HIGHER ORDER MULTIVARIATE INFERENCE USING APPROXIMATION METHODS

By

YAOSHI WU

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 and Biostatistics

Written under the direction of

John Kolassa

And approved by

New Brunswick, New Jersey

OCTOBER, 2015

ABSTRACT OF THE DISSERTATION

HIGHER ORDER MULTIVARIATE INFERENCE USING APPROXIMATION METHODS

By

YAOSHI WU

Dissertation Director:

John Kolassa

The bivariate confidence region of two parameters of interest conditioning on one nuisance parameter is evaluated to the error of $O_p(\frac{1}{n})$ using Laplace's approximation. A joint matching prior in the Bayesian framework is developed using Laplace's approximation to establish a mathematical equivalence such that the posterior quartiles coincide with the confidence region defined as in a frequency approach. Quadratic saddlepoint approximation is developed to evaluate bivariate conditional and unconditional tail probability to the error of $O_p(\frac{1}{n})$ using saddlepoint approximation. The quadratic saddlepoint approximation is extended to three dimension to evaluate trivariate unconditional tail probability.

CONTENTS

Part I

INTRODUCTION OF HIGHER ORDER METHODS FOR EVALUATING TAIL PROBABILITY AND BIVARIATE CONFIDENCE REGION

Higher-order statistical inference in terms of p-values, defined as the tail probability of a test statistic, and confidence regions for multi-parameters, are developed using approximation methods, for cases in which tail probability and confidence region do not have closed forms. Tail areas considered in the thesis are intersections of half-spaces. We apply Laplace's approximation method to evaluate bivariate confidence region of two parameters of interest with one nuisance parameter as presented in Part II, and saddlepoint approximation method to evaluate higher oder tail probability as described in Part III.

Laplace's method is a mathematical technique to evaluate integrals in the form of $\int e^{Cf(\underset{\sim}{x})} d\underset{\sim}{x}$ with C being a constant. By expanding $f(\underset{\sim}{x})$ at point $\underset{\sim}{x}_0$ where $f(\underset{\sim}{x})$ reaches its global maximum using Taylor's Theorem, the integral can be approximated by a Gaussian integral. In Part II, I use Laplace's approximation method to develop a prior distribution in the Bayesian framework to establish a mathematical equivalence such that the posterior quartiles coincide with the confidence limit points defined as in a frequency approach, i.e., $Pr((\theta_1, \theta_2) < (h_1(S, \alpha), h_2(S, \alpha))|(\theta_1, \theta_2, \theta_3)) = \alpha$ when $(h_1(S, \alpha), h_2(S, \alpha))$ are the joint bivariate quantiles of the posterior distribution at level of α. A prior distribution that satisfies this mathematical equivalence is called a matching prior. More details regarding the concepts of matching prior are provided in Chapter 2.1.

Saddlepoint approximations are derived by inverting the cumulant generating function $K_T(\underset{\sim}{\xi}) = \log(E[\exp(\underset{\sim}{\xi}^{'}\underset{\sim}{T})])$ giving the probability density in terms of the cumulant generating function. More details are provided in Chapter 4.1 as a background introduction of various saddlepoint approximation methods. In Part III, we develop an approach using saddlepoint approximation to express the higher order tail probability approximation in the forms of normal density functions and normal distribution functions, based on a transformation of variables such that the integrand of the saddlepoint approximation expression can be expressed as a exponential function of linear and quadratic terms of the integrated variables.

Part II

BIVARIATE CONFIDENCE REGION USING LAPLACE'S APPROXIMATION

1. BACKGROUND ON LAPLACE'S APPROXIMATION

One of the approaches to remove the effect of nuisance parameter is to perform conditional inference using the adjusted profile likelihood (Kolassa, 2004), this method and other approaches involved with direct derivation of the conditional probability density function may be complicated in many circumstances. The approach presented here involves a matching prior using Bayesian probability inference (Welch and Peers, 1963, and Peers, 1965) to obtain an equivalent conditional probability without direct derivation of the conditional density. Laplace's method is used when developing the matching prior. The research is focused on the problem of two parameters of interest with one nuisance parameter.

To evaluate an univariate integral in the form of $\int e^{Cf(x)}dx$, with C being constant, one may expand $f(x)$ using Taylor's Theorem as $f(x) = f(x_0) + f(x_0)^{'}(x - x_0) + \frac{1}{2}f(x_0)^{''}(x - x_0)^2 + O((x - x_0)^3)$, where $f(x)$ has a global maximum at x_0 and the second derivative $f(x_0)^{''} < 0$. The first derivative of $f(x)$ vanishes at x_0, and $f(x)$ can be approximated to the quadratic order as $f(x) \approx f(x_0) + \frac{1}{2}f(x_0)^{''}(x - x_0)^2$. Therefore, $\int e^{Cf(x)}dx$ can be approximated by $e^{Cf(x_0)} \int e^{\frac{1}{2}Cf(x_0)^{''}(x-x_0)^2}dx$, where $\int e^{\frac{1}{2}Cf(x_0)^{''}(x-x_0)^2}dx$ with $f(x_0)^{''} < 0$ is a simple Gaussian integral. Similarly, Laplace's approximation can be implemented to evaluate multivariate integral in the form of $\int e^{Cf(\underset{\sim}{x})}d\underset{\sim}{x}$ with $\underset{\sim}{x}$ being a vector.

2. METHODOLOGY

2.1 Matching Prior

There exists a formal mathematical equivalence between Bayesian solutions (posterior probability conditioning on the sample) and confidence theory solutions (confidence point conditioning on the parameter) by finding a proper prior distribution in the framework of single parameter of interest and none or multiple nuisance parameters. Welch and Peers (1963) developed the matching prior for a single parameter of interest with no nuisance parameters, and Peers (1965) extended the single parameter problem to multiple parameter problem with one parameter of interest and all others being nuisance parameters. We will use a similar approach to extend single parameter of interest in presence of other parameters to two parameters of interest in presence of one nuisance parameter. Peers (1965) provided a joint prior $w(t)$ such that the quantile $h(S,\alpha)$ of the posterior distribution satisfying

$$\frac{\int^{h(S,\alpha)}\int\ldots\int p(S,t)w(t)dt}{\int\int\ldots\int p(S,t)w(t)dt} = \alpha \tag{2.1}$$

also satisfies $Pr(\theta_1 < h(S,\alpha))|\underset{\sim}{\theta}) = \alpha$, where θ_1 is the parameter of interest, and $\underset{\sim}{\theta}$ is the vector of all parameters. Write $p(S,t) = \exp(L(S,t)), w(t) = \exp(\psi(t))$, where $L(S,t)$ is the log likelihood function and $\psi(t)$ is the log prior density function. Define a monotonic function of θ_1

$$r(S,\theta_1) = \frac{I(S,\theta_1)}{I(S,\underset{\sim}{\theta})} = \frac{\int^{\theta_1}\int\ldots\int \exp(L(S,t)+\psi(t))dt}{\int\int\ldots\int \exp(L(S,t)+\psi(t))dt}; \tag{2.2}$$

hence $\forall S, Pr(\theta_1 < h(S,\alpha))|\underset{\sim}{\theta}) = Pr(r(S,\theta_1) < r(S,h(S,\alpha))|\underset{\sim}{\theta})$, and note that $r(S,h(S,\alpha)) = \alpha$ implies $Pr(r(S,\theta_1) < \alpha|\underset{\sim}{\theta}) = \alpha$. Write $r(S,\theta_1) = N(z(S,\theta_1))$, using Laplace's approximation, where $N(x)$ is the standard normal integral $(2\pi)^{-\frac{1}{2}}\int^{x}\exp(-\frac{1}{2}u^2)du$. Then $Pr(r(S,\theta_1) < \alpha|\underset{\sim}{\theta}) = \alpha$ becomes

$$Pr(N(z(S,\theta_1)) < \alpha|\underset{\sim}{\theta}) = \alpha.$$

Therefore the problem reduces to finding a function $\psi(t)$ such that $z(S,\theta_1)$ follows the standard normal distribution in repeated sampling (Peers, 1965). Use the moment generating function $E(\exp(tz))$, and set the coefficient of the linear term of t equal to 0 such that $E(\exp(tz))$ is the moment generating function of a standard normal distribution. The differential equation of ψ with respect to θ is therefore obtained, and by solving the differential equation, we can find function of ψ and hence the matching prior.

Addressing the problem of two parameters of interest (θ_1,θ_2) with one nuisance parameter θ_3, we will extend the approach from single parameter of interest to two parameters of interest by letting the two coefficients of the linear terms of $\underset{\sim}{t} = (t_1,t_2)'$ of the bivariate moment generating function $E(\exp(\underset{\sim}{t}'\underset{\sim}{z}))$ equal to 0 such that $E(\exp(\underset{\sim}{t}'\underset{\sim}{z}))$ is the moment generating function of a bivariate normal distribution of $\underset{\sim}{z} = (z_1,z_2)'$; therefore a system of two differential equations of ψ with respect to θ will be obtained, and the joint matching prior can be found by solving the two differential equations.

Extend formula (13) in Peers (1965) to obtain

$$\begin{aligned}\mathbf{I}(S,\theta_1) = \quad & \int^{\theta_1}\int\ldots\int\exp\Big(L(S,T) + \tfrac{1}{2}\textstyle\sum_{i,j}(t_i-T_i)(t_j-T_j)\frac{\partial^2 L(S,T)}{\partial T_i \partial T_j} \\ & +\tfrac{1}{6}\textstyle\sum_{i,j,k}(t_i-T_i)(t_j-T_j)(t_k-T_k)\frac{\partial^3 L(S,T)}{\partial T_i\partial T_j\partial T_k} + \Psi(T) \\ & +\textstyle\sum_i(t_i-T_i)\frac{\partial\Psi(T)}{\partial T_i} + \mathbf{O_p}(n^{-1})\Big)dt,\end{aligned} \tag{2.3}$$

where T_i is the unique maximum likelihood estimator of θ_i, and

$$\begin{aligned}\mathbf{I}(S,\theta_1,\theta_2) = \quad & \int^{\theta_1}\int^{\theta_2}\int \exp\Big(L(S,T)+\tfrac{1}{2}\textstyle\sum_{i,j=1,2,3}(t_i-T_i)(t_j-T_j)\frac{\partial^2 L(S,T)}{\partial T_i \partial T_j} \\ & +\tfrac{1}{6}\textstyle\sum_{i,j,k=1,2,3}(t_i-T_i)(t_j-T_j)(t_k-T_k)\frac{\partial^3 L(S,T)}{\partial T_i\partial T_j\partial T_k}+\Psi(T) \\ & +\textstyle\sum_i(t_i-T_i)\frac{\partial\Psi(T)}{\partial T_i}+\mathbf{O_p}(n^{-1})\Big)dt_1dt_2dt_3.\end{aligned} \tag{2.4}$$

Following the notation used in Peers (1965), define

$$\begin{aligned} & v_{ij} = -n^{-1}\tfrac{\partial^2 L}{\partial T_i \partial T_j}, \\ & v^{ij} = \text{the } (i,j) \text{ element of } (v_{ij})^{-1}, \\ & v_{ijk} = n^{-\frac{3}{2}}\tfrac{\partial^3 L}{\partial T_i \partial T_j \partial T_k}, \\ & u_i = n^{\frac{1}{2}}(t_i - T_i)(v^{ii})^{-\frac{1}{2}}, \\ & x_i = n^{\frac{1}{2}}(\theta_i - T_i)(v^{ii})^{-\frac{1}{2}}, \text{for } i=1,2, \\ & w_i = n^{-\frac{1}{2}}(v^{ii})^{\frac{1}{2}}\tfrac{\partial\psi}{\partial T_i}, w_{ij} = (v^{ii}v^{jj})^{\frac{1}{2}}v_{ij}, \\ & w_{ijk} = (v^{ii}v^{jj}v^{kk})^{\frac{1}{2}}v_{ijk}. \end{aligned} \tag{2.5}$$

Define $Q(\underset{\sim}{u})$ to be the exponent in (2.3). Then

$$Q(\underset{\sim}{u}) = -\frac{1}{2}\sum_{i,j} w_{ij}u_iu_j + \frac{1}{6}\sum_{i,j,k} w_{ijk}u_iu_ju_k + \sum_i w_iu_i + \mathbf{O}(n^{-1}).$$

Apply formula (19) in Peers (1965)

$$r(S,\theta_1) = \frac{(2\pi)^{-\frac{3}{2}}||w_{ij}||^{\frac{1}{2}}\int^x\int\ldots\int e^{Q(\underset{\sim}{u})}du}{(2\pi)^{-\frac{3}{2}}||w_{ij}||^{\frac{1}{2}}\int\int\ldots\int e^{Q(\underset{\sim}{u})}du} \tag{2.6}$$

to the problem of two parameters of interest with one nuisance parameter, to obtain

$$\begin{aligned} r(S,\theta_1,\theta_2) \propto \quad & (2\pi)^{-\frac{3}{2}}||w_{ij}||^{\frac{1}{2}}\int\int^{x_1}\int^{x_2}\exp\Big(-\tfrac{1}{2}\textstyle\sum_{i,j=1,2,3}w_{ij}u_iu_j \\ & +\tfrac{1}{6}\textstyle\sum_{i,j,k=1,2,3}w_{ijk}u_iu_ju_k+\sum_{i=1}^{3}w_iu_i+\mathbf{O_p}(n^{-1})\Big)du_1du_2du_3.\end{aligned} \tag{2.7}$$

To write (2.7) in terms of bivariate normal integrals, first expand the integrand using

Taylor's expansion except for the u_i^2 terms in the exponent(details are given in Section 2.3). The expansion of (2.7) can be expressed as the summation of the first three moments of a truncated multinormal distribution truncated by the upper bounds $(x_1, x_2) < (\infty, \infty)$. Tallis (1961) provides explicit expressions for the first and second moment, and no results for higher-order moments are available in the literature; In the following section, we will derive an explicit expression for the third moment.

2.2 Third Moment of a Truncated Multinormal Distribution

An explicit formula for the third moment of a truncated multinormal distribution is given in the following theorem.

Theorem 2.1. *Suppose $\underset{\sim}{x}$ follows a truncated multinormal distribution with the truncation by the lower bounds $-\infty \leq a_s < \infty$, then the third moment of truncated multinormal distribution*

$$\begin{aligned}
&\alpha E(X_iX_jX_k) = \left(\rho_{ij}\tfrac{\partial \Phi_d(R)}{\partial t_k} + \rho_{ik}\tfrac{\partial \Phi_d(R)}{\partial t_j} + \rho_{kj}\tfrac{\partial \Phi_d(R)}{\partial t_i} + \tfrac{\partial^3 \Phi_d(R)}{\partial t_k \partial t_j \partial t_i}\right)\Big|_{\underset{\sim}{t}=\underset{\sim}{0}} \\
&= \textstyle\sum_{l=1}^{d}(\rho_{ij}\rho_{kl} + \rho_{ik}\rho_{jl} + \rho_{kj}\rho_{il} + (1+a_l^2)\rho_{il}\rho_{jl}\rho_{kl})\phi(a_l)\Phi_{d-1}(A_{ls}; R_l) \\
&+ \textstyle\sum_{l=1}^{d}\rho_{il}\sum_{q\neq l}\big(\rho_{jl}(\rho_{kq} - \rho_{lq}\rho_{lk})a_l + (\rho_{jq} - \rho_{lq}\rho_{lj})(\rho_{lk}(a_l + \rho^{lq}a_q) + \\
&\rho_{qk}(a_q + \rho^{lq}a_l))\big)\phi(a_l, a_q; \rho_{lq})\Phi_{d-2}(A^l_{qs}; R_{lq}) + \textstyle\sum_{l=1}^{d}\rho_{il}\sum_{q\neq l}(\rho_{jq} - \rho_{lq}\rho_{lj}) \\
&\cdot \textstyle\sum_{r\neq l,q}(\rho_{rk} - \beta_{rl.q}\rho_{lk} - \beta_{rq.l}\rho_{qk})\phi(a_l, a_q, a_r; \rho_{lqr})\Phi_{d-3}(A^{lq}_{rs}; R_{lqr}),
\end{aligned} \tag{2.8}$$

where $\alpha = Pr(W_1 > a_1, W_2 > a_2, \cdots, W_d > a_d)$, $W_s (s = 1, \cdots, d)$ has the standard normal distribution with correlation matrix $R = (\rho_{ij})$,
$\Phi_d(R) = (2\pi)^{-\frac{d}{2}}|R|^{-\frac{d}{2}}\ {}^{(d)}\!\int_{a_s}^{\infty} e^{-\frac{1}{2}\underset{\sim}{x}' R^{-1}\underset{\sim}{x}} dx_s$, ${}^{(d)}\!\int$ denotes the d-dimensional integral, A_{ls}, A^l_{qs}, A^{lq}_{rs}, and β's are defined in the proof.

Proof. Tallis (1961) showed that the moment generating function of a truncated multinormal distribution can be written as

$$\alpha m(\underset{\sim}{t}) = e^T \Phi_d(R),$$

where $T = \frac{1}{2}\underset{\sim}{t}' R \underset{\sim}{t}$, $b_s = a_s - \xi_s$, $\underset{\sim}{\xi} = R\underset{\sim}{t}$, α and $\Phi_d(R)$ are specified in the theorem. Then the third moment is

$$\alpha E(X_i X_j X_k) \quad = \alpha \frac{\partial^3 m(\underset{\sim}{t})}{\partial t_k \partial t_j \partial t_i}\Big|_{\underset{\sim}{t}=\underset{\sim}{0}},$$

and

$$\begin{aligned}
\alpha \frac{\partial^3 m(\underset{\sim}{t})}{\partial t_k \partial t_j \partial t_i} &= \frac{\partial^2}{\partial t_k \partial t_j}\left(\frac{\partial e^T}{\partial t_i}\Phi_d(R) + e^T \frac{\partial \Phi_d(R)}{\partial t_i}\right) \\
&= \frac{\partial}{\partial t_k}\left(\frac{\partial^2 e^T}{\partial t_j \partial t_i}\Phi_d(R) + \frac{\partial e^T}{\partial t_i}\frac{\partial \Phi_d(R)}{\partial t_j} + \frac{\partial e^T}{\partial t_j}\frac{\partial \Phi_d(R)}{\partial t_i} + e^T \frac{\partial^2 \Phi_d(R)}{\partial t_j \partial t_i}\right) \qquad (2.9) \\
&= \frac{\partial^3 e^T}{\partial t_k \partial t_j \partial t_i}\Phi_d(R) + e^T \frac{\partial^3 \Phi_d(R)}{\partial t_k \partial t_j \partial t_i} \\
&+ \frac{\partial^2 e^T}{\partial t_j \partial t_i}\frac{\partial \Phi_d(R)}{\partial t_k} + \frac{\partial^2 e^T}{\partial t_k \partial t_i}\frac{\partial \Phi_d(R)}{\partial t_j} + \frac{\partial^2 e^T}{\partial t_i \partial t_k}\frac{\partial \Phi_d(R)}{\partial t_j} \\
&+ \frac{\partial e^T}{\partial t_i}\frac{\partial^2 \Phi_d(R)}{\partial t_k \partial t_j} + \frac{\partial e^T}{\partial t_j}\frac{\partial^2 \Phi_d(R)}{\partial t_k \partial t_i} + \frac{\partial e^T}{\partial t_k}\frac{\partial^2 \Phi_d(R)}{\partial t_j \partial t_i}.
\end{aligned}$$

Since

$$\begin{aligned}
&\frac{\partial e^T}{\partial t_i}\Big|_{\underset{\sim}{t}=\underset{\sim}{0}} = e^T \sum_{l=1}^d \rho_{il} t_l \Big|_{\underset{\sim}{t}=\underset{\sim}{0}} = 0, \\
&e^T|_{\underset{\sim}{t}=\underset{\sim}{0}} = 1, \\
&\frac{\partial^2 e^T}{\partial t_j \partial t_i}\Big|_{\underset{\sim}{t}=\underset{\sim}{0}} = \rho_{ij}, \text{ and} \\
&\frac{\partial^3 e^T}{\partial t_k \partial t_j \partial t_i} = \left(\frac{\partial^2 e^T}{\partial t_k \partial t_j}\sum_l \rho_{il} t_l + \frac{\partial e^T}{\partial t_j}\rho_{ik} + \frac{\partial e^T}{\partial t_k}\rho_{ij}\right)\Big|_{\underset{\sim}{t}=\underset{\sim}{0}} = 0,
\end{aligned}$$

therefore, (2.9) can be simplified as

$$\alpha E(X_i X_j X_k) \quad = \left(\rho_{ij}\frac{\partial \Phi_d(R)}{\partial t_k} + \rho_{ik}\frac{\partial \Phi_d(R)}{\partial t_j} + \rho_{kj}\frac{\partial \Phi_d(R)}{\partial t_i} + \frac{\partial^3 \Phi_d(R)}{\partial t_k \partial t_j \partial t_i}\right)\Bigg|_{\underset{\sim}{t}=\underset{\sim}{0}}. \qquad (2.10)$$

Denote $\phi_d(x_s; R)$ as the probability density function of standard d-dimensional multinormal distribution with correlation coefficient matrix $R = (\rho_{ij})$, we have

$$\phi_d(x_s, x_l = b_l; R) = \phi(b_l)\phi_{d-1}(y_s; R_l), s \neq l,$$

$$\phi_d(x_s, x_l = b_l, x_q = b_q; R) = \phi_2(b_l, b_q; \rho_{lq})\phi_{d-2}(z_s; R_{lq}), s \neq l, q,$$

where

$$y_s = (x_s - \rho_{sl} b_l)/\sqrt{1 - \rho_{sl}^2},$$

$$z_s = (x_s - \beta_{sl.q} b_l - \beta_{sq.l} b_q)/(\sqrt{1 - \rho_{sl}^2}\sqrt{1 - \rho_{sq.l}^2}),$$

$\beta_{sl.q}$ and $\beta_{sq.l}$ are the partial regression coefficients of X_s on X_q and X_l respectively, $\rho_{sq.l}$ is the partial correlation coefficient between X_s and X_q for fixed X_l, and R_l, R_{lq} are the correlation coefficient matrices without x_l, x_l and x_q,respectively.

Write cumulative distribution function $\Phi_d(R)$ as the integral of probability density function ϕ_d

$$\Phi_d(R) =^{(d)} \int_{b_s}^{\infty} \phi_d(x_s; R) dx_s,$$

and

$$^{(d-1)} \int_{b_s}^{\infty} \phi_d(x_s, b_l; R) dx_s = \phi(b_l)\Phi_{d-1}(B_{ls}; R_l), s \neq l,$$

$$^{(d-2)} \int_{b_s}^{\infty} \phi_d(x_s, b_l, b_q; R) dx_s = \phi_2(b_l, b_q; \rho_{lq})\Phi_{d-2}(B_{qs}^l; R_{lq}), s \neq l, q,$$

$$^{(d-3)} \int_{b_s}^{\infty} \phi_d(x_s, b_l, b_q, b_r; R) dx_s = \phi_3(b_l, b_q, b_r; \rho_{lqr})\Phi_{d-3}(B_{rs}^{lq}; R_{lqr}), s \neq l, q, r,$$

where

$$B_{ls} = (b_s - \rho_{ls} b_l)/\sqrt{1 - \rho_{sl}^2} \text{ and}$$

$$B_{qs}^l = (b_s - \beta_{sl.q} b_l - \beta_{sq.l} b_q)/(\sqrt{1 - \rho_{sl}^2}\sqrt{1 - \rho_{sq.l}^2}).$$

Therefore

$$\frac{\partial\Phi_d(R)}{\partial t_i} = \sum_l \rho_{il}\phi(b_l)\Phi_{d-1}(B_{ls};R_l),$$
$$\frac{\partial^2\Phi_d(R)}{\partial t_j\partial t_i} = \sum_l \rho_{il}\big(\frac{\partial\phi(b_l)}{\partial t_j}\Phi_{d-1}(B_{ls};R_l)) + \phi(b_l)\frac{\partial\Phi_{d-1}(B_{ls};R_l)}{\partial t_j}\big)$$
$$= \sum_l \rho_{il}\rho_{jl}b_l\phi(b_l)\Phi_{d-1}(B_{ls};R_l)+$$
$$\sum_l \rho_{il}\sum_{q\neq l}(\rho_{jq}-\rho_{lq}\rho_{lj})\phi_2(b_l,b_q;\rho_{lq})\Phi_{d-2}(B^l_{qs};R_{lq}))$$
$$\frac{\partial^3\Phi_d(R)}{\partial t_k\partial t_j\partial t_i} = \sum_l \rho_{il}\rho_{jl}\frac{\partial b_l\phi(b_l)\Phi_{d-1}(B_{ls};R_l)}{\partial t_k}+$$
$$\sum_l \rho_{il}\sum_{q\neq l}(\rho_{jq}-\rho_{lq}\rho_{lj})\frac{\partial\phi_2(b_l,b_q;\rho_{lq})\Phi_{d-2}(B^l_{qs};R_{lq})}{\partial t_k}.$$

In the first term of the third differentiation of $\Phi_d(R)$,

$$\frac{\partial b_l\phi(b_l)\Phi_{d-1}(B_{ls};R_l)}{\partial t_k} = \frac{\partial b_l}{\partial t_k}\phi(b_l)\Phi_{d-1}(B_{ls};R_l) + b_l\frac{\partial\phi(b_l)\Phi_{d-1}(B_{ls};R_l)}{\partial t_k}$$
$$= \rho_{lk}\phi(b_l)\Phi_{d-1}(B_{ls};R_l) + \rho_{lk}b_l^2\phi(b_l)\Phi_{d-1}(B_{ls};R_l)$$
$$+\sum_{q\neq l}\phi(b_l,b_q;\rho_{lq})\Phi_{d-2}(B^l_{qs};R_{lq})(\rho_{kq}-\rho_{lq}\rho_{lk}).$$

In the second term of the third differentiation of $\Phi_d(R)$,

$$\frac{\partial\phi(b_l,b_q;\rho_{lq})}{\partial t_k} = \phi(b_l,b_q;\rho_{lq})\frac{\partial-\frac{1}{2}\underset{\sim}{b}'R_2^{-1}\underset{\sim}{b}}{\partial t_k}$$

since

$$\frac{\partial-\frac{1}{2}\underset{\sim}{b}'R_2^{-1}\underset{\sim}{b}}{\partial t_k} = (\rho_{lk}\underset{\sim}{e_1}' + \rho_{qk}\underset{\sim}{e_q}')R_2^{-1}\underset{\sim}{b} = \rho_{lk}(b_l+\rho^{lq}b_q) + \rho_{qk}(b_q+\rho^{lq}b_l)$$

and

$$\phi_2(b_l,b_q;\rho_{lq})\frac{\partial\Phi_{d-2}(B^l_{qs};R_{lq})}{\partial t_k} = \sum_{r\neq l,q}(\rho_{rk}-\beta_{rl.q}\rho_{lk}-\beta_{rq.l}\rho_{qk})\phi_3(b_l,b_q,b_r;\rho_{lqr})$$
$$\cdot\Phi_{d-3}(B^{lq}_{rs};R_{lqr});$$

therefore,

$$\frac{\partial^3\Phi_d(R)}{\partial t_k \partial t_j \partial t_i} = \sum_{l=1}^{d}(1+b_l^2)\rho_{il}\rho_{jl}\rho_{kl}\phi(b_l)\Phi_{d-1}(B_{ls};R_l)$$
$$+\sum_{l=1}^{d}\rho_{il}\sum_{q\neq l}\big(\rho_{jl}(\rho_{kq}-\rho_{lq}\rho_{lk})b_l+(\rho_{jq}-\rho_{lq}\rho_{lj})(\rho_{lk}(b_l+\rho^{lq}b_q)+\rho_{qk}(b_q+\rho^{lq}b_l))\big)$$
$$\phi(b_l,b_q;\rho_{lq})\Phi_{d-2}(B_{qs}^{l};R_{lq})$$
$$+\sum_{l=1}^{d}\rho_{il}\sum_{q\neq l}(\rho_{jq}-\rho_{lq}\rho_{lj})$$
$$\cdot\sum_{r\neq l,q}(\rho_{rk}-\beta_{rl.q}\rho_{lk}-\beta_{rq.l}\rho_{qk})\phi(b_l,b_q,b_r;\rho_{lqr})\Phi_{d-3}(B_{rs}^{lq};R_{lqr}).$$

Substitute the first and third derivatives of $\Phi_d(R)$ with respect to t into (2.9) and take $\underset{\sim}{t}=\underset{\sim}{0}$, we obtain (2.8). $\square$

Corollary 2.1. *The third moment of truncated multinormal distribution truncated by the upper bounds $-\infty < x_s \leq \infty$ is given by*

$$\begin{aligned}
&\alpha E(X_iX_jX_k) = -1\\
&\times[\sum_{l=1}^{d}(\rho_{ij}\rho_{kl}+\rho_{ik}\rho_{jl}+\rho_{kj}\rho_{il}+(1+x_l^2)\rho_{il}\rho_{jl}\rho_{kl})\phi(x_l)N_{d-1}(A_{ls};R_l)\\
&-\sum_{l=1}^{d}\rho_{il}\sum_{q\neq l}\big(\rho_{jl}(\rho_{kq}-\rho_{lq}\rho_{lk})x_l+(\rho_{jq}-\rho_{lq}\rho_{lj})(\rho_{lk}(x_l+\rho^{lq}x_q)+\\
&\rho_{qk}(x_q+\rho^{lq}x_l))\big)\phi(x_l,x_q;\rho_{lq})N_{d-2}(A_{qs}^{l};R_{lq})+\sum_{l=1}^{d}\rho_{il}\sum_{q\neq l}(\rho_{jq}-\rho_{lq}\rho_{lj})\\
&\cdot\sum_{r\neq l,q}(\rho_{rk}-\beta_{rl.q}\rho_{lk}-\beta_{rq.l}\rho_{qk})\phi(x_l,x_q,x_r;\rho_{lqr})N_{d-3}(A_{rs}^{lq};R_{lqr})]
\end{aligned} \tag{2.11}$$

where $N_d(R)=(2\pi)^{-\frac{d}{2}}|R|^{-\frac{d}{2}}\ {}^{(d)}\int_{-\infty}^{x_s}e^{-\frac{1}{2}\underset{\sim}{x}'R^{-1}\underset{\sim}{x}}dx_s$, ${}^{(d)}\int$ denotes the d-dimensional integral, R is the correlation matrix, ρ_{ij} is the (i,j) elements of the correlation matrix, A_{ls}, A_{qs}^{l} and A_{rs}^{lq} are defined in Theorem 2.1.

Proof. Let $a_s = -x_s$ for Theorem 2.1, then (2.8) becomes

$$\begin{aligned}
&\alpha E X_i X_j X_k = \\
&\textstyle\sum_{l=1}^{d} (\rho_{ij}\rho_{kl} + \rho_{ik}\rho_{jl} + \rho_{kj}\rho_{il} + (1 + x_l^2)\rho_{il}\rho_{jl}\rho_{kl})\phi(x_l)\Phi_{d-1}(-A_{ls}; R_l) \\
&- \textstyle\sum_{l=1}^{d} \rho_{il} \sum_{q\neq l} \big(\rho_{jl}(\rho_{kq} - \rho_{lq}\rho_{lk})x_l + (\rho_{jq} - \rho_{lq}\rho_{lj})(\rho_{lk}(x_l + \rho^{lq}x_q) + \rho_{qk}(x_q + \rho^{lq}x_l))\big) \\
&\phi(x_l, x_q; \rho_{lq})\Phi_{d-2}(-A^l_{qs}; R_{lq}) \\
&+ \textstyle\sum_{l=1}^{d} \rho_{il} \sum_{q\neq l} (\rho_{jq} - \rho_{lq}\rho_{lj}) \\
&\cdot \textstyle\sum_{r\neq l,q} (\rho_{rk} - \beta_{rl.q}\rho_{lk} - \beta_{rq.l}\rho_{qk})\phi(x_l, x_q, x_r; \rho_{lqr})\Phi_{d-3}(-A^{lq}_{rs}; R_{lqr}).
\end{aligned} \tag{2.12}$$

Note that the integrals ${}^{(d)}\int_{-x_s}^{\infty} ={}^{(d)} \int_{-\infty}^{x_s}$ of the normal density with mean 0 by symmetry, hence $\Phi_{d-1}(-A_{ls}; R_l) = N_{d-1}(A_{ls}; R_l)$, $\Phi_{d-2}(-A^l_{qs}; R_{lq}) = N_{d-2}(A^l_{qs}; R_{lq})$, $\Phi_{d-3}(-A^{lq}_{rs}; R_{lqr}) = N_{d-3}(A^{lq}_{rs}; R_{lqr})$, where $N_d(.)$ denotes the d-dimensional normal integral ${}^{(d)}\int_{-\infty}^{x_s} \phi_d(u_s)du_s$.

And since $x_i x_j x_k$ is an odd function, the integral of $x_i x_j x_k$ over the range of $-\infty$ to x_s has the negative sign of the integral over $-x_s$ and ∞, which gives (2.11). □

2.3 *Differential Equations*

A system of differential equations will be derived in this section using the approach described in section 2. We will assume the parameters of θ_1, θ_2, and θ_3 are orthogonal to each other, and that the log likelihood function L is bounded in probability by n. The orthogonality of two parameters is defined as

$$E(\frac{\partial^2 L}{\partial\theta_i \partial\theta_j}) = 0, \text{for } i, j = 1, 2, 3 \text{ and } i \neq j. \tag{2.13}$$

And as pointed out by Cox (1987), we can write $\frac{\partial^2 l}{\partial T_i \partial T_j} = E[\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}] + \frac{Z_{ij}}{\sqrt{n}}$ for $i, j = 1, 2, 3$,where $l = \frac{1}{n}L$ (the log likelihood per observation), and Z_{ij} are random variables of mean 0 and $O_p(1)$ as $n \to \infty$. Then by (2.13), $\frac{\partial^2 l}{\partial T_i \partial T_j}$ is $O_p(\frac{1}{\sqrt{n}})$ for $i \neq j$ and $O_p(1)$ for $i = j$; hence w_{ij} is $O_p(n^{-\frac{1}{2}})$ for $i \neq j$ and w_{ii} is $O_p(1)$. Also note that w_i and w_{ijk}

are $O_p(n^{-\frac{1}{2}})$.

To $O_p(n^{-\frac{1}{2}})$, we apply Taylor's expansion for (2.7) for the linear, quadratic (for $i \neq j$), and cubic terms. Hence (2.7) becomes the integral with respect to a multi-normal distribution with independent random variables as follows:

$$\begin{aligned} r(S,\theta_1,\theta_2) \;& \propto (2\pi)^{-\frac{3}{2}}||w_{ij}||^{\frac{1}{2}} \int \int^{x_1} \int^{x_2} e^{-\frac{1}{2}\sum_{i=1}^{3} w_{ii}u_i^2} \\ & \big(1 + \tfrac{1}{6}\textstyle\sum_{i,j,k=1,2,3} w_{ijk}u_iu_ju_k - \sum_{1\le i<j\le 3} w_{ij}u_iu_j + \sum_{i=1}^{3} w_iu_i\big) \\ & du_1du_2du_3 \\ & = \alpha\big(1 + \tfrac{1}{6}\textstyle\sum_{i,j,k=1,2,3} w_{ijk}E(X_iX_jX_k) - \sum_{1\le i<j\le 3} w_{ij}E(X_iX_j) \\ & + \textstyle\sum_{i=1}^{3} w_iE(X_i)\big). \end{aligned} \tag{2.14}$$

Tallis (1961) gives the first and second moment of the truncated multivariate normal distribution truncated by lower bounds as

$$\begin{aligned} \alpha E(X_i) \;& = \textstyle\sum_{l=1}^{d} \rho_{il}\phi(a_l)\Phi_{d-1}(A_{ls};R_l) \\ \alpha E(X_iX_j) \;& = \rho_{ij}\alpha + \textstyle\sum_{q=1}^{d} \rho_{qi}\rho_{qj}a_q\phi(a_q)\Phi_{d-1}(A_{qs};R_q) \\ & + \textstyle\sum_{q=1}^{d}\big(\rho_{qi}\sum_{r\neq q}\phi(a_q,a_r;\rho_{qr})\Phi_{d-2}(A^q_{rs};R_{qr})(\rho_{rj}-\rho_{qr}\rho_{qj})\big). \end{aligned} \tag{2.15}$$

Since x_i is an odd function, the first moment of truncated multinormal distribution truncated by the upper bounds will take the negative sign; on the other hand, the second moment will remain the same sign since x_ix_j is a even function. Let $d = 3$, $x_3 = \infty$, $\rho_{ij} = 0$ when $i \neq j$, and use Tallis's expression and Corollary 2.1 to obtain the first three moments of the truncated multi-normal distribution truncated by the upper bounds x_1 and x_2 as

$$\alpha E(X_i) = -\rho_{ii}\phi(x_i)N_2(A_{is};R_i), \tag{2.16}$$

$$\begin{aligned} \alpha E(X_iX_3) \;& = 0, \text{ for } i = 1,2, \\ \alpha E(X_1X_2) \;& = \rho_{11}\rho_{22}\phi(x_1,x_2)N_1(A^1_{23};R_{12}), \end{aligned} \tag{2.17}$$

and

$$
\begin{aligned}
E(X_1X_2X_3) = & \quad 0, \\
E(X_3^3) = & \quad 0, \\
E(X_i^2X_3) = & \quad 0, \text{for } i = 1,2, \\
\alpha E(X_i^3) = & \quad -(3\rho_{ii}^2 + \rho_{ii}^3(1+x_i^2))\phi(x_i)N_2(A_{is}, R_i), i = 1,2, \\
\alpha E(X_i^2X_j) = & \quad -[\rho_{ii}\rho_{jj}\phi(x_j)N_2(A_{js}, R_j) - \phi(x_1,x_2)\rho_{ii}^2\rho_{jj}x_iN_1(A_{j3}^i; R_{ij})], \\
& \quad \text{for } i,j = 1,2 \text{ and } i \neq j.
\end{aligned}
\tag{2.18}
$$

Also by definition, $A_{is} = (x_j, \infty)$ for $i = 1, 2$ and $j \neq i$, and $A_{j3}^i = \infty$ for $i, j = 1, 2$ and $i \neq j$. Substitute (2.16), (2.17), (2.18), $\alpha = N_2(x_1, x_2)$, $N_2(x_i, \infty) = N(x_i)$ for $i = 1, 2$ and $N_1(\infty; R_{ij}) = 1$ into (2.14) to obtain

$$
\begin{aligned}
r(S,\theta_1,\theta_2) = & \quad N_2(x_1,x_2) - \tfrac{1}{6}\textstyle\sum_{i=1}^{2} w_{iii}(3\rho_{ii}^2 + \rho_{ii}^3(1+x_i^2))\phi(x_i)N(x_j) \\
& -\tfrac{1}{2}\textstyle\sum_{i=1,j\neq i}^{2} w_{iij}\big(\rho_{ii}\rho_{jj}\phi(x_j)N(x_i) - \phi(x_1,x_2)\rho_{ii}^2\rho_{jj}x_i\big) \\
& -\textstyle\sum_{i=1}^{2} w_i\rho_{ii}\phi(x_i)N(x_j) - w_{12}\rho_{11}\rho_{22}\phi(x_1,x_2).
\end{aligned}
\tag{2.19}
$$

Note that $\frac{\partial N_2(x_1,x_2)}{\partial x_i} = \phi(x_i)N(x_j)$ for $i, j = 1, 2$ and $j \neq i$, $\frac{\partial^2 N_2(x_1,x_2)}{\partial x_2 \partial x_1} = \phi(x_1, x_2)$. Then (2.19) becomes

$$
\begin{aligned}
r(S,\theta_1,\theta_2) = & \quad N_2(x_1,x_2) - \tfrac{1}{6}\textstyle\sum_{i=1,j\neq i}^{2} w_{iii}(3\rho_{ii}^2 + \rho_{ii}^3(1+x_i^2))\frac{\partial N_2(x_1,x_2)}{\partial x_i} \\
& -\tfrac{1}{2}\textstyle\sum_{i=1,j\neq i}^{2} w_{iij}\big(\rho_{ii}\rho_{jj}\frac{\partial N_2(x_1,x_2)}{\partial x_j} - \rho_{ii}^2\rho_{jj}x_i\frac{\partial^2 N_2(x_1,x_2)}{\partial x_2\partial x_1}\big) \\
& -\textstyle\sum_{i=1}^{2} w_i\rho_{ii}\frac{\partial N_2(x_1,x_2)}{\partial x_i} - w_{12}\rho_{11}\rho_{22}\frac{\partial^2 N_2(x_1,x_2)}{\partial x_2\partial x_1} \\
& = N_2(x_1 - \tfrac{1}{6}w_{111}(3\rho_{11}^2 + \rho_{11}^3(1+x_1^2)) - \tfrac{1}{2}w_{221}\rho_{22}\rho_{11} - w_1\rho_{11}, \\
& x_2 - \tfrac{1}{6}w_{222}(3\rho_{22}^2 + \rho_{22}^3(1+x_2^2)) - \tfrac{1}{2}w_{112}\rho_{11}\rho_{22} - w_2\rho_{22}) \\
& = N_2(z_1,z_2)
\end{aligned}
\tag{2.20}
$$

where

$$
\begin{aligned}
(z_1, z_2) = \; & \big(x_1 - \tfrac{1}{6} w_{111}(3\rho_{11}^2 + \rho_{11}^3(1 + x_1^2)) - \tfrac{1}{2} w_{221}\rho_{22}\rho_{11} - w_1\rho_{11}, \\
& x_2 - \tfrac{1}{6} w_{222}(3\rho_{22}^2 + \rho_{22}^3(1 + x_2^2)) - \tfrac{1}{2} w_{112}\rho_{11}\rho_{22} - w_2\rho_{22}\big).
\end{aligned} \tag{2.21}
$$

Since $\rho_{ii} = w_{ii}^{-1}$, (2.21) becomes

$$
\begin{aligned}
(z_1, z_2) = \; & \big(x_1 - \tfrac{1}{6} w_{111}(3w_{11}^{-2} + w_{11}^{-3}(1 + x_1^2)) - \tfrac{1}{2} w_{221}w_{22}^{-1}w_{11}^{-1} - w_1 w_{11}^{-1}, \\
& x_2 - \tfrac{1}{6} w_{222}(3w_{22}^{-2} + w_{22}^{-3}(1 + x_2^2)) - \tfrac{1}{2} w_{112}w_{11}^{-1}w_{22}^{-1} - w_2 w_{22}^{-1}\big).
\end{aligned} \tag{2.22}
$$

Define quantities analogous to the v's by (Peers, 1965)

$$
\begin{aligned}
y_i &= n^{-\frac{1}{2}} \tfrac{\partial L}{\partial \theta_i}, \text{for } i = 1, 2, \cdots, d, \\
y_{ij} &= -n^{-1} \tfrac{\partial^2 L}{\partial \theta_i \partial \theta_j}, \text{for } i, j = 1, 2, \cdots, d, \\
y_{ijk} &= n^{-\frac{3}{2}} \tfrac{\partial^3 L}{\partial \theta_i \partial \theta_j \partial \theta_k}, \text{for } i, j, k = 1, 2, \cdots, d.
\end{aligned} \tag{2.23}
$$

As Peers (1965) pointed out the joint cumulant generating function of the y's can be written as

$$
\log E(e^{t_a y_a + t_{ab} y_{ab} + t_{abc} y_{abc}}) = \frac{\kappa_{a|b}}{2!} t_a t_b + \kappa_{ab} t_{ab} + n^{-\frac{1}{2}} (\frac{\kappa_{a|b|c}}{3!} t_a t_b t_c + \kappa_{a|bc} t_a t_{bc} + \kappa_{abc} t_{abc}), \tag{2.24}
$$

where

$$
\begin{aligned}
\kappa_{ij} &= E(-\tfrac{\partial^2 l}{\partial \theta_i \partial \theta_j}), \\
\kappa_{ijk} &= E(\tfrac{\partial^3 l}{\partial \theta_i \partial \theta_j \partial \theta_k}), \\
\kappa_{i|j} &= E(\tfrac{\partial l}{\partial \theta_i} \tfrac{\partial l}{\partial \theta_j}), \\
\kappa_{i|j|k} &= E(\tfrac{\partial l}{\partial \theta_i} \tfrac{\partial l}{\partial \theta_j} \tfrac{\partial l}{\partial \theta_k}), \\
\kappa_{i|jk} &= E(-\tfrac{\partial l}{\partial \theta_i} \tfrac{\partial^2 l}{\partial \theta_j \partial \theta_k}),
\end{aligned} \tag{2.25}
$$

and l denotes the log-likelihood function of a single observation. The κ's are in fact the cumulants of y's.

Hence the joint MGF of the y's about their means (Peers, 1953) is

$$M(t_1, t_2, \cdots, t_d, t_{11}, t_{12}, \cdots, t_{dd}) = e^{\frac{\kappa_{a|b}}{2!} t_a t_b + \kappa_{ab} t_{ab} + n^{-\frac{1}{2}} (\frac{\kappa_{a|b|c}}{3!} t_a t_b t_c + \kappa_{a|bc} t_a t_{bc}) + O_p(n^{-1})}. \tag{2.26}$$

Also employ the following expansions:

$$\begin{aligned}
0 = n^{-\frac{1}{2}} \frac{\partial L}{\partial T_i} = \quad & n^{-\frac{1}{2}} \Bigg(\frac{\partial L}{\partial \theta_i} + \textstyle\sum_j (T_j - \theta_j) \frac{\partial^2 L}{\partial \theta_i \partial \theta_j} + \\
& \tfrac{1}{2} \textstyle\sum_{j,k} (T_j - \theta_j)(T_k - \theta_k) \frac{\partial^3 L}{\partial \theta_i \partial \theta_j \partial \theta_k} \Bigg), \text{for } i = 1, 2, \cdots, d, \\
-n^{-1} \frac{\partial^2 L}{\partial T_i \partial T_j} = \quad & -n^{-1} \left(\frac{\partial^2 L}{\partial \theta_i \partial \theta_j} + \textstyle\sum_k (T_k - \theta_k) \frac{\partial^3 L}{\partial \theta_i \partial \theta_j \partial \theta_k} \right), \text{for } i, j = 1, 2, \cdots, d, \\
n^{-\frac{3}{2}} \frac{\partial^3 L}{\partial T_i \partial T_j \partial T_k} = \quad & n^{-\frac{3}{2}} \frac{\partial^3 L}{\partial \theta_i \partial \theta_j \partial \theta_k}, \text{for } i, j, k = 1, 2, \cdots, d.
\end{aligned} \tag{2.27}$$

Then from (2.5) we find

$$\begin{aligned}
x_i = \quad & -y_i (y^{ii})^{\frac{1}{2}} - \tfrac{1}{2} (y^{ii})^{\frac{1}{2}} \textstyle\sum_{j,k=1,2,3} y_j y_k y_{ijk} y^{jj} y^{kk} + \tfrac{1}{2} y_i (y^{ii})^{\frac{3}{2}} \sum_{j=1}^{3} y_j y_{iij} y^{jj} \\
& + A_i(y^{ab}; a \neq b), \text{ for i=1,2}, \\
v^{ii} = \quad & y^{ii} + y^{ii} \textstyle\sum_{j,k=1,2,3} y^{jj} y^{kk} y^k y_{ijk}, \text{ for } i = 1, 2, 3, \\
v_{ijk} = \quad & y_{ijk}, \text{for } i, j, k = 1, 2, 3,
\end{aligned} \tag{2.28}$$

where

$$\begin{aligned}
A_i(y^{ab}; a \neq b) = \quad & -\textstyle\sum_{w \neq i} y^{iw} y_w (y^{ii})^{-\frac{1}{2}} - \sum_{w \neq i, j \neq l, k \neq m} \frac{1}{2} y^{iw} y_{wjk} y^{jl} y_l y^{km} y_m (y^{ii})^{-\frac{1}{2}} \\
& + \textstyle\sum_{w,k,j \neq i, l \neq m} \frac{1}{2} y^{iw} y_w y^{ij} y^{ik} y_{jkl} y^{lm} y_m (y^{ii})^{-\frac{1}{2}}
\end{aligned}$$

contains all the terms of y^{ab} for $a \neq b$. Note that y_{ij} is $O_p(n^0)$ but y_i, y_{ijk} are $O_p(n^{-\frac{1}{2}})$, and we will not consider terms beyond $O_p(n^{-\frac{1}{2}})$. Substitute (2.28), (2.5) into (2.22) to obtain

$$\begin{aligned}(z_1, z_2) = \quad & \big(- y_1(y^{11})^{\frac{1}{2}} - \tfrac{1}{2}(y^{11})^{\frac{1}{2}} \textstyle\sum_{j,k=1,2,3} y_j y_k y_{1jk} y^{jj} y^{kk} \\ & + \tfrac{1}{2} y_1 (y^{11})^{\frac{3}{2}} \textstyle\sum_{j=1}^{3} y_j y_{11j} y^{jj} + \tfrac{2}{3} y_{111} (y^{11})^{\frac{3}{2}} + \tfrac{1}{2} y_{221} y^{22} (y^{11})^{\frac{1}{2}} \\ & + n^{-\frac{1}{2}} (y^{11})^{\frac{1}{2}} \tfrac{\partial \psi}{\partial \theta_1} + A_1(y^{ab}; a \neq b), \\ & - y_2(y^{22})^{\frac{1}{2}} - \tfrac{1}{2}(y^{22})^{\frac{1}{2}} \textstyle\sum_{j,k=1,2,3} y_j y_k y_{2jk} y^{jj} y^{kk} \\ & + \tfrac{1}{2} y_2 (y^{22})^{\frac{3}{2}} \textstyle\sum_{j=1}^{3} y_j y_{22j} y^{jj} + \tfrac{2}{3} y_{222} (y^{22})^{\frac{3}{2}} + \tfrac{1}{2} y_{112} y^{11} (y^{22})^{\frac{1}{2}} \\ & + n^{-\frac{1}{2}} (y^{22})^{\frac{1}{2}} \tfrac{\partial \psi}{\partial \theta_2} + A_2(y^{ab}; a \neq b)\big). \end{aligned} \tag{2.29}$$

The moment generating function of $\underset{\sim}{z}' = (z_1, z_2)$ is then given symbolically by

$$\begin{aligned} & E(e^{\underset{\sim}{t}' \underset{\sim}{z}}) = M(D_1, D_2, \ldots, D_{11}, D_{12} \ldots, D_{33}) \\ & \cdot exp\Bigg[t_1 \Bigg(- \eta_1 (\eta^{11})^{\frac{1}{2}} - \tfrac{1}{2} (\eta^{11})^{\frac{1}{2}} \textstyle\sum_{j,k=1,2,3} \eta_j \eta_k \eta_{1jk} \eta^{jj} \eta^{kk} \\ & + \tfrac{1}{2} \eta_1 (\eta^{11})^{\frac{3}{2}} \textstyle\sum_{j=1}^{3} \eta_j \eta_{11j} \eta^{jj} + \tfrac{2}{3} \eta_{111} (\eta^{11})^{\frac{3}{2}} + \tfrac{1}{2} \eta_{221} \eta^{22} (\eta^{11})^{\frac{1}{2}} + n^{-\frac{1}{2}} (\eta^{11})^{\frac{1}{2}} \tfrac{\partial \psi}{\partial \theta_1} \\ & + A_1(\eta^{ab}; a \neq b) \Bigg) + \\ & t_2 \Bigg(- \eta_2 (\eta^{22})^{\frac{1}{2}} - \tfrac{1}{2} (\eta^{22})^{\frac{1}{2}} \textstyle\sum_{j,k=1,2,3} \eta_j \eta_k \eta_{2jk} \eta^{jj} \eta^{kk} + \tfrac{1}{2} \eta_2 (\eta^{22})^{\frac{3}{2}} \textstyle\sum_{j=1}^{3} \eta_j \eta_{22j} \eta^{jj} \\ & + \tfrac{2}{3} \eta_{222} (\eta^{22})^{\frac{3}{2}} + \tfrac{1}{2} \eta_{112} \eta^{11} (\eta^{22})^{\frac{1}{2}} + n^{-\frac{1}{2}} (\eta^{22})^{\frac{1}{2}} \tfrac{\partial \psi}{\partial \theta_2} + A_2(\eta^{ab}; a \neq b) \Bigg) \Bigg], \end{aligned} \tag{2.30}$$

where D's denote differentiations with respect to corresponding η's. After performing all the differentiations we replace η_i by 0, η^{ij} by κ^{ij} and η_{ijk} by $n^{-\frac{1}{2}}\kappa_{ijk}$. Since by the assumption of orthogonality $\kappa^{ab} = 0$ for $a \neq b$, then $A_i(\kappa^{ab}; a \neq b) = 0$ for $i = 1, 2$. Use the identities of Bartlett (1953) as follows:

$$\begin{aligned} & \kappa_{i|j} - \kappa_{ij} = 0, \\ & \kappa_{i|j|k} - \kappa_{i|jk} - \kappa_{j|ki} - \kappa_{k|ij} + \kappa_{ijk} = 0, \\ & \kappa_{i|jk} - \kappa_{ijk} - \tfrac{\partial \kappa_{jk}}{\partial \theta_i} = 0, \text{for } i, j, k = 1, 2, \ldots, d. \end{aligned} \tag{2.31}$$

And to $O_p(n^{-\frac{1}{2}})$, we have

$$E(e^{\underset{\sim}{t}'\underset{\sim}{z}}) = \exp\{\sum_{i=1,j\neq i}^{2}\big(-t_i n^{-\frac{1}{2}}(\tfrac{2}{3}(\kappa^{ii})^{\frac{3}{2}}\tfrac{\partial\kappa_{ii}}{\partial\theta_i} - \tfrac{1}{2}\kappa^{jj}(\kappa^{ii})^{\frac{1}{2}}\tfrac{\partial\kappa_{ji}}{\partial\theta_j} - (\kappa^{ii})^{\frac{1}{2}}\tfrac{\partial\psi}{\partial\theta_i}) + \tfrac{1}{2}t_i^2\big)\}. \tag{2.32}$$

Therefore, in order to have $\underset{\sim}{z}' = (z_1, z_2)$ following a standard bivariate normal distribution, $\psi(\underset{\sim}{\theta})$ is chosen to satisfy the following two partial differential equations:

$$\begin{cases} \frac{2}{3}(\kappa^{11})^{\frac{3}{2}}\frac{\partial\kappa_{11}}{\partial\theta_1} - \frac{1}{2}\kappa^{22}(\kappa^{11})^{\frac{1}{2}}\frac{\partial\kappa_{21}}{\partial\theta_2} - (\kappa^{11})^{\frac{1}{2}}\frac{\partial\psi}{\partial\theta_1} = 0 \\ \frac{2}{3}(\kappa^{22})^{\frac{3}{2}}\frac{\partial\kappa_{22}}{\partial\theta_2} - \frac{1}{2}\kappa^{11}(\kappa^{22})^{\frac{1}{2}}\frac{\partial\kappa_{12}}{\partial\theta_1} - (\kappa^{22})^{\frac{1}{2}}\frac{\partial\psi}{\partial\theta_2} = 0 \end{cases} \tag{2.33}$$

Note that $\kappa_{ij} = 0$ if $i \neq j$ by orthogonality and $\kappa^{ii} = \kappa_{ii}^{-1}$ for $i = 1, 2$. After some simplification, (2.33) becomes

$$\begin{cases} \frac{2}{3}\kappa_{11}^{-1}\frac{\partial\kappa_{11}}{\partial\theta_1} - \frac{\partial\psi}{\partial\theta_1} = 0 \\ \frac{2}{3}\kappa_{22}^{-1}\frac{\partial\kappa_{22}}{\partial\theta_2} - \frac{\partial\psi}{\partial\theta_2} = 0 \end{cases} \tag{2.34}$$

Hence we have

$$\psi(\theta_1, \theta_2) = \frac{2}{3}(\log(\kappa_{11}) + \log(\kappa_{22})). \tag{2.35}$$

Therefore the bivariate matching prior is a power product of fisher information of θ_1 and θ_2:

$$\begin{aligned} w(\theta_1, \theta_2) &= e^{\psi(\theta_1,\theta_2)} \\ &= (\kappa_{11}\kappa_{22})^{\frac{2}{3}} \\ &= \big(E(-\tfrac{\partial^2 l}{\partial\theta_1^2})E(-\tfrac{\partial^2 l}{\partial\theta_2^2})\big)^{\frac{2}{3}} \end{aligned} \tag{2.36}$$

2.4 Bivariate Confidence Region

The quantity $r(S, \theta_1, \theta_2)$ is now written as a bivariate normal integral $N_2(\underset{\sim}{z})$ where $\underset{\sim}{z}' = (z_1, z_2)$ follows a standard bivariate normal distribution; hence in order to have $Pr(r(S, \theta_1, \theta_2) < \alpha | (\theta_1, \theta_2, \theta_3)) = \alpha$ we need to make $Pr((z_1, z_2) <$

$(\xi_1, \xi_2)|(\theta_1, \theta_2, \theta_3)) = \alpha$ where (ξ_1, ξ_2) satisfies $N(\xi_1)N(\xi_2) = \alpha$. Substitute (2.35) to (2.22) and use the definitions in (2.5) we have

$$\begin{aligned}(z_1, z_2) = \ & \big(x_1 - \tfrac{1}{6}(v^{11})^{\frac{3}{2}}v_{111}(4 + x_1^2) - \tfrac{1}{2}v^{22}(v^{11})^{\frac{1}{2}}v_{221} \\ & -n^{-\frac{1}{2}}(v^{11})^{\frac{1}{2}}\tfrac{\partial}{\partial T_1}\tfrac{2}{3}(\log(\kappa_{11}) + \log(\kappa_{22})), \\ & x_2 - \tfrac{1}{6}(v^{22})^{\frac{3}{2}}v_{222}(4 + x_2^2) - \tfrac{1}{2}v^{11}(v^{22})^{\frac{1}{2}}v_{112} \\ & -n^{-\frac{1}{2}}(v^{22})^{\frac{1}{2}}\tfrac{\partial}{\partial T_2}\tfrac{2}{3}(\log(\kappa_{11}) + \log(\kappa_{22}))\big)\end{aligned} \tag{2.37}$$

We first solve (2.37) for x_i and note that $z_i^2 = x_i^2$ by correcting to $O_p(n^{-\frac{1}{2}})$.

$$\begin{aligned}x_i = \ & z_i + \tfrac{1}{6}(v^{ii})^{\frac{3}{2}}v_{iii}(4 + z_i^2) + \tfrac{1}{2}v^{jj}(v^{ii})^{\frac{1}{2}}v_{jji} + n^{-\frac{1}{2}}(v^{ii})^{\frac{1}{2}} \times \\ & \tfrac{\partial}{\partial T_i}\tfrac{2}{3}(\log(\kappa_{ii}) + \log(\kappa_{jj})), \text{for } i, j = 1, 2 \text{ and } i \neq j.\end{aligned} \tag{2.38}$$

Substitute $x_i = n^{\frac{1}{2}}(h_i(S, \alpha) - T_i)(v^{ii})^{-\frac{1}{2}}$ and $z_i = \xi_i$ into (2.38), solve for $h_i(S, \alpha)$, and obtain the bivariate confidence region for the problem of two parameters of interest and one nuisance parameter as the following:

$$\begin{aligned}(h_1(S, \alpha), h_2(S, \alpha)) = \ & (T_1 + (-\tfrac{\partial^2 L}{\partial T_1^2})^{-\frac{1}{2}}\xi_1 + \tfrac{1}{6}(\xi_1^2 + 4)(-\tfrac{\partial^2 L}{\partial T_1^2})^{-2}\tfrac{\partial^3 L}{\partial T_1^3} \\ & + \tfrac{1}{2}(-\tfrac{\partial^2 L}{\partial T_1^2})^{-1}(-\tfrac{\partial^2 L}{\partial T_2^2})^{-1}\tfrac{\partial^3 L}{\partial T_1 \partial T_2^2} \\ & + (-\tfrac{\partial^2 L}{\partial T_1^2})^{-1}\tfrac{\partial}{\partial T_1}\tfrac{2}{3}(\log(\kappa_{11}) + \log(\kappa_{22})), \\ & T_2 + (-\tfrac{\partial^2 L}{\partial T_2^2})^{-\frac{1}{2}}\xi_2 + \tfrac{1}{6}(\xi_2^2 + 4)(-\tfrac{\partial^2 L}{\partial T_2^2})^{-2}\tfrac{\partial^3 L}{\partial T_2^3} \\ & + \tfrac{1}{2}(-\tfrac{\partial^2 L}{\partial T_2^2})^{-1}(-\tfrac{\partial^2 L}{\partial T_1^2})^{-1}\tfrac{\partial^3 L}{\partial T_2 \partial T_1^2} \\ & + (-\tfrac{\partial^2 L}{\partial T_2^2})^{-1}\tfrac{\partial}{\partial T_2}\tfrac{2}{3}(\log(\kappa_{11}) + \log(\kappa_{22})))\end{aligned} \tag{2.39}$$

such that

$$Pr(\theta_1 < h_1(S, \alpha), \theta_2 < h_2(S, \alpha)|(\theta_1, \theta_2, \theta_3)) = \alpha. \tag{2.40}$$

3. RESULTS AND DISCUSSION

3.1 Univariate Case

Welch and Peers (1963) provided expression of the confidence point $h(S,\alpha)$ for the univariate case, which gives

$$\begin{aligned} h(S,\alpha) = \;& T + \xi(-\tfrac{\partial^2 L}{\partial T^2})^{-\frac{1}{2}} + \tfrac{1}{6}(\xi^2+2)(\tfrac{\partial^3 L}{\partial T^3})(-\tfrac{\partial^2 L}{\partial T^2})^{-2} \\ & + \tfrac{1}{2}\tfrac{\partial \log(\kappa_2(T))}{\partial T}(-\tfrac{\partial^2 L}{\partial T^2})^{-1} + O_p(n^{-\frac{3}{2}}). \end{aligned} \tag{3.1}$$

Welch and Peers did not conduct any simulation to verify their results. We conduct the simulation for the univariate counterpart of Example 2 in Section 3.1.

Suppose we observe n independent univariate random variables y_i following a normal distribution of mean μ and variance σ^2 for $i = 1, \ldots, n$. The parameters of interest is σ^2, the nuisance parameter is μ. Therefore the bivariate confidence regions $h(S,\alpha)$ at confidence level α for (σ_1^2, σ_2^2) is given using (3.1):

$$h(S,\alpha) = (1 + \tfrac{5}{3n} + \sqrt{\tfrac{2}{n}}\xi + \tfrac{4}{3n}\xi^2)\bar{S}, \tag{3.2}$$

where $\bar{S} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{\mu})^2$ and $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i$; $\bar{S}$ and μ are the MLEs for σ^2 and μ, respectively. Here ξ satisfies $N(\xi) = \alpha$.

The MC simulation is performed to confirm (3.2); 10,000 samples are generated from a univariate random normal distribution $N(8, 20^2)$, and each sample contains 100 data points. The simulation results presented in Table 3.1 shows that the approximation obtained by Welch and Peers performs very well as the observed confidence levels are very close to the true values for $n = 100$.

Tab. 3.1: Simulation Result for Univariate Case for $n = 100$

True α	Observed α
0.01	0.014
0.1	0.1094
0.5	0.5087
0.975	0.9706

MC simulations are also conducted for $n = 30$ and $n = 10$, the results are presented in Table 3.2 and Table 3.3, respectively.

Tab. 3.2: Simulation Result for Univariate Case for $n = 30$

True α	Observed α
0.01	0.0348
0.1	0.1159
0.5	0.4918
0.975	0.9636

Table 3.2 show that for a small sample size of 30 , Welch and Peers's approximation works very well as the observed confidence levels are very close to the true confidence levels.

Tab. 3.3: Simulation Result for Univariate Case for $n = 10$

True α	Observed α
0.01	0.1987
0.1	0.1762
0.5	0.4942
0.975	0.9472

Table 3.3 shows that the approximation obtained by Welch and Peers performs well except for $\alpha = 0.01$, the discrepancies are $O_p(\frac{1}{n})$ with $n = 10$, i.e., bounded by 0.1 in probability.

3.2 Bivariate Case

Two specific examples will be provided in this chapter to demonstrate the application of (2.39) to obtain the bivariate confidence region at confidence level α.

Suppose that we observe n independent random samples $y_{1i}, i = 1, \ldots, n$ following a normal distribution with mean μ and variance σ_1^2. We would like to find a confidence region for (μ, σ_1^2) jointly. The traditional approach from a frequentist point of view may involve working on the joint density function of the sample mean and sample variance. Although sample mean and variance are independently distributed it is not clear if the confidence region may be a function of sample or variance or not, not to mention the computational difficulties. However, if associated with each y_{1i}, we also observe another random sample y_{2i} which is independent of y_{1i} and follows a normal distribution of mean 0 and variance σ_2^2, (2.40) can be directly implemented. In fact, it will be shown that we do not even need the additional observed y_{2i} since it is completely ancillary to (μ, σ_1^2).

Let $\theta_1 = \mu, \theta_2 = \sigma_1^2, \theta_3 = \sigma_2^2$, and let (θ_1, θ_2) are the parameters of interest and θ_3 is the nuisance parameter. It's easy to verify that $(\theta_1, \theta_2, \theta_3)$ are orthogonal parameters. The log of joint likelihood function of $(\theta_1, \theta_2, \theta_3)$ based on the observed (y_{1i}, y_{2i}) is given as

$$L(\theta_1, \theta_2, \theta_3) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}(\log\theta_2 + \log\theta_3) - \sum_{i=1}^{n} \frac{(y_{1i} - \theta_1)^2}{2\theta_2} - \sum_{i=1}^{n} \frac{y_{2i}^2}{2\theta_3} \tag{3.3}$$

and the maximum likelihood estimator T_1, T_2, and T_3 for θ_1, θ_2 and θ_3, respectively are

$$\begin{aligned} T_1 = &\ \tfrac{1}{n}\textstyle\sum_{i=1}^{n} y_{1i} = \bar{y}_1, \\ T_2 = &\ \tfrac{1}{n}\textstyle\sum_{i=1}^{n} (y_{1i} - \bar{y}_1)^2 = \bar{S}_1, \\ T_3 = &\ \tfrac{1}{n}\textstyle\sum_{i=1}^{n} y_{2i}^2 = \bar{S}_2. \end{aligned} \tag{3.4}$$

Substitute (3.3) and (3.4) to (2.39), we have

$$(h_1(S,\alpha), h_2(S,\alpha)) = (\bar{y_1} + \xi_1\sqrt{\tfrac{\bar{S_1}}{n}}, (1 + \tfrac{8}{3n} + \sqrt{\tfrac{2}{n}}\xi_2 + \tfrac{4}{3n}\xi_2^2)\bar{S_1}), \tag{3.5}$$

where ξ_1, ξ_2 satisfy $N(\xi_1)N(\xi_2) = \alpha$ and $Pr(\mu < h_1(S,\alpha), \sigma_1^2 < h_2(S,\alpha)|(\mu, \sigma_1^2, \sigma_2^2)) = \alpha$.

As shown in (3.5), the bivariate region does not involve y_{2i} at all,i.e.,(3.5) holds for all samples regardless of the nuisance parameter; hence

$$Pr(\mu < h_1(S,\alpha), \sigma_1^2 < h_2(S,\alpha)|(\mu, \sigma_1^2)) = \alpha. \tag{3.6}$$

Monto Carlo (MC) simulation is conducted to confirm (3.5); 10,000 samples are generated from a random normal distribution of mean $\mu = 4$ and $\sigma = 20$, and each sample contains 100 data points. Then calculate the bivariate confidence region $(h_1(S,\alpha), h_2(S,\alpha))$ using (3.5) for each sample at confidence level $\alpha = 0.01, 0.1, 0.5, 0.975$ (ξ_1 and ξ_2 are chosen to be equal); hence for each true confidence level α, 10,000 observed bivariate confidence region are generated. And the observed confidence level is calculated by counting the percentage of $(h_1(S,\alpha), h_2(S,\alpha)) > (\mu, \sigma^2)$. The simulation results show that the approximation obtained in Chapter 2 performs very well as the observed confidence levels are very close to the true confidence levels for $n = 100$ as shown in Table 3.4.

Tab. 3.4: Simulation Result for Example One for $n = 100$

True α	Observed α
0.01	0.0093
0.1	0.1063
0.5	0.5145
0.975	0.9718

MC simulations for $n = 30$ and $n = 10$ are also conducted for Example One and the results are presented in Table 3.5 and Table 3.6, respectively.

Tab. 3.5: Simulation Result for Example One for $n = 30$

True α	Observed α
0.01	0.0103
0.1	0.1115
0.5	0.5118
0.975	0.9609

Tab. 3.6: Simulation Result for Example One for $n = 10$

True α	Observed α
0.01	0.0176
0.1	0.1258
0.5	0.4970
0.975	0.9228

Table 3.5 and Table 3.6 show that even for a very small sample size of 30 or 10, the approximation works very well as the observed confidence levels are very close to the true confidence levels.

Suppose we observe n independent bivariate random vectors (y_{1i}, y_{2i}) following a bivariate normal distribution of mean vector (μ, μ) and variance matrix $\Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$ for $i = 1, \ldots, n$. The parameters of interest are (σ_1^2, σ_2^2), the nuisance parameter is μ. It is easy to verify that $(\mu, \sigma_1^2, \sigma_2^2)$ are orthogonal parameters. Therefore the bivariate confidence regions $(h_1(S, \alpha), h_2(S, \alpha))$ at confidence level α for (σ_1^2, σ_2^2) are given using (2.39):

$$\begin{aligned}(h_1(S,\alpha), h_2(S,\alpha)) = \ & ((1 + \tfrac{8}{3n} + \sqrt{\tfrac{2}{n}}\xi_1 + \tfrac{4}{3n}\xi_1^2)\bar{S}_1, \\ & (1 + \tfrac{8}{3n} + \sqrt{\tfrac{2}{n}}\xi_2 + \tfrac{4}{3n}\xi_2^2)\bar{S}_2),\end{aligned} \tag{3.7}$$

where $\bar{S}_k = \frac{1}{n}\sum_{i=1}^{n}(y_{ki} - \hat{\mu})^2$ for $k = 1, 2$ and $\hat{\mu} = \frac{1}{2n}\sum_{i=1}^{n}\sum_{k=1}^{2} y_{ki}$; $\bar{S}_1, \bar{S}_2$, and μ are the MLEs for σ_1^2, σ_2^2,and μ, respectively. ξ_1 and ξ_2 satisfy $N(\xi_1)N(\xi_2) = \alpha$.

Similar MC simulation is performed to confirm (3.7);10,000 samples are generated from a bivariate random normal distribution

$$N\left(\begin{pmatrix} 8 \\ 8 \end{pmatrix}, \begin{pmatrix} 20^2 & 0 \\ 0 & 35^2 \end{pmatrix}\right),$$

and each sample contains 100 data points. ξ_1 and ξ_2 are chosen to be the same. The simulation results presented in Table 3.7 show that the approximation obtained in Chapter 2 performs very well since the observed confidence levels are very close to the true confidence levels for $n = 100$.

Tab. 3.7: Simulation Result for Example Two for $n = 100$

True α	Observed α
0.01	0.0159
0.1	0.1317
0.5	0.5437
0.975	0.9765

MC simulations are also conducted for Example Two for $n = 30$ and $n = 10$, the results are presented in Table 3.8 and Table 3.9.

Tab. 3.8: Simulation Result for Example Two for $n = 30$

True α	Observed α
0.01	0.0370
0.1	0.1707
0.5	0.5713
0.975	0.9617

Table 3.8 show that for a small sample size of 30 , the approximation works very well as the observed confidence levels are very close to the true confidence levels.

Tab. 3.9: Simulation Result for Example Two for $n = 10$

True α	Observed α
0.01	0.1178
0.1	0.2286
0.5	0.5695
0.975	0.9327

The observed confidence levels for $n = 10$ as shown in Table 3.9 are close to the true values, but not as good as that for $n = 100$ or $n = 30$, the discrepancies are $O_p(\frac{1}{n})$ with $n = 10$, i.e., bounded by 0.1 in probability.

In addition, if the parameters of interest are (μ, σ_i^2) and the nuisance parameter is σ_j^2 for $i \neq j$. At confidence level α, the bivariate region is given using (2.39):

$$\begin{aligned}(h_1(S,\alpha), h_2(S,\alpha)) = \;& (\hat{\mu} + \xi_1\sqrt{\tfrac{\bar{S}_1\bar{S}_2}{n(\bar{S}_1+\bar{S}_2)}}, (1 - \tfrac{8}{3n} + \sqrt{\tfrac{2}{n}}\xi_2 - \tfrac{4}{3n}\xi_2^2)\bar{S}_i) \\ & \text{for i=1,2.}\end{aligned} \tag{3.8}$$

The approximation method developed in Chapter 2 applies under the assumption that the three parameters are pairwise orthogonal. As the number of parameters increase, the assumption of orthogonality become a lot more difficult to hold. Therefore, to extend the approach of matching prior to obtain a higher dimensional confidence region may not be helpful.

Part III

QUADRATIC SADDLEPOINT APPROXIMATION OF TAIL PROBABILITY

4. BACKGROUND ON SADDLEPOINT APPROXIMATION METHOD

Consider the following data on 63 case-control pairs of women with endometrial cancer presented by Stokes, Davis and Koch (1995). They evaluated the impact of three risk factors, gall bladder disease, hypertension and nonestrogen drug use on the occurrence of endometrial cancer. A logistic regression was used to model the probability of endometrial cancer π_j: $\pi_j = \exp(\theta_0 + \sum_{i=1}^{3} \theta_i W_{ij})/(1+\exp(\theta_0 + \sum_{i=1}^{3} \theta_i W_{ij}))$, where W_{1j}, W_{2j}, and W_{3j} are indicators for gall bladder disease, hypertension and nonestrogen drug use in individual j, respectively. The endometrial cancer data were obtained from a case-control study, the predictor is the case-control status and the responses are presence of the risk factors. The likelihood for these data is equivalent to that of the logistic regression specified earlier, the response variable in the logistic regression can be taken as unity and the predictor variables are the difference of the risk factor between the case member and the control member of the matched pairs. Table 4.1 presents the number of pairs with each configuration of differences of the three risk factors. Let $\underset{\sim}{z_j}^T$ be the row vector containing the top three entries in column j of Table 4.1, and let Z be the matrix whose rows are $\underset{\sim}{z_j}^T$. Let $\underset{\sim}{T} = Z^T \underset{\sim}{1}$, where $\underset{\sim}{1}$ is a column vector with entries that are all 1.

Tab. 4.1: Differences between cases and controls for endometrial cancer data

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

We test the null hypothesis that none of these potential risk factors increases the risk of cancer ($\underset{\sim}{\theta} = \underset{\sim}{0}$) vs. the alternative hypothesis that at least one of these potential risk factors actually increases the risk of cancer ($\theta_j \geq 0 \ \forall \ j$ and $\theta_j > 0$ for some j). The test statistics is obtained by comparing the minimum p-value from the three univariate one-sided conditional tests to its null distribution which is in the form of the tail probability $P(T_j \geq t_j)$. The endometrial cancer data is one of the examples in practice that request the evaluation of conditional or unconditional tail probability.

To derive an exact form of conditional or unconditional tail probability even for a simple distribution often can be very difficult; hence many approximation techniques are developed to evaluate the tail probability. Our research is based on one of the approximation techniques, known as saddlepoint approximation, for evaluating the multivariate tail probability for both conditional and unconditional case. The advantage of the saddlepoint approximation technique is that, while the derivation is often quite mathematically sophisticated, the calculation on the other hand requires merely trivial manipulation of the output of a widely available model fitting algorithm. We will provide some background of saddlepoint methods for tail probability approximation in Section 1.1, discuss some existing approximations in Section 1.2, and describe our proposed approximation and its difficulties in Section 1.3.

4.1 Concept of Saddlepoint Approximation

Mathematically, saddlepoint approximations are derived by inverting the cumulant generating function $K_T(\underset{\sim}{\xi}) = \log(E[\exp(\underset{\sim}{\xi}^{'}\underset{\sim}{T})])$ giving the density of probability in terms of the cumulant generating function. $\underset{\sim}{\xi}$ denotes a column vector with entries $\xi_1, \cdots, \xi_d$. When a d-dimensional random vector $\underset{\sim}{T}$ has a density $f_T(\underset{\sim}{t})$, then

$$f_T(\underset{\sim}{t}) = (2\pi i)^{-d} \int_{-i\infty}^{i\infty} \cdots \int_{-i\infty}^{i\infty} \exp(K_T(\underset{\sim}{\xi}) - \underset{\sim}{\xi}^{'}\underset{\sim}{t}) d\underset{\sim}{\xi}. \tag{4.1}$$

The saddlepoint approximation for the density is obtained by choosing the path of the integration to run through $\hat{\underset{\sim}{\xi}}$, for $\hat{\underset{\sim}{\xi}}$ satisfying the saddlepoint equation

$$K'_T(\hat{\underset{\sim}{\xi}}) = \underset{\sim}{t}. \tag{4.2}$$

Let $E_{\underset{\sim}{\xi}}[\underset{\sim}{T}]$ represent the expectation taken with respect to $f_{\underset{\sim}{T}}(\underset{\sim}{t}, \underset{\sim}{\xi})$. An approximation to any member of the family $f_{\underset{\sim}{T}}(\underset{\sim}{t}, \underset{\sim}{\xi})$ yields an approximation to $f_{\underset{\sim}{T}}(\underset{\sim}{t})$. The member of the family that might be approximated most accurately is the $\underset{\sim}{\xi}$ such that (4.2) holds.

In practice, we are usually interested in the statistical inference about the mean of independent and identically distributed random vectors $\underset{\sim}{Y}$, when $\underset{\sim}{T}$ arises as the sum of n independent and identically distributed summands $\underset{\sim}{Y}$. Then let $\bar{\underset{\sim}{t}} = \underset{\sim}{t}/n$.

Using the independence of the summands $\underset{\sim}{Y}$, $K_T(\underset{\sim}{\xi}) = nK_Y(\underset{\sim}{\xi})$. The saddlepoints are obtained by solving $K'_Y(\hat{\underset{\sim}{\xi}}) = \bar{\underset{\sim}{t}}$, giving the saddlepoint approximations for the mean density

$$f_{\bar{T}}(\bar{\underset{\sim}{t}}) = (\tfrac{n}{2\pi i})^d \int_{\hat{\xi}_1 - i\infty}^{\hat{\xi}_1 + i\infty} \cdots \int_{\hat{\xi}_d - i\infty}^{\hat{\xi}_d + i\infty} \exp[n(K_Y(\underset{\sim}{\xi}) - \underset{\sim}{\xi}'\bar{\underset{\sim}{t}})] d\underset{\sim}{\xi}. \tag{4.3}$$

Similarly, integrating through the saddlepoints gives the saddlepoint approximation for the conditional tail probability of the mean variables

$$\begin{aligned} &P(\bar{T}_1 \geq \bar{t}_1, \cdots, \bar{T}_m \geq \bar{t}_m | \bar{T}_{m+1} = \bar{t}_{m+1}, \cdots, \bar{T}_d = \bar{t}_d) = (2\pi i)^{-d} \times \\ &\int_{\hat{\xi}_1 - i\infty}^{\hat{\xi}_1 + i\infty} \cdots \int_{\hat{\xi}_d - i\infty}^{\hat{\xi}_d + i\infty} \exp[n(K_Y(\underset{\sim}{\xi}) - \underset{\sim}{\xi}'\bar{\underset{\sim}{t}})][\textstyle\prod_{j=1}^m \xi_j]^{-1} d\underset{\sim}{\xi} / f_{\bar{T}_{m+1}, \cdots, \bar{T}_d}(\bar{t}_{m+1}, \cdots, \bar{t}_d). \end{aligned} \tag{4.4}$$

The unconditional tail probability approximation is simply obtained by dropping the denominator of (4.4) and increase the number of the product of ξ_j to d.

Saddlepoint methods can be widely implemented for different distributions to obtain a good approximation for the tail probability of a mean random vectors; besides that, saddlepoint approximation can also be applied to evaluate p-value, conditional or unconditional power in the framework of hypothesis testing provided the cumulant generating function of the test statistics exists, and especially when the sample size is small, normal approximation is no longer valid, or the closed forms of p-value or power do not exist.

4.2 Existing Approximations

Robinson (1982) provided a saddlepoint approximation for the univariate case in the form of

$$\begin{aligned} P(T_n \geq x_n) &= Q_n(u) \int_{x_n}^{\infty} e^{-uy} dV_n(y) \\ &= (2\pi\sigma_n^2)^{-1/2} Q_n(u) \int_{x_n}^{\infty} \exp(-uy - (y-m_n)^2/2\sigma_n^2) dy \\ &+ Q_n(u) \int_{x_n}^{\infty} e^{-uy} d(V_n(y) - G_n(y)) \\ &= A_1 + A_2, \end{aligned} \tag{4.5}$$

where the leading term is $Q_n(u) = \int_{-\infty}^{\infty} e^{uy} dF_n(y)$ with F_n being the distribution function of T_n, $V_n(x) = Q_n(u)^{-1} \int_{-\infty}^{x} e^{uy} dF_n(y)$, $A_1 = Q_n(u) \exp(-\mu m_n + \mu^2 \sigma_n^2/2)(1 - \Phi(\mu\sigma_n))$, $|A_2| \leq 2 Q_n(u) \exp(-\mu m_n) O_p(\frac{1}{\sqrt{n}})$, $m_n(u)$ and $\sigma_n^2(u)$ are the mean and variance of V_n. u is chosen such that $m_n(u) = x_n$.

It is worth noting that the above approximation (4.5) applies only under the condition

$$x_n > 0. \tag{4.6}$$

Kolassa (2003) extended the saddlepoint approximation to the multivariate case in form of

$$P(\underset{\sim}{T} \geq \underset{\sim}{t}) = Q(\underset{\sim}{t}) + \exp(n[K(\hat{\underset{\sim}{\xi}}) - \textstyle\sum_j \hat{\xi}_j t_j]) E_2(\underset{\sim}{t})/n, \tag{4.7}$$

and the leading term $Q(\underset{\sim}{t})$ is defined as

$$\begin{aligned}
&\exp(n[K(\hat{\underset{\sim}{\xi}}) - \textstyle\sum_j \hat{\xi}_j t_j + \sum_{jk} \hat{\xi}_j K^{jk}(\hat{\underset{\sim}{\xi}})\hat{\xi}_k/2]) \\
&\times\big[I(nK''(\hat{\underset{\sim}{\xi}}), \underset{\sim}{s}, \underset{\sim}{0}, nK''(\hat{\underset{\sim}{\xi}})\hat{\underset{\sim}{\xi}}, \hat{\underset{\sim}{\xi}}) \\
&+\tfrac{1}{6}\textstyle\sum_{jkl} K^{jkl}(\hat{\underset{\sim}{\xi}}) I(nK''(\hat{\underset{\sim}{\xi}}), \underset{\sim}{s}, \underset{\sim}{e_j} + \underset{\sim}{e_k} + \underset{\sim}{e_l}, nK''(\hat{\underset{\sim}{\xi}})\hat{\underset{\sim}{\xi}}, \hat{\underset{\sim}{\xi}})/\sqrt{n}\big].
\end{aligned} \tag{4.8}$$

The quantities I are determined as the linear combinations of derivatives of multivariate tail probability, and are given by Kolassa (2003), p278-279. The leading terms for both (4.5) and (4.7) are such that the resulting approximation is not in the form of a normal tail probability evaluated at some data-defined quantity, and further more if we try to approximate $P(X \geq x_n)$ with $x_n < 0$ by doing $P(X \geq x_n) = 1 - P(Y \geq y_n)$ for $Y = -X$ and $y_n = -x_n$, the result is valid but not of form that extends by naively applying (4.5) and (4.7) when constraint (4.6) is omitted.

Wang (1990) provided a saddlepoint approximation for the bivariate case in the form of the univariate and bivariate normal distribution and density functions. Wang's approximation will be reviewed in Section 5.

4.3 Proposed Approximation

As shown in Section 1.2, to avoid the leading terms that arise in (4.5) and (4.7), the numerator of (4.4) can be expressed by the change of variables $\{\underset{\sim}{\xi} \rightarrow \underset{\sim}{\omega}\}$ suggested by Kolassa and Li (2010) as

$$\begin{aligned}
&(2\pi i)^{-d} \textstyle\int_{\hat{\omega}_1 - i\infty}^{\hat{\omega}_1 + i\infty} \cdots \int_{\hat{\omega}_d - i\infty}^{\hat{\omega}_d + i\infty} \exp(n[\underset{\sim}{\omega}'\underset{\sim}{\omega}/2 - \hat{\underset{\sim}{\omega}}'\underset{\sim}{\omega}]) \\
&\times\big[\tfrac{1}{\prod_{k=1}^{m}(\omega_j - \tilde{\omega_j})} + \textstyle\sum_{j=1}^{m} \tfrac{G^{(j)}(\underset{\sim}{\omega})}{\prod_{k \neq j}(\omega_j - \tilde{\omega_j})}\big] d\underset{\sim}{\omega},
\end{aligned} \tag{4.9}$$

where ω_j and $\hat{\omega}_j$ are defined as

$$\begin{aligned}(\omega_j - \hat{\omega}_j)^2/2 = \quad & \min_{\underset{\sim}{\gamma}}(K_Y(\underset{\sim}{\gamma}) - \underset{\sim}{\gamma}'\underset{\sim}{\bar{t}}|\gamma_i = \xi_i, \forall i \le j) \\ & -\min_{\underset{\sim}{\gamma}}(K_Y(\underset{\sim}{\rho}) - \underset{\sim}{\rho}'\underset{\sim}{\bar{t}}|\rho_i = \xi_i, \forall i < j), \\ \hat{\omega}_j^2/2 = \quad & \min_{\underset{\sim}{\gamma}}(K_Y(\underset{\sim}{\gamma}) - \underset{\sim}{\gamma}'\underset{\sim}{\bar{t}}|\gamma_i = 0, \forall i \le j) \\ & -\min_{\underset{\sim}{\gamma}}(K_Y(\underset{\sim}{\rho}) - \underset{\sim}{\rho}'\underset{\sim}{\bar{t}}|\rho_i = 0, \forall i < j),\end{aligned} \tag{4.10}$$

so that the contribution of variables of integration corresponding to the unconditioned components of $\underset{\sim}{T}$ to the exponent in (4.4) is exactly quadratic, and $\tilde{\omega_j}(\omega_1, \cdots, \omega_{j-1})$ is defined to satisfy $\xi_j(\omega_1, \cdots, \omega_{j-1}, \tilde{\omega_j}) = 0$ for $j = 1, \cdots, m$ (Kolassa and Li (2010)).

Let $G(\underset{\sim}{\omega}) = \prod_{j=1}^m(\omega_j - \tilde{\omega}_j(\omega_1, \cdots, \omega_{j-1}))[\prod_{j=1}^m \xi_j]^{-1}\frac{d}{d\underset{\sim}{\omega}}\underset{\sim}{\xi}'$; note that $G(0) = 1$. For any index $j \in 1, \cdots, m$, define

$$\begin{aligned}&G^{(j)}(\underset{\sim}{\omega}) = \\ &(G(\omega_1, \cdots, \omega_j, 0, \cdots, 0) - G(\omega_1, \cdots, \omega_{j-1}, \tilde{\omega}_j, 0, \cdots, 0))/(\omega_j - \tilde{\omega}_j),\end{aligned} \tag{4.11}$$

so that $\frac{G^{(j)}(\underset{\sim}{\omega})}{\prod_{k \ne j}(\omega_j - \tilde{\omega_j})}$ is analytic.

The quantities $G^{(j)}(\underset{\sim}{\omega})$ are functions only of $\omega_1, \cdots, \omega_j$. When $j = 1$ the integral factors into an integral involving the first component of $\underset{\sim}{\omega}$, and an integral involving all other components. The integral involving with $G^{(1)}$ can be evaluated using Watson's Lemma, since the integrands are analytic functions (Watson's Lemma will be described in Section 2.). The integrals of the other terms in the bracket of (4.9) are very difficult to evaluate since the integrands are not analytic. We will develop a new approximation technique for the first-term integral to the error of order $O_p(1/n)$ by expanding the quantities in the numerator of the first term using

$$\tilde{\omega_j}(\boldsymbol{\omega_j}) = \textstyle\sum_{k<j} \omega_k \hat{a_j}^k + \sum_{k,l<j} \omega_k(\omega_l - \hat{\omega}_l) a_j^{k;l}(\boldsymbol{\omega_j^*}), \tag{4.12}$$

where

$$\begin{aligned} a_j^k(\boldsymbol{\omega_j}) &= (\tilde{\omega_j}(\omega_1, \cdots, \omega_k, 0, \cdots) - \tilde{\omega_j}(\omega_1, \cdots, \omega_{k-1}, 0, \cdots))/\omega_k, \\ a_j^{k;l}(\boldsymbol{\omega_j}) &= \tfrac{\partial a_j^k(\boldsymbol{\omega_j})}{\partial \omega_l} \end{aligned} \tag{4.13}$$

for $k, l \leq j$, and therefore obtain an approximation for the conditional and unconditional tail probability to the error of order $O_p(1/n)$.

Particularly, for the bivariate case, using the definition of (4.13) gives

$$\begin{aligned} a_2^1(\omega_1) &= \tilde{\omega_2}(\omega_1)/\omega_1, \\ a_2^{1;1}(\omega_1) &= -\tfrac{\tilde{\omega_2}(\omega_1)}{\omega_1^2} + \tfrac{1}{\omega_1}\tfrac{\partial \tilde{\omega_2}(\omega_1)}{\partial \omega_1}, \end{aligned} \tag{4.14}$$

and using definition (4.10) gives

$$\begin{aligned} &\tilde{\omega_2}(\hat{\omega}_1) = \hat{\omega}_2 - \operatorname{sgn}(\hat{\xi}_2)\sqrt{2\big((K_Y(\hat{\gamma}_1, 0) - \hat{\gamma}_1\bar{t}_1) - (K_Y(\hat{\xi}_1, \hat{\xi}_2) - \hat{\xi}_1\bar{t}_1 - \hat{\xi}_2\bar{t}_2)\big)}, \\ &\hat{\omega}_1 = \operatorname{sgn}(\hat{\gamma}_1)\sqrt{2\big((K_Y(0, \hat{\gamma}_2) - \hat{\gamma}_2\bar{t}_2) - (K_Y(\hat{\xi}_1, \hat{\xi}_2) - \hat{\xi}_1\bar{t}_1 - \hat{\xi}_2\bar{t}_2)\big)}, \\ &\hat{\omega}_2 = \operatorname{sgn}(\hat{\gamma}_2)\sqrt{2(\hat{\gamma}_2\bar{t}_2 - K_Y(0, \hat{\gamma}_2))}, \end{aligned} \tag{4.15}$$

where $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are the saddlepoints when fixing $\xi_2 = 0$ and $\xi_1 = 0$, respectively. The derivative of $\tilde{\omega_2}(\omega_1)$ with respect to ω_1 evaluated at $\hat{\omega}_1$ is given in Lemma 4.1.

Lemma 4.1.

$$\tfrac{\partial \tilde{\omega_2}}{\partial \omega_1}\big|_{\hat{\omega}_1} = \tfrac{K_Y^1(\hat{\xi}_1, 0) - \bar{t}_1}{\tilde{\omega}_2(\hat{\omega}_1) - \hat{\omega}_2}\sqrt{\tfrac{1}{K_Y^{11}(\hat{\xi}_1, \hat{\xi}_2) - K_Y^{12}(\hat{\xi}_1, \hat{\xi}_2)(K_Y^{22}(\hat{\xi}_1, \hat{\xi}_2))^{-1}K_Y^{21}(\hat{\xi}_1, \hat{\xi}_2)}}. \tag{4.16}$$

Proof. Let subscript k represents component k of a vector, and superscript l represents the first derivative with respect to argument l (for example, ξ_k^l represents $\frac{\partial \xi_k}{\partial \omega_l}$); let K_Y^i represents the first derivative of the cumulant function $K_Y(\xi_1, \xi_2)$ with respect to argument i (for example, K_Y^1 represents $\frac{\partial K_Y(\xi_1, \xi_2)}{\partial \xi_1}$), and let K_Y^{ij} represents the second derivative with respect to argument i and j of the cumulant function K_Y (for example,

K_Y^{11} represents $\frac{\partial^2 K_Y(\xi_1,\xi_2)}{\partial \xi_1^2}$). By definition, $\tilde{\omega}_2(\omega_1)$ satisfies $\xi_2(\omega_1, \tilde{\omega}_2(\omega_1)) = 0$, and

$$\xi_2^1(\omega_1, \tilde{\omega}_2(\omega_1)) + \xi_2^2(\omega_1, \tilde{\omega}_2(\omega_1))\tilde{\omega}_2^1(\omega_1) = 0. \tag{4.17}$$

So

$$\tilde{\omega}_2^1(\omega_1) = -\xi_2^1(\omega_1, \tilde{\omega}_2(\omega_1))/\xi_2^2(\omega_1, \tilde{\omega}_2(\omega_1)). \tag{4.18}$$

Differentiating

$$(\underset{\sim}{\omega} - \hat{\underset{\sim}{\omega}})^T(\underset{\sim}{\omega} - \hat{\underset{\sim}{\omega}})/2 = K_Y(\underset{\sim}{\xi}) - \underset{\sim}{\xi}^T \bar{\underset{\sim}{t}} - (K_Y(\hat{\underset{\sim}{\xi}}) - \hat{\underset{\sim}{\xi}}^T \bar{\underset{\sim}{t}}) \tag{4.19}$$

with respect to ω_2 on both sides gives $(\omega_2 - \hat{\omega}_2) = (K_Y^2 - \bar{t}_2)\xi_2^2$; hence

$$\xi_2^2(\hat{\omega}_1, \tilde{\omega}_2(\hat{\omega}_1)) = \frac{\tilde{\omega}_2(\hat{\omega}_1) - \hat{\omega}_2}{K_Y^2(\hat{\xi}_1, 0) - \bar{t}_2}. \tag{4.20}$$

Differentiating (4.19) with respect to ω_1 on both sides gives $(\omega_1 - \hat{\omega}_1) = (K_Y^1 - \bar{t}_1)\xi_1^1 + (K_Y^2 - \bar{t}_2)\xi_2^1$, and evaluating at $\tilde{\xi}_2(\xi_1)$ (the value of ξ_2 minimizing the saddlepoint equation as a function of ξ_1) sets $K_Y^2 - \bar{t}_2 = 0$ and gives $(\omega_1 - \hat{\omega}_1) = (K_Y^1(\xi_1, \tilde{\xi}_2(\xi_1)) - \bar{t}_1)\xi_1^1$. Differentiating again with respect to ω_1 gives

$$(K_Y^{11}(\xi_1, \tilde{\xi}_2(\xi_1)) + K_Y^{12}(\xi_1, \tilde{\xi}_2(\xi_1))\tilde{\xi}_2^1(\xi_1))(\xi_1^1)^2 + (K_Y^1(\xi_1, \tilde{\xi}_2(\xi_1)) - \bar{t}_1)\frac{d\xi_1^1}{d\omega_1} = 1,$$

and evaluating at $\hat{\omega}_1$ sets $K_Y^1(\hat{\xi}_1, \tilde{\xi}_2(\hat{\xi}_1)) - \bar{t}_1 = 0$, since $\tilde{\xi}_2(\hat{\xi}_1) = \hat{\xi}_2$, then $\xi_1^1(\hat{\omega}_1) = \sqrt{\frac{1}{K_Y^{11}(\hat{\xi}_1,\hat{\xi}_2) + K_Y^{12}(\hat{\xi}_1,\hat{\xi}_2)\tilde{\xi}_2^1(\hat{\xi}_1)}}$. The derivative $\tilde{\xi}_2^1(\hat{\xi}_1)$ can be obtained by differentiating $K_Y^2 - \bar{t}_2 = 0$ with respect to ξ_1 which gives $K_Y^{21}(\xi_1, \tilde{\xi}_2(\xi_1)) + K_Y^{22}(\xi_1, \tilde{\xi}_2(\xi_1))\tilde{\xi}_2^1(\xi_1) = 0$, then $\tilde{\xi}_2^1(\hat{\xi}_1) = -K_Y^{21}(\hat{\xi}_1, \hat{\xi}_2)/K_Y^{22}(\hat{\xi}_1, \hat{\xi}_2)$. Hence

$$\begin{aligned}
\xi_2^1(\hat{\omega}_1, \tilde{\omega}_2(\hat{\omega}_1)) &= \tfrac{(\omega_1-\hat{\omega}_1)-(K_Y^1-\bar{t}_1)\xi_1^1}{K_Y^2-\bar{t}_2}|_{\hat{\omega}_1} \\
&= -\tfrac{(K_Y^1(\hat{\xi}_1,0)-\bar{t}_1)\xi_1^1(\hat{\omega}_1)}{K_Y^2(\hat{\xi}_1,0)-\bar{t}_2} \\
&= -\tfrac{K_Y^1(\hat{\xi}_1,0)-\bar{t}_1}{K_Y^2(\hat{\xi}_1,0)-\bar{t}_2}\sqrt{\tfrac{1}{K_Y^{11}(\hat{\xi}_1,\hat{\xi}_2)-K_Y^{12}(\hat{\xi}_1,\hat{\xi}_2)(K_Y^{22}(\hat{\xi}_1,\hat{\xi}_2))^{-1}K_Y^{21}(\hat{\xi}_1,\hat{\xi}_2)}}.
\end{aligned} \tag{4.21}$$

$\square$

5. APPROXIMATION FOR THE UNCONDITIONAL BIVARIATE TAIL PROBABILITY

5.1 Methodology

In this section, we will evaluate the tail probability (4.9) to the error of $O_p(\frac{1}{n})$ for the bivariate case term by term, and we will compare our approximation to Wang's approximation in Section 5. The integral of the first term in (4.9) for the bivariate case can be expressed as

$$\begin{aligned}
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2])\tfrac{1}{(\omega_1-\tilde{\omega_1})(\omega_2-\tilde{\omega_2})}\tfrac{d\omega_1 d\omega_2}{(2\pi i)^2}\\
&=\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2])\tfrac{1}{\omega_1(\omega_2-a_2^1\omega_1-a_2^{11}\omega_1(\omega_1-\hat{\omega}_1))}\tfrac{d\omega_1 d\omega_2}{(2\pi i)^2}\\
&+O_p(\tfrac{1}{n}),
\end{aligned}\tag{5.1}$$

where $a_2^1 = a_2^1(\hat{\omega}_1)$ and $a_2^{11} = a_2^{1;1}(\hat{\omega}_1)$.

Split the fraction in RHS of (5.1) and cancel ω_1 in the numerator and denominator to obtain

$$\begin{aligned}
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\tfrac{\exp(n[\omega_1^2/2+\omega_2^2/2-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2])}{\omega_1(\omega_2-a_2^1\omega_1)}\tfrac{d\omega_1 d\omega_2}{(2\pi i)^2}\\
&+\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\tfrac{\exp(n[\omega_1^2/2+\omega_2^2/2-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2])a_2^{11}(\omega_1-\hat{\omega}_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\omega_1(\omega_1-\hat{\omega}_1))}\tfrac{d\omega_1 d\omega_2}{(2\pi i)^2}.
\end{aligned}\tag{5.2}$$

The integrand of the second integral of (5.2) contains a nonlinear term of ω_1 in its denominator, $a_2^{11}\omega_1(\omega_1-\hat{\omega}_1)$, which is the major obstacle for evaluating the integral of (5.2). Our strategy is to replace the nonlinear term by $a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1)$ and show that the error induced by the linear replacement is $O_p(\frac{1}{n})$. The results are presented through Theorem 5.1 to Theorem 5.3. We evaluate the remaining integral without the nonlinearity of ω_1 in the denominator. The results are presented in Lemma 5.1, Lemma 5.2, and Theorem 5.4.

The denominator in the second integral, $(\omega_2 - a_2^1\omega_1)(\omega_2 - a_2^1\omega_1 - a_2^{11}\omega_1(\omega_1 - \hat{\omega}_1))$ is a product of two factors. The second factor in the product equals to the first factor plus a quadratic term of ω_1. Expressing the denominator such that it is exactly equal to a square requires finding a such that

$$(\omega_2 - a_2^1\omega_1)(\omega_2 - a_2^1\omega_1 - a_2^{11}\omega_1(\omega_1 - \hat{\omega}_1)) = (\omega_2 - a_2^1\omega_1 - a\omega_1(\omega_1 - \hat{\omega}_1))^2, \qquad (5.3)$$

where a is the root of the quadratic equation $A(\omega_1, \hat{\omega}_1)a^2 + B(\omega_1, \omega_2)a + C(\omega_1, \omega_2) = 0$ for $A(\omega_1, \hat{\omega}_1) = \omega_1(\omega_1 - \hat{\omega}_1)$, $B(\omega_1, \omega_2) = -(\omega_2 - a_2^1\omega_1)$, and $C(\omega_1, \omega_2) = a_2^{11}(\omega_2 - a_2^1\omega_1)$, so $a = a(A, B, C)$ is a function of (A, B, C). Reparameterize the second integral with $z_1 = \omega_1$ and $z_2 = \omega_2 - a_2^1\omega_1 - a\omega_1(\omega_1 - \hat{\omega}_1)$, to give

$$\begin{aligned}&\int_{\hat{\omega}_1 - i\infty}^{\hat{\omega}_1 + i\infty}\int_{\hat{\omega}_2 - i\infty}^{\hat{\omega}_2 + i\infty} \frac{\exp(n[\omega_1^2/2 + \omega_2^2/2 - \hat{\omega}_1\omega_1 - \hat{\omega}_2\omega_2])}{\omega_1(\omega_2 - a_2^1\omega_1)} \frac{d\omega_1 d\omega_2}{(2\pi i)^2} \\ &+ \int_{\hat{z}_1 - i\infty}^{\hat{z}_1 + i\infty}\int_{\hat{z}_2 - i\infty}^{\hat{z}_2 + i\infty} \frac{\exp(n[L(z_1, z_2)])a_2^{11}(z_1 - \hat{z}_1)}{z_2^2} \frac{dz_1 dz_2}{(2\pi i)^2},\end{aligned} \qquad (5.4)$$

where $L(z_1, z_2) = \frac{1}{2}(1 + (a_2^1 + a\hat{z}_1)^2)(z_1 - \hat{z}_1)^2 + \frac{1}{2}(z_2 - \hat{z}_2)^2 + (a_2^1 + a\hat{z}_1)(z_1 - \hat{z}_1)(z_2 - \hat{z}_2) - \frac{1}{2}(\hat{\omega}_1^2 + \hat{\omega}_2^2) + Q_a(z_1, z_2)$,and $Q_a(z_1, z_2) = a(z_1 - \hat{z}_1)^2(z_2 - \hat{z}_2) + a(a_2^1 + a\hat{z}_1)(z_1 - \hat{z}_1)^3 + \frac{1}{2}a^2(z_1 - \hat{z}_1)^4$ containing cubic and higher terms of $z_1 - \hat{z}_1$ and $z_2 - \hat{z}_2$ with a in the coefficients. We will show that dropping $Q_a(z_1, z_2)$ will give an error of $O_p(\frac{1}{n})$. Observe that there is a z_2^2 in the denominator of the second integral in (5.2), Kolassa (2003) showed that $|\int_{\hat{z}_2 - i\infty}^{\hat{z}_2 + i\infty} \frac{1}{z_2} dz_2|$ is bounded by a finite constant independent of $\hat{z}_2$; we will show in Theorem 5.1 that $|\int_{\hat{z}_2 - i\infty}^{\hat{z}_2 + i\infty} \frac{1}{z_2^2} dz_2|$ is also bounded by a finite constant independent of $\hat{z}_2$, and this conclusion will be used in Theorem 5.2 to show the error is $O_p(\frac{1}{n})$ by dropping $Q_a(z_1, z_2)$.

Theorem 5.1.

$$|\int_{\hat{z}_2-i\infty}^{\hat{z}_2+i\infty}\frac{1}{z_2^2}dz_2| \le 2\sqrt{n}\sum_{k=0}^{\infty}\frac{1}{k^2} \tag{5.5}$$

Proof. Make the change of variable $z_2 = \hat{z}_2 + i\frac{y}{\sqrt{n}}$ to give

$$\begin{aligned}
&|\int_{\hat{z}_2-i\infty}^{\hat{z}_2+i\infty}\frac{1}{z_2^2}dz_2| = |\int_{-\infty}^{\infty}\frac{\hat{z}_2^2-y^2/n}{(\hat{z}_2^2+y^2/n)^2}\frac{dy}{\sqrt{n}}| \\
&\le \int_{-\infty}^{\infty}\frac{\hat{z}_2^2}{(\hat{z}_2^2+y^2/n)^2}\frac{dy}{\sqrt{n}} + \int_{-\infty}^{\infty}\frac{y^2/n}{(\hat{z}_2^2+y^2/n)^2}\frac{dy}{\sqrt{n}} \\
&= \int_{-\infty}^{\infty}\frac{1}{\hat{z}_2^2+y^2/n}\frac{dy}{\sqrt{n}} \\
&= \sqrt{n}\int_{-\infty}^{\infty}\frac{1}{n\hat{z}_2^2+y^2}dy.
\end{aligned} \tag{5.6}$$

Let $a = (\frac{\sqrt{n}|\hat{z}_2|}{\sqrt{n}|\hat{z}_2|+1})^2$ to give

$$\begin{aligned}
&\sqrt{n}\int_{-\infty}^{\infty}\frac{1}{n\hat{z}_2^2+y^2}dy = \sqrt{n}\int_{-\infty}^{\infty}\frac{1}{y^2+a(\sqrt{n}|\hat{z}_2|+1)^2}dy \\
&\le \sqrt{n}\int_{-\infty}^{\infty}\frac{\sqrt{n}|\hat{z}_2|+1}{y^2+a(\sqrt{n}|\hat{z}_2|+1)^2}dy \\
&= 2\sqrt{n}\int_{0}^{\infty}\frac{\sqrt{n}|\hat{z}_2|+1}{y^2+a(\sqrt{n}|\hat{z}_2|+1)^2}dy;
\end{aligned} \tag{5.7}$$

the last equality in (5.7) is due to the symmetry of the integrand. Divide the range of the integrand in (5.7) into intervals $(k(\sqrt{n}|\hat{z}_2|+1), (k+1)(\sqrt{n}|\hat{z}_2|+1)]$ of equal length $\sqrt{n}|\hat{z}_2|+1$ for $k = 0, 1, \cdots$, the maximum of the integrand in each interval is $\frac{\sqrt{n}|\hat{z}_2|+1}{(a+k^2)(\sqrt{n}|\hat{z}_2|+1)^2}$, hence the integral over each interval is bounded by the maximum multiplied by the length of the interval $\sqrt{n}|\hat{z}_2|+1$, i.e, $\frac{1}{a+k^2} \le \frac{1}{k^2}$, therefore the entire integral of (5.7) is bounded by $\sum_{k=0}^{\infty}\frac{1}{k^2}$. □

The contribution of $Q_a(z_1, z_2)$ which contains the cubic and higher order terms of $z_1 - \hat{z}_1$ and $z_2 - \hat{z}_2$ to the integral (5.4) is $O_p(\frac{1}{n})$. The proof is presented in Theorem 5.2.

Theorem 5.2. *Dropping* $Q_a(z_1, z_2)$ *from* $L(z_1, z_2)$ *in the second integral of* (5.2) *gives* *an error of* $O_p(\frac{1}{n})$,*i.e.,*

$$\int_{\hat{z}_1-i\infty}^{\hat{z}_1+i\infty}\int_{\hat{z}_2-i\infty}^{\hat{z}_2+i\infty}\frac{\exp(n[L(z_1,z_2)])a_2^{11}(z_1-\hat{z}_1)}{z_2^2}\frac{dz_1dz_2}{(2\pi i)^2}= \\ \int_{\hat{z}_1-i\infty}^{\hat{z}_1+i\infty}\int_{\hat{z}_2-i\infty}^{\hat{z}_2+i\infty}\frac{\exp(n[L^*(z_1,z_2)])a_2^{11}(z_1-\hat{z}_1)}{z_2^2}\frac{dz_1dz_2}{(2\pi i)^2}(1+O_p(\frac{1}{\sqrt{n}})), \tag{5.8}$$

where $L^*(z_1,z_2)=\frac{1}{2}(1+(a_2^1+a\hat{z}_1)^2)(z_1-\hat{z}_1)^2+\frac{1}{2}(z_2-\hat{z}_2)^2+(a_2^1+a\hat{z}_1)(z_1-\hat{z}_1)(z_2-\hat{z}_2)-\frac{1}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2)$.

Proof. Note that $L(z_1,z_2)=L^*(z_1,z_2)+Q_a(z_1,z_2)$ and use Taylor expansion for $Q_a(z_1,z_2)$ about $(\hat{z}_1,\hat{z}_2)$, since all the linear and quadratic terms vanished. The RHS of (5.8) becomes

$$\int_{\hat{z}_1-i\infty}^{\hat{z}_1+i\infty}\int_{\hat{z}_2-i\infty}^{\hat{z}_2+i\infty}\frac{\exp(n[L^*(z_1,z_2)])}{z_2^2}a_2^{11}(z_1-\hat{z}_1)\times \\ (1+n\sum_{l\geq 2,m\geq 1}(z_1-\hat{z}_1)^l(z_2-\hat{z}_2)^m q^{lm}(\hat{z}_1,\hat{z}_2))\frac{dz_1dz_2}{(2\pi i)^2} \\ =\int_{\hat{z}_1-i\infty}^{\hat{z}_1+i\infty}\int_{\hat{z}_2-i\infty}^{\hat{z}_2+i\infty}\frac{\exp(n[L^*(z_1,z_2)])}{z_2^2}a_2^{11}(z_1-\hat{z}_1)\frac{dz_1dz_2}{(2\pi i)^2}+ \\ n\int_{\hat{z}_1-i\infty}^{\hat{z}_1+i\infty}\int_{\hat{z}_2-i\infty}^{\hat{z}_2+i\infty}\frac{\exp(n[L^*(z_1,z_2)])}{z_2^2}a_2^{11}\sum_{l\geq 2,m\geq 1}(z_1-\hat{z}_1)^l(z_2-\hat{z}_2)^m q^{lm}(\hat{z}_1,\hat{z}_2) \\ (z_1-\hat{z}_1)\frac{dz_1dz_2}{(2\pi i)^2}. \tag{5.9}$$

Make the change of variables $z_1=\hat{z}_1+i\frac{x}{\sqrt{n}}$ and $z_2=\hat{z}_2+i\frac{y}{\sqrt{n}}$ in (5.9). Note that $\exp(-n\frac{1}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2))\leq 1$, for the first term by Theorem 5.1 the L_2 norm of $\frac{1}{z_2}$,i.e.,$|\int_{\hat{z}_2-i\infty}^{\hat{z}_2+i\infty}\frac{1}{z_2^2}dz_2|$ is $O_p(\sqrt{n})$ which cancels out the factor of $\frac{1}{\sqrt{n}}$ contributed from dz_1, since $z_1-\hat{z}_1$ contributes a factor of $\frac{1}{\sqrt{n}}$. Hence the first term is $O_p(\frac{1}{\sqrt{n}})$; for the second term, the factor of $\frac{1}{n}$ contributed from $(z_1-\hat{z}_1)dz_1$ cancels the leading n, and $(z_1-\hat{z}_1)^l(z_2-\hat{z}_2)^m$ for $l\geq 2, m\geq 1$ contributes a factor of $\frac{1}{n\sqrt{n}}$, and by Theorem 5.1 the second term is bounded by $\frac{1}{n}$. □

By reparameterizing RHS of (5.8) to the variables of $\{\omega_1,\omega_2\}$, Theorem 5.3 will show that replacing the nonlinearity of ω_1,$a_2^{11}\omega_1(\omega_1-\hat{\omega}_1)$, in the denominator in (5.2) by the linear term $a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1)$ will induce an error in the order of $O_p(\frac{1}{n})$.

Theorem 5.3.

$$\begin{aligned}
&\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\frac{\exp(n[\omega_1^2/2+\omega_2^2/2-\hat\omega_1\omega_1-\hat\omega_2\omega_2])}{\omega_1(\omega_2-a_2^1\omega_1)}\frac{d\omega_1 d\omega_2}{(2\pi i)^2}\\
&+\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\frac{\exp(n[\omega_1^2/2+\omega_2^2/2-\hat\omega_1\omega_1-\hat\omega_2\omega_2])a_2^{11}(\omega_1-\hat\omega_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\omega_1(\omega_1-\hat\omega_1))}\frac{d\omega_1 d\omega_2}{(2\pi i)}\\
&=\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\frac{\exp(n[\omega_1^2/2+\omega_2^2/2-\hat\omega_1\omega_1-\hat\omega_2\omega_2])}{\omega_1(\omega_2-a_2^1\omega_1)}\frac{d\omega_1 d\omega_2}{(2\pi i)^2}\\
&+\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\frac{\exp(n[\omega_1^2/2+\omega_2^2/2-\hat\omega_1\omega_1-\hat\omega_2\omega_2])a_2^{11}(\omega_1-\hat\omega_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat\omega_1(\omega_1-\hat\omega_1))}\frac{d\omega_1 d\omega_2}{(2\pi i)^2}(1+O_p(\tfrac{1}{\sqrt{n}})).
\end{aligned} \tag{5.10}$$

Proof. Since by Theorem 5.2, the second integral of (5.2) can be written as

$$\begin{aligned}
&\int_{\hat z_1-i\infty}^{\hat z_1+i\infty}\int_{\hat z_2-i\infty}^{\hat z_2+i\infty}\frac{\exp(n[\frac{1}{2}(1+(a_2^1+a\hat z_1)^2)(z_1-\hat z_1)^2+\frac{1}{2}(z_2-\hat z_2)^2+(a_2^1+a\hat z_1)(z_1-\hat z_1)(z_2-\hat z_2)-\frac{1}{2}(\hat\omega_1^2+\hat\omega_2^2)])}{z_2^2}\\
&a_2^{11}(z_1-\hat z_1)\frac{dz_1dz_2}{(2\pi i)^2}(1+O_p(\tfrac{1}{\sqrt{n}})).
\end{aligned} \tag{5.11}$$

Reparameterize (5.11) by $z_1=\omega_1$ and $z_2=\omega_2-a_2^1\omega_1-a\hat\omega_1(\omega_1-\hat\omega_1)$ to give the numerator of the integrand of (5.11) as $\exp(n[\omega_1^2/2+\omega_2^2/2-\hat\omega_1\omega_1-\hat\omega_2\omega_2])a_2^{11}(\omega_1-\hat\omega_1)$ and the denominator as $(\omega_2-a_2^1\omega_1-a\hat\omega_1(\omega_1-\hat\omega_1))^2$. Since $a=a(A,B,C)$ satisfies (5.3) where $A(\omega_1,\hat\omega_1)=\hat\omega_1(\omega_1-\hat\omega_1)$ for this reparameterization, hence $(\omega_2-a_2^1\omega_1-a\hat\omega_1(\omega_1-\hat\omega_1))^2=(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat\omega_1(\omega_1-\hat\omega_1))$. □

We will now evaluate RHS of (5.10) through reparameterization of $\{\omega_1,\omega_2\}$ such that the integrals of RHS of (5.10) can be expressed as bivariate normal survival function or the products of normal density function and normal survival function. Lemma 5.1 gives the expression for the first integral of RHS of (5.10), and Lemma 5.2 gives the expression for the second integral.

Lemma 5.1. *Let $\bar\Phi_2((t_1,t_2),\Omega)$ denotes the bivariate normal survival function of mean vector $\underset{\sim}{0}$ and variance-covariance matrix Ω beyond point (t_1,t_2), the first integral of RHS of* (5.10) *can be expressed as*

$$\begin{aligned}
&\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\frac{\exp(n[\omega_1^2/2+\omega_2^2/2-\hat\omega_1\omega_1-\hat\omega_2\omega_2])}{\omega_1(\omega_2-a_2^1\omega_1)}\frac{d\omega_1 d\omega_2}{(2\pi i)^2}\\
&=\bar\Phi_2((\hat\omega_1+a_2^1\hat\omega_2,\hat\omega_2),\Sigma/n),
\end{aligned} \tag{5.12}$$

where $\Sigma = \begin{pmatrix} \varsigma^2 & a_2^1 \\ a_2^1 & 1 \end{pmatrix}$ *and* $\varsigma = \sqrt{1+(a_2^1)^2}$.

Proof. Reparameterize the first integral of the RHS of (5.10) such that the denominator of the integrand is the product of two variables. Let $\{\omega_1 \to u_1, \omega_2 - a_2^1\omega_1 \to u_2\}$ and $\varsigma = \sqrt{1+(a_2^1)^2}$ to give

$$\begin{aligned}&\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty} \tfrac{\exp(n[\omega_1^2/2+\omega_2^2/2-\hat\omega_1\omega_1-\hat\omega_2\omega_2])}{\omega_1(\omega_2-a_2^1\omega_1)}\tfrac{d\omega_1 d\omega_2}{(2\pi i)^2} \\ &= \int_{\hat u_1-i\infty}^{\hat u_1+i\infty}\int_{\hat u_2-i\infty}^{\hat u_2+i\infty} \tfrac{\exp(n[\varsigma^2 u_1^2/2+u_2^2/2+a_2^1u_1u_2-a_2^1(\hat u_2+a_2^1\hat u_1)u_1-a_2^1\hat u_1u_2-\hat u_1u_1-\hat u_2u_2])}{u_1u_2}\tfrac{du_1du_2}{(2\pi i)^2}.\end{aligned} \tag{5.13}$$

Use the saddlepoint approximation of the normal tail probability of mean variables (4.4) for $d=2$, (5.13) is the bivariate normal survival function of mean $(-a_2^1(\hat u_2 + a_2^1\hat u_1), -a_2^1\hat u_1)$, variance-covariance matrix $\Sigma/n = \begin{pmatrix} \varsigma^2 & a_2^1 \\ a_2^1 & 1 \end{pmatrix}/n$ evaluated at $(\hat u_1, \hat u_2)$. Hence the first integral of the RHS of (5.10) is

$$\begin{aligned}&\bar\Phi_2((\varsigma^2\hat u_1 + a_2^1\hat u_2, \hat u_2 + a_2^1\hat u_1), \Sigma/n) \\ &= \bar\Phi_2((\hat\omega_1 + a_2^1\hat\omega_2, \hat\omega_2), \Sigma/n).\end{aligned} \tag{5.14}$$

$\square$

Lemma 5.2. *The second integral of RHS of* (5.10) *can be expressed as*

$$\begin{aligned}&\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty} \tfrac{\exp(n[\omega_1^2/2+\omega_2^2/2-\hat\omega_1\omega_1-\hat\omega_2\omega_2])a_2^{11}(\omega_1-\hat\omega_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat\omega_1(\omega_1-\hat\omega_1))}\tfrac{d\omega_1d\omega_2}{(2\pi i)^2} \\ &= \mathrm{sgn}(a_2^{11}\hat\omega_1)\tfrac{1}{\sqrt{n}a_2^{11}\hat\omega_1^2}e^{-\frac{n}{2}(\hat\omega_1^2+\hat\omega_2^2-\hat u_2^2)}[\sigma_1^{-1}\phi(\tfrac{1+\varsigma_1}{\sqrt{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat u_2)\bar\Phi(\sqrt{\tfrac{\varsigma_2^2-\varsigma_1^2}{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat u_2) - \\ &\sigma_2^{-1}\phi(\tfrac{\varsigma_1}{\varsigma_2}\sqrt{n}\hat u_2)\bar\Phi(\sqrt{\tfrac{\varsigma_2^2-\varsigma_1^2}{\varsigma_2^2}}\sqrt{n}\hat u_2)],\end{aligned} \tag{5.15}$$

where $\sigma_1^2 = 1+2\varsigma_1+\varsigma_2^2$, $\sigma_2^2 = \varsigma_2^2$, $\varsigma_1 = \frac{a_2^1}{a_2^{11}\hat\omega_1}$, $\varsigma_2 = \frac{\varsigma}{a_2^{11}\hat\omega_1}$, $\hat u_1 = \hat\omega_1$, *and* $\hat u_2 = \hat\omega_2 - a_2^1\hat\omega_1$.

Proof. Reparameterize the second integral of the RHS of (5.10) such that the denominator of the integrand is the product of two variables. Let $\{\omega_2 - a_2^1\omega_1 -$

$a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1)\to u_1, \omega_2-a_2^1\omega_1\to u_2\}$, and let $\varsigma_1=\frac{a_2^1}{a_2^{11}\hat{\omega}_1}$, $\varsigma_2=\frac{\varsigma}{a_2^{11}\hat{\omega}_1}$, $L(u_1,u_2)=\varsigma_2^2u_1^2/2+(1+2\varsigma_1+\varsigma_2^2)u_2^2/2-(\varsigma_1+\varsigma_2^2)u_1u_2-(1+\varsigma_1)\hat{u}_2u_2+\varsigma_1\hat{u}_2u_1$ to give

$$\begin{aligned}&\mathrm{sgn}(a_2^{11}\hat{\omega}_1)\tfrac{1}{a_2^{11}\hat{\omega}_1^2}e^{-\frac{n}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2-\hat{u}_2^2)}\int_{\hat{u}_1-i\infty}^{\hat{u}_1+i\infty}\int_{\hat{u}_2-i\infty}^{\hat{u}_2+i\infty}\tfrac{\exp(nL(u_1,u_2))(u_2-u_1)}{u_1u_2}\tfrac{du_1du_2}{(2\pi i)^2}\\&=\mathrm{sgn}(a_2^{11}\hat{\omega}_1)\tfrac{1}{a_2^{11}\hat{\omega}_1^2}e^{-\frac{n}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2-\hat{u}_2^2)}\Big[\int_{\hat{u}_1-i\infty}^{\hat{u}_1+i\infty}\int_{\hat{u}_2-i\infty}^{\hat{u}_2+i\infty}\tfrac{\exp(nL(u_1,u_2))}{u_1}\tfrac{du_1du_2}{(2\pi i)^2}-\\&\int_{\hat{u}_1-i\infty}^{\hat{u}_1+i\infty}\int_{\hat{u}_2-i\infty}^{\hat{u}_2+i\infty}\tfrac{\exp(nL(u_1,u_2))}{u_2}\tfrac{du_1du_2}{(2\pi i)^2}\Big].\end{aligned}\tag{5.16}$$

Reparameterize (5.16) such that 1) the term of product of u_1 and u_2 in $L(u_1,u_2)$ vanishes in the numerator and 2)u_i for $i=1,2$ in the denominator is held unchanged, so the bivariate integral can be expressed as the product of two univariate integrals. Make change of variables for the first term in (5.16), let $a=(\varsigma_1+\varsigma_2^2)/(1+2\varsigma_1+\varsigma_2^2)$ and $\{u_1\to v_1, u_2-au_1\to v_2\}$ to give

$$\begin{aligned}&\mathrm{sgn}(a_2^{11}\hat{\omega}_1)\tfrac{1}{a_2^{11}\hat{\omega}_1^2}e^{-\frac{n}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2-\hat{u}_2^2)}\times\\&\int_{\hat{v}_1-i\infty}^{\hat{v}_1+i\infty}\tfrac{\exp(n[\frac{1}{2}(\varsigma_2^2-2a(\varsigma_1+\varsigma_2^2)+a^2(1+2\varsigma_1+\varsigma_2^2))v_1^2-(a(1+\varsigma_1)-\varsigma_1)\hat{u}_2v_1])}{v_1}\tfrac{dv_1}{2\pi i}\times\\&\int_{\hat{v}_2-i\infty}^{\hat{v}_2+i\infty}\exp(n[\tfrac{1}{2}(1+2\varsigma_1+\varsigma_2^2)v_2^2-(1+\varsigma_1)\hat{u}_2v_2])\tfrac{dv_2}{2\pi i}.\end{aligned}\tag{5.17}$$

Make change of variables for the second term in (5.16), let $b=(\varsigma_1+\varsigma_2^2)/\varsigma_2^2$ and $\{u_1-bu_2\to v_1, u_2\to v_2\}$ to give

$$\begin{aligned}&\mathrm{sgn}(a_2^{11}\hat{\omega}_1)\tfrac{1}{a_2^{11}\hat{\omega}_1^2}e^{-\frac{n}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2-\hat{u}_2^2)}\times\\&\int_{\hat{v}_1-i\infty}^{\hat{v}_1+i\infty}\exp(n[\tfrac{1}{2}\varsigma_2^2v_1^2+\varsigma_1\hat{u}_2v_1])\tfrac{dv_1}{2\pi i}\times\\&\int_{\hat{v}_2-i\infty}^{\hat{v}_2+i\infty}\tfrac{\exp(n[\frac{1}{2}(b^2\varsigma_2^2-2b(\varsigma_1+\varsigma_2^2)+1+2\varsigma_1+\varsigma_2^2)v_2^2-(1+\varsigma_1-b\varsigma_1)\hat{u}_2v_2])}{v_2}\tfrac{dv_2}{2\pi i}.\end{aligned}\tag{5.18}$$

Hence (5.16), the second integral of RHS of (5.10), can be expressed as

$$
\begin{aligned}
&\operatorname{sgn}(a_2^{11}\hat{\omega}_1)\tfrac{1}{a_2^{11}\hat{\omega}_1^2}e^{-\frac{n}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2-\hat{u}_2^2)}\times\\
&\Big[\textstyle\int_{\hat{v}_1-i\infty}^{\hat{v}_1+i\infty}\frac{\exp(n[\frac{1}{2}(\varsigma_2^2-2a(\varsigma_1+\varsigma_2^2)+a^2(1+2\varsigma_1+\varsigma_2^2))v_1^2-(a(1+\varsigma_1)-\varsigma_1)\hat{u}_2v_1])}{v_1}\frac{dv_1}{2\pi i}\times\\
&\textstyle\int_{\hat{v}_2-i\infty}^{\hat{v}_2+i\infty}\exp(n[\frac{1}{2}(1+2\varsigma_1+\varsigma_2^2)v_2^2-(1+\varsigma_1)\hat{u}_2v_2])\frac{dv_2}{2\pi i}\\
&-\\
&\textstyle\int_{\hat{v}_1-i\infty}^{\hat{v}_1+i\infty}\exp(n[\frac{1}{2}\varsigma_2^2v_1^2+\varsigma_1\hat{u}_2v_1])\frac{dv_1}{2\pi i}\times\\
&\textstyle\int_{\hat{v}_2-i\infty}^{\hat{v}_2+i\infty}\frac{\exp(n[\frac{1}{2}(b^2\varsigma_2^2-2b(\varsigma_1+\varsigma_2^2)+1+2\varsigma_1+\varsigma_2^2)v_2^2-(1+\varsigma_1-b\varsigma_1)\hat{u}_2v_2])}{v_2}\frac{dv_2}{2\pi i}\Big].
\end{aligned}
\tag{5.19}
$$

The univariate integrals in (5.19) with a denominator of u_1 or u_2 can be expressed as a standard normal survival function, and the univariate integrals without a denominator can be expressed as the standard normal density function multiplied by a constant.

Write

$$
\begin{aligned}
&\textstyle\int_{\hat{v}_1-i\infty}^{\hat{v}_1+i\infty}\frac{\exp(n[\frac{1}{2}(\varsigma_2^2-2a(\varsigma_1+\varsigma_2^2)+a^2(1+2\varsigma_1+\varsigma_2^2))v_1^2-(a(1+\varsigma_1)-\varsigma_1)\hat{u}_2v_1])}{v_1}\frac{dv_1}{2\pi i}\\
&=\textstyle\int_{\hat{v}_1-i\infty}^{\hat{v}_1+i\infty}\frac{\exp(n[\frac{1}{2}\sigma^2v_1^2+\mu v_1-\hat{v}_1v_1])}{v_1}\frac{dv_1}{2\pi i},
\end{aligned}
\tag{5.20}
$$

where $\sigma^2=\varsigma_2^2-2a(\varsigma_1+\varsigma_2^2)+a^2(1+2\varsigma_1+\varsigma_2^2)$ and $\mu=\hat{v}_1-(a(1+\varsigma_1)-\varsigma_1)\hat{u}_2$.

Use the saddlepoint approximation of the tail probability of the mean variables (4.4) for $d=1$. Expression (5.20) is the normal survival function evaluated at $\hat{v}_1$ of a normal distribution of mean μ and variance σ^2/n. Substitute a and convert to the standard normal survival function to give

$$
\begin{aligned}
&\textstyle\int_{\hat{v}_1-i\infty}^{\hat{v}_1+i\infty}\frac{\exp(n[\frac{1}{2}(\varsigma_2^2-2a(\varsigma_1+\varsigma_2^2)+a^2(1+2\varsigma_1+\varsigma_2^2))v_1^2-(a(1+\varsigma_1)-\varsigma_1)\hat{u}_2v_1])}{v_1}\frac{dv_1}{2\pi i}\\
&=\bar{\Phi}\Big(\sqrt{\tfrac{\varsigma_2^2-\varsigma_1^2}{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2\Big).
\end{aligned}
\tag{5.21}
$$

Write

$$
\begin{aligned}
&\textstyle\int_{\hat{v}_2-i\infty}^{\hat{v}_2+i\infty}\exp(n[\frac{1}{2}(1+2\varsigma_1+\varsigma_2^2)v_2^2-(1+\varsigma_1)\hat{u}_2v_2])\frac{dv_2}{2\pi i}\\
&=\textstyle\int_{\hat{v}_2-i\infty}^{\hat{v}_2+i\infty}\exp(n[\frac{1}{2}(1+2\varsigma_1+\varsigma_2^2)v_2^2+(\hat{v}_2-(1+\varsigma_1)\hat{u}_2)v_2-\hat{v}_2v_2])\frac{dv_2}{2\pi i}.
\end{aligned}
\tag{5.22}
$$

Use the saddlepoint approximation of the mean density (4.3) for $d = 1$, (5.22) is $\frac{1}{n}$ times a density function evaluated at $\hat{v}_2$ of a normal distribution of mean $\hat{v}_2-(1+\varsigma_1)\hat{u}_2$ and variance $(1 + 2\varsigma_1 + \varsigma_2^2)/n$, convert to the standard normal density to give

$$\begin{aligned}&\int_{\hat{v}_2-i\infty}^{\hat{v}_2+i\infty} \exp(n[\tfrac{1}{2}(1 + 2\varsigma_1 + \varsigma_2^2)v_2^2 - (1 + \varsigma_1)\hat{u}_2 v_2])\tfrac{dv_2}{2\pi i}\\ &= \tfrac{1}{n}\sigma_1^{-1}\phi(\tfrac{1+\varsigma_1}{\sqrt{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2),\end{aligned} \tag{5.23}$$

where $\sigma_1^2 = (1 + 2\varsigma_1 + \varsigma_2^2)/n$.

Let $\sigma_2^2 = \varsigma_2^2/n$ and use same arguments for the remaining two univariate integrals in (5.19) to express (5.19) in the form of normal functions as

$$\begin{aligned}&\mathrm{sgn}(a_2^{11}\hat{\omega}_1)\tfrac{1}{na_2^{11}\hat{\omega}_1^2}e^{-\frac{n}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2-\hat{u}_2^2)}[\sigma_1^{-1}\phi(\tfrac{1+\varsigma_1}{\sqrt{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\tfrac{\varsigma_2^2-\varsigma_1^2}{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)-\\ &\sigma_2^{-1}\phi(\tfrac{\varsigma_1}{\varsigma_2}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\tfrac{\varsigma_2^2-\varsigma_1^2}{\varsigma_2^2}}\sqrt{n}\hat{u}_2)].\end{aligned} \tag{5.24}$$

Factor out $\sqrt{n}$ from σ_1^{-1} and σ_2^{-1} in the bracket in (5.24) and re-define $\sigma_1^2 = 1+2\varsigma_1+\varsigma_2^2$ and $\sigma_2^2 = \varsigma_2^2$ to give

$$\begin{aligned}&\mathrm{sgn}(a_2^{11}\hat{\omega}_1)\tfrac{1}{\sqrt{n}a_2^{11}\hat{\omega}_1^2}e^{-\frac{n}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2-\hat{u}_2^2)}[\sigma_1^{-1}\phi(\tfrac{1+\varsigma_1}{\sqrt{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\tfrac{\varsigma_2^2-\varsigma_1^2}{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)-\\ &\sigma_2^{-1}\phi(\tfrac{\varsigma_1}{\varsigma_2}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\tfrac{\varsigma_2^2-\varsigma_1^2}{\varsigma_2^2}}\sqrt{n}\hat{u}_2)].\end{aligned} \tag{5.25}$$

□

Lemma 5.1 and Lemma 5.2 give the expressions of the first and second integral of RHS of (5.10) in the form of normal functions, respectively. Therefore, by (5.12) and (5.15) we have the following theorem.

Theorem 5.4. *To the error of $O_p(\frac{1}{n})$, the first term of the bivariate tail probability*

(5.1) *can be approximated by*

$$\bar{\Phi}_2((\hat{\omega}_1 + a_2^1\hat{\omega}_2, \hat{\omega}_2), \Sigma/n) + \mathrm{sgn}(a_2^{11}\hat{\omega}_1)\frac{1}{\sqrt{n}a_2^{11}\hat{\omega}_1^2}e^{-\frac{n}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2-\hat{u}_2^2)}\times$$
$$[\sigma_1^{-1}\phi(\frac{1+\varsigma_1}{\sqrt{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2) - \sigma_2^{-1}\phi(\frac{\varsigma_1}{\varsigma_2}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{\varsigma_2^2}}\sqrt{n}\hat{u}_2)], \tag{5.26}$$

where the quantity $\hat{\omega}_1$ *and* $\hat{\omega}_2$ *are the square root of twice the unconditional log likelihood ratio statistics for testing the corresponding* $\xi_j = 0, j = 1, 2,$ *assuming that* $\xi_l = 0$ *for* $l < j$*, but without restriction on* ξ_l *for* $l > j$*, and* $\sigma_1^2 = 1+2\varsigma_1+\varsigma_2^2$*,* $\sigma_2^2 = \varsigma_2^2$*,* $\varsigma_1 = \frac{a_2^1}{a_2^{11}\hat{\omega}_1}$*,* $\varsigma_2 = \frac{\varsigma}{a_2^{11}\hat{\omega}_1}$*,* $\varsigma = \sqrt{1+(a_2^1)^2}$*,* $\hat{u}_1 = \hat{\omega}_1$*,* $\hat{u}_2 = \hat{\omega}_2 - a_2^1\hat{\omega}_1$*,* $\Sigma = \begin{pmatrix} \varsigma^2 & a_2^1 \\ a_2^1 & 1 \end{pmatrix}$.

Proof. Since (5.12) equals to the first integral of RHS of (5.10) and (5.15) equals to second integral of the RHS of (5.10), summing (5.12) and (5.15) gives (5.26). □

Note that $\hat{\omega}_1 = 0$ is a singularity point for (5.26). Lemma 5.3 will show that the singularity point is removable.

Lemma 5.3. $\hat{\omega}_1 = 0$ *is a removable singularity point for*

$$f(\hat{\omega}_1) = \frac{1}{a_2^{11}\hat{\omega}_1^2}\times$$
$$[\sigma_1^{-1}\phi(\frac{1+\varsigma_1}{\sqrt{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2) - \sigma_2^{-1}\phi(\frac{\varsigma_1}{\varsigma_2}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{\varsigma_2^2}}\sqrt{n}\hat{u}_2)]. \tag{5.27}$$

Proof. Use Riemann's theorem, it is to show $\lim_{\hat{\omega}_1\to 0}\hat{\omega}_1 f(\hat{\omega}_1) = 0$. Since by mean value theorem $\lim_{\hat{\omega}_1\to 0} a_2^{11}(\epsilon\hat{\omega}_1)\hat{\omega}_1 = \lim_{\hat{\omega}_1\to 0}(a_2^1(\hat{\omega}_1) - a_2^1(0)) = 0$, it is equivalent to show that $\lim_{a_2^{11}\hat{\omega}_1\to 0}\hat{\omega}_1 f(\hat{\omega}_1) = 0$. Write

$$\hat{\omega}_1 f(\hat{\omega}_1) = a_2^{11}\hat{\omega}_1\frac{f(\hat{\omega}_1)}{a_2^{11}} = (a_2^{11}\hat{\omega}_1\sigma_1)^{-1}\phi(\frac{1+\varsigma_1}{\sqrt{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)$$
$$-(a_2^{11}\hat{\omega}_1\sigma_2)^{-1}\phi(\frac{\varsigma_1}{\varsigma_2}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{\varsigma_2^2}}\sqrt{n}\hat{u}_2), \tag{5.28}$$

since $\varsigma_2^2 = (1+\frac{1}{(a_2^1)^2})\varsigma_1^2$. The product of the density function and the survival function in the second term of (5.28)

$$\phi(\frac{\varsigma_1}{\varsigma_2}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{\varsigma_2^2}}\sqrt{n}\hat{u}_2) = \phi(\sqrt{\frac{(a_2^1)^2}{1+(a_2^1)^2}}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{1}{1+(a_2^1)^2}}\sqrt{n}\hat{u}_2),$$

and the limitation of the product of the density function and the survival function in the first term

$\lim_{a_2^{11}\hat{\omega}_1\to 0}\phi(\frac{1+\varsigma_1}{\sqrt{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2) =$
$\lim_{a_2^{11}\hat{\omega}_1\to 0}\phi(\sqrt{\frac{(a_2^1)^2}{1+(a_2^1)^2}}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{1}{1+(a_2^1)^2}}\sqrt{n}\hat{u}_2);$
hence the products in the two terms of (5.28) are equal as $\hat{\omega}_1 \to 0$. And

$$\lim_{\hat{\omega}_1\to 0} a_2^{11}\hat{\omega}_1\sigma_1 = \lim_{a_2^{11}\hat{\omega}_1\to 0}\sqrt{a_2^{11}\hat{\omega}_1 + 2a_2^1 a_2^{11}\hat{\omega}_1 + \varsigma^2} = \lim_{a_2^{11}\hat{\omega}_1\to 0}\varsigma = \lim_{a_2^{11}\hat{\omega}_1\to 0} a_2^{11}\hat{\omega}_1\sigma_2,$$

the two constant multiplier of the two terms in (5.28) are also equal at the singularity point. Hence (5.28) goes to 0 at the singularity point of $\hat{\omega}_1 = 0$. □

The integrals of the second and third term in (4.9) to the error of $O_p(\frac{1}{n})$ are much easier to perform using Watson's Lemma and presented in Theorem 5.5 and 5.6, respectively.

Watson's Lemma states that if $g(\omega)$ is analytic in a neighborhood of $\omega = \hat{\omega}$ then $i^{-1}\sqrt{\frac{n}{2\pi}}\int_{\hat{\omega}-i\infty}^{\hat{\omega}+i\infty}\exp(n(\omega-\hat{\omega})^2/2)g(\omega)d\omega = \sum_{j=0}^{\infty}\frac{(-1)^j g^{(2j)}(\hat{\omega})}{(2n)^j j!}$. Hence to the error of $O_p(\frac{1}{n})$, $\int_{\hat{\omega}-i\infty}^{\hat{\omega}+i\infty}\exp(n(\omega^2/2-\hat{\omega}\omega))g(\omega)\frac{d\omega}{2\pi i} = \frac{\phi(\sqrt{n}\hat{\omega})}{\sqrt{n}}g(\hat{\omega})$.

Theorem 5.5. *To the error of $O_p(\frac{1}{n})$, the second term of the bivariate tail probability (4.9) can be approximated by*

$$\frac{1}{\sqrt{n}}\bar{\Phi}(\sqrt{n}(\hat{\omega}_2-\tilde{\omega}_2(\hat{\omega}_1)))\phi(\sqrt{n}\hat{\omega}_1)(\frac{1}{\hat{z}_1}-\frac{1}{\hat{\omega}_1})e^{n(\tilde{\omega}_2(\hat{\omega}_1)^2/2-\hat{\omega}_2\tilde{\omega}_2(\hat{\omega}_1))}, \tag{5.29}$$

where $\hat{\omega}_1$ and $\hat{\omega}_2$ are the square roots of twice the unconditional log likelihood ratio statistics defined as in Theorem 5.4, and

$\hat{z}_1 = \hat{\xi}_1\sqrt{\frac{K_Y^{11}(\hat{\xi}_1,\hat{\xi}_2)K_Y^{22}(\hat{\xi}_1,\hat{\xi}_2)-K_Y^{12}(\hat{\xi}_1,\hat{\xi}_2)K_Y^{21}(\hat{\xi}_1,\hat{\xi}_2)}{K_Y^{22}(\hat{\xi}_1,\hat{\xi}_2)}}.$

Proof. Since

$$\frac{G^{(1)}(\omega_1,0)}{\omega_2-\tilde{\omega}_2(\omega_1)} = \frac{1}{\omega_2-\tilde{\omega}_2(\omega_1)}\left(\frac{1}{\xi_1}\frac{d\xi_1}{d\omega_1} - \frac{1}{\omega_1}\right) \tag{5.30}$$

is an analytic function at $\hat{\omega}_1$ and 0, by Watson's Lemma

$$\begin{aligned}
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty} \exp(n[\omega_1^2/2+\omega_2^2/2-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2])\frac{G^{(1)}(\omega_1,0)}{\omega_2-\tilde{\omega}_2(\omega_1)}\frac{d\omega_1 d\omega_2}{(2\pi i)^2} = \\
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty} \exp(n[\omega_1^2/2+\omega_2^2/2-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2])\frac{1}{\omega_2-\tilde{\omega}_2(\omega_1)}\left(\frac{1}{\xi_1}\frac{d\xi_1}{d\omega_1}-\frac{1}{\omega_1}\right) \\
&\times\frac{d\omega_1 d\omega_2}{(2\pi i)^2} = \\
&\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty} e^{n(\omega_2^2/2-\hat{\omega}_2\omega_2)}\Big[\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty} e^{n(\omega_1^2/2-\hat{\omega}_1\omega_1)}\frac{1}{\omega_2-\tilde{\omega}_2(\omega_1)}\left(\frac{1}{\xi_1}\frac{d\xi_1}{d\omega_1}-\frac{1}{\omega_1}\right)\frac{d\omega_1}{2\pi i}\Big]\frac{d\omega_2}{2\pi i} \\
&= \int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty} e^{n(\omega_2^2/2-\hat{\omega}_2\omega_2)}\frac{\phi(\sqrt{n}\hat{\omega}_1)}{\sqrt{n}}\Big[\frac{1}{\omega_2-\tilde{\omega}_2(\hat{\omega}_1)}\left(\frac{1}{\hat{z}_1}-\frac{1}{\hat{\omega}_1}\right)+O_p\left(\frac{1}{n}\right)\Big]\frac{d\omega_2}{2\pi i} \\
&= \int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty} e^{n(\omega_2^2/2-\hat{\omega}_2\omega_2)}\frac{1}{\omega_2-\tilde{\omega}_2(\hat{\omega}_1)}\frac{d\omega_2}{2\pi i}\frac{\phi(\sqrt{n}\hat{\omega}_1)}{\sqrt{n}}\left(\frac{1}{\hat{z}_1}-\frac{1}{\hat{\omega}_1}\right)+O_p\left(\frac{1}{n}\right),
\end{aligned} \tag{5.31}$$

where $\hat{z}_1 = \hat{\xi}_1\left(\frac{d\xi_1}{d\omega_1}(\hat{\omega}_1)\right)^{-1} = \hat{\xi}_1\sqrt{\frac{K_Y^{11}(\hat{\xi}_1,\hat{\xi}_2)K_Y^{22}(\hat{\xi}_1,\hat{\xi}_2)-K_Y^{12}(\hat{\xi}_1,\hat{\xi}_2)K_Y^{21}(\hat{\xi}_1,\hat{\xi}_2)}{K_Y^{22}(\hat{\xi}_1,\hat{\xi}_2)}}.$
Make the change of variable $u = \omega_2 - \tilde{\omega}_2(\hat{\omega}_1)$ for (5.31) to give

$$\begin{aligned}
&e^{n(\tilde{\omega}_2(\hat{\omega}_1)^2/2-\hat{\omega}_2\tilde{\omega}_2(\hat{\omega}_1))}\int_{\hat{u}-i\infty}^{\hat{u}+i\infty} e^{n(u^2/2-\hat{u}u)}\frac{1}{u}\frac{du}{2\pi i}\frac{\phi(\sqrt{n}\hat{\omega}_1)}{\sqrt{n}}\left(\frac{1}{\hat{z}_1}-\frac{1}{\hat{\omega}_1}\right)+O_p\left(\frac{1}{n}\right) \\
&= \frac{1}{\sqrt{n}}\bar{\Phi}(\sqrt{n}(\hat{\omega}_2-\tilde{\omega}_2(\hat{\omega}_1)))\phi(\sqrt{n}\hat{\omega}_1)\left(\frac{1}{\hat{z}_1}-\frac{1}{\hat{\omega}_1}\right)e^{n(\tilde{\omega}_2(\hat{\omega}_1)^2/2-\hat{\omega}_2\tilde{\omega}_2(\hat{\omega}_1))}+O_p\left(\frac{1}{n}\right).
\end{aligned} \tag{5.32}$$

□

Theorem 5.6. *To the error of $O_p(\frac{1}{n})$, the third term of the bivariate tail probability* (4.9) *can be approximated by*

$$\frac{1}{\sqrt{n}}\bar{\Phi}(\sqrt{n}\hat{\omega}_1)\phi(\sqrt{n}\hat{\omega}_2)\left(\frac{1}{\hat{z}_2}-\frac{1}{\hat{\omega}_2}\right), \tag{5.33}$$

where $\hat{\omega}_1$ and $\hat{\omega}_2$ are the square roots of twice the unconditional log likelihood ratio statistics defined as in Theorem 5.4, $\hat{z}_2 = \check{\xi}_2\sqrt{K_Y^{22}(0,\check{\xi}_2)}$ and $\check{\xi}_2 = \xi_2(0,\hat{\omega}_2)$.

Proof. Write $G^{(2)}(\omega_1,\omega_2) = G^{(2)}(0,\omega_2)+(G^{(2)}(\omega_1,\omega_2)-G^{(2)}(0,\omega_2))$, where $G^{(2)}(0,\omega_2) =$

$\frac{1}{\xi_2(0,\omega_2)}\frac{d\xi_2}{d\omega_2}(0,\omega_2)-\frac{1}{\omega_2}$. The second term will integrate to the order of $O_p(\frac{1}{n})$, hence to the error of $O_p(\frac{1}{n})$ and use Watson's Lemma since $G^{(2)}(0,\omega_2)$ is analytic at $\hat{\omega}_2$, integrating the first term $G^{(2)}(0,\omega_2)$ gives

$$\begin{aligned}
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\exp(n[\omega_1^2/2+\omega_2^2/2-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2])\frac{G^{(2)}(0,\omega_2)}{\omega_1}\frac{d\omega_1 d\omega_2}{(2\pi i)^2}= \\
&\Big[\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}e^{n(\omega_2^2/2-\hat{\omega}_2\omega_2)}G^{(2)}(0,\omega_2)\frac{d\omega_2}{2\pi i}\Big]\Big[\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}e^{n(\omega_1^2/2-\hat{\omega}_1\omega_1)}\frac{1}{\omega_1}\frac{d\omega_1}{2\pi i}\Big] \\
&=\frac{1}{\sqrt{n}}\phi(\sqrt{n}\hat{\omega}_2)G^{(2)}(0,\hat{\omega}_2)\bar{\Phi}(\sqrt{n}\hat{\omega}_1),
\end{aligned} \tag{5.34}$$

where $G^{(2)}(0,\hat{\omega}_2)=\frac{1}{\xi_2(0,\hat{\omega}_2)}\frac{d\xi_2}{d\omega_2}(0,\hat{\omega}_2)-\frac{1}{\hat{\omega}_2}=\frac{1}{\hat{z}_2}-\frac{1}{\hat{\omega}_2}$, and $\check{\xi}_2=\xi_2(0,\hat{\omega}_2)$ is the root of equation $K_Y(0,\check{\xi}_2)-\check{\xi}_2\bar{t}_2-(K_Y(\hat{\xi}_1,\hat{\xi}_2)-\hat{\underset{\sim}{\xi}}^T\underset{\sim}{\bar{t}})-\hat{\omega}_1^2/2=0$. □

Combining Theorem 5.4, 5.5 and 5.6 gives the approximation to the tail probability of (4.9) to the error of $O_p(\frac{1}{n})$, the result is presented in the Theorem 5.7.

Theorem 5.7. *To the error of $O_p(\frac{1}{n})$, the bivariate tail probability (4.9) can be approximated by*

$$\begin{aligned}
&\bar{\Phi}_2((\hat{\omega}_1+a_2^1\hat{\omega}_2,\hat{\omega}_2),\Sigma/n)+\mathrm{sgn}(a_2^{11}\hat{\omega}_1)\frac{1}{\sqrt{n}a_2^{11}\hat{\omega}_1^2}e^{-\frac{n}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2-\hat{u}_2^2)}\times \\
&[\sigma_1^{-1}\phi(\frac{1+\varsigma_1}{\sqrt{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)-\sigma_2^{-1}\phi(\frac{\varsigma_1}{\varsigma_2}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{\varsigma_2^2}}\sqrt{n}\hat{u}_2)] \\
&+\frac{1}{\sqrt{n}}\bar{\Phi}(\sqrt{n}(\hat{\omega}_2-\tilde{\omega}_2(\hat{\omega}_1)))\phi(\sqrt{n}\hat{\omega}_1)(\frac{1}{\hat{z}_1}-\frac{1}{\hat{\omega}_1})e^{n(\tilde{\omega}_2(\hat{\omega}_1)^2/2-\hat{\omega}_2\tilde{\omega}_2(\hat{\omega}_1))} \\
&+\frac{1}{\sqrt{n}}\bar{\Phi}(\sqrt{n}\hat{\omega}_1)\phi(\sqrt{n}\hat{\omega}_2)(\frac{1}{\hat{z}_2}-\frac{1}{\hat{\omega}_2}),
\end{aligned} \tag{5.35}$$

where the quantity $\hat{\omega}_1$ and $\hat{\omega}_2$ are the square root of twice the unconditional log likelihood ratio statistics for testing the corresponding $\xi_j=0, j=1,2$, assuming that $\xi_l=0$ for $l<j$, but without restriction on ξ_l for $l>j$, $\sigma_1^2=1+2\varsigma_1+\varsigma_2^2$, $\sigma_2^2=\varsigma_2^2$, $\varsigma_1=\frac{a_2^1}{a_2^{11}\hat{\omega}_1}$, $\varsigma_2=\frac{\varsigma}{a_2^{11}\hat{\omega}_1}$, $\varsigma=\sqrt{1+(a_2^1)^2}$, $\hat{u}_1=\hat{\omega}_1$, $\hat{u}_2=\hat{\omega}_2-a_2^1\hat{\omega}_1$, $\Sigma=\begin{pmatrix}\varsigma^2 & a_2^1\\ a_2^1 & 1\end{pmatrix}$, $\hat{z}_1=\hat{\xi}_1\sqrt{\frac{K_Y^{11}(\hat{\xi}_1,\hat{\xi}_2)K_Y^{22}(\hat{\xi}_1,\hat{\xi}_2)-K_Y^{12}(\hat{\xi}_1,\hat{\xi}_2)K_Y^{21}(\hat{\xi}_1,\hat{\xi}_2)}{K_Y^{22}(\hat{\xi}_1,\hat{\xi}_2)}}$, $\hat{z}_2=\check{\xi}_2\sqrt{K_Y^{22}(0,\check{\xi}_2)}$, and $\check{\xi}_2=\xi_2(0,\hat{\omega}_2)$.

Proof. Omitted. □

5.2 Results and Conclusion

In the first example, we consider the bivariate random vector $(Y_1, Y_2) = (X_1 + X_2, X_1 + X_3)$, where (X_1, X_2, X_3) are identically independent distributed random variables following exponential distribution with rate equal to 1 of density function $f(x) = e^{-x}$. A Monto Carlo (MC) simulation of sample size 100000 was conducted using R to evaluate the saddlepoint approximation (5.35) given in Section 2 for the bivariate tail probability $P(\bar{Y}_1 \geq t_1, \bar{Y}_2 \geq t_2)$ when $n = 5$. The relative errors of the approximation compared to the exact values for different values of (t_1, t_2) are plotted as the contour lines in the following contour plot (Figure 5.1). Figure 5.1 shows that in general the saddlepoint approximation given in (5.35) performs very well as the relative errors are very small and stable, and the approximation is remarkably good when (t_1, t_2) is in the rectangle of $((0, 1.5), (0, 1.5))$ as the relative errors are consistently near 0.

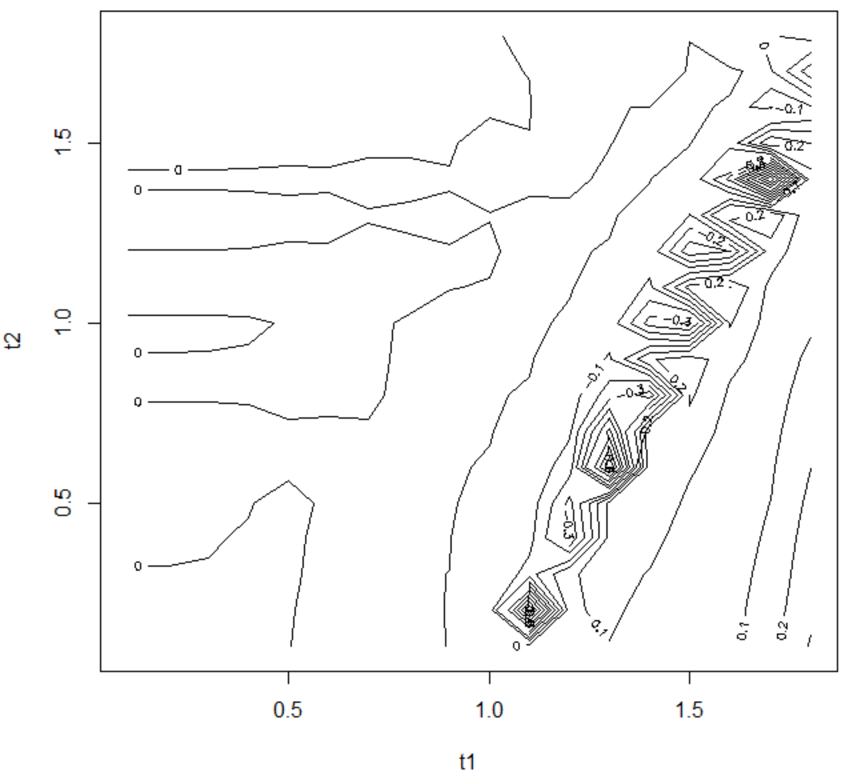

Figure 5.1: Relative errors of the bivariate saddlepoint approximation (5.35) compared to the exact value in the continuous case.
(a) The maximum relative error is 0.31.

In the second example, we consider a discrete case that (X_1, X_2, X_3) are identically independent distributed unit lattice random variables following binomial distribution of the probability function $P(X = x) = \binom{x}{N} p^x (1-p)^{N-x}$ with $N = 10$ and $p = 0.2$. Similarly, a Monto Carlo (MC) simulation of sample size 100000 was conducted to evaluate the saddlepoint approximation for the bivariate tail probability $P(\bar{Y}_1 \geq t_1, \bar{Y}_2 \geq t_2)$ for $n = 5$ where the bivariate random vector $(Y_1, Y_2) = (X_1 + X_2, X_1 + X_3)$. The relative errors of the approximation compared to the exact values are plotted for different values of (t_1, t_2) as contour lines in Figure 5.2. Figure 5.2 shows that in general the approximation performs very well and the best performance of the approximation is observed for smaller (t_1, t_2) similarly to the continuous case in the first example.

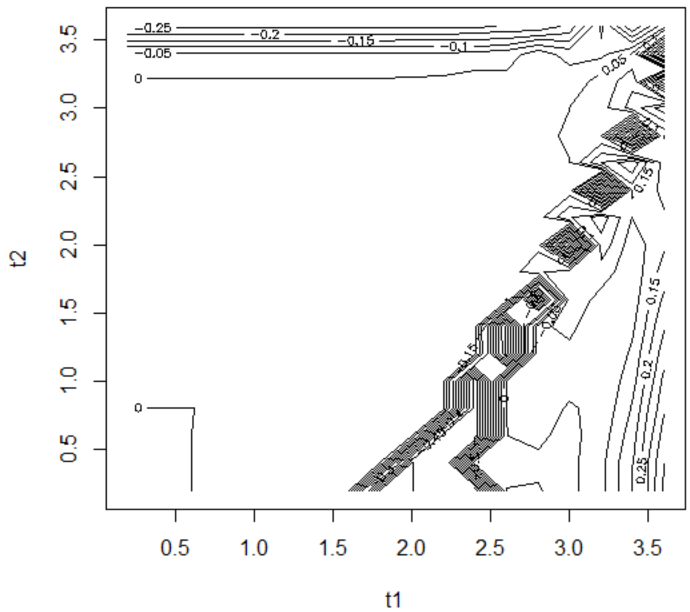

Figure 5.2: Relative errors of the bivariate saddlepoint approximation (5.35) compared to the exact value in the unit lattice case.
(a) The maximum relative error is 0.22.

5.3 *Comparison to Wang's Approximation*

Wang (1990) gave a saddlepoint approximation for the cumulative distribution function of the sample mean of n independent bivariate random vectors using Lugannani and Rice's saddlepoint formula and the standard bivariate normal distribution function. Using our notation to express Wang's approximation for the general continuous case gives

$$\begin{aligned}
&P(\bar{T}_1 \le \bar{t}_1, \bar{T}_2 \le \bar{t}_2) = \Phi(\sqrt{n}x_1, \sqrt{n}y_1, \rho_1) \\
&+\tfrac{1}{\sqrt{n}}\Phi(\sqrt{n}\hat{\omega}_2)\phi(\sqrt{n}\hat{z})(\tfrac{1}{\hat{z}} - \tfrac{1}{\hat{\xi}_2 G}) \\
&+\tfrac{1}{\sqrt{n}}\Phi(\sqrt{n}\hat{z})\phi(\sqrt{n}x_1)(\tfrac{1}{\hat{\omega}_2} - \tfrac{1}{\hat{\xi}_1\sqrt{K_Y^{22}(\hat{\xi}_1,\hat{\xi}_2)}}) \\
&+\tfrac{1}{\sqrt{2\pi n}}\exp\{n[K_Y(\hat{\xi}_1, \hat{\xi}_2) - \hat{\xi}_1\bar{t}_1 - \hat{\xi}_2\bar{t}_2]\}(\tfrac{1}{\hat{\omega}_2} - \tfrac{1}{\hat{\xi}_1\sqrt{K_Y^{22}(\hat{\xi}_1,\hat{\xi}_2)}})(\tfrac{1}{\hat{z}} - \tfrac{1}{\hat{\xi}_2 G}),
\end{aligned} \tag{5.36}$$

where

$$
\begin{aligned}
&w(\xi_2) = \mathrm{sgn}(\tilde{\xi}_1)\sqrt{2[h_{\xi_2}(0) - h_{\xi_2}(\tilde{\xi}_1)]},\\
&h_{\xi_2}(\xi_1) = K_Y(\xi_1, \xi_2) - \xi_1\bar{t}_1 \text{ for each fixed } \xi_2,\\
&\tilde{\xi}_1 \text{ is the minimizer of } h_{\xi_2}(\xi_1),\\
&\hat{z}^2 = -2(K_Y(0, \hat{\xi}_2) - \hat{\xi}_2\bar{t}_2)\\
&x_1 = w(0),\\
&b = \tfrac{\hat{\omega}_2 - w(0)}{\hat{z}},\\
&y_1 = (\hat{z} - bx_1)/(1 + b^2)^{1/2},\\
&\rho_1 = -b/(1 + b^2)^{1/2},\\
&\text{and } G = \sqrt{K^{22}(\hat{\xi}_1, \hat{\xi}_2) - (K^{12}(\hat{\xi}_1, \hat{\xi}_2))^2/K^{11}(\hat{\xi}_1, \hat{\xi}_2)}.
\end{aligned}
\tag{5.37}
$$

Use the same bivariate random vector $(Y_1, Y_2) = (X_1 + X_2, X_1 + X_3)$ as in Section 3 (example one), where (X_1, X_2, X_3) are identically independent distributed random variables following exponential distribution with rate equal to 1 of density function $f(x) = e^{-x}$. Conduct the same Monto Carlo (MC) simulation of sample size 100000 to evaluate Wang's approximation (5.36) for tail probability $P(\bar{Y}_1 \geq t_1, \bar{Y}_2 \geq t_2)$ when $n = 5$ by doing $1 - (5.36)$. The relative errors of Wang's approximation compared to the exact values are plotted in Figure 5.3). Figure 5.3 shows that Wang's approximation (5.36) results in much bigger relative errors compared to our approximation (5.35) as shown in Figure 5.1 (Section 3).

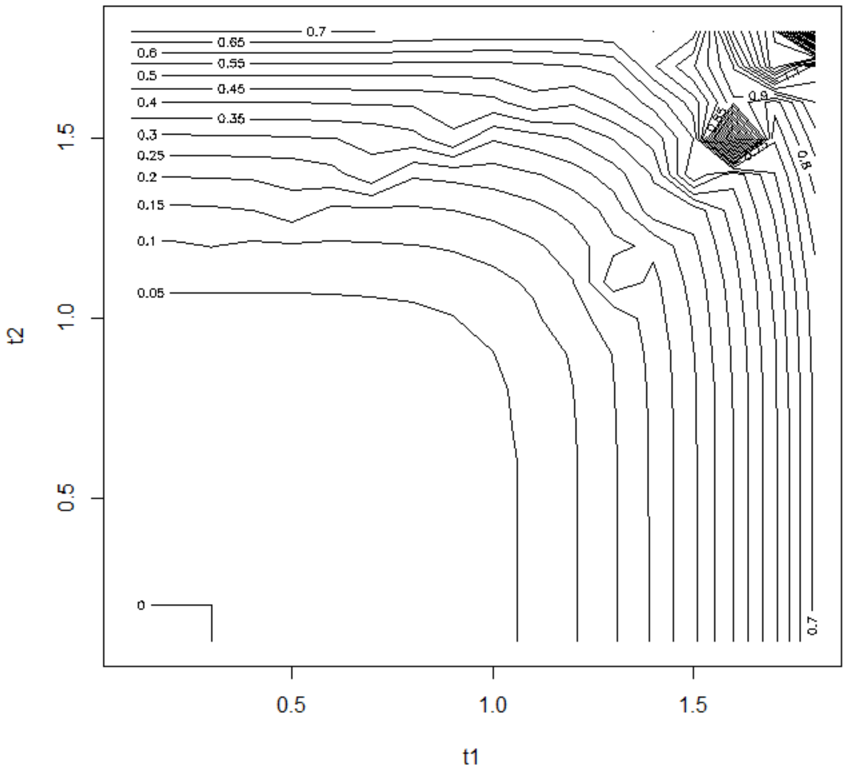

Figure 5.3: Relative errors of Wang's bivariate approximation (5.35) compared to the exact value in the continuous case.
(a) The maximum relative error is 2.18.

6. APPROXIMATION FOR THE CONDITIONAL BIVARIATE TAIL PROBABILITY

6.1 Methodology

In this section, we will consider the approximation for the conditional bivariate tail probability, defined as

$$P(\bar{T}_1 > \bar{t}_1, \bar{T}_2 > \bar{t}_2 | \bar{T}_3 = \bar{t}_3) \tag{6.1}$$

Kolassa and Li (2010) pointed out the multivariate conditional probability approximation is simply equal to the multivariate unconditional tail probability multiplied by $1 + O_p(\frac{1}{n})$ and summarized in Theorem 6.1.

Theorem 6.1. *To the error of $O_p(\frac{1}{n})$, the conditional bivariate tail probability $P(\bar{T}_1 > \bar{t}_1, \bar{T}_2 > \bar{t}_2 | \bar{T}_3 = \bar{t}_3)$ can be approximated by*

$$\int_{\hat{\omega}_1 - i\infty}^{\hat{\omega}_1 + i\infty} \int_{\hat{\omega}_2 - i\infty}^{\hat{\omega}_2 + i\infty} \exp(n[\tfrac{\omega_1^2 + \omega_2^2}{2} - \hat{\omega}_1 \omega_1 - \hat{\omega}_2 \omega_2]) \frac{G(\omega_1, \omega_2)}{(\omega_1 - \tilde{\omega_1})(\omega_2 - \tilde{\omega_2})} \frac{d\omega_1 d\omega_2}{(2\pi i)^2}, \tag{6.2}$$

where the quantity $\hat{\omega}_1$ and $\hat{\omega}_2$ are the square roots of twice the conditional log likelihood ratio statistics for testing the corresponding $\xi_j = 0, j = 1, 2,$ assuming that $\xi_l = 0$ for $l < j$, but without restriction on ξ_l for $l > j$, and conditioning on $\bar{T}_3 = \bar{t}_3$.

Proof. Similarly as Kolassa and Li (2010) showed for the multivariate conditional case, the numerator of the conditional bivariate tail probability

$$\begin{aligned}
&\frac{n}{(2\pi i)^3}\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\
&\times\frac{G(\omega_1,\omega_2)}{(\omega_1-\tilde{\omega_1})(\omega_2-\tilde{\omega_2})}\frac{d\xi_3(\omega_1,\omega_2,\omega_3)}{d\omega_3}|_{(\omega_1=0,\omega_2=0,\omega_3=\hat{\omega}_3)}d\omega_1 d\omega_2 d\omega_3\\
&=\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2])\frac{G(\omega_1,\omega_2)}{(\omega_1-\tilde{\omega_1})(\omega_2-\tilde{\omega_2})}\frac{d\omega_1 d\omega_2}{(2\pi i)^2}\\
&\times\mathbf{J}_3,
\end{aligned}\tag{6.3}$$

where $\mathbf{J}_3 = \frac{n}{2\pi i}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\frac{\omega_3^2}{2}-\hat{\omega}_3\omega_3])d\omega_3 = f_{T_3}(t_3)(1+O_p(\frac{1}{n}))$. Hence by the definition of the conditional probability $\mathbf{J}_3$ is canceled out from the numerator and denominator, (6.2) is achieved. □

By Theorem 6.1, we can directly use Theorem 5.7 to approximate the conditional bivariate tail probability except the only difference is that $\hat{\omega}_1$ and $\hat{\omega}_2$ are the square root of twice the conditional log likelihood ratio statistics.

6.2 Results and Conclusion

In the first example, we consider the same bivariate continuous random vector (Y_1, Y_2) as in Section 3 (example one), and let $Y_3 = X_1$ be the third variable that (Y_1, Y_2) is conditioned on. A Monto Carlo (MC) simulation was conducted to evaluate the saddlepoint approximation for the conditional bivariate tail probability $P(\bar{Y}_1 \geq t_1, \bar{Y}_2 \geq t_2 | \bar{Y}_3 = t_3)$ for $n = 5$. The relative errors for different values of t_3 are plotted as the contour lines in Figure 6.1. Figure 6.1 shows that in general the approximation performs well for different values of t_3, and better performance is observed when t_3 is smaller.

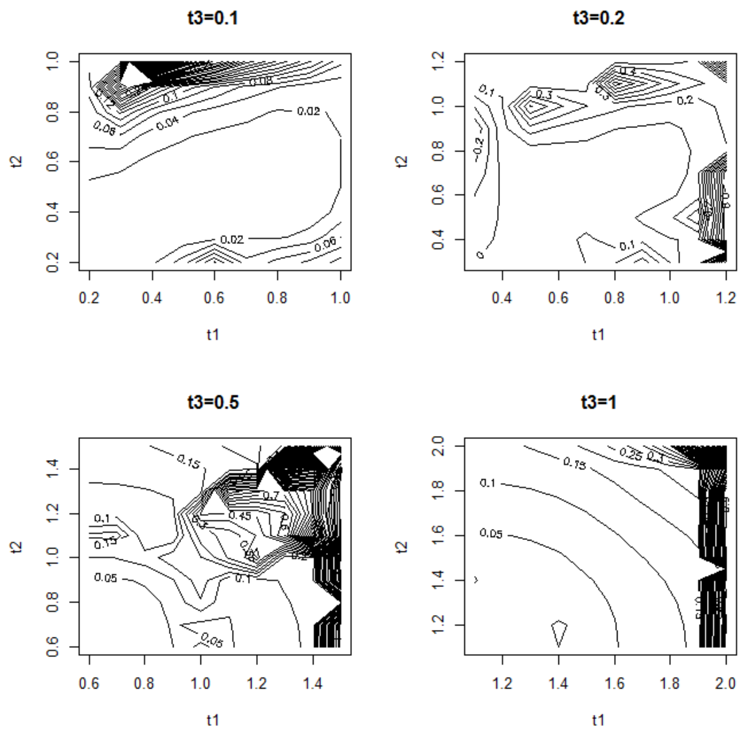

Figure 6.1: Relative errors of the bivariate saddlepoint approximation (5.35) of the conditional tail probability compared to the exact value for the continuous case.

In the second example, we consider the same bivariate lattice random vector (Y_1, Y_2) as in Section 3 (example two), and let $Y_3 = X_1$ be the third variable that (Y_1, Y_2) is conditioning on. A Monto Carlo (MC) simulation was conducted to evaluate the saddlepoint approximation for the conditional bivariate tail probability $P(\bar{Y}_1 \geq t_1, \bar{Y}_2 \geq t_2 | \bar{Y}_3 = t_3)$ for $n = 5$. The relative errors for different values of t_3 are plotted as the contour lines in Figure 6.2. Figure 6.2 shows that the approximation performs well for different values of t_3.

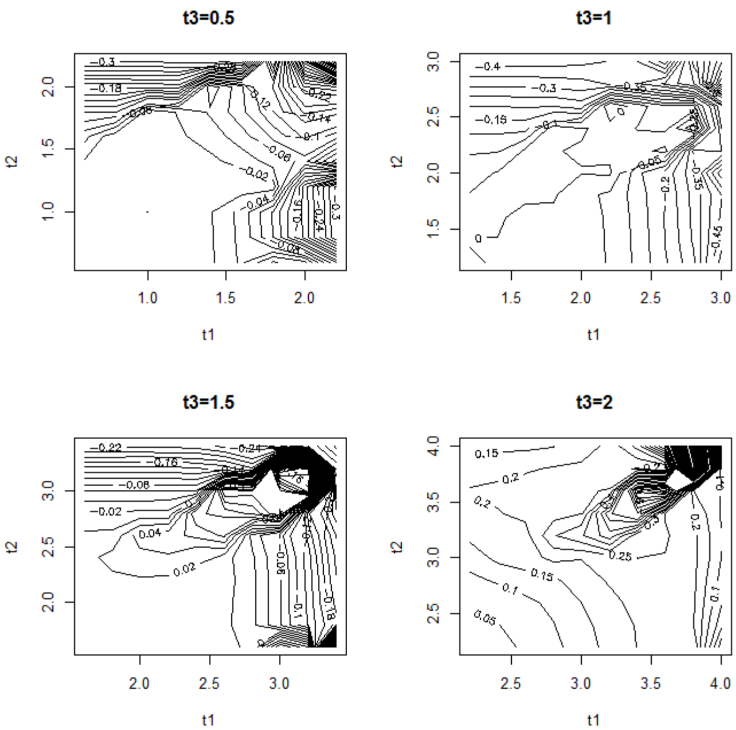

Figure 6.2: Relative errors of the bivariate saddlepoint approximation (5.35) of the conditional tail probability compared to the exact value for the discrete case.

7. APPROXIMATION FOR TRIVARIATE UNCONDITIONAL TAIL PROBABILITY

In this chapter, we will extend the bivariate unconditional tail probability approximation to three dimension with error of $O_p(\frac{1}{n})$. The integral of the first term in (4.9) for the trivariate case can be expressed as

$$\begin{aligned}
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\
&\times\tfrac{1}{(\omega_1-\tilde{\omega}_1)(\omega_2-\tilde{\omega}_2(\omega_1))(\omega_3-\tilde{\omega}_3(\omega_1,\omega_2))}\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}\\
&=\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\
&\times\tfrac{1}{\omega_1[\omega_2-a_2^1\omega_1-a_2^{11}\omega_1(\omega_1-\hat{\omega}_1)][\omega_3-a_3^2\omega_2-a_3^1\omega_1-\sum_{i,j=1,2}a_3^{ij}\omega_i(\omega_j-\hat{\omega}_j)]}\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}\\
&+O_p(\tfrac{1}{n}),
\end{aligned} \tag{7.1}$$

where $a_k^i=a_k^i(\underset{\sim}{\hat{\omega}}_{k-1})$ for $1\le i<k\le 3$, $a_k^{ij}=a_k^{i;j}(\underset{\sim}{\hat{\omega}}_{k-1})$ for $1\le i,j<k\le 3$, and $\underset{\sim}{\hat{\omega}}_{k-1}=(\hat{\omega}_1,\cdots,\hat{\omega}_{k-1})^T$.

As in the bivariate case, the nonlinear terms appear in the denominators of the integrand of (7.1) are the major obstacle for evaluating the integral. Theorem 7.1 will show that replacing $a_2^{11}\omega_1$ by $a_2^{11}\hat{\omega}_1$ in the second denominator of the integrand of (7.1), and replacing $\sum_{i,j=1,2}a_3^{ij}\omega_i(\omega_j-\hat{\omega}_j)$ by $\sum_{i,j=1,2}a_3^{ij}\hat{\omega}_i(\omega_j-\hat{\omega}_j)$ in the third denominator will induce an error of $O_p(\frac{1}{n})$.

Theorem 7.1.

$$\begin{aligned}
&\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat\omega_1\omega_1-\hat\omega_2\omega_2-\hat\omega_3\omega_3])\\
&\times\tfrac{1}{\omega_1[\omega_2-a_2^1\omega_1-a_2^{11}\omega_1(\omega_1-\hat\omega_1)][\omega_3-a_3^2\omega_2-a_3^1\omega_1-\sum_{i,j=1,2}a_3^{ij}\omega_i(\omega_j-\hat\omega_j)]}\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}\\
&=\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat\omega_1\omega_1-\hat\omega_2\omega_2-\hat\omega_3\omega_3])\\
&\times\big(\tfrac{1}{\omega_1(\omega_2-a_2^1\omega_1)}+\tfrac{a_2^{11}(\omega_1-\hat\omega_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat\omega_1(\omega_1-\hat\omega_1))}\big)\tfrac{1}{\omega_3-a_3^2\omega_2-a_3^1\omega_1-\sum_{i,j=1,2}a_3^{ij}\hat\omega_i(\omega_j-\hat\omega_j)}\\
&\times\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}+O_p(\tfrac{1}{n}).
\end{aligned}\tag{7.2}$$

Proof. We will prove the theorem by two steps. To save some notation, let $B(\omega_1,\omega_2)=\frac{1}{\omega_1(\omega_2-a_2^1\omega_1)}+\frac{a_2^{11}(\omega_1-\hat\omega_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat\omega_1(\omega_1-\hat\omega_1))}$. Step 1: To show replacing

$$\frac{1}{\omega_1[\omega_2-a_2^1\omega_1-a_2^{11}\omega_1(\omega_1-\hat\omega_1)]}$$

by $B(\omega_1,\omega_2)$, with the third denominator held unchanged, will induce an error of $O_p(\frac{1}{n})$. Step 2: Based on the result of step 1, to show replacing $\sum_{i,j=1,2}a_3^{ij}\omega_i(\omega_j-\hat\omega_j)$ by $\sum_{i,j=1,2}a_3^{ij}\hat\omega_i(\omega_j-\hat\omega_j)$ in the third denominator of the integral with the replacement of $B(\omega_1,\omega_2)$ will induce an error of $O_p(\frac{1}{n})$.

Write (7.1) as

$$\begin{aligned}
&\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2}{2}-\hat\omega_1\omega_1-\hat\omega_2\omega_2])\tfrac{1}{(\omega_1-\tilde\omega_1)(\omega_2-\tilde\omega_2(\omega_1))}\\
&\times\big[\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\tfrac{\exp(n[\omega_3^2/2-\hat\omega_3\omega_3])}{\omega_3-\tilde\omega_3(\omega_1,\omega_2)}\tfrac{d\omega_3}{2\pi i}\big]\tfrac{d\omega_1 d\omega_2}{(2\pi i)^2}.
\end{aligned}\tag{7.3}$$

The inner integral with respect to ω_3 is a function of (ω_1,ω_2). Let $h(\omega_1,\omega_2)=\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\frac{\exp(n[\omega_3^2/2-\hat\omega_3\omega_3])}{\omega_3-\tilde\omega_3(\omega_1,\omega_2)}\frac{d\omega_3}{2\pi i}$, and for each fixed (ω_1,ω_2), reparameterize the inner integral with $\{\omega_3-\tilde\omega_3(\omega_1,\omega_2)\to\omega_3^*\}$ to give

$$\begin{aligned}
&h(\omega_1,\omega_2)\\
&=e^{\tilde\omega_3(\omega_1,\omega_2)^2/2-\hat\omega_3\tilde\omega_3(\omega_1,\omega_2)}\int_{\hat\omega_3^*-i\infty}^{\hat\omega_3^*+i\infty}\tfrac{\exp(n[\tilde\omega_3(\omega_1,\omega_2)^2/2-(\hat\omega_3-\tilde\omega_3(\omega_1,\omega_2))\omega_3^*])}{\omega_3^*}\tfrac{d\omega_3^*}{2\pi i}\\
&=e^{\tilde\omega_3(\omega_1,\omega_2)^2/2-\hat\omega_3\tilde\omega_3(\omega_1,\omega_2)}\bar\Phi(\hat\omega_3-\tilde\omega_3(\omega_1,\omega_2)).
\end{aligned}\tag{7.4}$$

Note that $\bar{\Phi}(\hat{\omega}_3 - \tilde{\omega}_3(\omega_1, \omega_2)) = O_p(1)$ and substituting $h(\omega_1, \omega_2)$ to (7.3) gives

$$\begin{aligned}&\int_{\hat{\omega}_1 - i\infty}^{\hat{\omega}_1 + i\infty} \int_{\hat{\omega}_2 - i\infty}^{\hat{\omega}_2 + i\infty} \exp(n[\tfrac{\omega_1^2 + \omega_2^2 + \tilde{\omega}_3^2}{2} - \hat{\omega}_1\omega_1 - \hat{\omega}_2\omega_2 - \hat{\omega}_3\tilde{\omega}_3]) \tfrac{1}{(\omega_1 - \tilde{\omega}_1)(\omega_2 - \tilde{\omega}_2(\omega_1))} \\ &\times O_p(1) \tfrac{d\omega_1 d\omega_2}{(2\pi i)^2}.\end{aligned} \tag{7.5}$$

Since $\tilde{\omega}_3(\omega_1, \omega_2)$ satisfies $\xi_3(\omega_1, \omega_2, \tilde{\omega}_3) = 0$, and by definitions of $(\omega_i - \hat{\omega}_i)^2/2$ and $\hat{\omega}_i^2/2$ for $i = 1, 2, 3$

$$\begin{aligned}&\tfrac{\omega_1^2 + \omega_2^2 + \tilde{\omega}_3^2}{2} - \hat{\omega}_1\omega_1 - \hat{\omega}_2\omega_2 - \hat{\omega}_3\tilde{\omega}_3 \\ &= (\omega_1 - \hat{\omega}_1)^2/2 - \hat{\omega}_1^2/2 + (\omega_2 - \hat{\omega}_2)^2/2 - \hat{\omega}_2^2/2 + (\tilde{\omega}_3 - \hat{\omega}_3)^2/2 - \hat{\omega}_3^2/2 \\ &= K_Y(\xi_1, \xi_2, 0) - \xi_1\bar{t}_1 - \xi_2\bar{t}_2,\end{aligned} \tag{7.6}$$

therefore (7.5) is equivalent to the bivariate case for which we have proved in Theorem 5.3 in Section 2 that replacing $\frac{1}{\omega_1[\omega_2 - a_2^1\omega_1 - a_2^{11}\omega_1(\omega_1 - \hat{\omega}_1)]}$ by $B(\omega_1, \omega_2)$ induces an error of $O_p(\frac{1}{n})$. Hence Step 1 is proved. To prove Step 2, let $r(\omega_1, \omega_2, \omega_3) = \omega_3 - a_3^2\omega_2 - a_3^1\omega_1 - \sum_{i,j} a_3^{ij}\hat{\omega}_i(\omega_j - \hat{\omega}_j)$, and write the third denominator as

$$\begin{aligned}&\tfrac{1}{\omega_3 - a_3^2\omega_2 - a_3^1\omega_1 - \sum_{i,j=1,2} a_3^{ij}\omega_i(\omega_j - \hat{\omega}_j)} \\ &= \tfrac{1}{\omega_3 - a_3^2\omega_2 - a_3^1\omega_1 - \sum_{i,j} a_3^{ij}\hat{\omega}_i(\omega_j - \hat{\omega}_j) - \sum_{i,j} a_3^{ij}(\omega_i - \hat{\omega}_i)(\omega_j - \hat{\omega}_j)} \\ &= \tfrac{1}{\omega_3 - a_3^2\omega_2 - a_3^1\omega_1 - \sum_{i,j} a_3^{ij}\hat{\omega}_i(\omega_j - \hat{\omega}_j)} + \tfrac{\sum_{i,j} a_3^{ij}(\omega_i - \hat{\omega}_i)(\omega_j - \hat{\omega}_j)}{r(r - \sum_{i,j} a_3^{ij}(\omega_i - \hat{\omega}_i)(\omega_j - \hat{\omega}_j))}.\end{aligned} \tag{7.7}$$

Substituting (7.7) to (7.1), with the replacement of $a_2^{11}\hat{\omega}_1$ gives

$$\begin{aligned}
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\
&\times\Big(\tfrac{1}{\omega_1(\omega_2-a_2^1\omega_1)}+\tfrac{a_2^{11}(\omega_1-\hat{\omega}_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1))}\Big)\tfrac{1}{\omega_3-a_3^2\omega_2-a_3^1\omega_1-\sum_{i,j=1,2}a_3^{ij}\hat{\omega}_i(\omega_j-\hat{\omega}_j)}\\
&\times\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}\\
&+\\
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\
&\times\Big(\tfrac{1}{\omega_1(\omega_2-a_2^1\omega_1)}+\tfrac{a_2^{11}(\omega_1-\hat{\omega}_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1))}\Big)\tfrac{1}{r(r-\sum_{i,j}a_3^{ij}(\omega_i-\hat{\omega}_i)(\omega_j-\hat{\omega}_j))}\\
&\times\textstyle\sum_{i,j}a_3^{ij}(\omega_i-\hat{\omega}_i)(\omega_j-\hat{\omega}_j)\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}.
\end{aligned}\tag{7.8}$$

To show the second integral of (7.8) is $O_p(\frac{1}{n})$, we use the same arguments as in Section 2 by letting $(r^*)^2 = r(r-\sum_{i,j}a_3^{ij}(\omega_i-\hat{\omega}_i)(\omega_j-\hat{\omega}_j))$, Theorem 5.1 shows that for all fixed (ω_1,ω_2), the univariate integral of $\frac{1}{(r^*)^2}$ as a function of ω_3 is bounded by $\sqrt{n}$. And since

$$\begin{aligned}
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\
&\times\Big(\tfrac{1}{\omega_1(\omega_2-a_2^1\omega_1)}+\tfrac{a_2^{11}(\omega_1-\hat{\omega}_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1))}\Big)\textstyle\sum_{i,j=1,2}a_3^{ij}(\omega_i-\hat{\omega}_i)(\omega_j-\hat{\omega}_j)\\
&\times\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}\\
&=\Big[\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\omega_3^2/2-\hat{\omega}_3\omega_3])\tfrac{d\omega_3}{2\pi i}\Big]\Big[\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2])\\
&\times B(\omega_1,\omega_2)\textstyle\sum_{i,j=1,2}a_3^{ij}(\omega_i-\hat{\omega}_i)(\omega_j-\hat{\omega}_j)\tfrac{d\omega_1 d\omega_2}{(2\pi i)^2}\Big]\\
&=\tfrac{1}{\sqrt{n}}e^{-\frac{1}{2}n\hat{\omega}_3^2}\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2])B(\omega_1,\omega_2)\\
&\times\textstyle\sum_{i,j=1,2}a_3^{ij}(\omega_i-\hat{\omega}_i)(\omega_j-\hat{\omega}_j)\tfrac{d\omega_1 d\omega_2}{(2\pi i)^2},
\end{aligned}\tag{7.9}$$

the leading factor of $\frac{1}{\sqrt{n}}$ from (7.9) cancels out $\sqrt{n}$ contributed by the integral of $\frac{1}{(r^*)^2}$. Note that $\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\exp(n[\frac{\omega_1^2+\omega_2^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2])B(\omega_1,\omega_2)\frac{d\omega_1 d\omega_2}{(2\pi i)^2}$ is $O_p(1)$ as shown in Section 2. Make the change of variables $\omega_j=\hat{\omega}_j+i\frac{z_j}{\sqrt{n}}$ for $j=1,2$. The two terms of $\omega_1-\hat{\omega}_1,\omega_2-\hat{\omega}_2$ contribute a factor of $\frac{1}{n}$, hence the second integral of (7.8) is bounded by $\frac{1}{n}$. Step 2 is proved. $\square$

We will now evaluate the integral of the RHS of (7.2). Splitting the third denominator to two terms with the first term containing only linearity of ω's in the denominator gives

$$\begin{aligned}&\frac{1}{\omega_3-a_3^2\omega_2-a_3^1\omega_1-\sum_{i,j=1,2}a_3^{ij}\hat{\omega}_i(\omega_j-\hat{\omega}_j)}\\&=\frac{1}{\omega_3-a_3^2\omega_2-a_3^1\omega_1}+\frac{\sum_{i,j=1,2}a_3^{ij}\hat{\omega}_i(\omega_j-\hat{\omega}_j)}{(\omega_3-a_3^2\omega_2-a_3^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1-\sum_{i,j=1,2}a_3^{ij}\hat{\omega}_i(\omega_j-\hat{\omega}_j))}.\end{aligned}\tag{7.10}$$

To save some notation, let $\omega_3^\dagger(\omega_1,\omega_2)=\sum_{i,j=1,2}a_3^{ij}\hat{\omega}_i(\omega_j-\hat{\omega}_j)$, and substitute (7.10) to the RHS of (7.2) to give

$$\begin{aligned}&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\&\times\Big(\frac{1}{\omega_1(\omega_2-a_2^1\omega_1)}+\frac{a_2^{11}(\omega_1-\hat{\omega}_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1))}\Big)\frac{1}{\omega_3-a_3^2\omega_2-a_3^1\omega_1-\omega_3^\dagger(\omega_1,\omega_2)}\\&\times\frac{d\omega_1d\omega_2d\omega_3}{(2\pi i)^3}\\&=\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\&\times\frac{1}{\omega_1(\omega_2-a_2^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1)}\frac{d\omega_1d\omega_2d\omega_3}{(2\pi i)^3}\\&+\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\&\times\frac{\omega_3^\dagger(\omega_1,\omega_2)}{\omega_1(\omega_2-a_2^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1-\omega_3^\dagger(\omega_1,\omega_2))}\frac{d\omega_1d\omega_2d\omega_3}{(2\pi i)^3}\\&+\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\&\times\frac{a_2^{11}(\omega_1-\hat{\omega}_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1))(\omega_3-a_3^2\omega_2-a_3^1\omega_1)}\frac{d\omega_1d\omega_2d\omega_3}{(2\pi i)^3}\\&+\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\&\times\frac{a_2^{11}(\omega_1-\hat{\omega}_1)\omega_3^\dagger(\omega_1,\omega_2)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1))(\omega_3-a_3^2\omega_2-a_3^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1-\omega_3^\dagger(\omega_1,\omega_2))}\\&\times\frac{d\omega_1d\omega_2d\omega_3}{(2\pi i)^3}.\end{aligned}\tag{7.11}$$

We will evaluate the RHS of (7.11) term by term and the results are presented in Lemma 7.1 to 7.4, respectively.

Lemma 7.1.

$$\begin{aligned}
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\
&\times\tfrac{1}{\omega_1(\omega_2-a_2^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1)}\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}\\
&=\bar{\Phi}_3((A_1^{-T}\underset{\sim}{\hat{\omega}})^T,\Sigma_1),
\end{aligned} \tag{7.12}$$

where $\bar{\Phi}_3((A^{-T}\underset{\sim}{\hat{\omega}})^T,\Sigma)$ *is the trivariate normal survival function beyond point* $(A^{-T}\underset{\sim}{\hat{\omega}})^T$ *of a trivariate normal distribution of mean* $\underset{\sim}{0}$ *and variance* $\Sigma=(A^TA)^{-1}$*, and*

$$A=\begin{pmatrix}1 & 0 & 0\\ -a_2^1 & 1 & 0\\ -a_3^1 & -a_3^2 & 1\end{pmatrix}. \tag{7.13}$$

Proof. Reparameterize the integral of LHS of (7.12) such that the denominator of the integrand is the product of three variables. Let $\underset{\sim}{z}=(z_1,z_2,z_3)^T$, $\underset{\sim}{\omega}=(\omega_1,\omega_2,\omega_3)^T$. Make the change of variables by letting $\{\omega_1\to z_1,\omega_2-a_2^1\omega_1\to z_2,\omega_3-a_3^2\omega_2-a_3^1\omega_1\to z_3\}$, or, in matrix form of $\underset{\sim}{z}=A\underset{\sim}{\omega}$ to give

$$\begin{aligned}
&\int_{\hat{z}_1-i\infty}^{\hat{z}_1+i\infty}\int_{\hat{z}_2-i\infty}^{\hat{z}_2+i\infty}\int_{\hat{z}_3-i\infty}^{\hat{z}_3+i\infty}\exp(n[\tfrac{1}{2}\underset{\sim}{z}^T(A^TA)^{-1}\underset{\sim}{z}-(A^{-T}\underset{\sim}{\hat{\omega}})^T\underset{\sim}{z}])\\
&\times\tfrac{1}{z_1z_2z_3}\tfrac{dz_1dz_2dz_3}{(2\pi i)^3}\\
&=\bar{\Phi}_3((A^{-T}\underset{\sim}{\hat{\omega}})^T,\Sigma),
\end{aligned} \tag{7.14}$$

where $\Sigma=(A^TA)^{-1}$, $A^{-T}=(A^{-1})^T$, and $\bar{\Phi}_3((A^{-T}\underset{\sim}{\hat{\omega}})^T,\Sigma)$ is the trivariate normal survival function beyond point $(A^{-T}\underset{\sim}{\hat{\omega}})^T$ of a trivariate normal distribution of mean $\underset{\sim}{0}$ and variance Σ. □

Lemma 7.2.

$$\begin{aligned}
&\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat\omega_1\omega_1-\hat\omega_2\omega_2-\hat\omega_3\omega_3])\\
&\times\tfrac{\omega_3^\dagger(\omega_1,\omega_2)}{\omega_1(\omega_2-a_2^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1-\omega_3^\dagger(\omega_1,\omega_2))}\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}\\
&=\\
&-\textstyle\sum_{j=1,2}(a_3^{1j}+a_2^1a_3^{2j})e^{-\frac{1}{2}n(\underset{\sim}{c}^T\Sigma_1\underset{\sim}{c}+\underset{\sim}{\hat\omega}^TA_1^{-1}\underset{\sim}{c})}|A_1|\times\\
&\left[(\underset{\sim}{e}_j^TA_1^{-1}\underset{\sim}{c})\hat\omega_j\bar\Phi_3(\underset{\sim}{c}^T\Sigma_1+\underset{\sim}{\hat\omega}^TA_1^{-1},\Sigma_1)+\tfrac{1}{n}\tfrac{\partial}{\partial t_j}\bar\Phi_3(\underset{\sim}{c}^T\Sigma_1+\underset{\sim}{t}^TA_1^{-1},\Sigma_1)|_{\underset{\sim}{\hat\omega}}\right]\\
&-\textstyle\sum_{j=1,2}a_3^{2j}e^{-\frac{1}{2}n(\underset{\sim}{c}^T\Sigma_2\underset{\sim}{c}+\underset{\sim}{\hat\omega}^TA_2^{-1}\underset{\sim}{c})}|A_2|\times\\
&\left[(\underset{\sim}{e}_j^TA_2^{-1}\underset{\sim}{c})\hat\omega_j\bar\Phi_3(\underset{\sim}{c}^T\Sigma_2+\underset{\sim}{\hat\omega}^TA_2^{-1},\Sigma_2)+\tfrac{1}{n}\tfrac{\partial}{\partial t_j}\bar\Phi_3(\underset{\sim}{c}^T\Sigma_2+\underset{\sim}{t}^TA_2^{-1},\Sigma_2)|_{\underset{\sim}{\hat\omega}}\right]\\
&+O_p(\tfrac{1}{n}),
\end{aligned}\tag{7.15}$$

where $\underset{\sim}{e}_j$ *is the unit vector of the* j^{th} *element being 1 and all others being 0,* $\underset{\sim}{c}=(0,0,\sum_{i,j=1,2}a_3^{ij}\hat\omega_i\hat\omega_j)^T$,

$$A_1=\begin{pmatrix}-a_2^1 & 1 & 0\\ -a_3^1 & -a_3^2 & 1\\ -a_3^1-a_3^{11}\hat\omega_1-a_3^{12}\hat\omega_2 & -a_3^2-a_3^{12}\hat\omega_1-a_3^{22}\hat\omega_2 & 1\end{pmatrix},\tag{7.16}$$

$$A_2=\begin{pmatrix}1 & 0 & 0\\ -a_3^1 & -a_3^2 & 1\\ -a_3^1-a_3^{11}\hat\omega_1-a_3^{12}\hat\omega_2 & -a_3^2-a_3^{12}\hat\omega_1-a_3^{22}\hat\omega_2 & 1\end{pmatrix}\tag{7.17}$$

and $\Sigma_i=(A_i^TA_i)^{-1}$ *for* $i=1,2$.

Proof. We will first show that replacing $a_3^{1j}\hat\omega_1$ by $a_3^{1j}\omega_j$ in the numerator of the integrand will induce an error of $O_p(\frac{1}{n})$. The error after replacing $a_3^{1j}\hat\omega_1$ by $a_3^{1j}\omega_j$ is

$$\begin{aligned}
&\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat\omega_1\omega_1-\hat\omega_2\omega_2-\hat\omega_3\omega_3])\\
&\times\tfrac{\sum_j a_3^{1j}(\omega_1-\hat\omega_1)(\omega_j-\hat\omega_j)}{\omega_1(\omega_2-a_2^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1-\sum_{i,j=1,2}a_3^{ij}\hat\omega_i(\omega_j-\hat\omega_j))}\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}.
\end{aligned}\tag{7.18}$$

Using the same arguments as in Theorem 7.1, let $(r^*)^2 = r(r - \sum_{i,j} a_3^{ij}(\omega_i - \hat{\omega}_i)(\omega_j - \hat{\omega}_j))$. By Theorem 5.1, for all fixed (ω_1, ω_2), the univariate integral of $\frac{1}{(r^*)^2}$ as a function of ω_3 is bounded by $\sqrt{n}$, which cancels out the factor $\frac{1}{\sqrt{n}}$ contributed by the integral $\int_{\hat{\omega}_3 - i\infty}^{\hat{\omega}_3 + i\infty} \exp(n[\omega_3^2/2 - \hat{\omega}_3\omega_3])\frac{d\omega_3}{2\pi i}$. Note that $\int_{\hat{\omega}_1 - i\infty}^{\hat{\omega}_1 + i\infty}\int_{\hat{\omega}_2 - i\infty}^{\hat{\omega}_2 + i\infty} \exp(n[\frac{\omega_1^2+\omega_2^2}{2} - \hat{\omega}_1\omega_1 - \hat{\omega}_2\omega_2])\frac{1}{\omega_1(\omega_2 - a_2^1\omega_1)}\frac{d\omega_1 d\omega_2}{(2\pi i)^2} = O_p(1)$, and make the change of variables $\omega_j = \hat{\omega}_j + i\frac{z_j}{\sqrt{n}}$ for $j = 1, 2$, the two terms of $\omega_1 - \hat{\omega}_1, \omega_j - \hat{\omega}_j$ contribute a factor of $\frac{1}{n}$, hence (7.18) is bounded by $\frac{1}{n}$. Similarly the error of replacing $a_3^{2j}\hat{\omega}_2$ by $a_3^{2j}\omega_2$ in the numerator of the integrand will also induce an error of $O_p(\frac{1}{n})$. We can then cancel out ω_1 or $\omega_2 - a_2^1\omega_1$ (by letting $\omega_2 = \omega_2 - a_2^1\omega_1 + a_2^1\omega_1$ in the numerator) from the denominator; hence to the order of $O_p(\frac{1}{n})$, the RHS of (7.15) becomes

$$
\begin{aligned}
&\int_{\hat{\omega}_1 - i\infty}^{\hat{\omega}_1 + i\infty}\int_{\hat{\omega}_2 - i\infty}^{\hat{\omega}_2 + i\infty}\int_{\hat{\omega}_3 - i\infty}^{\hat{\omega}_3 + i\infty} \exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2} - \hat{\omega}_1\omega_1 - \hat{\omega}_2\omega_2 - \hat{\omega}_3\omega_3]) \\
&\times \frac{\sum_{j=1,2}(a_3^{1j} + a_2^1 a_3^{2j})(\omega_j - \hat{\omega}_j)}{(\omega_2 - a_2^1\omega_1)(\omega_3 - a_3^2\omega_2 - a_3^1\omega_1)(\omega_3 - a_3^2\omega_2 - a_3^1\omega_1 - \omega_3^\dagger(\omega_1,\omega_2))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} \\
&+ \int_{\hat{\omega}_1 - i\infty}^{\hat{\omega}_1 + i\infty}\int_{\hat{\omega}_2 - i\infty}^{\hat{\omega}_2 + i\infty}\int_{\hat{\omega}_3 - i\infty}^{\hat{\omega}_3 + i\infty} \exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2} - \hat{\omega}_1\omega_1 - \hat{\omega}_2\omega_2 - \hat{\omega}_3\omega_3]) \\
&\times \frac{\sum_{j=1,2} a_3^{2j}(\omega_j - \hat{\omega}_j)}{\omega_1(\omega_3 - a_3^2\omega_2 - a_3^1\omega_1)(\omega_3 - a_3^2\omega_2 - a_3^1\omega_1 - \omega_3^\dagger(\omega_1,\omega_2))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} \\
&+ O_p(\tfrac{1}{n}).
\end{aligned}
\tag{7.19}
$$

In order to express (7.19) in terms of normal functions, we need to further remove $\omega_j - \hat{\omega}_j$ from the numerators. This can be done by defining a function $S_j(t_1, t_2, t_3)$ for $j = 1, 2$ as

$$
\begin{aligned}
S_j(t_1,t_2,t_3) &= -\tfrac{1}{n}(a_3^{1j} + a_2^1 a_3^{2j})e^{-\frac{1}{2}n\sum_{i=1}^3 \hat{\omega}_i^2}\int_{\hat{\omega}_1 - i\infty}^{\hat{\omega}_1 + i\infty}\int_{\hat{\omega}_2 - i\infty}^{\hat{\omega}_2 + i\infty}\int_{\hat{\omega}_3 - i\infty}^{\hat{\omega}_3 + i\infty} \\
&\times \frac{\exp(n[(\omega_1 - t_1)^2 + (\omega_2 - t_2)^2 + (\omega_3 - t_3)^2]/2)}{(\omega_2 - a_2^1\omega_1)(\omega_3 - a_3^2\omega_2 - a_3^1\omega_1)(\omega_3 - a_3^2\omega_2 - a_3^1\omega_1 - \omega_3^\dagger(\omega_1,\omega_2))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} \\
&= -\tfrac{1}{n}(a_3^{1j} + a_2^1 a_3^{2j})e^{-\frac{1}{2}n\sum_{i=1}^3(\hat{\omega}_i^2 - t_i^2)}\int_{\hat{\omega}_1 - i\infty}^{\hat{\omega}_1 + i\infty}\int_{\hat{\omega}_2 - i\infty}^{\hat{\omega}_2 + i\infty}\int_{\hat{\omega}_3 - i\infty}^{\hat{\omega}_3 + i\infty} \\
&\times \frac{\exp(n[\frac{\omega_1^2+\omega_2^2+\omega_3^2}{2} - t_1\omega_1 - t_2\omega_2 - t_3\omega_3])}{(\omega_2 - a_2^1\omega_1)(\omega_3 - a_3^2\omega_2 - a_3^1\omega_1)(\omega_3 - a_3^2\omega_2 - a_3^1\omega_1 - \omega_3^\dagger(\omega_1,\omega_2))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}.
\end{aligned}
\tag{7.20}
$$

The first term of (7.19) is $\sum_{j=1,2}\frac{\partial}{\partial t_j}S_j(t_1,t_2,t_3)|_{(t_1,t_2,t_3)=(\hat{\omega}_1,\hat{\omega}_2,\hat{\omega}_3)}$.

Reparameterize (7.20) by letting $\underset{\sim}{z} = A_1\underset{\sim}{w} + \underset{\sim}{c}$ for $j = 1, 2$ to give

$$\begin{aligned} &S_j(t_1, t_2, t_3) = -\tfrac{1}{n}(a_3^{1j} + a_2^1 a_3^{2j})e^{-\frac{1}{2}n(\hat{\underset{\sim}{\omega}}^T\hat{\underset{\sim}{\omega}} - \underset{\sim}{t}^T\underset{\sim}{t} + \underset{\sim}{c}^T\Sigma_1\underset{\sim}{c} + \underset{\sim}{t}^T A_1^{-1}\underset{\sim}{c})}|A_1| \\ &\times\bar{\Phi}_3(\underset{\sim}{c}^T\Sigma_1 + \underset{\sim}{t}^T A_1^{-1}, \Sigma_1), \end{aligned} \tag{7.21}$$

where $\Sigma_1 = (A_1^T A_1)^{-1}$. Let $\underset{\sim}{e}_j$ denotes the unit vector with the j^{th} element being 1 and all others being 0, the first term of (7.19) is

$$\begin{aligned} &-\textstyle\sum_{j=1,2}(a_3^{1j} + a_2^1 a_3^{2j})e^{-\frac{1}{2}n(\underset{\sim}{c}^T\Sigma_1\underset{\sim}{c} + \hat{\underset{\sim}{\omega}}^T A_1^{-1}\underset{\sim}{c})}|A_1|\times \\ &\big[(\underset{\sim}{e}_j^T A_1^{-1}\underset{\sim}{c})\hat{\omega}_j\bar{\Phi}_3(\underset{\sim}{c}^T\Sigma_1 + \hat{\underset{\sim}{\omega}}^T A_1^{-1}, \Sigma_1) + \tfrac{1}{n}\tfrac{\partial}{\partial t_j}\bar{\Phi}_3(\underset{\sim}{c}^T\Sigma_1 + \underset{\sim}{t}^T A_1^{-1}, \Sigma_1)|_{\hat{\underset{\sim}{\omega}}}\big]. \end{aligned} \tag{7.22}$$

Similarly, the second term of (7.23) is

$$\begin{aligned} &-\textstyle\sum_{j=1,2} a_3^{2j} e^{-\frac{1}{2}n(\underset{\sim}{c}^T\Sigma_2\underset{\sim}{c} + \hat{\underset{\sim}{\omega}}^T A_2^{-1}\underset{\sim}{c})}|A_2|\times \\ &\big[(\underset{\sim}{e}_j^T A_2^{-1}\underset{\sim}{c})\hat{\omega}_j\bar{\Phi}_3(\underset{\sim}{c}^T\Sigma_2 + \hat{\underset{\sim}{\omega}}^T A_2^{-1}, \Sigma_2) + \tfrac{1}{n}\tfrac{\partial}{\partial t_j}\bar{\Phi}_3(\underset{\sim}{c}^T\Sigma_2 + \underset{\sim}{t}^T A_2^{-1}, \Sigma_2)|_{\hat{\underset{\sim}{\omega}}}\big]. \end{aligned} \tag{7.23}$$

□

Lemma 7.3.

$$\begin{aligned} &\textstyle\int_{\hat{\omega}_1 - i\infty}^{\hat{\omega}_1 + i\infty}\int_{\hat{\omega}_2 - i\infty}^{\hat{\omega}_2 + i\infty}\int_{\hat{\omega}_3 - i\infty}^{\hat{\omega}_3 + i\infty}\exp(n[\frac{\omega_1^2 + \omega_2^2 + \omega_3^2}{2} - \hat{\omega}_1\omega_1 - \hat{\omega}_2\omega_2 - \hat{\omega}_3\omega_3]) \\ &\times\tfrac{a_2^{11}(\omega_1 - \hat{\omega}_1)}{(\omega_2 - a_2^1\omega_1)(\omega_2 - a_2^1\omega_1 - a_2^{11}\hat{\omega}_1(\omega_1 - \hat{\omega}_1))(\omega_3 - a_3^2\omega_2 - a_3^1\omega_1)}\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} \\ &= -a_2^{11}e^{-\frac{1}{2}n(\underset{\sim}{c}^T\Sigma\underset{\sim}{c} + \hat{\underset{\sim}{\omega}}^T A^{-1}\underset{\sim}{c})}\times \\ &\big[(\underset{\sim}{e}_1^T A^{-1}\underset{\sim}{c})\hat{\omega}_1\bar{\Phi}_3(\underset{\sim}{c}^T\Sigma + \hat{\underset{\sim}{\omega}}^T A^{-1}, \Sigma) + \tfrac{1}{n}\tfrac{\partial}{\partial t_1}\bar{\Phi}_3(\underset{\sim}{c}^T\Sigma + \underset{\sim}{t}^T A^{-1}, \Sigma)|_{\hat{\underset{\sim}{\omega}}}\big], \end{aligned} \tag{7.24}$$

where

$$A = \begin{pmatrix} 1 & 0 & 0 \\ -(a_2^1 + a_2^{11}\hat{\omega}_1) & 1 & 0 \\ -a_3^1 & -a_3^2 & 1 \end{pmatrix}, \tag{7.25}$$

and $\underset{\sim}{c} = (0, a_2^{11}\hat{\omega}_1^2, 0)^T$.

Proof. Let

$$\begin{aligned} S(t_1,t_2,t_3) &= -\tfrac{1}{n}a_2^{11}e^{-\frac{1}{2}n\sum_{i=1}^{3}\hat{\omega}_i^2}\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty} \\ &\times \tfrac{\exp(n[(\omega_1-t_1)^2+(\omega_2-t_2)^2+(\omega_3-t_3)^2]/2)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1))(\omega_3-a_3^2\omega_2-a_3^1\omega_1)}\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} \\ &= -\tfrac{1}{n}a_2^{11}e^{-\frac{1}{2}n\sum_{i=1}^{3}(\hat{\omega}_i^2-t_i^2)}\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty} \\ &\times \tfrac{\exp(n[\frac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-t_1\omega_1-t_2\omega_2-t_3\omega_3])}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1))(\omega_3-a_3^2\omega_2-a_3^1\omega_1)}\tfrac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} \end{aligned} \tag{7.26}$$

such that the RHS of (7.24) is $\frac{\partial}{\partial t_1}S(t_1,t_2,t_3)|_{(t_1,t_2,t_3)=(\hat{\omega}_1,\hat{\omega}_2,\hat{\omega}_3)}$.

Reparameterize (7.26) by letting $\underset{\sim}{z} = A\underset{\sim}{w} + \underset{\sim}{c}$ for

$$A = \begin{pmatrix} 1 & 0 & 0 \\ -(a_2^1 + a_2^{11}\hat{\omega}_1) & 1 & 0 \\ -a_3^1 & -a_3^2 & 1 \end{pmatrix}, \tag{7.27}$$

and $\underset{\sim}{c} = (0, a_2^{11}\hat{\omega}_1^2, 0)^T$ to give

$$S(t_1,t_2,t_3) = -\tfrac{1}{n}a_2^{11}e^{-\frac{1}{2}n(\underset{\sim}{\hat{\omega}}^T\underset{\sim}{\hat{\omega}}-\underset{\sim}{t}^T\underset{\sim}{t}+\underset{\sim}{c}^T\Sigma\underset{\sim}{c}+\underset{\sim}{t}^T A^{-1}\underset{\sim}{c})}\bar{\Phi}_3(\underset{\sim}{c}^T\Sigma + \underset{\sim}{t}^T A^{-1}, \Sigma), \tag{7.28}$$

where $\Sigma = (A^T A)^{-1}$. Let $\underset{\sim}{e}_1 = (1,0,0)^T$, the RHS of (7.24) is

$$\begin{aligned} &-a_2^{11}e^{-\frac{1}{2}n(\underset{\sim}{c}^T\Sigma\underset{\sim}{c}+\underset{\sim}{\hat{\omega}}^T A^{-1}\underset{\sim}{c})}\times \\ &\left[(\underset{\sim}{e}_1^T A^{-1}\underset{\sim}{c})\hat{\omega}_1\bar{\Phi}_3(\underset{\sim}{c}^T\Sigma + \underset{\sim}{\hat{\omega}}^T A^{-1}, \Sigma) + \tfrac{1}{n}\tfrac{\partial}{\partial t_1}\bar{\Phi}_3(\underset{\sim}{c}^T\Sigma + \underset{\sim}{t}^T A^{-1}, \Sigma)|_{\underset{\sim}{\hat{\omega}}}\right]. \end{aligned} \tag{7.29}$$

□

Lemma 7.4.

$$\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat\omega_1\omega_1-\hat\omega_2\omega_2-\hat\omega_3\omega_3])$$
$$\times\frac{a_2^{11}(\omega_1-\hat\omega_1)\omega_3^\dagger(\omega_1,\omega_2)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat\omega_1(\omega_1-\hat\omega_1))(\omega_3-a_3^2\omega_2-a_3^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1-\omega_3^\dagger(\omega_1,\omega_2))}$$
$$\times\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}$$
$$=\sum_{j=1,2}(a_3^{1j}\hat\omega_1+a_3^{2j}\hat\omega_2)e^{-\frac{1}{2}n(\underset{\sim}{c_1}^T\Sigma_1\underset{\sim}{c_1}+\underset{\sim}{\hat\omega}^TA_1^{-1}\underset{\sim}{c_1})}|A_1|\times \tag{7.30}$$
$$\big[(\underset{\sim}{e_j}^TA_1^{-1}\underset{\sim}{c_1})\hat\omega_j\bar\Phi_3(\underset{\sim}{c_1}^T\Sigma_1+\underset{\sim}{\hat\omega}^TA_1^{-1},\Sigma_1)+\tfrac{1}{n}\tfrac{\partial}{\partial t_j}\bar\Phi_3(\underset{\sim}{c_1}^T\Sigma_1+\underset{\sim}{t}^TA_1^{-1},\Sigma_1)|_{\underset{\sim}{\hat\omega}}\big]$$
$$-\sum_{j=1,2}(a_3^{1j}\hat\omega_1+a_3^{2j}\hat\omega_2)e^{-\frac{1}{2}n(\underset{\sim}{c_2}^T\Sigma_2\underset{\sim}{c_2}+\underset{\sim}{\hat\omega}^TA_2^{-1}\underset{\sim}{c_2})}|A_2|\times$$
$$\big[(\underset{\sim}{e_j}^TA_2^{-1}\underset{\sim}{c_2})\hat\omega_j\bar\Phi_3(\underset{\sim}{c_2}^T\Sigma_2+\underset{\sim}{\hat\omega}^TA_2^{-1},\Sigma_2)+\tfrac{1}{n}\tfrac{\partial}{\partial t_j}\bar\Phi_3(\underset{\sim}{c_2}^T\Sigma_2+\underset{\sim}{t}^TA_2^{-1},\Sigma_2)|_{\underset{\sim}{\hat\omega}}\big],$$

where $\underset{\sim}{e_i}$ *is the unit vector of the* i^{th} *element being 1 and all others being 0,* $\underset{\sim}{c_1}=(0,0,\sum_{i,j=1,2}a_3^{ij}\hat\omega_i\hat\omega_j)^T$, $\underset{\sim}{c_2}=(0,a_2^{11}\hat\omega_1^2,\sum_{i,j=1,2}a_3^{ij}\hat\omega_i\hat\omega_j)^T$,

$$A_1=\begin{pmatrix}-a_2^1 & 1 & 0\\ -a_3^1 & -a_3^2 & 1\\ -a_3^1-a_3^{11}\hat\omega_1-a_3^{12}\hat\omega_2 & -a_3^2-a_3^{12}\hat\omega_1-a_3^{22}\hat\omega_2 & 1\end{pmatrix}, \tag{7.31}$$

$$A_2=\begin{pmatrix}-(a_2^1+a_2^{11}\hat\omega_1) & 1 & 0\\ -a_3^1 & -a_3^2 & 1\\ -a_3^1-a_3^{11}\hat\omega_1-a_3^{12}\hat\omega_2 & -a_3^2-a_3^{12}\hat\omega_1-a_3^{22}\hat\omega_2 & 1\end{pmatrix}, \tag{7.32}$$

and $\Sigma_i=(A_i^TA_i)^{-1}$ *for* $i=1,2$.

Proof. Substitute $\frac{a_2^{11}(\omega_1-\hat{\omega}_1)}{(\omega_2-a_2^1\omega_1)(\omega_2-a_2^1\omega_1-a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1))} = \frac{1}{\omega_2-a_2^1\omega_1-a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1)} - \frac{1}{\omega_2-a_2^1\omega_1}$ to give

$$\begin{aligned}
&-\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\
&\times\frac{\sum_{j=1,2}(a_3^{1j}\hat{\omega}_1+a_3^{2j}\hat{\omega}_2)(\omega_j-\hat{\omega}_j)}{(\omega_2-a_2^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1-\omega_3^\dagger(\omega_1,\omega_2))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}\\
&+\\
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n[\tfrac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-\hat{\omega}_1\omega_1-\hat{\omega}_2\omega_2-\hat{\omega}_3\omega_3])\\
&\times\frac{\sum_{j=1,2}(a_3^{1j}\hat{\omega}_1+a_3^{2j}\hat{\omega}_2)(\omega_j-\hat{\omega}_j)}{(\omega_2-a_2^1\omega_1-a_2^{11}\hat{\omega}_1(\omega_1-\hat{\omega}_1))(\omega_3-a_3^2\omega_2-a_3^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1-\omega_3^\dagger(\omega_1,\omega_2))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}.
\end{aligned} \tag{7.33}$$

Let

$$\begin{aligned}
&S_j(t_1,t_2,t_3) = \tfrac{1}{n}(a_3^{1j}\hat{\omega}_1+a_3^{2j}\hat{\omega}_2)e^{-\frac{1}{2}n\sum_{i=1}^3\hat{\omega}_i^2}\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\\
&\times\frac{\exp(n[(\omega_1-t_1)^2+(\omega_2-t_2)^2+(\omega_3-t_3)^2]/2)}{(\omega_2-a_2^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1-\omega_3^\dagger(\omega_1,\omega_2))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}\\
&=\tfrac{1}{n}(a_3^{1j}\hat{\omega}_1+a_3^{2j}\hat{\omega}_2)e^{-\frac{1}{2}n\sum_{i=1}^3(\hat{\omega}_i^2-t_i^2)}\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\\
&\times\frac{\exp(n[\frac{\omega_1^2+\omega_2^2+\omega_3^2}{2}-t_1\omega_1-t_2\omega_2-t_3\omega_3])}{(\omega_2-a_2^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1)(\omega_3-a_3^2\omega_2-a_3^1\omega_1-\omega_3^\dagger(\omega_1,\omega_2))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}
\end{aligned} \tag{7.34}$$

such that the first term of (7.33) is $\sum_{j=1,2}\frac{\partial}{\partial t_j}S(t_1,t_2,t_3)|_{(t_1,t_2,t_3)=(\hat{\omega}_1,\hat{\omega}_2,\hat{\omega}_3)}$. Reparameterize (7.34) by letting $\underset{\sim}{z} = A_1\underset{\sim}{w} + \underset{\sim}{c_1}$ for

$$A_1 = \begin{pmatrix} -a_2^1 & 1 & 0 \\ -a_3^1 & -a_3^2 & 1 \\ -a_3^1 - a_3^{11}\hat{\omega}_1 - a_3^{12}\hat{\omega}_2 & -a_3^2 - a_3^{12}\hat{\omega}_1 - a_3^{22}\hat{\omega}_2 & 1 \end{pmatrix}, \tag{7.35}$$

and $\underset{\sim}{c_1} = (0, 0, \sum_{i,j=1,2} a_3^{ij}\hat{\omega}_i\hat{\omega}_j)^T$ to give

$$\begin{aligned}
&S_j(t_1,t_2,t_3) = \tfrac{1}{n}(a_3^{1j}\hat{\omega}_1+a_3^{2j}\hat{\omega}_2)e^{-\frac{1}{2}n(\underset{\sim}{\hat{\omega}}^T\underset{\sim}{\hat{\omega}}-\underset{\sim}{t}^T\underset{\sim}{t}+\underset{\sim}{c_1}^T\Sigma_1\underset{\sim}{c_1}+\underset{\sim}{t}^TA_1^{-1}\underset{\sim}{c_1})}|A_1|\\
&\times\bar{\Phi}_3(\underset{\sim}{c_1}^T\Sigma_1+\underset{\sim}{t}^TA_1^{-1},\Sigma_1),
\end{aligned} \tag{7.36}$$

where $\Sigma_1 = (A_1^TA_1)^{-1}$. Let $\underset{\sim}{e}_j$ denotes the unit vector with the j^{th} element being 1 and all others being 0, the first term of (7.33) is

$$\begin{aligned}&\textstyle\sum_{j=1,2}(a_3^{1j}\hat{\omega}_1 + a_3^{2j}\hat{\omega}_2)e^{-\frac{1}{2}n(\underset{\sim}{c_1}^T\Sigma_1\underset{\sim}{c_1}+\underset{\sim}{\hat{\omega}}^T A_1^{-1}\underset{\sim}{c_1})}|A_1|\times \\ &\big[(\underset{\sim}{e_j}^T A_1^{-1}\underset{\sim}{c_1})\hat{\omega}_j\bar{\Phi}_3(\underset{\sim}{c_1}^T\Sigma_1 + \underset{\sim}{\hat{\omega}}^T A_1^{-1}, \Sigma_1) + \tfrac{1}{n}\tfrac{\partial}{\partial t_j}\bar{\Phi}_3(\underset{\sim}{c_1}^T\Sigma_1 + \underset{\sim}{t}^T A_1^{-1}, \Sigma_1)|_{\underset{\sim}{\hat{\omega}}}\big].\end{aligned} \tag{7.37}$$

Similarly, the second term of (7.33) is

$$\begin{aligned}&\textstyle -\sum_{j=1,2}(a_3^{1j}\hat{\omega}_1 + a_3^{2j}\hat{\omega}_2)e^{-\frac{1}{2}n(\underset{\sim}{c_2}^T\Sigma_2\underset{\sim}{c_2}+\underset{\sim}{\hat{\omega}}^T A_2^{-1}\underset{\sim}{c_2})}|A_2|\times \\ &\big[(\underset{\sim}{e_j}^T A_2^{-1}\underset{\sim}{c_2})\hat{\omega}_j\bar{\Phi}_3(\underset{\sim}{c_2}^T\Sigma_2 + \underset{\sim}{\hat{\omega}}^T A_2^{-1}, \Sigma_2) + \tfrac{1}{n}\tfrac{\partial}{\partial t_j}\bar{\Phi}_3(\underset{\sim}{c_2}^T\Sigma_2 + \underset{\sim}{t}^T A_2^{-1}, \Sigma_2)|_{\underset{\sim}{\hat{\omega}}}\big],\end{aligned} \tag{7.38}$$

where

$$A_2 = \begin{pmatrix} -(a_2^1 + a_2^{11}\hat{\omega}_1) & 1 & 0 \\ -a_3^1 & -a_3^2 & 1 \\ -a_3^1 - a_3^{11}\hat{\omega}_1 - a_3^{12}\hat{\omega}_2 & -a_3^2 - a_3^{12}\hat{\omega}_1 - a_3^{22}\hat{\omega}_2 & 1 \end{pmatrix}, \tag{7.39}$$

$\underset{\sim}{c_2} = (0, a_2^{11}\hat{\omega}_1^2, \sum_{i,j=1,2} a_3^{ij}\hat{\omega}_i\hat{\omega}_j)^T$ and $\Sigma_2 = (A_2^T A_2)^{-1}$. □

Sum up the results of Lemma 7.1 to Lemma 7.4 to give Theorem 7.2.

Theorem 7.2. *To the error of $O_p(\frac{1}{n})$, the integral of the first term in* (4.9) *of the saddlepoint approximation for the trivariate unconditional tail probability is*

$$\begin{aligned}&\textstyle\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty} \exp(n[\frac{\omega_1^2+\omega_2^2+\omega_3^2}{2} - \hat{\omega}_1\omega_1 - \hat{\omega}_2\omega_2 - \hat{\omega}_3\omega_3]) \\ &\times\frac{1}{\omega_1[\omega_2 - a_2^1\omega_1 - a_2^{11}\omega_1(\omega_1-\hat{\omega}_1)][\omega_3 - a_3^2\omega_2 - a_3^1\omega_1 - \sum_{i,j=1,2} a_3^{ij}\omega_i(\omega_j-\hat{\omega}_j)]}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} \\ &= \\ &\bar{\Phi}_3((A_0^{-T}\underset{\sim}{\hat{\omega}})^T, \Sigma_0) \\ &+ \\ &\textstyle\sum_{k=1}^4\sum_{j=1}^2 b_{kj}e^{-\frac{1}{2}n(\underset{\sim}{c_k}^T\Sigma_k\underset{\sim}{c_k}+\underset{\sim}{\hat{\omega}}^T A_k^{-1}\underset{\sim}{c_k})}|A_k|\times \\ &\big[(\underset{\sim}{e_j}^T A_k^{-1}\underset{\sim}{c_k})\hat{\omega}_j\bar{\Phi}_3(\underset{\sim}{c_k}^T\Sigma_k + \underset{\sim}{\hat{\omega}}^T A_k^{-1}, \Sigma_k) + \tfrac{1}{n}\tfrac{\partial}{\partial t_j}\bar{\Phi}_3(\underset{\sim}{c_k}^T\Sigma_k + \underset{\sim}{t}^T A_k^{-1}, \Sigma_k)|_{\underset{\sim}{\hat{\omega}}}\big] \\ &+O_p(\tfrac{1}{n}),\end{aligned} \tag{7.40}$$

where $b_{11} = -a_2^{11}$, $b_{12} = 0$, $b_{2j} = a_3^{1j}\hat{\omega}_1 + a_3^{2j}\hat{\omega}_2 - a_3^{1j} - a_2^1 a_3^{2j}$, $b_{3j} = -a_3^{2j}$, $b_{4j} = \sum_{j=1}^{2}(a_3^{1j}\hat{\omega}_1 + a_3^{2j}\hat{\omega}_2)$ *for* $j = 1, 2$, $\underset{\sim}{e}_j$ *is the unit vector of the* j^{th} *element being 1 and all others being 0,* $\underset{\sim}{c}_1 = (0, a_2^{11}\hat{\omega}_1^2, 0)^T$, $\underset{\sim}{c}_2 = \underset{\sim}{c}_3 = (0, 0, \sum_{i,j=1,2} a_3^{ij}\hat{\omega}_i\hat{\omega}_j)^T$,
$\underset{\sim}{c}_4 = (0, a_2^{11}\hat{\omega}_1^2, \sum_{i,j=1,2} a_3^{ij}\hat{\omega}_i\hat{\omega}_j)^T$, $\underset{\sim}{\hat{\omega}} = (\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3)^T$,

$$A_0 = \begin{pmatrix} 1 & 0 & 0 \\ -a_2^1 & 1 & 0 \\ -a_3^1 & -a_3^2 & 1 \end{pmatrix} \tag{7.41}$$

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ -(a_2^1 + a_2^{11}\hat{\omega}_1) & 1 & 0 \\ -a_3^1 & -a_3^2 & 1 \end{pmatrix}, \tag{7.42}$$

$$A_2 = \begin{pmatrix} -a_2^1 & 1 & 0 \\ -a_3^1 & -a_3^2 & 1 \\ -a_3^1 - a_3^{11}\hat{\omega}_1 - a_3^{12}\hat{\omega}_2 & -a_3^2 - a_3^{12}\hat{\omega}_1 - a_3^{22}\hat{\omega}_2 & 1 \end{pmatrix}, \tag{7.43}$$

$$A_3 = \begin{pmatrix} 1 & 0 & 0 \\ -a_3^1 & -a_3^2 & 1 \\ -a_3^1 - a_3^{11}\hat{\omega}_1 - a_3^{12}\hat{\omega}_2 & -a_3^2 - a_3^{12}\hat{\omega}_1 - a_3^{22}\hat{\omega}_2 & 1 \end{pmatrix}, \tag{7.44}$$

$$A_4 = \begin{pmatrix} -(a_2^1 + a_2^{11}\hat{\omega}_1) & 1 & 0 \\ -a_3^1 & -a_3^2 & 1 \\ -a_3^1 - a_3^{11}\hat{\omega}_1 - a_3^{12}\hat{\omega}_2 & -a_3^2 - a_3^{12}\hat{\omega}_1 - a_3^{22}\hat{\omega}_2 & 1 \end{pmatrix}, \tag{7.45}$$

and $\Sigma_k = (A_k^T A_k)^{-1}$ *for* $k = 0, \cdots, 4$.

Proof. Omitted. □

Note that $(\hat{\omega}_1, \hat{\omega}_2) = (0, 0)$ are the singularity points for (7.40) since $|A_k|(\hat{\omega}_1, \hat{\omega}_2) = 0$ at $(\hat{\omega}_1, \hat{\omega}_2) = (0, 0)$ for $k \geq 2$ such that the inverse of A_k doesn't exist. However, Lemma 7.5 will show that the singularity points are removable.

Lemma 7.5. *The singularity points of* $(\hat{\omega}_1, \hat{\omega}_2) = (0,0)$ *are removable for*

$$f(\hat{\omega}_1, \hat{\omega}_2) = e^{-\frac{1}{2}n(\underset{\sim}{c_k}^T \Sigma_k \underset{\sim}{c_k} + \underset{\sim}{\hat{\omega}}^T A_k^{-1} \underset{\sim}{c_k})}$$
$$|A_k| \big[(\underset{\sim}{e_j}^T A_k^{-1} \underset{\sim}{c_k}) \hat{\omega}_j \bar{\Phi}_3(\underset{\sim}{c_k}^T \Sigma_k + \underset{\sim}{\hat{\omega}}^T A_k^{-1}, \Sigma_k) + \frac{1}{n} \frac{\partial}{\partial t_j} \bar{\Phi}_3(\underset{\sim}{c_k}^T \Sigma_k + \underset{\sim}{t}^T A_k^{-1}, \Sigma_k)|_{\underset{\sim}{\hat{\omega}}} \big], \tag{7.46}$$

for $k \geq 2$.

Proof. Use Riemann's theorem, it is to show $\lim_{|A_k| \to 0} |A_k| f(\hat{\omega}_1, \hat{\omega}_2) = 0$. Substitute $A_k^{-1} = \frac{adj A_k}{|A_k|}$, where $adj A_k$ denotes the adjoint of A_k, to give

$$|A_k| f(\hat{\omega}_1, \hat{\omega}_2) = |A_k| e^{-\frac{1}{2}n(\underset{\sim}{c_k}^T \Sigma_k \underset{\sim}{c_k} + \underset{\sim}{\hat{\omega}}^T A_k^{-1} \underset{\sim}{c_k})}$$
$$| \big[(\underset{\sim}{e_j}^T (adj A_k) \underset{\sim}{c_k}) \hat{\omega}_j \bar{\Phi}_3(\underset{\sim}{c_k}^T \Sigma_k + \underset{\sim}{\hat{\omega}}^T A_k^{-1}, \Sigma_k) + \frac{1}{n} |A_k| \frac{\partial}{\partial t_j} \bar{\Phi}_3(\underset{\sim}{c_k}^T \Sigma_k + \underset{\sim}{t}^T A_k^{-1}, \Sigma_k)|_{\underset{\sim}{\hat{\omega}}} \big]. \tag{7.47}$$

Note that $|A_k|(\hat{\omega}_1, \hat{\omega}_2) = c_k \hat{\omega}_1 + d_k \hat{\omega}_2$ for some constants c_k and d_k, then write $A_k^{-1} = \frac{1}{|A_k|}(\underset{\sim}{r_1}, \underset{\sim}{r_2}, \underset{\sim}{r_3})^T = \frac{1}{c_k \hat{\omega}_1 + d_k \hat{\omega}_2}(\underset{\sim}{r_{k1}}, \underset{\sim}{r_{k2}}, \underset{\sim}{r_{k3}})^T$ where $\underset{\sim}{r_{ki}}$ denotes the i^{th} row vector corresponding to A_k^{-1}. Hence for $k = 2$, $\lim_{(\hat{\omega}_1, \hat{\omega}_2) \to (0,0)} \underset{\sim}{\hat{\omega}}^T A_2^{-1} \underset{\sim}{c_2} = \lim_{(\hat{\omega}_1, \hat{\omega}_2) \to (0,0)} \hat{\omega}_3 \underset{\sim}{r_{23}}^T \underset{\sim}{c_2} = \hat{\omega}_3 \lim_{(\hat{\omega}_1, \hat{\omega}_2) \to (0,0)} \frac{(a_2^1 a_3^2 + a_2^1 + a_2^{11} \hat{\omega}_1) \sum a_3^{ij} \hat{\omega}_i \hat{\omega}_j}{c_2 \hat{\omega}_1 + d_2 \hat{\omega}_2} = 0$. Similarly, $\lim_{(\hat{\omega}_1, \hat{\omega}_2) \to (0,0)} \underset{\sim}{\hat{\omega}}^T A_k^{-1} \underset{\sim}{c_k} = 0$ for $k > 2$. Also, because
$\lim_{(\hat{\omega}_1, \hat{\omega}_2) \to (0,0)} e^{-\frac{1}{2}n \underset{\sim}{c_k}^T \Sigma_k \underset{\sim}{c_k}} = 1$ and both $\bar{\Phi}_3(\underset{\sim}{c_k}^T \Sigma_k + \underset{\sim}{\hat{\omega}}^T A_k^{-1}, \Sigma_k)$ and
$\frac{\partial}{\partial t_j} \bar{\Phi}_3(\underset{\sim}{c_k}^T \Sigma_k + \underset{\sim}{t}^T A_k^{-1}, \Sigma_k)|_{\underset{\sim}{\hat{\omega}}}$ are bounded, hence (7.47) $\to 0$ as $|A_k| \to 0$. □

As in Section 2, the integrals of the rest terms in (4.9) for the trivariate case involving with $G^{(j)}(\omega_1, \omega_2, \omega_3)$ can be evaluated using Watson's Lemma, since the integrands are analytic functions. Theorem 7.3 to 7.5 present the results of integrals involving with $G^{(j)}$ for $j = 1, 2,$ and 3, respectively. Let $\breve{\omega}_2 = \tilde{\omega}_2(\hat{\omega}_1)$ and $\breve{\omega}_3 = \tilde{\omega}_3(\hat{\omega}_1, \hat{\omega}_2)$.

Theorem 7.3. *To the error of* $O_p(\frac{1}{n})$, *the integral of the trivariate tail probability of*

(4.9) *involving with* $G^{(1)}$

$$\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\exp(n\sum_{j=1}^{3}(\omega_j^2/2-\hat\omega_j\omega_j))\frac{G^{(1)}(\omega_1,0,0)}{(\omega_2-\tilde\omega_2(\omega_1))(\omega_3-\tilde\omega_3(\omega_1,\omega_2))} \times\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} \tag{7.48}$$

can be approximated by

$$\frac{1}{\sqrt{n}}\phi(\sqrt{n}\hat\omega_1)\bar\Phi_2(\sqrt{n}(\hat\omega_2-\breve\omega_2),\sqrt{n}(\hat\omega_3-\breve\omega_3))(\frac{1}{\hat z_1}-\frac{1}{\hat\omega_1})e^{n\sum_{j=2}^{3}(\breve\omega_j^2/2-\hat\omega_j\breve\omega_j)}, \tag{7.49}$$

where $\hat z_1=\hat\xi_1\sqrt{K_Y^{11}(\hat\xi_1,\hat\xi_2,\hat\xi_3)-K_Y^{12}(\hat\xi_1,\hat\xi_2,\hat\xi_3)(K_Y^{22}(\hat\xi_1,\hat\xi_2,\hat\xi_3))^{-1}K_Y^{21}(\hat\xi_1,\hat\xi_2,\hat\xi_3)}$.

Proof. We will first show that replacing ω_2 of $\tilde\omega_3(\omega_1,\omega_2)$ by $\hat\omega_2$ in the numerator of the integrand in the denominator of the integrand of (7.48) will induce an error of $O_p(\frac{1}{n})$. The error after replacing ω_2 by $\hat\omega_2$ is

$$\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\exp(n\sum_{j=1}^{3}(\omega_j^2/2-\hat\omega_j\omega_j)) \times\frac{\tilde\omega_3(\omega_1,\omega_2)-\tilde\omega_3(\omega_1,\hat\omega_2)}{(\omega_2-\tilde\omega_2(\omega_1))(\omega_3-\tilde\omega_3(\omega_1,\omega_2))(\omega_3-\tilde\omega_3(\omega_1,\hat\omega_2))}G^{(1)}(\omega_1,0,0)\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}. \tag{7.50}$$

Use the same arguments as in Theorem 7.1. Let $(r^*)^2=(\omega_3-\tilde\omega_3(\omega_1,\omega_2))(\omega_3-\tilde\omega_3(\omega_1,\hat\omega_2))$. By Theorem 5.1, for all fixed (ω_1,ω_2), the univariate integral of $\frac{1}{(r^*)^2}$ as a function of ω_3 is bounded by $\sqrt{n}$, which cancels out the factor $\frac{1}{\sqrt{n}}$ contributed by the integral $\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\exp(n[\omega_3^2/2-\hat\omega_3\omega_3])\frac{d\omega_3}{2\pi i}$. Note that for all fixed ω_1,

$$\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\frac{\exp(n[\omega_2^2/2-\hat\omega_2\omega_2])}{\omega_2-\tilde\omega_2(\omega_1)}\frac{d\omega_2}{2\pi i}=O_p(1),$$

and by Watson's lemma,

$$\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\exp(n[\omega_1^2/2-\hat\omega_1\omega_1])G^{(1)(\omega_1,0,0)}\frac{d\omega_1}{2\pi i}=\frac{1}{\sqrt{n}}\phi(\sqrt{n}\hat\omega_1)G^{(1)}(\hat\omega_1,0,0)$$

contributing a factor of $\frac{1}{\sqrt{n}}$. Make the change of variables $\omega_2=\hat\omega_2+i\frac{z}{\sqrt{n}}$. The term

$\tilde{\omega}_3(\omega_1, \omega_2) - \tilde{\omega}_3(\omega_1, \hat{\omega}_2) = -(a_3^2 + a_3^{12} + a_3^{22})(\omega_2 - \hat{\omega}_2)$ contribute another factor of $\frac{1}{\sqrt{n}}$, hence (7.50) is bounded by $\frac{1}{n}$. Since

$$\frac{G^{(1)}(\omega_1,0,0)}{(\omega_2-\tilde{\omega}_2(\omega_1))(\omega_3-\tilde{\omega}_3(\omega_1,\hat{\omega}_2))} = \frac{1}{(\omega_2-\tilde{\omega}_2(\omega_1))(\omega_3-\tilde{\omega}_3(\omega_1,\hat{\omega}_2))}\left(\frac{1}{\xi_1}\frac{d\xi_1}{d\omega_1} - \frac{1}{\omega_1}\right) \tag{7.51}$$

is an analytic function at $\hat{\omega}_1$. Hence by Watson's Lemma

$$\begin{aligned}
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty} \exp(n\textstyle\sum_{j=1}^3(\omega_j^2/2 - \hat{\omega}_j\omega_j))\frac{G^{(1)}(\omega_1,0,0)}{(\omega_2-\tilde{\omega}_2(\omega_1))(\omega_3-\tilde{\omega}_3(\omega_1,\hat{\omega}_2))}\\
&\times\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} =\\
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty} \exp(n\textstyle\sum_{j=1}^3(\omega_j^2/2 - \hat{\omega}_j\omega_j))\frac{1}{(\omega_2-\tilde{\omega}_2(\omega_1))(\omega_3-\tilde{\omega}_3(\omega_1,\hat{\omega}_2))}\\
&\times\left(\frac{1}{\xi_1}\frac{d\xi_1}{d\omega_1} - \frac{1}{\omega_1}\right)\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} =\\
&\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty} e^{n\sum_{j=2}^3(\omega_j^2/2-\hat{\omega}_j\omega_j)}\Big[\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty} e^{n(\omega_1^2/2-\hat{\omega}_1\omega_1)}\frac{1}{(\omega_2-\tilde{\omega}_2(\hat{\omega}_1))(\omega_3-\tilde{\omega}_3(\omega_1,\hat{\omega}_2))}\\
&\times\left(\frac{1}{\xi_1}\frac{d\xi_1}{d\omega_1} - \frac{1}{\omega_1}\right)\frac{d\omega_1}{2\pi i}\Big]\frac{d\omega_2 d\omega_3}{(2\pi i)^2} =\\
&\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty} e^{n\sum_{j=2}^3(\omega_j^2/2-\hat{\omega}_j\omega_j)}\frac{\phi(\sqrt{n}\hat{\omega}_1)}{\sqrt{n}}\frac{1}{(\omega_2-\tilde{\omega}_2(\hat{\omega}_1))(\omega_3-\tilde{\omega}_3(\hat{\omega}_1,\hat{\omega}_2))}\left(\frac{1}{\hat{z}_1} - \frac{1}{\hat{\omega}_1}\right)\frac{d\omega_2 d\omega_3}{(2\pi i)^2}\\
&+O_p(\tfrac{1}{n}),
\end{aligned} \tag{7.52}$$

where $\hat{z}_1 = \hat{\xi}_1(\frac{d\xi_1}{d\omega_1}(\hat{\omega}_1))^{-1} =$
$\hat{\xi}_1\sqrt{K_Y^{11}(\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3) - K_Y^{12}(\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3)(K_Y^{22}(\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3))^{-1}K_Y^{21}(\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3)}$.

Make the change of variable $\{\omega_2 - \tilde{\omega}_2(\hat{\omega}_1), \omega_3 - \tilde{\omega}_3(\hat{\omega}_1, \hat{\omega}_2)\} \to \{u, v\}$ for (7.52) to give

$$\begin{aligned}
&e^{n\sum_{j=2}^3(\breve{\omega}_j^2/2-\hat{\omega}_j\breve{\omega}_j)}\int_{\hat{u}-i\infty}^{\hat{u}+i\infty}\int_{\hat{v}-i\infty}^{\hat{v}+i\infty} e^{n(u^2/2+v^2/2-\hat{u}u-\hat{v}v)}\frac{1}{uv}\frac{dudv}{(2\pi i)^2}\frac{\phi(\sqrt{n}\hat{\omega}_1)}{\sqrt{n}}\left(\frac{1}{\hat{z}_1} - \frac{1}{\hat{\omega}_1}\right)\\
&+O_p(\tfrac{1}{n})\\
&= \frac{1}{\sqrt{n}}\phi(\sqrt{n}\hat{\omega}_1)\bar{\Phi}_2(\sqrt{n}(\hat{\omega}_2 - \breve{\omega}_2), \sqrt{n}(\hat{\omega}_3 - \breve{\omega}_3))\left(\frac{1}{\hat{z}_1} - \frac{1}{\hat{\omega}_1}\right)e^{n\sum_{j=2}^3(\breve{\omega}_j^2/2-\hat{\omega}_j\breve{\omega}_j)}\\
&+O_p(\tfrac{1}{n}).
\end{aligned} \tag{7.53}$$

□

Theorem 7.4. *To the error of $O_p(\frac{1}{n})$, the integral of the trivariate tail probability*

(4.9) *involving with* $G^{(2)}$

$$\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\exp(n\sum_{j=1}^3(\omega_j^2/2-\hat\omega_j\omega_j))\frac{G^{(2)}(\omega_1,\omega_2,0)}{\omega_1(\omega_3-\tilde\omega_3(\omega_1,\omega_2))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} \tag{7.54}$$

can be approximated by

$$\frac{1}{\sqrt{n}}\phi(\sqrt{n}\hat\omega_2)\bar\Phi_2(\sqrt{n}\hat\omega_1,\sqrt{n}(\hat\omega_3-\check\omega_3))(\frac{1}{\hat z_2}-\frac{1}{\hat\omega_2})e^{n(\check\omega_3^2-\hat\omega_2\check\omega_3)}, \tag{7.55}$$

where $\hat z_2=\check\xi_2\sqrt{K_Y^{22}(0,\check\xi_2,0)}$ *and* $\check\xi_2=\xi_2(0,\hat\omega_2)$.

Proof. Write $G^{(2)}(\omega_1,\omega_2,0)=G^{(2)}(0,\omega_2,0)+(G^{(2)}(\omega_1,\omega_2,0)-G^{(2)}(0,\omega_2,0))$, where $G^{(2)}(0,\omega_2,0)=\frac{1}{\xi_2(0,\omega_2)}\frac{d\xi_2}{d\omega_2}(0,\omega_2)-\frac{1}{\omega_2}$. The second term will integrate to the order of $O_p(\frac{1}{n})$, further more replacing $\frac{G^{(2)}(0,\omega_2,0)}{\omega_3-\tilde\omega_3(\omega_1,\omega_2)}$ by $\frac{G^{(2)}(0,\omega_2,0)}{\omega_3-\tilde\omega_3(\hat\omega_1,\omega_2)}$ will also induce an error of $O_p(\frac{1}{n})$ since the error after the replacement is

$$\begin{aligned}&\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\exp(n\sum_{j=1}^3(\omega_j^2/2-\hat\omega_j\omega_j))\frac{\tilde\omega_3(\omega_1,\omega_2)-\tilde\omega_3(\hat\omega_1,\omega_2)}{\omega_1(\omega_3-\tilde\omega_3(\omega_1,\omega_2))(\omega_3-\tilde\omega_3(\hat\omega_1,\omega_2))}\\&\times G^{(2)}(0,\omega_2,0)\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}.\end{aligned} \tag{7.56}$$

Use the same arguments as in Theorem 7.3. Let $(r^*)^2=(\omega_3-\tilde\omega_3(\omega_1,\omega_2))(\omega_3-\tilde\omega_3(\hat\omega_1,\omega_2))$. By Theorem 5.1, for all fixed (ω_1,ω_2), the univariate integral of $\frac{1}{(r^*)^2}$ as a function of ω_3 is bounded by $\sqrt{n}$ which cancels out the factor $\frac{1}{\sqrt{n}}$ contributed by the integral $\int_{\hat\omega_3-i\infty}^{\hat\omega_3+i\infty}\exp(n[\omega_3^2/2-\hat\omega_3\omega_3])\frac{d\omega_3}{2\pi i}$. Note that $\int_{\hat\omega_1-i\infty}^{\hat\omega_1+i\infty}\frac{\exp(n[\omega_1^2/2-\hat\omega_1\omega_1])}{\omega_1}\frac{d\omega_1}{2\pi i}=O_p(1)$, and by Watson's lemma,

$$\int_{\hat\omega_2-i\infty}^{\hat\omega_2+i\infty}\exp(n[\omega_2^2/2-\hat\omega_2\omega_2])G^{(2)(0,\omega_2,0)}\frac{d\omega_2}{2\pi i}=\frac{1}{\sqrt{n}}\phi(\sqrt{n}\hat\omega_2)G^{(2)}(0,\hat\omega_2,0)$$

contributing a factor of $\frac{1}{\sqrt{n}}$. Make the change of variables $\omega_1=\hat\omega_1+i\frac{z}{\sqrt{n}}$. The term $\tilde\omega_3(\omega_1,\omega_2)-\tilde\omega_3(\hat\omega_1,\omega_2)=-(a_3^1+a_3^{11}+a_3^{21})(\omega_1-\hat\omega_1)$ contribute another factor of $\frac{1}{\sqrt{n}}$, hence (7.56) is bounded by $\frac{1}{n}$. Hence, to error $O_p(\frac{1}{n})$, use Watson's Lemma, since

$\frac{G^{(2)}(0,\omega_2,0)}{\omega_3-\tilde{\omega}_3(\hat{\omega}_1,\omega_2)}$ is analytic at $\hat{\omega}_2$. Integrating the first term $G^{(2)}(0,\omega_2,0)$ gives

$$\begin{aligned}
&\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n\textstyle\sum_{j=1}^3(\omega_j^2/2-\hat{\omega}_j\omega_j))\frac{G^{(2)}(0,\omega_2,0)}{\omega_1(\omega_3-\tilde{\omega}_3(\hat{\omega}_1,\omega_2))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}=\\
&\Big[\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}e^{n(\omega_3^2/2-\hat{\omega}_3\omega_3)}\Big(\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}e^{n(\omega_2^2/2-\hat{\omega}_2\omega_2)}\frac{G^{(2)}(0,\omega_2,0)}{\omega_3-\tilde{\omega}_3(\hat{\omega}_1,\omega_2)}\frac{d\omega_2}{2\pi i}\Big)\frac{d\omega_3}{2\pi i}\Big]\\
&\times\Big[\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}e^{n(\omega_1^2/2-\hat{\omega}_1\omega_1)}\frac{1}{\omega_1}\frac{d\omega_1}{2\pi i}\Big] \qquad (7.57)\\
&=\frac{1}{\sqrt{n}}\phi(\sqrt{n}\hat{\omega}_2)G^{(2)}(0,\hat{\omega}_2)\bar{\Phi}(\sqrt{n}\hat{\omega}_1)\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}e^{n(\omega_3^2/2-\hat{\omega}_3\omega_3)}\frac{1}{\omega_3-\tilde{\omega}_3(\hat{\omega}_1,\hat{\omega}_2)}\frac{d\omega_3}{2\pi i}\\
&=\frac{1}{\sqrt{n}}\phi(\sqrt{n}\hat{\omega}_2)G^{(2)}(0,\hat{\omega}_2)\bar{\Phi}_2(\sqrt{n}\hat{\omega}_1,\sqrt{n}(\hat{\omega}_3-\breve{\omega}_3))e^{n(\breve{\omega}_3^2-\hat{\omega}_2\breve{\omega}_3)},
\end{aligned}$$

where $G^{(2)}(0,\hat{\omega}_2,0)=\frac{1}{\xi_2(0,\hat{\omega}_2)}\frac{d\xi_2}{d\omega_2}(0,\hat{\omega}_2)-\frac{1}{\hat{\omega}_2}=\frac{1}{\hat{z}_2}-\frac{1}{\hat{\omega}_2}$, and $\breve{\xi}_2=\xi_2(0,\hat{\omega}_2)$ is the root of equation $K_Y(0,\breve{\xi}_2,0)-\breve{\xi}_2\bar{t}_2-(K_Y(\hat{\xi}_1,\hat{\xi}_2,\hat{\xi}_3)-\hat{\underset{\sim}{\xi}}^T\underset{\sim}{\bar{t}})-\hat{\omega}_1^2/2-\hat{\omega}_3^2/2=0$. □

Theorem 7.5. *To the error of $O_p(\frac{1}{n})$, the integral of the trivariate tail probability (4.9) involving with $G^{(3)}$*

$$\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty}\exp(n\textstyle\sum_{j=1}^3(\omega_j^2/2-\hat{\omega}_j\omega_j))\frac{G^{(3)}(\omega_1,\omega_2,\omega_3)}{\omega_1(\omega_2-\tilde{\omega}_2(\omega_1))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3} \qquad (7.58)$$

can be approximated by

$$\frac{1}{\sqrt{n}}\phi(\sqrt{n}\hat{\omega}_3)(\frac{1}{\hat{z}_3}-\frac{1}{\hat{\omega}_3})\Psi(\hat{\omega}_1,\hat{\omega}_2), \qquad (7.59)$$

where $\hat{z}_3=\breve{\xi}_3\sqrt{K_Y^{33}(0,0,\breve{\xi}_3)}$, $\breve{\xi}_3=\xi_3(0,0,\hat{\omega}_3)$, and

$$\begin{aligned}
&\Psi(\hat{\omega}_1,\hat{\omega}_2)=\bar{\Phi}_2((\hat{\omega}_1+a_2^1\hat{\omega}_2,\hat{\omega}_2),\Sigma/n)+\mathrm{sgn}(a_2^{11}\hat{\omega}_1)\frac{1}{\sqrt{n a_2^{11}\hat{\omega}_1^2}}e^{-\frac{n}{2}(\hat{\omega}_1^2+\hat{\omega}_2^2-\hat{u}_2^2)}\times\\
&[\sigma_1^{-1}\phi(\frac{1+\varsigma_1}{\sqrt{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{1+2\varsigma_1+\varsigma_2^2}}\sqrt{n}\hat{u}_2)-\sigma_2^{-1}\phi(\frac{\varsigma_1}{\varsigma_2}\sqrt{n}\hat{u}_2)\bar{\Phi}(\sqrt{\frac{\varsigma_2^2-\varsigma_1^2}{\varsigma_2^2}}\sqrt{n}\hat{u}_2)]
\end{aligned} \qquad (7.60)$$

defined as in Theorem 5.4.

Proof. Write

$$\begin{aligned}
G^{(3)}(\omega_1,\omega_2,\omega_3)=\;&G^{(3)}(0,0,\omega_3)+(G^{(3)}(0,\omega_2,\omega_3)-G^{(2)}(0,0,\omega_3))\\
&+(G^{(3)}(\omega_1,\omega_2,\omega_3)-G^{(2)}(0,\omega_2,\omega_3)),
\end{aligned}$$

where $G^{(3)}(0,0,\omega_3) = \frac{1}{\xi_3(0,0,\omega_3)}\frac{d\xi_3}{d\omega_3}(0,0,\omega_3) - \frac{1}{\omega_3}$. The second and the third term will integrate to the order of $O_p(\frac{1}{n})$. Since $G^{(3)}(0,0,\omega_3)$ is analytic at $\hat{\omega}_3$. Hence by Watson's Lemma since , integrating the first term $G^{(3)}(0,0,\omega_3)$ gives

$$
\begin{aligned}
&\textstyle\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty} \exp(n\sum_{j=1}^{3}(\omega_j^2/2-\hat{\omega}_j\omega_j))\frac{G^{(3)}(0,0,\omega_3)}{\omega_1(\omega_2-\tilde{\omega}_2(\omega_1))}\frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi i)^3}\\
&=\\
&\textstyle\big[\int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty} \exp(n\sum_{j=1}^{2}(\omega_j^2/2-\hat{\omega}_j\omega_j))\frac{1}{\omega_1(\omega_2-\tilde{\omega}_2(\omega_1))}\frac{d\omega_1 d\omega_2}{(2\pi i)^2}\big] \qquad (7.61)\\
&\textstyle\times\big[\int_{\hat{\omega}_3-i\infty}^{\hat{\omega}_3+i\infty} e^{n(\omega_3^2/2-\hat{\omega}_3\omega_3)}G^{(3)}(0,0,\omega_3)\frac{d\omega_3}{2\pi i}\big]\\
&= \tfrac{1}{\sqrt{n}}\phi(\sqrt{n}\hat{\omega}_3)G^{(3)}(0,0,\hat{\omega}_3)\Psi(\hat{\omega}_1,\hat{\omega}_2),
\end{aligned}
$$

where $\Psi(\hat{\omega}_1,\hat{\omega}_2) = \int_{\hat{\omega}_1-i\infty}^{\hat{\omega}_1+i\infty}\int_{\hat{\omega}_2-i\infty}^{\hat{\omega}_2+i\infty}\frac{\exp(n\sum_{j=1}^{2}(\omega_j^2/2-\hat{\omega}_j\omega_j))}{\omega_1(\omega_2-\tilde{\omega}_2(\omega_1))}\frac{d\omega_1 d\omega_2}{(2\pi i)^2}$ is the first term of the bivariate tail probability which is given by Theorem 5.4,

$$G^{(3)}(0,0,\hat{\omega}_3) = \frac{1}{\xi_3(0,0,\hat{\omega}_3)}\frac{d\xi_3}{d\omega_3}(0,0,\hat{\omega}_3) - \frac{1}{\hat{\omega}_3} = \frac{1}{\hat{z}_3} - \frac{1}{\hat{\omega}_3},$$

and $\check{\xi}_3 = \xi_3(0,0,\hat{\omega}_3)$ is the root of equation $K_Y(0,0,\check{\xi}_3) - \check{\xi}_3\bar{t}_3 - (K_Y(\hat{\xi}_1,\hat{\xi}_2,\hat{\xi}_3) - \hat{\underset{\sim}{\xi}}^T\underset{\sim}{\bar{t}}) - \hat{\omega}_1^2/2 - \hat{\omega}_2^2/2 = 0$. □

Combining Theorem 7.2, 7.3, 7.4 and 7.5 gives the approximation to the trivariate tail probability of (7.1) to the error of $O_p(\frac{1}{n})$, which is presented in the Theorem 7.6.

Theorem 7.6. *To the error of $O_p(\frac{1}{n})$, the unconditional trivariate tail probability*

(7.1) can be approximated by

$$
\begin{aligned}
&\bar{\Phi}_3((A_0^{-T}\underset{\sim}{\hat{\omega}})^T, \Sigma_0) \\
&+ \\
&\textstyle\sum_{k=1}^4 \sum_{j=1}^2 b_{kj} e^{-\frac{1}{2}n(\underset{\sim}{c_k}^T \Sigma_k \underset{\sim}{c_k} + \underset{\sim}{\hat{\omega}}^T A_k^{-1} \underset{\sim}{c_k})} |A_k| \times \\
&\left[(\underset{\sim}{e_j}^T A_k^{-1} \underset{\sim}{c_k})\hat{\omega}_j \bar{\Phi}_3(\underset{\sim}{c_k}^T \Sigma_k + \underset{\sim}{\hat{\omega}}^T A_k^{-1}, \Sigma_k) + \frac{1}{n}\frac{\partial}{\partial t_j}\bar{\Phi}_3(\underset{\sim}{c_k}^T \Sigma_k + \underset{\sim}{t}^T A_k^{-1}, \Sigma_k)|_{\underset{\sim}{\hat{\omega}}}\right] \qquad (7.62) \\
&+\tfrac{1}{\sqrt{n}}\phi(\sqrt{n}\hat{\omega}_1)\bar{\Phi}_2(\sqrt{n}(\hat{\omega}_2 - \check{\omega}_2), \sqrt{n}(\hat{\omega}_3 - \check{\omega}_3))(\tfrac{1}{\hat{z}_1} - \tfrac{1}{\hat{\omega}_1})e^{n\sum_{j=2}^3(\check{\omega}_j^2/2 - \hat{\omega}_j\check{\omega}_j)} \\
&+\tfrac{1}{\sqrt{n}}\phi(\sqrt{n}\hat{\omega}_2)\bar{\Phi}_2(\sqrt{n}\hat{\omega}_1, \sqrt{n}(\hat{\omega}_3 - \check{\omega}_3))(\tfrac{1}{\hat{z}_2} - \tfrac{1}{\hat{\omega}_2})e^{n(\check{\omega}_3^2 - \hat{\omega}_2\check{\omega}_3)} \\
&+\tfrac{1}{\sqrt{n}}\phi(\sqrt{n}\hat{\omega}_3)(\tfrac{1}{\hat{z}_3} - \tfrac{1}{\hat{\omega}_3})\Psi(\hat{\omega}_1, \hat{\omega}_2),
\end{aligned}
$$

where the quantity $\hat{\omega}_1$, $\hat{\omega}_2$, and $\hat{\omega}_3$ are the square roots of twice the unconditional log likelihood ratio statistics for testing the corresponding $\xi_j = 0, j = 1, 2, 3$, assuming that $\xi_l = 0$ for $l < j$, but without restriction on ξ_l for $l > j$. The matrices A_k for $k = 0, \cdots, 4$ are defined as in Theorem 7.2,

$$\hat{z}_1 = \hat{\xi}_1\sqrt{K_Y^{11}(\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3) - K_Y^{12}(\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3)(K_Y^{22}(\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3))^{-1}K_Y^{21}(\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3)},$$

$\hat{z}_2 = \check{\xi}_2\sqrt{K_Y^{22}(0, \check{\xi}_2, 0)}$, $\hat{z}_3 = \check{\xi}_3\sqrt{K_Y^{33}(0, 0, \check{\xi}_3)}$, $\check{\xi}_2 = \xi_2(0, \hat{\omega}_2)$, $\check{\xi}_3 = \xi_2(0, 0, \hat{\omega}_3)$, and $\Psi(\hat{\omega}_1, \hat{\omega}_2)$ is defined as in Theorem 5.1.

Proof. Applying Theorem 7.2, 7.3, 7.4 and 7.5 to each integral of (4.9) for the trivariate case gives the approximation to the trivariate tail probability to the error of $O_p(\frac{1}{n})$ as (7.62). $\square$

As shown in this chapter, one may use mathematical induction to extend the quadratic saddlepoint approximation to any dimension.

Part IV

CONCLUSION

Approximation methods play a critical role for obtaining higher order multivariate inference in terms of confidence region and tail probability. Statistical inference based on confidence regions and tail probabilities are most frequently used for statistical hypothesis testing and for estimation of sample size, power, or other statistical quantities. Hence one would like to have the error of the approximation bounded by a small constant, and the constant boundary of the error is expected to have an inverse relationship with sample size; i.e., the error becomes smaller as the number of subjects in the observed sample increases. The method of Laplace for evaluating bivariate confidence regions, as described in Part II, and the method of quadratic saddlepoint approximation for evaluating conditional or unconditional tail probability, as described in Part III, both result in an error to the order of $O_p(\frac{1}{n})$, with n being the sample size. The precision of our approximations is significantly improved compared to other approximations available in the literature of error $O_p(\frac{1}{\sqrt{n}})$. When the sample size is small, applying our approximations will result in drawing a more accurate conclusion for hypothesis testing and obtaining less biased estimation of power determination or other statistical problems for higher order multivariate inference.

The limitation of Laplace's approximation to determine a matching prior for obtaining the bivariate confidence region in Part II may begin with the assumption of orthogonality of the parameters. The approach requires that all parameters, including parameters of interest and nuisance parameters, are pairwisely orthogonal. The orthogonality assumption may be held up to three parameters; however, become extremely difficult to meet as the number of parameters exceeding three. Therefore, it is not meaningful to extend this approach to dimensions of more than three.

In Part III, we expressed higher order conditional or unconditional tail probability in the forms of normal density functions and normal distribution functions using the quadratic saddlepoint approximation method. The approach can be extended to any dimension as shown in Chapter 4.4. However, the approximation results in more complexity of involving substantially more terms to calculate as the dimension increases, which may be a limitation of the quadratic saddlepoint approximation method for high dimension.

These theoretical limitations for both approximation methods have minimal impact on the applications of these approximation methods. As described in the beginning of this chapter, our approximations are developed to have smaller error and in expression of only normal functions. These characteristics give the approximations as presented in Part II and Part III the benefit for obtaining higher order multivariate inference assessed by confidence region and tail probability in terms of higher precision and easier implementation.

BIBLIOGRAPHY

[1] Kolassa,J. and Li,J.(2010). Multivariate Saddlepoint Tail Probability Approximations. *The Annals of Statistics* **31** 274-286.

[2] Kolassa,J. (2004). Approximate multivariate conditional inference using the adjusted profile likelihood. *The Canadian Journal of Statistics* **32** 5-14.

[3] Kolassa,J. (2003). Multivariate saddlepoint tail probability approximations. *The Annals of Statistics* **31** 274-286.

[4] Welch,B.L. and Peers, H.W. (1963). On formulae for confidence points based on integrals of weighted likelihoods. *Journal of the Royal Statistical Society* **25** 318-329.

[5] Peers, H.W. (1965). On confidence points and Bayesian probability points in the case of several parameters. *Journal of the Royal Statistical Society* **27** 9-16.

[6] Tallis, G.M. (1961). The moment generating function of the truncated multinormal distribution. *Journal of the Royal Statistical Society* **23** 223-229.

[7] Bartlett, M.S. (1953). Approximate confidence intervals. II. More than one parameter. *Biometrika* **40** 306-317.

[8] Cox, D.R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society* **49** 1-39.

[9] Robinson, J. (1982). Saddlepoint approximation for permutation tests and confidence intervals. *Journal of the Royal Statistical Society* **44** 91-101.

[10] Wang, Suojin (1990). Saddlepoint approximation for Bivariate Distributions. *Journal of Applied Probability* **27** 586-597.

[11] Stokes, M. E., DAVIS, C. S. and KOCH, G. G. (1995). Categorical Data Analysis Using the SAS System. SAS Institute, Cary, NC.