}\left(\frac{1}{\tilde{\tau}_{(1)}(0)\sqrt{K^{22}_{(1)}(0,\tilde{\tau}_{(1)}(0))}}-\frac{1}{\hat{w}_2}\right)\bar{\Phi}(-\sqrt{n}\hat{w}_{1(1)})\\
=&\frac{\phi(\sqrt{n}\hat{w}_2)}{\sqrt{n}}\left(\frac{1}{\tilde{\tau}_{(1)}(0)\sqrt{K^{22}_{(1)}(0,\tilde{\tau}_{(1)}(0))}}-\frac{1}{\hat{w}_2}\right)(\bar{\Phi}(\sqrt{n}\hat{w}_{1(1)})+\bar{\Phi}(-\sqrt{n}\hat{w}_{1(1)}))\\
=&\frac{\phi(\sqrt{n}\hat{w}_2)}{\sqrt{n}}\left(\frac{1}{\tilde{\tau}_{(1)}(0)\sqrt{K^{22}_{(1)}(0,\tilde{\tau}_{(1)}(0))}}-\frac{1}{\hat{w}_2}\right)
\end{aligned}
$$

Last we show that $I^{\{1\}}_{(1)}+I^{\{1\}}_{(2)}=0$. We can check (3.1.11), (3.1.18) and (3.1.23) using facts (5.2.1a), (5.2.1d),(5.2.1e) and (5.2.1i). □

5.3 An example

We revisit the unconditional continuous example of Section 3.2. The results are shown in the table below. The results for the cases $(\bar{y}_1, \bar{y}_2) = (2.5, 3.0)$ and $(\bar{y}_1, \bar{y}_2) = (3.0, 4.0)$ are the special cases that $\hat{w}_1 = 0$ we mentioned above and omitted here. The relative error of the proposed method, i.e., "P. approx." are listed. We can see that the results are better than normal approximation and the approximation in [Kolassa 2003].

Table 5.1: Results of saddlepoint approximation compared with other approximations in the continuous case.

$\bar{y}_1$	$\bar{y}_2$	P. approx.	K. approx	N. approx.	Exact	Relative Error
2.5	2.5	9.12×10^{-2}	8.98×10^{-2}	9.65×10^{-2}	9.22×10^{-2}	-1.08%
2.5	3.5	1.41×10^{-2}	1.41×10^{-2}	6.54×10^{-3}	1.41×10^{-2}	0.00%
2.5	4.0	3.91×10^{-3}	3.99×10^{-3}	6.69×10^{-3}	3.93×10^{-3}	-0.51%
3.0	3.0	2.20×10^{-2}	2.14×10^{-2}	1.46×10^{-2}	2.22×10^{-2}	-0.90%
3.0	3.5	8.97×10^{-3}	8.73×10^{-3}	3.52×10^{-3}	8.96×10^{-3}	0.11%
3.5	3.5	4.40×10^{-3}	4.25×10^{-3}	1.09×10^{-3}	4.40×10^{-3}	0.00%
3.5	4.0	1.67×10^{-3}	1.61×10^{-3}	1.78×10^{-4}	1.66×10^{-3}	0.60%
4.0	4.0	7.69×10^{-4}	7.34×10^{-4}	3.88×10^{-5}	7.58×10^{-4}	1.45%

5.4 Higher dimensional extension

Now we extend the method to higher dimensions. In this section, we will use the tensor notation. As shown in (2.0.11), we have

$$\frac{n^{d-d_0}}{(2\pi i)^d}\int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0}(w_j - \tilde{w}_j(\mathbf{w}_{j-1}))} G(\boldsymbol{\tau})\, d\mathbf{w} \sim \sum_{|t|\le 1, t\subset U} I^t.$$

The first term is

$$I^{\emptyset} = \frac{n^{d-d_0}}{(2\pi i)^d}\int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0}(w_j - \tilde{w}_j(\mathbf{w}_{j-1}))}\, d\mathbf{w}. \tag{5.4.1}$$

Let $a_j^k(\mathbf{w}_k) = \frac{\tilde{w}_j(\mathbf{w}_k, \mathbf{0}_{-k}) - \tilde{w}_j(\mathbf{w}_{k-1}, \mathbf{0}_{-(k-1)})}{w_k}$, for $k < j$. The function $a_j^k(\mathbf{w}_k)$ is analytic, and $\tilde{w}_j(\mathbf{w}_{j-1}) = \sum_{k=1}^{j-1} a_j^k(\mathbf{w}_k) w_k$. For instance, $\tilde{w}_3(w_1, w_2) = a_3^1(w_1)w_1 + a_3^2(w_1, w_2)w_2$, where $a_3^1(w_1) = \frac{\tilde{w}_3(w_1,0)}{w_1}$ and $a_3^2(w_1, w_2) = \frac{\tilde{w}_3(w_1,w_2) - \tilde{w}_3(w_1,0)}{w_2}$. Here the superscripts do not represent derivatives, but represent the result of a finite differencing operation. We will use a semicolon to separate such indices and derivatives, e.g., $a_3^{2;1}$ is the partial derivative of $a_3^2(w_1, w_2)$ with respect to its first argument. We now extend Theorem (5.1.1) to higher dimension in the following theorem:

Theorem 5.4.1.

$$I^{\emptyset} = \frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0}(w_j - \sum_{k=1}^{j-1} b_j^k w_k - \sum_{l\le k<j} b_j^{kl} w_k(w_l - \hat{w}_l))} d\mathbf{w}\left(1 + O(\frac{1}{n})\right), \tag{5.4.2}$$

for $b_j^{kl} = \hat{a}_j^{k;l} + \sum_{m=k}^{j-1} \alpha_{kl}\hat{a}_j^{m;kl} w_m$, $l \le k$, $b_j^k = \hat{a}_j^k - \sum_{l\le k\le m<j} \alpha_{kl}\hat{a}_j^{m;kl} w_m w_l$, *where the notation* $\hat{a}$ *represents a function* a *evaluated at* $\hat{\mathbf{w}}$, *and*

$$\alpha_{kl} = \begin{cases} 1, & k > l \\ \frac{1}{2}, & k = l \\ 0, & k < l \end{cases}.$$

Proof. We follow the same idea as in the proof of Theorem (5.1.1). By the argument in [Kolassa 2003], we only need to prove that the coefficients of the first three degrees of Taylor expansions of the exponents of both sides agree.

Choose a reparameterization $\mathbf{w}(\mathbf{u})$, such that $\mathbf{w}(\hat{\mathbf{u}}) = \hat{\mathbf{w}}$. Let $g(\mathbf{u}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} - \hat{\mathbf{w}}^T\mathbf{w}$. Then $g(\hat{\mathbf{u}}) = -\frac{1}{2}\hat{\mathbf{w}}^T\hat{\mathbf{w}}$, $g^r(\mathbf{u}) = (\mathbf{w} - \hat{\mathbf{w}})^T\mathbf{w}^r$, and $g^r(\hat{\mathbf{u}}) = 0$. Furthermore, $g^{rs}(\mathbf{u}) = (\mathbf{w} - \hat{\mathbf{w}})^T\mathbf{w}^{rs} + (\mathbf{w}^s)^T\mathbf{w}^r$, and $g^{rs}(\hat{\mathbf{u}}) = (\hat{\mathbf{w}}^s)^T\hat{\mathbf{w}}^r$. Also, $g^{rst}(\mathbf{u}) = (\mathbf{w} - \hat{\mathbf{w}})^T\mathbf{w}^{rst} + (\mathbf{w}^r)^T\mathbf{w}^{st} + (\mathbf{w}^s)^T\mathbf{w}^{rt} + (\mathbf{w}^t)^T\mathbf{w}^{rs}$, and $g^{rst}(\hat{\mathbf{u}}) = (\hat{\mathbf{w}}^r)^T\hat{\mathbf{w}}^{st} + (\hat{\mathbf{w}}^s)^T\hat{\mathbf{w}}^{rt} + (\hat{\mathbf{w}}^t)^T\hat{\mathbf{w}}^{rs}$. The superscripts on $\mathbf{w}$ denotes derivatives with respect to $\mathbf{u}$. In the following, the superscripts on $\mathbf{u}$ denotes derivatives with respect to $\mathbf{w}$.

For the left hand side, if $u_j = w_j - \sum_{k=1}^{j-1} a_j^k(\mathbf{w}_k)w_k$, then $u_j^j = 1$, and also, $u_j^r = -a_j^r(\mathbf{w}_r) - \sum_{k=r}^{j-1} a_j^{k;r}(\mathbf{w}_k)w_k$, for $r < j$, and $\hat{u}_j^r = -\hat{a}_j^r - \sum_{k=r}^{j-1} \hat{a}_j^{k;r}\hat{w}_k$. Furthermore, $u_j^{rs} = -a_j^{r;s}(\mathbf{w}_r) - a_j^{s;r}(\mathbf{w}_s) - \sum_{k=\max(r,s)}^{j-1} a_j^{k;rs}(\mathbf{w}_k)w_k$, and $\hat{u}_j^{rs} = -\hat{a}_j^{r;s} - \hat{a}_j^{s;r} - \sum_{k=\max(r,s)}^{j-1} \hat{a}_j^{k;rs}\hat{w}_k$. Note that $a_j^{r;s}(\mathbf{w}_r) = 0$ if $r < s$, so either $a_j^{r;s}(\mathbf{w}_r) = 0$ or $a_j^{r;s}(\mathbf{w}_r) = 0$ unless $r = s$.

For the right hand side, if $u_j = w_j - \sum_{k=1}^{j-1} b_j^k w_k - \sum_{1\le l\le k<j} b_j^{kl} w_k(w_l - \hat{w}_l)$, then $u_j^j = 1$, and also $u_j^r = -b_j^r - \sum_{l=1}^{r} b_j^{rl}(w_l - \hat{w}_l) - \sum_{l=r}^{j-1} b_j^{kr} w_k$, for $r < j$, and $\hat{u}_j^r = -b_j^r - \sum_{k=r}^{j-1} b_j^{kr}\hat{w}_k$. Furthermore, $u_j^{rs} = -b_j^{rs} - b_j^{sr}$, and $\hat{u}_j^{rs} = -b_j^{rs} - b_j^{sr}$. Also note that only one of b_j^{rs} and b_j^{sr} is valid unless $r = s$.

The approximation is valid if $-b_j^{rs} - b_j^{sr} = -\hat{a}_j^{r;s} - \hat{a}_j^{s;r} - \sum_{k=\max(r,s)}^{j-1} \hat{a}_j^{k;rs}\hat{w}_k$, which is satisfied if $b_j^{rs} = \hat{a}_j^{r;s} + \sum_{k=r}^{j-1} \alpha_{rs}\hat{a}_j^{k;rs}\hat{w}_k$, for $r \geq s$, and if $-b_j^r - \sum_{l=1}^{r} b_j^{rl}(w_l - \hat{w}_l) - \sum_{l=r}^{j-1} b_j^{kr} w_k = -\hat{a}_j^r - \sum_{k=r}^{j-1} \hat{a}_j^{k;r}\hat{w}_k$, which is satisfied if $b_j^r = \hat{a}_j^r + \sum_{m=r}^{j-1}(\hat{a}_j^{m;r} - b_j^{mr})\hat{w}_m = \hat{a}_j^r - \sum_{r \leq m \leq k < j} \alpha_{mr}\hat{a}_j^{k;mr}\hat{w}_k\hat{w}_m$. □

Next we change variables to $s_j = w_j - \sum_{k=1}^{j-1} b_j^k w_k$. Then $(w_j - \sum_{k=1}^{j-1} b_j^k w_k - \sum_{l \leq k < j} b_j^{kl} w_k(w_l - \hat{w}_l)) = s_j - \sum_{l<j,k<j} c_j^{kl} s_k(s_l - \hat{s}_l))$, for some c_j^{kl} computable from b_j^k and b_j^{kl}, where $\hat{s}_l$ is the value of s_l corresponding to $w_l = \hat{w}_l$. We denote the new integral by

$$f(c_j^{kl}, k<j, l<j) = \frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{s}}-i\mathbf{K}}^{\hat{\mathbf{s}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}(\mathbf{s})^T\mathbf{w}(\mathbf{s}) - \hat{\mathbf{w}}^T\mathbf{w}(\mathbf{s})])}{\prod_{j=1}^{d_0}(s_j - \sum_{l \leq j, k<j} c_j^{kl} s_k(s_l - \hat{s}_l))} d\mathbf{s}, \tag{5.4.3}$$

so that we view it as a function of all c_j^{kl}, for $k<j$, $l<j$. We use Taylor expansion to obtain

$$\begin{aligned} f(c_j^{kl}, k<j, l<j) =& f(\mathbf{0}) + \sum_{l<j,k<j} c_j^{kl} f_j^{kl}(\mathbf{0}) + \sum_{k,l,m,n<j} \frac{c_j^{kl}c_j^{mn}}{2} f_j^{kl,mn}(c_j^{rs*}, r<j, s<j) \\ =& I_0 + \sum_{l<j,k<j} I_j^{kl} + E, \end{aligned} \tag{5.4.4}$$

where f_j^{kl} denotes the partial derivative of f with respect to c_j^{kl}, $f_j^{kl,mn}$ denotes the double partial derivative of f with respect to c_j^{kl} and c_j^{mn}, and c_j^{rs*} are some values between 0 and c_j^{rs}.

The first integral

$$I_0 = \frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{s}}-i\mathbf{K}}^{\hat{\mathbf{s}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}(\mathbf{s})^T\mathbf{w}(\mathbf{s}) - \hat{\mathbf{w}}^T\mathbf{w}(\mathbf{s})])}{\prod_{j=1}^{d_0} s_j} d\mathbf{s} \tag{5.4.5}$$

is a multivariate normal tail probability and easy to compute. Next we consider I_j^{kl}. We have

$$I_j^{kl} = c_j^{kl} \frac{n^{d-d_0}}{(2\pi i)^d} \int_{\hat{\mathbf{s}}-i\mathbf{K}}^{\hat{\mathbf{s}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}(\mathbf{s})^T\mathbf{w}(\mathbf{s}) - \hat{\mathbf{w}}^T\mathbf{w}(\mathbf{s})])(s_l - \hat{s}_l)}{s_j/s_k \prod_{m=1}^{d_0} s_m} d\mathbf{s}, \tag{5.4.6}$$

which is of relative error $O(n^{-\frac{1}{2}})$ by [Kolassa 2003]. I_j^{kl} can be integrated out analytically. First, we have

$$\begin{aligned} I_j^{kl} =& c_j^{kl}\frac{n^{d-d_0}}{(2\pi i)^d}\int_{\hat{\mathbf{s}}-i\mathbf{K}}^{\hat{\mathbf{s}}+i\mathbf{K}}\frac{\exp(n[\frac{1}{2}\mathbf{w}(\mathbf{s})^T\mathbf{w}(\mathbf{s})-\hat{\mathbf{w}}^T\mathbf{w}(\mathbf{s})])s_l s_k}{s_j\prod_{m=1}^{d_0}s_m}\,d\mathbf{s}- \\ & c_j^{kl}\hat{s}_l\frac{n^{d-d_0}}{(2\pi i)^d}\int_{\hat{\mathbf{s}}-i\mathbf{K}}^{\hat{\mathbf{s}}+i\mathbf{K}}\frac{\exp(n[\frac{1}{2}\mathbf{w}(\mathbf{s})^T\mathbf{w}(\mathbf{s})-\hat{\mathbf{w}}^T\mathbf{w}(\mathbf{s})])s_k}{s_j\prod_{m=1}^{d_0}s_m}\,d\mathbf{s} \\ =& I_{j1}^{kl}-I_{j2}^{kl} \end{aligned} \tag{5.4.7}$$

Since the transformation from $\mathbf{w}$ to $\mathbf{s}$ is linear, $\frac{1}{2}\mathbf{w}(\mathbf{s})^T\mathbf{w}(\mathbf{s})-\hat{\mathbf{w}}^T\mathbf{w}(\mathbf{s})$ is still quadratic. Both I_{j1}^{kl} and I_{j2}^{kl} can be viewed as integration or differentiation of multivariate tail probabilities with respect to covariates l, j and k, times some constants, and, while algebraically complex, are not hard to obtain. And finally we consider the error terms. For some constant C and C',

$$\begin{aligned} f_j^{kl,mn} =& C\int_{\hat{\mathbf{s}}-i\mathbf{K}}^{\hat{\mathbf{s}}+i\mathbf{K}}\frac{\exp(n[\frac{1}{2}\mathbf{w}(\mathbf{s})^T\mathbf{w}(\mathbf{s})-\hat{\mathbf{w}}^T\mathbf{w}(\mathbf{s})])s_k s_m(s_l-\hat{s}_l)(s_n-\hat{s}_n)}{(s_j-\sum\limits_{l<j,k<j}c_j^{kl*}s_k(s_l-\hat{s}_l))^2\prod_{p=1}^{d_0}(s_p-\sum\limits_{l<p,k<p}c_p^{kl*}s_k(s_l-\hat{s}_l))}\,d\mathbf{s} \\ =& \frac{C'}{n}\int_{\hat{\mathbf{t}}-i\mathbf{K}}^{\hat{\mathbf{t}}+i\mathbf{K}}\frac{\exp(Q_n(\mathbf{t}))t_k t_m(t_l-\hat{t}_l)(t_n-\hat{t}_n)}{(t_j-\sum\limits_{l<j,k<j}\frac{c_j^{kl*}}{\sqrt{n}}t_k(t_l-\hat{t}_l))^2\prod_{p=1}^{d_0}(t_p-\sum\limits_{l<p,k<p}\frac{c_p^{kl*}}{\sqrt{n}}t_k(t_l-\hat{t}_l))}\,d\mathbf{t}, \end{aligned} \tag{5.4.8}$$

where $t_j=\sqrt{n}s_j$, $Q_n(\mathbf{t})$ is a quadratic function of $\mathbf{t}$, and the coefficients of t_j^2 do not contain n. Therefore, though may not be strictly relative, the error term diminish at the rate of n^{-1}.

5.5 A three-dimensional example

The definition of $\hat{\mathbf{w}}$ and $\mathbf{w}$ follows that of Section 4.2. First of all, we have

$$\begin{aligned} I^{\emptyset} =& \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty}\frac{\exp(n[\frac{1}{2}w_1^2+\frac{1}{2}w_2^2+\frac{1}{2}w_3^2-\hat{w}_1w_1-\hat{w}_2w_2-\hat{w}_3w_3])}{(2\pi i)^3w_1(w_2-\tilde{w}_2(w_1))(w_3-\tilde{w}_3(w_1,w_2))}\,d\mathbf{w} \\ \sim& f(c_2^{11},c_3^{11},c_3^{21},c_3^{12},c_3^{22}). \end{aligned} \tag{5.5.1}$$

First of all, we need to compute $\hat{a}_2^1$, $\hat{a}_2^{1;1}$, $\hat{a}_2^{1;11}$, $\hat{a}_3^1$, $\hat{a}_3^{1;1}$, $\hat{a}_3^{1;11}$, $\hat{a}_3^2$, $\hat{a}_3^{2;1}$, $\hat{a}_3^{2;2}$, $\hat{a}_3^{2;11}$,$\hat{a}_3^{2;21}$ and $\hat{a}_3^{2;22}$. Terms $\hat{a}_2^1$, $\hat{a}_2^{1;1}$ and $\hat{a}_2^{1;11}$ are coefficients of Taylor expansion of $\tilde{w}_2(w_1)/w_1$. Terms $\hat{a}_3^1$, $\hat{a}_3^{1;1}$ and $\hat{a}_3^{1;11}$ are coefficients of Taylor expansion of $\tilde{w}_3(w_1,0)/w_1$. The rest are coefficients of Taylor expansion of $(\tilde{w}_3(w_1, w_2) - \tilde{w}_3(w_1, 0))/w_2$. More specifically,

$$
\begin{aligned}
\hat{a}_2^1 &= \frac{\tilde{w}_2(\hat{w}_1)}{\hat{w}_1} \\
\hat{a}_2^{1;1} &= \frac{\breve{w}_2' \hat{w}_1 - \breve{w}_2}{\hat{w}_1^2} \\
\hat{a}_2^{1;11} &= \frac{\breve{w}_2'' \hat{w}_1^2 - 2\breve{w}_2' \hat{w}_1 + 2\breve{w}_2}{\hat{w}_1^3} \\
\hat{a}_3^1 &= \frac{\tilde{w}_3(\hat{w}_1, 0)}{\hat{w}_1} \\
\hat{a}_3^{1;1} &= \frac{\tilde{w}_3^1(\hat{w}_1, 0)\hat{w}_1 - \tilde{w}_3(\hat{w}_1, 0)}{\hat{w}_1^2} \\
\hat{a}_3^{1;11} &= \frac{\tilde{w}_3^{11}(\hat{w}_1, 0) - 2\tilde{w}_3^1(\hat{w}_1, 0)\hat{w}_1 + 2\tilde{w}_3(\hat{w}_1, 0)}{\hat{w}_1^3} \\
\hat{a}_3^2 &= \frac{\breve{w}_3 - \tilde{w}_3(\hat{w}_1, 0)}{\hat{w}_2} \\
\hat{a}_3^{2;1} &= \frac{\breve{w}_3^1 - \tilde{w}_3^1(\hat{w}_1, 0)}{\hat{w}_2} \\
\hat{a}_3^{2;2} &= \frac{\breve{w}_3^2 \hat{w}_2 - \breve{w}_3 + \tilde{w}_3(\hat{w}_1, 0)}{\hat{w}_2^2} \\
\hat{a}_3^{2;11} &= \frac{\breve{w}_3^{11} - \tilde{w}_3^{11}(\hat{w}_1, 0)}{\hat{w}_2} \\
\hat{a}_3^{2;21} &= \frac{\breve{w}_3^{12} \hat{w}_2 - \breve{w}_3^1 + \tilde{w}_3(\hat{w}_1, 0)}{\hat{w}_2^2} \\
\hat{a}_3^{2;22} &= \frac{\breve{w}_3^{22} \hat{w}_2^2 - 2\breve{w}_3^2 \hat{w}_2 + 2\breve{w}_3 - 2\tilde{w}_3(\hat{w}_1, 0)}{\hat{w}_2^3}.
\end{aligned}
\tag{5.5.2}
$$

Then

$$
\begin{aligned}
\hat{b}_2^1 &= \hat{a}_2^1 - \frac{1}{2}\hat{a}_2^{1;11}\hat{w}_1^2 = \breve{w}_2' - \frac{1}{2}\breve{w}_2''\hat{w}_1 \\
\hat{b}_2^{1;1} &= \hat{a}_2^{1;1} + \frac{1}{2}\hat{a}_2^{1;11}\hat{w}_1 = \frac{1}{2}\breve{w}_2'' \\
\hat{b}_3^1 &= \hat{a}_3^1 - \frac{1}{2}\hat{a}_3^{1;11}\hat{w}_1^2 - \frac{1}{2}\hat{a}_3^{2;11}\hat{w}_1\hat{w}_2 - \frac{1}{2}\hat{a}_3^{2;21}\hat{w}_2^2 = \breve{w}_3^1 - \frac{1}{2}\breve{w}_3^{11}\hat{w}_1 - \breve{w}_3^{12}\hat{w}_2 \\
\hat{b}_3^{11} &= \hat{a}_3^{1;1} + \frac{1}{2}\hat{a}_3^{1;11}\hat{w}_1 + \frac{1}{2}\hat{a}_3^{2;11}\hat{w}_2 = \frac{1}{2}\breve{w}_3^{11} \\
\hat{b}_3^2 &= \hat{a}_3^2 - \frac{1}{2}\hat{a}_3^{2;22}\hat{w}_2^2 = \breve{w}_3^2 - \frac{1}{2}\breve{w}_3^{22}\hat{w}_2 \\
\hat{b}_3^{21} &= \hat{a}_3^{2;1} + \hat{a}_3^{2;21}\hat{w}_2 = \breve{w}_3^{12} \\
\hat{b}_3^{22} &= \hat{a}_3^{2;2} + \frac{1}{2}\hat{a}_3^{2;22}\hat{w}_2 = \frac{1}{2}\breve{w}_3^{22}.
\end{aligned}
\tag{5.5.3}
$$

Note that b_j^r and b_j^{rs} no longer involve $\tilde{w}_3(\hat{w}_1, 0)$ and the corresponding derivatives. Next we need to compute c_j^{kl}.

Let $s_1 = w_1$, $s_2 = w_2 - b_2^1 w_1$ and $s_3 = w_3 - b_3^1 w_1 - b_3^2 w_2$. We have

$$
\begin{aligned}
& w_1(w_2 - b_2^1 w_1 - b_2^{11} w_1(w_1 - \hat{w}_1))\cdot \\
& (w_3 - b_3^1 w_1 - b_3^2 w_2 - b_3^{11}(w_1 - \hat{w}_1)w_1 - b_3^{21} w_2(w_1 - \hat{w}_1) - b_3^{22} w_2(w_2 - \hat{w}_2)) \\
=& s_1(s_2 - b_2^{11} s_1(s_1 - \hat{s}_1))\cdot \\
& (s_3 - b_3^{11} s_1(s_1 - \hat{s}_1) - b_3^{21}(s_2 + b_2^1 s_1)(s_1 - \hat{s}_1) - b_3^{22}(s_2 + b_2^1 s_1 - \hat{s}_2 - b_2^1 \hat{s}_1)(s_2 + b_2^1 s_1)) \\
=& s_1(s_2 - c_2^{11} s_1(s_1 - \hat{s}_1))\cdot \\
& (s_3 - c_3^{11} s_1(s_1 - \hat{s}_1) - c_3^{21} s_2(s_1 - \hat{s}_1) - c_3^{22} s_2(s_2 - \hat{s}_2) - c_3^{12} s_1(s_2 - \hat{s}_2)).
\end{aligned}
$$

Compare coefficients to obtain $c_2^{11} = b_2^{11}$, $c_3^{11} = b_3^{11} + b_3^{21} b_2^1 + b_3^{22}(b_2^1)^2$, $c_3^{21} = b_3^{21} + b_3^{22} b_2^1$, $c_3^{22} = b_3^{22}$ and $c_3^{12} = b_3^{22} b_2^1$. Then

$$
\begin{aligned}
& f(c_2^{11}, c_3^{11}, c_3^{11}, c_3^{22}, c_3^{12}) \\
=& f(0,0,0,0,0) + c_2^{11} f_2^{11}(0,0,0,0,0) + c_3^{11} f_3^{11}(0,0,0,0,0) + c_3^{21} f_3^{21}(0,0,0,0,0) + \\
& c_3^{22} f_3^{22}(0,0,0,0,0) + c_3^{12} f_3^{12}(0,0,0,0,0) + E \\
=& I_0 + I_2^{11} + I_3^{11} + I_3^{11} + I_3^{22} + I_3^{12} + E.
\end{aligned}
\tag{5.5.4}
$$

For the main term, we have

$$
\begin{aligned}
I_0 =& \int_{\hat{\mathbf{w}} - i\infty}^{\hat{\mathbf{w}} + i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 + \frac{1}{2}w_3^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2 - \hat{w}_3 w_3])}{(2\pi i)^3 w_1(w_2 - b_2^1 w_1)(w_3 - b_3^1 w_1 - b_3^2 w_2)}\, d\mathbf{w} \\
=& \bar{\Phi}(x, y, z, \boldsymbol{\rho}),
\end{aligned}
\tag{5.5.5}
$$

where $x = \frac{\sqrt{n}}{\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1 b_3^2)^2}}(\hat{w}_1 + b_2^1 \hat{w}_2 + (b_3^1 + b_2^1 b_3^2)\hat{w}_3)$, $y = \frac{\sqrt{n}}{\sqrt{1+(b_3^2)^2}}(\hat{w}_2 + b_3^2 \hat{w}_3)$, $z = \sqrt{n}\hat{w}_3$, and $\boldsymbol{\rho}$ is the covariance matrix with $\rho_{12} = \frac{b_2^1 + b_3^1 b_3^2 + b_2^1 (b_3^2)^2}{\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1 b_3^2)^2}\sqrt{1+(b_3^2)^2}}$, $\rho_{13} = \frac{b_3^1 + b_2^1 b_3^2}{\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1 b_3^2)^2}}$ and $\rho_{23} = \frac{b_3^2}{\sqrt{1+(b_3^2)^2}}$.

Similarly,

$$\begin{aligned}
I_2^{11} =& c_2^{11} \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 + \frac{1}{2}w_3^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2 - \hat{w}_3 w_3])(w_1 - \hat{w}_1)}{(2\pi i)^3 (w_2 - b_2^1 w_1)^2 (w_3 - b_3^1 w_1 - b_3^2 w_2)} \, d\mathbf{w} \\
=& \frac{c_2^{11}\sqrt{1+(b_3^2)^2}}{\sqrt{n}(1+(b_2^1)^2 + (b_3^1 + b_2^1 b_3^2)^2)} \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T \boldsymbol{\rho} \mathbf{v} - xv_1 - yv_2 - zv_3)(v_1 - \hat{v}_1)}{(2\pi i)^3 v_2^2 v_3} \, d\mathbf{v} \\
=& \frac{c_2^{11}\sqrt{1+(b_3^2)^2}}{\sqrt{n}(1+(b_2^1)^2 + (b_3^1 + b_2^1 b_3^2)^2)} \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T \boldsymbol{\rho} \mathbf{v} - xv_1 - yv_2 - zv_3)v_1}{(2\pi i)^3 v_2^2 v_3} \, d\mathbf{v} - \\
& \frac{c_2^{11}\sqrt{1+(b_3^2)^2}\hat{v}_1}{\sqrt{n}(1+(b_2^1)^2 + (b_3^1 + b_2^1 b_3^2)^2)} \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T \boldsymbol{\rho} \mathbf{v} - xv_1 - yv_2 - zv_3)}{(2\pi i)^3 v_2^2 v_3} \, d\mathbf{v} \\
=& I_{21}^{11} - I_{22}^{11},
\end{aligned} \tag{5.5.6}$$

where $\hat{v}_1 = \sqrt{n}\sqrt{1+(b_2^1)^2 + (b_3^1 + b_2^1 b_3^2)^2}\hat{w}_1$. We can obtain I_{21}^{11} and I_{22}^{11} by considering them as functions of $(x, y, z, \boldsymbol{\rho})$ and solving the differential equations $I_{21}^{11;223}(x, y, z, \boldsymbol{\rho}) = \frac{c_2^{11}\sqrt{1+(b_3^2)^2}}{\sqrt{n}(1+(b_2^1)^2+(b_3^1+b_2^1 b_3^2)^2)}\phi^1(x, y, z, \boldsymbol{\rho})$ and $I_{22}^{11;223}(x, y, z, \boldsymbol{\rho}) = \frac{c_2^{11}\sqrt{1+(b_3^2)^2}\hat{v}_1}{\sqrt{n}(1+(b_2^1)^2+(b_3^1+b_2^1 b_3^2)^2)}\phi(x, y, z, \boldsymbol{\rho})$. Likewise,

$$\begin{aligned}
I_3^{11} =& c_3^{11} \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 + \frac{1}{2}w_3^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2 - \hat{w}_3 w_3])(w_1 - \hat{w}_1)}{(2\pi i)^3 (w_2 - b_2^1 w_1)(w_3 - b_3^1 w_1 - b_3^2 w_2)^2} \, d\mathbf{w} \\
=& \frac{c_3^{11}}{\sqrt{n}(1+(b_2^1)^2 + (b_3^1 + b_2^1 b_3^2)^2)} \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T \boldsymbol{\rho} \mathbf{v} - xv_1 - yv_2 - zv_3)(v_1 - \hat{v}_1)}{(2\pi i)^3 v_2 v_3^2} \, d\mathbf{v} \\
=& \frac{c_3^{11}}{\sqrt{n}(1+(b_2^1)^2 + (b_3^1 + b_2^1 b_3^2)^2)} \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T \boldsymbol{\rho} \mathbf{v} - xv_1 - yv_2 - zv_3)v_1}{(2\pi i)^3 v_2 v_3^2} \, d\mathbf{v} - \\
& \frac{c_3^{11}\hat{v}_1}{\sqrt{n}(1+(b_2^1)^2 + (b_3^1 + b_2^1 b_3^2)^2)} \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T \boldsymbol{\rho} \mathbf{v} - xv_1 - yv_2 - zv_3)}{(2\pi i)^3 v_2 v_3^2} \, d\mathbf{v} \\
=& I_{31}^{11} - I_{32}^{11},
\end{aligned} \tag{5.5.7}$$

Again, we can obtain I_{31}^{11} and I_{32}^{11} by considering them as functions of $(x, y, z, \boldsymbol{\rho})$ and solving $I_{31}^{11;233}(x, y, z, \boldsymbol{\rho}) = \frac{c_3^{11}}{\sqrt{n}(1+(b_2^1)^2+(b_3^1+b_2^1 b_3^2)^2)}\phi^1(x, y, z, \boldsymbol{\rho})$ and $I_{32}^{11;233}(x, y, z, \boldsymbol{\rho}) =$

$\frac{c_3^{11}\hat{v}_1}{\sqrt{n}(1+(b_2^1)^2+(b_3^1+b_2^1b_3^2)^2)} \cdot \phi(x,y,z,\boldsymbol{\rho})$. We also have

$$
\begin{aligned}
I_3^{21} =& c_3^{21} \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 + \frac{1}{2}w_3^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2 - \hat{w}_3 w_3])(w_1 - \hat{w}_1)}{(2\pi i)^3 w_1 (w_3 - b_3^1 w_1 - b_3^2 w_2)^2}\, d\mathbf{w} \\
=& \frac{c_3^{21}}{\sqrt{n}\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1b_3^2)^2}\sqrt{1+(b_3^2)^2}} \cdot \\
& \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T\boldsymbol{\rho}\mathbf{v} - xv_1 - yv_2 - zv_3)(v_1 - \hat{v}_1)}{(2\pi i)^3 v_1 v_3^2}\, d\mathbf{v} \\
=& \frac{c_3^{21}}{\sqrt{n}\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1b_3^2)^2}\sqrt{1+(b_3^2)^2}} \cdot \\
& \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T\boldsymbol{\rho}\mathbf{v} - xv_1 - yv_2 - zv_3)v_1}{(2\pi i)^3 v_1 v_3^2}\, d\mathbf{v} - \\
& \frac{c_3^{21}\hat{v}_1}{\sqrt{n}\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1b_3^2)^2}\sqrt{1+(b_3^2)^2}} \cdot \\
& \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T\boldsymbol{\rho}\mathbf{v} - xv_1 - yv_2 - zv_3)}{(2\pi i)^3 v_1 v_3^2}\, d\mathbf{v} \\
=& I_{31}^{21} - I_{32}^{21},
\end{aligned}
\tag{5.5.8}
$$

and we can obtain I_{31}^{21} and I_{32}^{21} by considering them as functions of $(x,y,z,\boldsymbol{\rho})$ and solving

$$
I_{31}^{21;133}(x,y,z,\boldsymbol{\rho}) = \frac{c_3^{21}}{\sqrt{n}\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1b_3^2)^2}\sqrt{1+(b_3^2)^2}}\phi^1(x,y,z,\boldsymbol{\rho})
$$

and

$$
I_{32}^{21;133}(x,y,z,\boldsymbol{\rho}) = \frac{c_3^{21}\hat{v}_1}{\sqrt{n}\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1b_3^2)^2}\sqrt{1+(b_3^2)^2}}\phi(x,y,z,\boldsymbol{\rho}).
$$

Next, we have

$$
\begin{aligned}
I_3^{22} =& c_3^{22} \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 + \frac{1}{2}w_3^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2 - \hat{w}_3 w_3])(s_2 - \hat{s}_2)}{(2\pi i)^3 w_1 (w_3 - b_3^1 w_1 - b_3^2 w_2)^2}\, d\mathbf{w} \\
=& \frac{c_3^{22}}{\sqrt{n}(1+(b_3^2)^2)} \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T\boldsymbol{\rho}\mathbf{v} - xv_1 - yv_2 - zv_3)(v_2 - \hat{v}_2)}{(2\pi i)^3 v_1 v_3^2}\, d\mathbf{v} \\
=& \frac{c_3^{22}}{\sqrt{n}(1+(b_3^2)^2)} \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T\boldsymbol{\rho}\mathbf{v} - xv_1 - yv_2 - zv_3)v_2}{(2\pi i)^3 v_1 v_3^2}\, d\mathbf{v} - \\
& \frac{c_3^{22}\hat{v}_2}{\sqrt{n}(1+(b_3^2)^2)} \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T\boldsymbol{\rho}\mathbf{v} - xv_1 - yv_2 - zv_3)}{(2\pi i)^3 v_1 v_3^2}\, d\mathbf{v} \\
=& I_{31}^{22} - I_{31}^{22},
\end{aligned}
\tag{5.5.9}
$$

and we can obtain I_{31}^{22} and I_{31}^{22} by considering them as functions of $(x,y,z,\boldsymbol{\rho})$ and solving $I_{31}^{22;133}(x,y,z,\boldsymbol{\rho}) = \frac{c_3^{22}}{\sqrt{n}(1+(b_3^2)^2)}\phi^2(x,y,z,\boldsymbol{\rho})$ and $I_{32}^{22;133}(x,y,z,\boldsymbol{\rho}) = \frac{c_3^{22}\hat{v}_2}{\sqrt{n}(1+(b_3^2)^2)}\phi(x,y,z,\boldsymbol{\rho})$. And finally we obtain We also have

$$
\begin{aligned}
I_3^{12} =& c_3^{12}\int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2+\frac{1}{2}w_2^2+\frac{1}{2}w_3^2-\hat{w}_1w_1-\hat{w}_2w_2-\hat{w}_3w_3])(s_2-\hat{s}_2)}{(2\pi i)^3(w_2-b_2^1w_1)(w_3-b_3^1w_1-b_3^2w_2)^2}\,d\mathbf{w} \\
=& \frac{c_3^{12}}{\sqrt{n}\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1b_3^2)^2}\sqrt{1+(b_3^2)^2}}\cdot \\
& \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T\boldsymbol{\rho}\mathbf{v}-xv_1-yv_2-zv_3)(v_2-\hat{v}_2)}{(2\pi i)^3v_2v_3^2}\,d\mathbf{v} \\
=& \frac{c_3^{12}}{\sqrt{n}\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1b_3^2)^2}\sqrt{1+(b_3^2)^2}}\cdot \\
& \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T\boldsymbol{\rho}\mathbf{v}-xv_1-yv_2-zv_3)v_2}{(2\pi i)^3v_2v_3^2}\,d\mathbf{v}- \\
& \frac{c_3^{12}\hat{v}_2}{\sqrt{n}\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1b_3^2)^2}\sqrt{1+(b_3^2)^2}}\cdot \\
& \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T\boldsymbol{\rho}\mathbf{v}-xv_1-yv_2-zv_3)}{(2\pi i)^3v_2v_3^2}\,d\mathbf{v} \\
=& I_{31}^{12}-I_{31}^{12},
\end{aligned}
\tag{5.5.10}
$$

and we can obtain I_{31}^{12} and I_{31}^{12} by considering them as functions of $(x,y,z,\boldsymbol{\rho})$ and solving $I_{31}^{12;233}(x,y,z,\boldsymbol{\rho}) = \frac{c_3^{22}}{\sqrt{n}\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1b_3^2)^2}\sqrt{1+(b_3^2)^2}}\phi^2(x,y,z,\boldsymbol{\rho})$ and $I_{32}^{12;233}(x,y,z,\boldsymbol{\rho}) = \frac{c_3^{22}\hat{v}_2}{\sqrt{n}\sqrt{1+(b_2^1)^2+(b_3^1+b_2^1b_3^2)^2}\sqrt{1+(b_3^2)^2}}\phi(x,y,z,\boldsymbol{\rho})$.

Next, we deal with $I^{\{1\}}$,$I^{\{2\}}$ and $I^{\{3\}}$. By definition,

$$
\begin{aligned}
I^{\{1\}} =& \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2+\frac{1}{2}w_2^2+\frac{1}{2}w_3^2-\hat{w}_1w_1-\hat{w}_2w_2-\hat{w}_3w_3])}{(2\pi i)^3(w_2-\tilde{w}_2(w_1)))(w_3-\tilde{w}_3(w_1,w_2))}\cdot \\
& \frac{G(\tau_1,0,0)-G(0,0,0)}{w_1}\,d\mathbf{w} \\
=& \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2+\frac{1}{2}w_2^2+\frac{1}{2}w_3^2-\hat{w}_1w_1-\hat{w}_2w_2-\hat{w}_3w_3])}{(2\pi i)^3(w_2-\tilde{w}_2(w_1)))(w_3-\tilde{w}_3(w_1,w_2))}h^{\{1\}}(w_1)\,d\mathbf{w},
\end{aligned}
\tag{5.5.11}
$$

where $h^{\{1\}}(w_1) = \frac{1}{\tau_1}\frac{d\tau_1}{dw_1} - \frac{1}{w_1}$ is analytic. Now we can do a change of variables similar to the one done in the two dimensional case. First of all, let $u_1 = w_1$, $u_2 = w_2 - \tilde{w}_2(w_1)$ and $u_3 = w_3 - \tilde{w}_3(w_1,w_2)$; let $\hat{u}_1 = \hat{w}_1$, $\hat{u}_2 = \hat{w}_2 - \tilde{w}_2(\hat{w}_1)$ and $\hat{u}_3 = \hat{w}_3 - \tilde{w}_3(\hat{w}_1,\hat{w}_2)$.

Then

$$I^{\{1\}} = \int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{\exp(n[g^{\{1\}}(u_1,u_2,u_3)])}{(2\pi i)^3 u_2 u_3} h^{\{1\}}(u_1)\,d\mathbf{u}$$
$$\dot{\sim} \int_{\hat{\mathbf{u}}-i\infty}^{\hat{\mathbf{u}}+i\infty} \frac{\exp(n[c_{00}^{\{1\}} + \frac{1}{2}(\mathbf{u}-\hat{\mathbf{u}})^T \mathbf{cc}^{\{1\}}(\mathbf{u}-\hat{\mathbf{u}}) - \mathbf{c}^{\{1\}}(\mathbf{u}-\hat{\mathbf{u}})])}{(2\pi i)^3 u_2 u_3} h^{\{1\}}(u_1)\,d\mathbf{u}, \tag{5.5.12}$$

where $g^{\{1\}}(u_1,u_2,u_3)$ is the exponent as a function of $\mathbf{u}$ after the change of variable, $c_{00}^{\{1\}} = g^{\{1\}}(\hat{\mathbf{u}})$, $\mathbf{cc}^{\{1\}}$ is the matrix such that its element $cc_{ij}^{\{1\}} = [g^{\{1\}}]^{ij}(\hat{\mathbf{u}})$ and $\mathbf{c}^{\{1\}}$ is the vector such that $c_i^{\{1\}} = [g^{\{1\}}]^i(\hat{\mathbf{u}})$. We can do a further change of variables $v_1 = \sqrt{n}\sqrt{c_{11}^{\{1\}}}u_1$, $v_2 = \sqrt{n}\sqrt{c_{22}^{\{1\}}}u_2$, $v_3 = \sqrt{n}\sqrt{c_{33}^{\{1\}}}u_3$, such that

$$I^{\{1\}} \dot{\sim} \frac{C^{\{1\}}}{\sqrt{n}\sqrt{c_{11}^{\{1\}}}} \int_{\hat{\mathbf{v}}-i\infty}^{\hat{\mathbf{v}}+i\infty} \frac{\exp(\frac{1}{2}\mathbf{v}^T \boldsymbol{\rho}^{\{1\}}\mathbf{v} - x^{\{1\}}v_1 - y^{\{1\}}v_2 - z^{\{1\}}v_3)}{(2\pi i)^3 v_2 v_3} h^*(v_1)\,d\mathbf{v}, \tag{5.5.13}$$

where $C^{\{1\}} = \exp(n[c_{00}^{\{1\}} + \frac{1}{2}\hat{\mathbf{u}}^T \mathbf{cc}^{\{1\}}\hat{\mathbf{u}}])$, $\boldsymbol{\rho}^{\{1\}}$ is the covariance matrix with $\rho_{ij}^{\{1\}} = c_{ij}^{\{1\}}/\sqrt{c_{ii}^{\{1\}}c_{jj}^{\{1\}}}$ and $x^{\{1\}} = \frac{\sqrt{n}(c_{11}^{\{1\}}\hat{u}_1 + c_{12}^{\{1\}}\hat{u}_2 + c_{13}^{\{1\}}\hat{u}_3)}{\sqrt{c_{11}^{\{1\}}}}$, $y^{\{1\}} = \frac{\sqrt{n}(c_{12}^{\{1\}}\hat{u}_1 + c_{22}^{\{1\}}\hat{u}_2 + c_{23}^{\{1\}}\hat{u}_3)}{\sqrt{c_{22}^{\{1\}}}}$, $z^{\{1\}} = \frac{\sqrt{n}(c_{13}^{\{1\}}\hat{u}_1 + c_{23}^{\{1\}}\hat{u}_2 + c_{33}^{\{1\}}\hat{u}_3)}{\sqrt{c_{33}^{\{1\}}}}$. We can not apply Watson's lemma directly but can use the same technique as in the two dimensional case. Let $t_1 = v_1 + \rho_{12}^{\{1\}}v_2 + \rho_{13}^{\{1\}}v_3$, $t_2 = \sqrt{1-(\rho_{12}^{\{1\}})^2}v_2$ and $t_3 = \sqrt{1-(\rho_{13}^{\{1\}})^2}v_3$. Do a change of variables to obtain

$$I^{\{1\}} \dot{\sim} \frac{C^{\{1\}}}{\sqrt{n}\sqrt{c_{11}^{\{1\}}}} \int_{\hat{\mathbf{t}}-i\infty}^{\hat{\mathbf{t}}+i\infty} \frac{\exp(Q(\mathbf{t}))}{(2\pi i)^3 t_2 t_3} h^{**}(t_1,t_2,t_3)\,d\mathbf{t}, \tag{5.5.14}$$

where $Q(\mathbf{t}) = \frac{1}{2}\mathbf{t}^T\mathbf{t} + \frac{\rho_{23}^{\{1\}} - \rho_{12}^{\{1\}}\rho_{13}^{\{1\}}}{\sqrt{\rho_{12}^{\{1\}}}\sqrt{\rho_{13}^{\{1\}}}} t_2 t_3 - x^{\{1\}}t_1 - \frac{y^{\{1\}} - \rho_{12}^{\{1\}}x^{\{1\}}}{\sqrt{1-(\rho_{12}^{\{1\}})^2}} t_2 - \frac{z^{\{1\}} - \rho_{13}^{\{1\}}x^{\{1\}}}{\sqrt{1-(\rho_{13}^{\{1\}})^2}} t_3$. We then obtain

$$
\begin{aligned}
I^{\{1\}} \dot{\sim} & \frac{C^{\{1\}}}{\sqrt{n}\sqrt{c_{11}^{\{1\}}}} \int_{\hat{\mathbf{t}}-i\infty}^{\hat{\mathbf{t}}+i\infty} \frac{\exp(Q(\mathbf{t}))}{(2\pi i)^3 t_2 t_3} h^{**}(t_1,0,0)\, d\mathbf{t} + \\
& \frac{C^{\{1\}}}{\sqrt{n}\sqrt{c_{11}^{\{1\}}}} \int_{\hat{\mathbf{t}}-i\infty}^{\hat{\mathbf{t}}+i\infty} \frac{\exp(Q(\mathbf{t}))}{(2\pi i)^3 t_3} \frac{h^{**}(t_1,t_2,0) - h^{**}(t_1,0,0)}{t_2}\, d\mathbf{t} + \\
& \frac{C^{\{1\}}}{\sqrt{n}\sqrt{c_{11}^{\{1\}}}} \int_{\hat{\mathbf{t}}-i\infty}^{\hat{\mathbf{t}}+i\infty} \frac{\exp(Q(\mathbf{t}))}{(2\pi i)^3 t_2} \frac{h^{**}(t_1,0,t_3) - h^{**}(t_1,0,0)}{t_3}\, d\mathbf{t} + \\
& \frac{C^{\{1\}}}{\sqrt{n}\sqrt{c_{11}^{\{1\}}}} \int_{\hat{\mathbf{t}}-i\infty}^{\hat{\mathbf{t}}+i\infty} \frac{\exp(Q(\mathbf{t}))}{(2\pi i)^3} \cdot \\
& \frac{h^{**}(t_1,t_2,t_3) - h^{**}(t_1,t_2,0) - h^{**}(t_1,0,t_3) + h^{**}(t_1,0,0)}{t_2 t_3}\, d\mathbf{t} \\
= & I_{00}^{\{1\}} + I_{10}^{\{1\}} + I_{01}^{\{1\}} + I_{11}^{\{1\}}.
\end{aligned}
\tag{5.5.15}
$$

First consider $I_{11}^{\{1\}}$. The expression $\frac{h^{**}(t_1,t_2,t_3) - h^{**}(t_1,t_2,0) - h^{**}(t_1,0,t_3) + h^{**}(t_1,0,0)}{t_2 t_3}$ is analytic, which absorbs both t_2 and t_3 in the denominator, and by Watson's lemma, $I_{11}^{\{1\}}$ has an error of $O(n^{-1})$ with regard to $I^{\{1\}}$ and therefore can be omitted. For $I_{01}^{\{1\}}$, since $\frac{h^{**}(t_1,0,t_3) - h^{**}(t_1,0,0)}{t_3}$ is analytic, t_3 is absorbed, but t_2 is still in the denominator and we still can not apply Watson's lemma. However, we can recursive apply the above technique of change of variables and decomposition. Since only t_3 is absorbed, this term has an error of $O(n^{-\frac{1}{2}})$ with regard to $I^{\{1\}}$, and therefore can be omitted. The same arguments holds for $I_{10}^{\{1\}}$. And finally since t_1 is not correlated with t_2 and t_3, it can be separated and by Watson's lemma, we have

$$
\begin{aligned}
I^{\{1\}} \dot{\sim} I_{00}^{\{1\}} = & C^{\{1\}} \int_{\hat{\mathbf{t}}-i\infty}^{\hat{\mathbf{t}}+i\infty} \frac{\exp(Q(\mathbf{t}))}{(2\pi i)^3 t_2 t_3} h^{**}(t_1,0,0)\, d\mathbf{t} \\
\dot{\sim} & \frac{C^{\{1\}} h^{\{1\}}(\hat{w}_1)}{\sqrt{n c_{11}^{\{1\}}}} \phi(x^{\{1\}}) \cdot \\
& \bar{\Phi}\left(\frac{y^{\{1\}} - \rho_{12}^{\{1\}} x^{\{1\}}}{\sqrt{1-(\rho_{12}^{\{1\}})^2}}, \frac{z^{\{1\}} - \rho_{13}^{\{1\}} x^{\{1\}}}{\sqrt{1-(\rho_{13}^{\{1\}})^2}}, \frac{\rho_{23}^{\{1\}} - \rho_{12}^{\{1\}}\rho_{13}^{\{1\}}}{\sqrt{1-(\rho_{12}^{\{1\}})^2}\sqrt{1-(\rho_{13}^{\{1\}})^2}} \right).
\end{aligned}
\tag{5.5.16}
$$

Similarly, we have

$$\begin{aligned}I^{\{2\}} =& \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 + \frac{1}{2}w_3^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2 - \hat{w}_3 w_3])}{(2\pi i)^3 w_1 (w_3 - \tilde{w}_3(w_1, w_2))} \cdot \\ & \frac{G(0,\tau_2,0) - G(0,0,0)}{w_2 - \tilde{w}_2(w_1)}\, d\mathbf{w} \\ \dot{\sim}& \frac{C^{\{2\}} h^{\{2\}}(\hat{w}_2)}{\sqrt{n c_{22}^{\{2\}}}} \phi(y^{\{2\}}) \cdot \\ & \bar{\Phi}\left(\frac{x^{\{2\}} - \rho_{12}^{\{1\}} y^{\{2\}}}{\sqrt{1-(\rho_{12}^{\{2\}})^2}}, \frac{z^{\{2\}} - \rho_{23}^{\{2\}} y^{\{2\}}}{\sqrt{1-(\rho_{23}^{\{2\}})^2}}, \frac{\rho_{13}^{\{2\}} - \rho_{12}^{\{2\}}\rho_{23}^{\{2\}}}{\sqrt{1-(\rho_{12}^{\{2\}})^2}\sqrt{1-(\rho_{23}^{\{2\}})^2}} \right),\end{aligned} \tag{5.5.17}$$

where now we have $g^{\{2\}}(u_1, u_2, u_3)$, the exponent as a function of $\mathbf{u}$ with the change of variable $u_1 = w_1$ $u_2 = w_2$ and $u_3 = w_3 - \tilde{w}_3(w_1, w_2)$. $c_{00}^{\{2\}} = g^{\{2\}}(\hat{\mathbf{u}})$, $\mathbf{cc}^{\{2\}}$ is the matrix such that its element $cc_{ij}^{\{2\}} = [g^{\{2\}}]^{ij}(\hat{\mathbf{u}})$, $\mathbf{c}^{\{2\}}$ is the vector such that $c_i^{\{2\}} = [g^{\{2\}}]^i(\hat{\mathbf{u}})$, $C^{\{2\}} = \exp(n[c_{00}^{\{2\}} + \frac{1}{2}\hat{\mathbf{u}}^T \mathbf{cc}^{\{2\}} \hat{\mathbf{u}}])$, $\boldsymbol{\rho}^{\{2\}}$ is the covariance matrix with $\rho_{ij}^{\{2\}} = c_{ij}^{\{2\}} / \sqrt{c_{ii}^{\{2\}} c_{jj}^{\{2\}}}$ and $x^{\{2\}} = \frac{\sqrt{n}(c_{11}^{\{2\}}\hat{u}_1 + c_{12}^{\{2\}}\hat{u}_2 + c_{13}^{\{2\}}\hat{u}_3)}{\sqrt{c_{11}^{\{2\}}}}$, $y^{\{2\}} = \frac{\sqrt{n}(c_{12}^{\{2\}}\hat{u}_1 + c_{22}^{\{2\}}\hat{u}_2 + c_{23}^{\{2\}}\hat{u}_3)}{\sqrt{c_{22}^{\{2\}}}}$, $z^{\{2\}} = \frac{\sqrt{n}(c_{13}^{\{2\}}\hat{u}_1 + c_{23}^{\{2\}}\hat{u}_2 + c_{33}^{\{2\}}\hat{u}_3)}{\sqrt{c_{33}^{\{2\}}}}$ and

$$\begin{aligned}h^{\{2\}}(\hat{w}_2) =& \left(\frac{1}{\tau_2}\frac{d\tau_2}{dw_2} - \frac{1}{w_2 - \tilde{w}_2(w_1)} \right)\Bigg|_{(0,\hat{w}_2)} = \frac{1}{\tilde{\tau}_2(0)}\, \frac{d\tau_2}{dw_2}\Bigg|_{(0,\hat{w}_2)} - \frac{1}{\hat{w}_2} \\ =& \frac{1}{\tilde{\tau}_2(0)\sqrt{K^{22}(0,\tilde{\tau}_2(0),\tilde{\tau}_3(0)) + K^{23}(0,\tilde{\tau}_2(0),\tilde{\tau}_3(0))\tilde{\tau}_3^2(0,\tilde{\tau}_2(0))}} - \frac{1}{\hat{w}_2}.\end{aligned}$$

And lastly, we obtain

$$\begin{aligned}I^{\{3\}} =& \int_{\hat{\mathbf{w}}-i\infty}^{\hat{\mathbf{w}}+i\infty} \frac{\exp(n[\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 + \frac{1}{2}w_3^2 - \hat{w}_1 w_1 - \hat{w}_2 w_2 - \hat{w}_3 w_3])}{(2\pi i)^3 w_1 (w_2 - \tilde{w}_2(w_1))} \cdot \\ & \frac{G(0,0,\tau_3) - G(0,0,0)}{w_3 - \tilde{w}_3(w_1, w_2)}\, d\mathbf{w} \\ \dot{\sim}& \frac{C^{\{3\}} h^{\{3\}}(\hat{w}_3)}{\sqrt{n c_{33}^{\{3\}}}} \phi(z^{\{3\}}) \cdot \\ & \bar{\Phi}\left(\frac{x^{\{3\}} - \rho_{13}^{\{3\}} z^{\{3\}}}{\sqrt{1-(\rho_{13}^{\{3\}})^2}}, \frac{y^{\{3\}} - \rho_{23}^{\{3\}} z^{\{3\}}}{\sqrt{1-(\rho_{23}^{\{3\}})^2}}, \frac{\rho_{12}^{\{3\}} - \rho_{13}^{\{3\}}\rho_{23}^{\{3\}}}{\sqrt{1-(\rho_{13}^{\{3\}})^2}\sqrt{1-(\rho_{23}^{\{3\}})^2}} \right),\end{aligned} \tag{5.5.18}$$

where again we have $g^{\{3\}}(u_1, u_2, u_3)$, the quadratic form of $\mathbf{w}$ substituted by $\mathbf{u}$ with the change of variable $u_1 = w_1$ $u_2 = w_2 - \tilde{w}_2(w_1)$ and $u_3 = w_3$. $c_{00}^{\{3\}} = g^{\{3\}}(\hat{\mathbf{u}})$, $\mathbf{cc}^{\{3\}}$ is the matrix such that its element $cc_{ij}^{\{3\}} = [g^{\{3\}}]^{ij}(\hat{\mathbf{u}})$, $\mathbf{c}^{\{3\}}$ is the vector such that $c_i^{\{3\}} = [g^{\{3\}}]^i(\hat{\mathbf{u}})$, $C^{\{3\}} = \exp(n[c_{00}^{\{3\}} + \frac{1}{2}\hat{\mathbf{u}}^T \mathbf{cc}^{\{3\}} \hat{\mathbf{u}}])$, $\boldsymbol{\rho}^{\{3\}}$ is the covariance matrix with $\rho_{ij}^{\{3\}} =$

$c_{ij}^{\{3\}}/\sqrt{c_{ii}^{\{3\}}c_{jj}^{\{3\}}}$ and $x^{\{3\}} = \frac{\sqrt{n}(c_{11}^{\{3\}}\hat{u}_1+c_{12}^{\{3\}}\hat{u}_2+c_{13}^{\{3\}}\hat{u}_3)}{\sqrt{c_{11}^{\{3\}}}}$, $y^{\{3\}} = \frac{\sqrt{n}(c_{12}^{\{3\}}\hat{u}_1+c_{22}^{\{3\}}\hat{u}_2+c_{23}^{\{3\}}\hat{u}_3)}{\sqrt{c_{22}^{\{3\}}}}$, $z^{\{3\}} = \frac{\sqrt{n}(c_{13}^{\{3\}}\hat{u}_1+c_{23}^{\{3\}}\hat{u}_2+c_{33}^{\{3\}}\hat{u}_3)}{\sqrt{c_{33}^{\{3\}}}}$ and

$$\begin{aligned} h^{\{3\}}(\hat{w}_2) = &\left(\frac{1}{\tau_3}\frac{d\tau_3}{dw_3} - \frac{1}{w_3 - \tilde{w}_3(w_1, w_2)}\right)\Bigg|_{(0,0,\hat{w}_3)} = \frac{1}{\tilde{\tau}_3(0,0)}\,\frac{d\tau_3}{dw_3}\Bigg|_{(0,0,\hat{w}_3)} - \frac{1}{\hat{w}_3} \\ =&\frac{1}{\tilde{\tau}_3(0,0)\sqrt{K^{33}(0,0,\tilde{\tau}_3(0,0))}} - \frac{1}{\hat{w}_3}. \end{aligned}$$

Let $X_i,\ i = 1, 2, 3, 4$ be independent and identically distributed random variables following the exponential distribution as in the first example. Consider the random vector (Y_1, Y_2, Y_3) with $Y_1 = X_0 + X_1,\ Y_2 = X_0 + X_2,\ Y_3 = X_0 + X_3$. We can calculate the cumulant generating function $K(\tau_1, \tau_2, \tau_3)$, $K^j(\tau_1, \tau_2, \tau_3)$, $K^{jk}(\tau_1, \tau_2, \tau_3)$ and $K^{jkl}(\tau_1, \tau_2, \tau_3)$. These then can be used to calculate $\tilde{\tau}_2(\tau_1)$, $\tilde{\tau}_3(\tau_1)$, $\tilde{\tau}_3(\tau_1, \tau_2)$, $\tilde{\tau}_2'(\tau_1)$, $\tilde{\tau}_3'(\tau_1)$, $\tilde{\tau}_3^1(\tau_1, \tau_2)$, $\tilde{\tau}_3^2(\tau_1, \tau_2)$, $\tilde{\tau}_2''(\tau_1)$, $\tilde{\tau}_3''(\tau_1)$, $\tilde{\tau}_3^{11}(\tau_1, \tau_2)$, $\tilde{\tau}_3^{12}(\tau_1, \tau_2)$ and $\tilde{\tau}_3^{22}(\tau_1, \tau_2)$. All the above quantities can then be used to calculate $\frac{d\tau_1}{dw_1}\Big|_{\hat{w}_1}$, $\frac{d^2\tau_1}{dw_1^2}\Big|_{\hat{w}_1}$, $\frac{d\tau_2}{dw_1}\Big|_{\hat{w}_1,\hat{w}_2}$, $\frac{d\tau_2}{dw_2}\Big|_{\hat{w}_1,\hat{w}_2}$, $\frac{d^2\tau_2}{dw_1^2}\Big|_{\hat{w}_1,\hat{w}_2}$, $\frac{d^2\tau_2}{dw_2^2}\Big|_{\hat{w}_1,\hat{w}_2}$ and $\frac{d^2\tau_2}{dw_1 dw_2}\Big|_{\hat{w}_1,\hat{w}_2}$. And then $\check{w}_2$, $\check{w}_2'$, $\check{w}_2''$, $\check{w}_3$, $\check{w}_3^1$, $\check{w}_3^2$, $\check{w}_3^{11}$, $\check{w}_3^{12}$ and $\check{w}_3^{22}$ can be calculated to obtain b_j^r, b_j^{rs} and c_j^{rs}. The results of approximation of $P(\bar{Y}_1 \geq \bar{y}_1, \bar{Y}_2 \geq \bar{y}_2, \bar{Y}_3 \geq \bar{y}_3)$ for $n = 4$ are shown in the table below. Here, "P. approx" stands for the results of the new method. "N. approx" stands for the results of the normal approximation. The relative error of the new method are calculated. We can see from the table that the normal approximation deteriorate at the far tail, while the saddlepoint approximation is quite stable over the range we considered.

Table 5.2: Results of saddlepoint approximation compared with normal approximations in the three dimensional case.

$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	P. approx	N. approx.	Exact	Relative Error
2.5	2.6	2.7	5.51×10^{-2}	4.83×10^{-2}	5.49×10^{-2}	0.36%
3	3.1	3.2	1.44×10^{-2}	6.51×10^{-3}	1.44×10^{-2}	0.00%
3.5	3.6	3.7	3.39×10^{-3}	4.66×10^{-4}	3.38×10^{-3}	0.30%
4	4.1	4.2	7.32×10^{-4}	1.71×10^{-5}	7.30×10^{-4}	0.27%

Chapter 6

Conclusion

Hypothesis testing requires the computation of the tail probability of sufficient statistics or sufficient statistics conditioned on others. Therefore, we would like to approximate the tail probability of a sufficient statistics, in particular, the mean of independent identically distributed random variables, for both conditional and unconditional distributions. The Edgeworth approximations, and the normal approximations as a special case, bound absolution error rather than relative error. The Edgeworth approximations may not work well in the tail area, and in the extreme case, can even be negative. The saddlepoint approximations bound the relative error. [Daniels 1987] summarizes two univariate tail probability methods, one of which is the Lugannai and Rice approximation. In the thesis, we extend the method to multivariate distributions.

In univariate case, the Lugannani and Rice approximation works as follows. It splits the inverse integral

$$\frac{1}{2\pi i}\int_{\hat{w}-i\infty}^{\hat{w}+i\infty}\frac{\exp(n[\frac{1}{2}w^2-\hat{w}w])}{\tau(w)}\frac{d\tau}{dw}\,d\mathbf{w}$$

into two parts. The main term, i.e.,

$$\frac{1}{2\pi i}\int_{\hat{w}-i\infty}^{\hat{w}+i\infty}\frac{\exp(n[\frac{1}{2}w^2-\hat{w}w])}{w}\,d\mathbf{w},$$

contains an integrant that has the singularity, but has explicit formula, i.e., the normal tail probability. The other, i.e.,

$$\frac{1}{2\pi i}\int_{\hat{w}-i\infty}^{\hat{w}+i\infty}\exp(n[\frac{1}{2}w^2-\hat{w}w])\left(\frac{1}{\tau}\frac{d\tau}{dw}-\frac{1}{w}\right)\,d\mathbf{w},$$

contains an integrant that's analytic (removable singularity), which can be approximated by applying Watson's Lemma. The extension of the idea to multivariate approximation turns out not so easy.

First, we define an appropriate mapping between the variables $\boldsymbol{\tau}$ and $\mathbf{w}$. The definition we use as in (2.0.2a) and (2.0.2b) utilizes the log likelihood ratio statistics. In the definition, $\boldsymbol{\tau}_j$ is only a function of $\mathbf{w}_j$ and vise versa, and therefore the Jacobian matrix is simplified to a product $\prod_{j=1}^{d} \frac{d\tau_j}{dw_j}$.

Second, except for the first variable w_1, the singular points for other variables depend on all previous results. The singular point corresponding to $\tau_j = 0$, in general is not $w_j = 0$, but $w_j = \tilde{w}_j(\mathbf{w}_{j-1})$, for some function $\tilde{w}_j$ that depends on all previous variables $\mathbf{w}_{j-1}$. Unlike the main term in the Lugannani and Rice approximation, where the denominator is simply w, in the multivariate scenario, the denominator becomes $\prod_{j=1}^{d_0}(w_j - \tilde{w}_j(\mathbf{w}_{j-1})))$, rather than simplify $\prod_{j=1}^{d_0} w_j$. Therefore, the main term, in general, would not be exact multivariate normal tail probability. In our first method, we use a change of variables and the main terms becomes a multivariate normal tail probability, times a coefficient. In the second method, we removed the coefficient by first reducing the function $\tilde{w}_j$ to the form $\sum_{l=1}^{j-1} a_j^l(\mathbf{w_l})w_l$, and then further to $\sum_{k,l} c_j^{kl} w_k(w_l - \hat{w}_l)$, and lastly viewing the integral as a function c_j^{kl} and expanding around $\mathbf{0}$.

The approximation for variables taking unit lattice values is similar. We need to take care of the continuity correct, which is straightforward. We also have $2\sinh(\tau/2)$ in the denominator of the integrant rather than τ, but this is not an issue since both have the same singularity property. The approximation for conditional distribution requires us to do some additional work for the conditioning variables. However, the above method is still valid with the introduction of the conditioning variables. And we incorporate the conditional case into the same framework.

An important property of the Lugannani and Rice's approximation is the reflexivity property. In saddlepoint approximation, the path of integral from $\mathbf{c} - i\mathbf{K}$ to $\mathbf{c} + i\mathbf{K}$ has to have positive real part, i.e., $\mathbf{c} > 0$. This positivity restriction in general require the saddlepoints $\boldsymbol{\tau}$ to be also positive. In case they do not satisfy the requirement, the Bool's law must be used. The reflexivity property enable us to apply the approximation formula directly even if the positivity restriction is violated. We proved the bivariate version of the reflexivity property for our second method.

The approximation constructed in my thesis is a multivariate analog of the approximation of Lugannani and Rice (1980). One possible future work is to construct an alternative approximation, which incorporates adjustments of $O(n^{-\frac{1}{2}})$ as corrections to the likelihood ratio statistic, in analogy with the work of [Jensen 1992]. I will start with a variation of the integral in (2.0.10) as the following:

$$
\begin{aligned}
&\frac{n^{d-d_0}}{(2\pi i)^d}\int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w}-\hat{\mathbf{w}}^T\mathbf{w}])}{\prod_{j=1}^{d_0}\rho(\tau_j(\mathbf{w}_j))}\prod_{j=1}^{d}\frac{d\tau_j}{dw_j}\,d\mathbf{w} \\
=&\frac{n^{d-d_0}}{(2\pi i)^d}\int_{\hat{\mathbf{w}}-i\mathbf{K}}^{\hat{\mathbf{w}}+i\mathbf{K}} \frac{\exp(n[\frac{1}{2}\mathbf{w}^T\mathbf{w}-\hat{\mathbf{w}}^T\mathbf{w}-\frac{1}{n}\log\prod_{j=1}^{d_0}\frac{dw_j}{d\tau_j}\frac{\rho(\tau_j(\mathbf{w}_j))}{w_j}])}{\prod_{j=1}^{d_0}w_j}G(\boldsymbol{\tau})\,d\mathbf{w},
\end{aligned}
\tag{6.0.1}
$$

If we could derive a saddlepoint approximation in the form of $\bar{\Phi}(w_1^*, w_2^*)$ in the bivariate case, for instance, then w_1^* and w_2^* can be viewed as adjusted likelihood ratio statistics and be used directly as test statistics for a group of hypothesis testing problems. Another advantage of Barndorff and Nielsen's method is that it is more compact.

The second topic involves integration within the curved angle. Take, for example, the bivariate case. It is relatively easy to obtain $P(\bar{X}_1 \geq \bar{x}_1, \bar{X}_2 \geq \bar{x}_2)$, a rectangle on the $(\bar{X}_1, \bar{X}_2)$ space, through integration over the rectangular area. It is also relatively east to obtain $P(\bar{X}_1 \geq \bar{x}_1, a\bar{X}_1 + b\bar{X}_2 \geq \bar{x}_2)$, an angular area, since we can do a linear transformation of variables to turn it into a rectangular area for the new variables. However, it remains a question in the (w_1^*, w_2^*) space, where and transformation is nonlinear, and the integration area become curve-angular. It then would be interesting to study the approximation over this integration and its properties.

References

[Daniels 1954] H.E. Daniels. Saddlepoint approximations in statistics. *Ann. Math. Statist.*. **25**, 631-645, 1954.

[Daniels 1987] H.E. Daniels. Tail probability approximations. *Internat. Statisti. Rev.* **55**, 37-46, 1987.

[Jensen 1992] J.L. Jensen. The modified signed likelihood statistic and saddlepoint approximations. *Biometrika.* **79**, 693–703.

[Kolassa 1997] J.E. Kolassa. *Series Approximation Methods in Statistics*, *Lecture Notes in Statistics*, 88, Springer, 2006.

[Kolassa 2003] J.E. Kolassa. Multivariate Saddlepoint Tail Probability Approximations. *Ann. Statist.* **31** 274-286, 2003

[Kolassa 2004] J.E. Kolassa. Approximate Multivariate Conditional Inference Using the Adjusted Profile Likelihood. *The Canadian J. of Statist.* **32** 5-14, 2004.

[Lugannani and Rice 1980] R. Lugannani and S. Rice. Saddle point approximation for the distribution of the sum of indepednent random variables. *Adv. in Appl. Probab.* **12** 475-490, 1980.

[Mathematica 5.0 2005] Wolfram Research, Inc. Mathematica. Version 5.0. Champaign, IL, 2003

[Reid 1988] N. Reid. Saddlepoint methods and statistical inference (with discussion). *Statistical Sci.* **3** 213-238, 1988.

[Robinson 1982] J. Robinson. Saddlepoint approximation for permutation tests and confidence intervals. *J. Roy. Statist. Soc. Ser. B.* **44** 91-101, 1982.

[Stokes *et al.* 1995] M.E. Stokes, C.S. David and G.G. Koch. Categorical Data Analysis using the SAS System. SAS Institute, Cary, North Carolina.

[Skovgaard 1987] I.M. Skovgaard. Saddlepoint Expansions for Conditional Distributions. *J. Appl. Probab.* **24** 875-887

[Wang 1991] S. Wang. Saddlepoint approximation for bivariate distribution. *J. Appl. Probab.* **27** 586-597, 1991.

[Watson 1948] Watson, G.N. *Theory of Bessel Functions.* Cambridge University Press.

[Yang and Kolassa 2002] Yang, B. and Kolassa, J.E. Saddlepoint approximation for the distribution function near the mean. *Ann. Inst. Statist. Math.* **54** 743-747, 2002

[Zhang and Kolassa 2008] Zhang, J. and Kolassa, J.E. Saddlepoint Approximation for the distribution of the modified signed root of likelihood ratio statistics near the mean. *Comm. Statist. Theory Methods.* **37** 194-203, 2008

Vita

Jixin Li

2009 Ph. D. in Statistics, Rutgers University

2002 M. Sc. in Computer Science, Rutgers University

2000 B. Eng. in Computer Science and Engineering, Zhejiang University

2005-08 Graduate assistant, Department of Statistics, Rutgers University

2000-05 Teaching assistant, Department of Computer Science, Rutgers University

2009 Kolassa, J.E. and Li, J. Multivariate Saddlepoint Approximations in Tail Probability and Conditional Inference. Submitted.

2008 Gunter, M.J., Hoover, D. R., Yu, H., Wassertheil-Smoller, S., Rohan, T.E., Manson, J.E., Li, J., Ho, G.Y.F., Xue, X., Anderson, G.L., Kaplan, R.C., Harris, T.G., Howard, B.V., Wylie-Rosett, J., Burk, R.D. and Strickler, H.D. Insulin, Insulin-Like Growth Factor-I, and Risk of Breast Cancer in Postmenopausal Women. J. Natl. Can-cer Inst., Accepted.

2008 Gunter, M.J., Hoover, D. R., Yu, H., Wassertheil-Smoller, S., Manson, J.E., Li, J., Har-ris, T.G., Rohan, T.E., Xue, X., Ho, G.Y.F., Einstein, M.H., Kaplan, R.C., Burk, R.D., Wylie-Rosett, J., Pollak, M.N., Anderson, G., Howard, B.V. and Strickler, H.D. A Prospective Evaluation of Insulin and Insulin-like Growth Factor-I as Risk Factors for Endometrial Cancer. Cancer Epidemiology Biomarkers and Prevention, 17, 921-929.

2008 Gunter, M.J., Hoover, D. R., Yu, H., Wassertheil-Smoller, S., Rohan, T.E., Manson, J.E., Howard, B.V., Wylie-Rosett, J., Anderson, G.L., Ho, G.Y.F., Kaplan, R.C., Li, J., Xue, X., Harris, T.G., Burk, R.D. and Strickler, H.D. Insulin, Insulin-Like Growth Factor-I, Endogenous Estradiol and Risk of Colorectal Cancer in Postmenopausal Women. Cancer Research, 68, 329-337.

2006 J. Li, I. Muchnik and D. Schneider, Influences on Breast Cancer Survival via SVM Classification in the SEER Database, *Discrete Methods in Epidemiology*, volume 70 of DIMACS. AMS