Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/160412

On the Asymptotic Evaluation of the Finite-Sample Risk of the Nearest Neighbor Classifier

Mohamed M. Rizk

Mathematics & Computer Science Department, Faculty of Science, Menoufia University, Shebin El-Kom, Egypt

Received: 3 Feb. 2022, Revised: 15 Mar. 2022, Accepted: 20 Mar. 2022
Published online: 1 Jul. 2022

Abstract: Based on an exact integral expression for the risk, an asymptotic evaluation of the conditional risk is derived for distributions with have unbounded supports, which, using Laplace's method. Then, by integrating these asymptotic expansions, we evaluate the asymptotic evaluation of the finite sample risk (the unconditional probability error). The finite sample risk for the Pareto and exponential distributions are discussed as typical.

Keywords: Pattern recognition, nearest neighbor classification, finite-sample risk, Laplace's method, Lagrange inversion formula, Pareto distribution, exponential distribution.

1 Introduction

The nearest neighbor rule is one of the nonparametric classification methods, the first study of this method was suggested by Fix and Hodges [1], [2], and continued by Cover and Hart [3] which they gave upper bounds for the infinite sample risk R_∞ of the nearest-neighbor classifiers such that, they proved $R^* \leq R_\infty \leq 2R^*(1-R^*)$, where R^* indicates to the Bayes error, Cover [4] has shown that $R_m = R_\infty + O(m^{-2})$, where R_m indicates to the finite sample risk, and m is the sample size, Wagner [5] and Fritz [6] treated convergence of the conditional error rate for the nearest neighbor rule, Fukunaga and Hummels [7] evaluated $R_m \sim R_\infty + B\frac{\Gamma(m+1)}{\Gamma(m+1+(2/d))}$, where Γ refers to the gamma function and B is a constant. Psaltis et al. [8] generalized the results of Cover [4] to d-dimensional which, they proved that $R_m \sim R_\infty + \sum_{k=2}^{\infty} c_k m^{-k/d}$ as $m \to \infty$, where the coefficients c_k are constants independent of the sample size m, this was extended to the case of multiple classes by Snapp and Venkatesh [9]. Irle and Rizk [10] found an asymptotic evaluation of the conditional risk $R_m(x)$ (the probability of error conditioned on the event that $X = x$) by using partial integration and Laplace's method.

There are many studies in different directions are available for nearest neighbor rules, in addition, several results on the convergence rates, e.g., Dasarathy [11], Devroye et al. [12], Fukunaga [13], Biau and Devroye [14], Döring et al. [15], and Zhao and Lai [16].

2 The finite sample risk

Let $(X_1,\theta_1),(X_2,\theta_2),\cdots,(X_m,\theta_m)$ be a sequence of m independent identically distributed random variables (i.i.d.), where X_i takes values in R^d, and its corresponding class θ_i in a finite set $T=\{1,2,...,C\}$. After reordering the data according to the increasing values of the Euclidean distance $\|X_i - x\|$ for fixed x, we obtain the reordered data sequence in the form $\left(X^{(1)},\theta^{(1)}\right),\left(X^{(2)},\theta^{(2)}\right),\cdots,\left(X^{(m)},\theta^{(m)}\right)$.

Suppose that δ is a mapping from R^d to $\{1,2,...,C\}$, it is called a classifier. The probability of misclassification is $P(\delta(X) \neq \theta)$. Given an i.i.d. training sequence $D_m = \left(\left(X^{(1)},\theta^{(1)}\right),\left(X^{(2)},\theta^{(2)}\right),\cdots,\left(X^{(m)},\theta^{(m)}\right)\right)$ and a new independent random variable (X,θ) with the same distribution as D_m, which $X \in R^d$ is an observed pattern, and we wish to predict its corresponding class θ. The nearest neighbor rule predicts θ by the label of the nearest neighbor X, i.e., by using a suitable tie-breaking, the nearest neighbor rule assigns X to a class $\theta^{(i)}$ with the property $\left\|X - X^{(i)}\right\| \leq \left\|X - X^{(j)}\right\|$ for all $i \neq j$.

* Corresponding author e-mail: mhm96@yahoo.com

In the following, suppose that $f = p_1 f_1 + p_2 f_2$ denotes the mixture density, where $\sum_{i=1}^{2} p_i = 1$, and the densities f_i are those of the class-conditional distributions F_i which are assumed absolutely continuous for each $i \in \{1,2\}$. Let S be its support in R^d, $B(\rho,x) = \{x' \in R^d : \|x-x'\| \le \rho\}$ is the closed ball of radius ρ at x, and $C = 2$, i.e. $T = \{1,2\}$.

The finite sample risk R_m can be written in integral form as follows, compare Irle and Rizk [10], and Snapp and Venkatesh [8]. Assume that X' is the nearest neighbor feature vector in the training sequence D_m to the random test vector X, and θ' indicates the corresponding class label to X'. Suppose $P(\theta \neq \theta'|X=x, X'=x')$ denotes the probability of misclassification θ by θ' given X and its nearest neighbor X'. Then

$$\begin{aligned}
&P\left(\theta \neq \theta'|X=x, X'=x'\right)\\
&= P\left(\theta = 1, \theta' = 2|X=x, X'=x'\right)\\
&\ +P\left(\theta = 2, \theta' = 1|X=x, X'=x'\right)\\
&= P(\theta = 1|X=x)\, P\left(\theta' = 2|X'=x'\right)\\
&\ +P(\theta = 2|X=x)\, P\left(\theta' = 1|X'=x'\right)\\
&= \frac{p_1 p_2}{f(x) f(x')}\left(f_1(x) f_2(x') + f_1(x') f_2(x)\right) \qquad (1)
\end{aligned}$$

By averaging $P(\theta \neq \theta'|X=x, X'=x')$ over X', we obtain

$$\begin{aligned}
R_m(X) &= P\left(\theta \neq \theta'|X=x\right)\\
&= \int_S P\left(\theta \neq \theta'|X=x, X'=x'\right) f_m\left(x'|x\right) dx', \qquad (2)
\end{aligned}$$

where $R_m(X)$ indicates to the conditional risk, and $f_m(x'|x)$ indicates to the conditional density of X' given $X=x$. Thus, with $\rho = |x'-x|$ we obtain $f_m(x'|x) = P($one of the X_j's $\in B(\rho,x)$, and all others $\notin B(\rho,x))f(x')$, that is

$$\begin{aligned}
f_m\left(x'|x\right) &= m\left(1 - P\left(X \in B\left(|x'-x|, x\right)\right)\right)^{m-1} f(x')\\
&= m\left(P\left(|X-x| > |x'-x|\right)\right)^{m-1} f(x'). \qquad (3)
\end{aligned}$$

Hence, by taking the expectation of $P(\theta \neq \theta'|X=x)$ with respect to $X=x$, we obtain

$$R_m = P\left(\theta \neq \theta'\right) = \int_S P\left(\theta \neq \theta'|X=x\right) f(x) dx, \qquad (4)$$

After substituting (1) and (3) in (2), substitute (2) in (4), thus

$$\begin{aligned}
R_m = m p_1 p_2 \int_S \int_S &\left(f_1(x) f_2(x') + f_1(x') f_2(x)\right) \times\\
&\left(P\left(|X-x| > |x'-x|\right)\right)^{m-1} dx' dx. \qquad (5)
\end{aligned}$$

In this paper, an asymptotic evaluation of the finite sample risk R_m was derived for support $S = (\beta,\infty), \beta \ge 0$, which, considering this in two cases for $\beta > 0$ and $\beta = 0$, respectively, i.e. for different supports $S = (\beta,\infty), \beta > 0$ and $S = (0,\infty)$, and using asymptotic expansion by Laplace's method, we find the upper bounds of R_m in these cases. We look at the error estimates, and give some contributions for which we compute upper bounds for the Pareto and exponential distributions as typical.

3 Methods

We start to evaluate the finite-sample risk R_m for a two-class pattern recognition problem for support $S = (\beta,\infty), \beta > 0$.

3.1 Support $S = (\beta,\infty), \beta > 0$:

From (5), the finite-sample risk R_m can be written in the following form:

$$\begin{aligned}
R_{m+1} &= (m+1) p_1 p_2 \int_S \int_S \left(f_1(x) f_2(x') + f_1(x') f_2(x)\right) \times\\
&\quad \left(P\left(|X-x| > |x'-x|\right)\right)^m dx' dx\\
&= (m+1) p_1 p_2 \int_{\beta_1}^{\infty} \int_{\beta_2}^{\infty} \left(f_1(x) f_2(x') + f_1(x') f_2(x)\right) \times\\
&\quad \left(P\left(|X-x| > |x'-x|\right)\right)^m dx' dx\\
&= (m+1) p_1 p_2 \int_{\beta_1}^{k} \int_{\beta_2}^{\infty} \left(f_1(x) f_2(x') + f_1(x') f_2(x)\right) \times\\
&\quad \left(P\left(|X-x| > |x'-x|\right)\right)^m dx' dx\\
&+ (m+1) p_1 p_2 \int_{k}^{\infty} \int_{\beta_2}^{\infty} \left(f_1(x) f_2(x') + f_1(x') f_2(x)\right) \times\\
&\quad \left(P\left(|X-x| > |x'-x|\right)\right)^m dx' dx\\
&= I + J, \qquad (6)
\end{aligned}$$

where the constants $\beta_1, \beta_2 > 0$, and $k = k(m)$ is a constant depending on m,

$$\begin{aligned}
I = (m+1) p_1 p_2 \int_{\beta_1}^{k} \int_{\beta_2}^{\infty} &\left(f_1(x) f_2(x') + f_1(x') f_2(x)\right) \times\\
&\left(P\left(|X-x| > |x'-x|\right)\right)^m dx' dx, \qquad (7)
\end{aligned}$$

and

$$\begin{aligned}
J = (m+1) p_1 p_2 \int_{k}^{\infty} \int_{\beta_2}^{\infty} &\left(f_1(x) f_2(x') + f_1(x') f_2(x)\right) \times\\
&\left(P\left(|X-x| > |x'-x|\right)\right)^m dx' dx. \qquad (8)
\end{aligned}$$

Firstly, we evaluate the asymptotic expansions for I and J. From (7)

$$\begin{aligned}
I &= (m+1) p_1 p_2 \times\\
&\quad \int_{\beta_1}^{k} f_1(x) \left[\int_{\beta_2}^{\infty} f_2(x') \left(P\left(|X-x| > |x'-x|\right)\right)^m dx'\right] dx\\
&+ (m+1) p_1 p_2 \times\\
&\quad \int_{\beta_1}^{k} f_2(x) \left[\int_{\beta_2}^{\infty} f_1(x') \left(P\left(|X-x| > |x'-x|\right)\right)^m dx'\right] dx\\
&= (m+1) p_1 p_2 \int_{\beta_1}^{k} \left[f_1(x) I_1(x) + f_2(x) I_2(x)\right] dx, \qquad (9)
\end{aligned}$$

where

$$I_1(x) = I_1 = \int_{\beta_2}^{\infty} f_2(x') \left(P\left(|X-x| > |x'-x|\right)\right)^m dx', \quad (10)$$

and

$$I_2(x) = I_2 = \int_{\beta_2}^{\infty} f_1(x') \left(P\left(|X-x| > |x'-x|\right)\right)^m dx'. \quad (11)$$

For evaluating I_1 and I_2:

$$\begin{aligned}
I_1 &= \int_{\beta_2}^{\infty} f_2(x') \left(P\left(|X-x| > |x'-x|\right)\right)^m dx' \\
&= \int_{\beta_2}^{x} f_2(x') \left(P\left(|X-x| > |x'-x|\right)\right)^m dx' \\
&+ \int_{x}^{\infty} f_2(x') \left(P\left(|X-x| > |x'-x|\right)\right)^m dx' \\
&= \int_{\beta_2}^{x} f_2(y) \left[P(X<y) + P(X > x+(x-y))\right]^m dy \\
&+ \int_{x}^{\infty} f_2(y) \left[P(X>y) + P(X < x-(y-x))\right]^m dy \\
&= \int_{0}^{x-\beta_2} f_2(x-\rho) \left[P(X<x-\rho) + P(X>x+\rho)\right]^m d\rho \\
&+ \int_{0}^{\infty} f_2(x+\rho) \left[P(X>x+\rho) + P(X<x-\rho)\right]^m d\rho \\
&= \int_{0}^{x-\beta_2} \left(f_2(x-\rho) + f_2(x+\rho)\right) \times \\
&\quad \left[P(X<x-\rho) + P(X>x+\rho)\right]^m d\rho \\
&+ \int_{x-\beta_2}^{\infty} f_2(x+\rho) \left(P(X>x+\rho)\right)^m d\rho \\
&= \int_{0}^{x-\beta_2} \left(f_2(x-\rho) + f_2(x+\rho)\right) \times \\
&\quad \left[1 - (F(x+\rho) - F(x-\rho))\right]^m d\rho \\
&+ \int_{x-\beta_2}^{\infty} f_2(x+\rho) \left[1 - F(x+\rho)\right]^m d\rho \\
&= I_1' + I_1'', \qquad (12)
\end{aligned}$$

where

$$\begin{aligned}
I_1' = \int_{0}^{x-\beta_2} &\left(f_2(x-\rho) + f_2(x+\rho)\right) \times \\
&\left[1 - (F(x+\rho) - F(x-\rho))\right]^m d\rho, \qquad (13)
\end{aligned}$$

and

$$I_1'' = \int_{x-\beta_2}^{\infty} f_2(x+\rho) \left[1 - F(x+\rho)\right]^m d\rho.$$

Similarly

$$\begin{aligned}
I_2 &= \int_{\beta_2}^{\infty} f_1(x') \left(P\left(|X-x| > |x'-x|\right)\right)^m dx' \\
&= \int_{0}^{x-\beta_2} \left(f_1(x-\rho) + f_1(x+\rho)\right) \times \\
&\quad \left[1 - (F(x+\rho) - F(x-\rho))\right]^m d\rho \\
&+ \int_{x-\beta_2}^{\infty} f_1(x+\rho) \left[1 - F(x+\rho)\right]^m d\rho \\
&= I_2' + I_2'', \qquad (14)
\end{aligned}$$

where

$$\begin{aligned}
I_2' = \int_{0}^{x-\beta_2} &\left(f_1(x-\rho) + f_1(x+\rho)\right) \times \\
&\left[1 - (F(x+\rho) - F(x-\rho))\right]^m d\rho, \qquad (15)
\end{aligned}$$

and

$$I_2'' = \int_{x-\beta_2}^{\infty} f_1(x+\rho) \left[1 - F(x+\rho)\right]^m d\rho.$$

Evaluating I_1'' and I_2''

$$\begin{aligned}
I_1'' &= \int_{x-\beta_2}^{\infty} f_2(x+\rho) \left[1 - F(x+\rho)\right]^m d\rho \\
&\le \frac{1}{p_2} \int_{x-\beta_2}^{\infty} f(x+\rho) \left[1 - F(x+\rho)\right]^m d\rho \\
&= \frac{-1}{p_2(m+1)} \int_{x-\beta_2}^{\infty} \frac{d}{d\rho} \left[1 - F(x+\rho)\right]^{m+1} d\rho \\
&= \frac{1}{p_2(m+1)} \left[1 - F(2x-\beta_2)\right]^{m+1} \\
&\le \frac{1}{p_2(m+1)} e^{-(m+1)F(2x-\beta_2)}, \qquad (16)
\end{aligned}$$

similarly

$$I_2'' \le \frac{1}{p_1(m+1)} e^{-(m+1)F(2x-\beta_2)}. \qquad (17)$$

In (13) and (15), we put

$$\begin{aligned}
P(x,\rho) &= -\log\left(1 - (F(x+\rho) - F(x-\rho))\right) \qquad (18) \\
\implies \quad & e^{-mP(x,\rho)} = \left(1 - (F(x+\rho) - F(x-\rho))\right)^m.
\end{aligned}$$

Then

$$\begin{aligned}
I_1' &= \int_{0}^{x-\beta_2} \left(f_2(x-\rho) + f_2(x+\rho)\right) e^{-mP(x,\rho)} d\rho \\
&= \int_{0}^{x-\beta_2} Q(x,\rho) e^{-mP(x,\rho)} d\rho, \qquad (19)
\end{aligned}$$

and

$$\begin{aligned}
I_2' &= \int_{0}^{x-\beta_2} \left(f_1(x-\rho) + f_1(x+\rho)\right) e^{-mP(x,\rho)} d\rho \\
&= \int_{0}^{x-\beta_2} \bar{Q}(x,\rho) e^{-mP(x,\rho)} d\rho, \qquad (20)
\end{aligned}$$

where $Q(x,\rho) = f_2(x-\rho) + f_2(x+\rho)$ and $\bar{Q}(x,\rho) = f_1(x-\rho) + f_1(x+\rho)$.
We now estimate the asymptotic expansions of the integrals of I_1' and I_2' in (19) and (20), respectively, by using Laplace's method for fixed x.

Lemma 3.1.1 Suppose that the functions $P(x,\rho), Q(x,\rho)$ and $\bar{Q}(x,\rho)$ are defined as above, and the following two conditions (i) and (ii) are held:
(i) The probability density functions f_i, for each $i \in \{1,2\}$, have a power series expansion.
(ii) For all sufficiently large m, the integrals of I_1' and I_2' converge absolutely throughout its range. Then

$$\begin{aligned}
I_1' &= \int_{0}^{x-\beta_2} Q(x,\rho) e^{-mP(x,\rho)} d\rho \\
&\sim e^{-mP(x,0)} \sum_{s=0}^{N-1} \frac{\Gamma(s+1) a_s}{m^{s+1}} + O\left(m^{-(N+1)}\right),
\end{aligned}$$

and

$$\begin{aligned}
I_2' &= \int_{0}^{x-\beta_2} \bar{Q}(x,\rho) e^{-mP(x,\rho)} d\rho \\
&\sim e^{-mP(x,0)} \sum_{s=0}^{N-1} \frac{\Gamma(s+1) a_s'}{m^{s+1}} + O\left(m^{-(N+1)}\right),
\end{aligned}$$

where a_s and a_s' are defined through the proof, and Γ refers to the gamma function such that $\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt$.

Proof: We evaluate the asymptotic expansions of I_1', such that the evaluation method proceeds in several steps: First, we estimate the asymptotic expansion of the functions $Q(x,\rho)$ and $P(x,\rho)$, respectively, which, by using the Taylor expansion for the functions $f_2(x+\rho)$ and $f_2(x-\rho)$, respectively, as $\rho \longrightarrow 0$, we obtain:

$$f_2(x+\rho)+f_2(x-\rho) = 2f_2(x)+2f_2''(x)\frac{\rho^2}{2!}+2f_2^{(4)}(x)\frac{\rho^4}{4!} +2f_2^{(6)}(x)\frac{\rho^6}{6!}+\cdots,$$

thus

$$Q(x,\rho) = \sum_{s=0}^{\infty}\frac{2f_2^{(2s)}(x)}{(2s)!}\rho^{2s} = \sum_{s=0}^{\infty} l_s(x)\rho^s, \quad (\rho \longrightarrow 0), \tag{21}$$

where $l_{2n} = \frac{2f_2^{(2n)}(x)}{(2n)!}$ and $l_{2n+1}=0$, for $n=0,1,2,\ldots,$ similarly, use the Taylor expansion for the functions $F(x+\rho)$ and $F(x-\rho)$, respectively, as $\rho \longrightarrow 0$, to obtain the asymptotic expansion of the function $P(x,\rho)$ as follows:

$$F(x+\rho) - F(x-\rho) = \frac{2f(x)\rho}{1!}+\frac{2f''(x)\rho^3}{3!} +\frac{2f^{(4)}(x)\rho^5}{5!}+\cdots = \sum_{s=0}^{\infty}\tilde{P}_s(x)\rho^{s+1}, \tag{22}$$

where $\tilde{P}_{2n}(x) = \frac{2f^{(2n)}(x)}{(2n+1)!}$ and $\tilde{P}_{2n+1}(x)=0$, for $n=0,1,2,\ldots.$ Since $-\log(1-t) = \sum_{s=1}^{\infty}\frac{t^s}{s}$ for $t\in(-1,1)$, and by using (22), we can evaluate the asymptotic expansion for the function $P(x,\rho)$, then

$$\begin{aligned} P(x,\rho) &= \sum_{s=1}^{\infty}\frac{\left(\sum_{\kappa=0}^{\infty}\tilde{P}_\kappa(x)\rho^{\kappa+1}\right)^s}{s} \\ &= \rho\tilde{P}_0(x)+\rho^2\left(\frac{\tilde{P}_0^2}{2}+\tilde{P}_1\right)+\rho^3\left(\frac{\tilde{P}_0^3}{3}+\tilde{P}_0\tilde{P}_1+\tilde{P}_2\right) \\ &+\cdots = \sum_{s=0}^{\infty}P_s(x)\rho^{s+1}, \quad (\rho\longrightarrow 0), \end{aligned} \tag{23}$$

where $P_0(x)=2f(x)$, $P_1(x)=2f^2(x)$, $P_2(x)=\frac{8}{3}f^3(x)+\frac{1}{3}f''(x),\ldots.$ Note that $P(x,0)=0$, and the differentiation of the equation (23) takes the form

$$P'(x,\rho) = \sum_{s=0}^{\infty}(s+1)P_s(x)\rho^s, \quad (\rho\longrightarrow 0). \tag{24}$$

Second, we change the variable of integration.
Suppose that a number u can be found close enough to 0 to ensure that in $(0,u]$, $P'(x,\rho)$ is continuous and nonnegative, and $Q(x,\rho)$ is continuous.
Note that, $P(x,\rho)$ is an increasing function in $(0,u)$, so we may take the function $\upsilon = P(x,\rho)-P(x,0)$ as new integration variable in this interval. Thus, the two functions υ and ρ are continuous of each other and

$$e^{mP(x,0)}\int_0^u Q(x,\rho)e^{-mP(x,\rho)}d\rho = \int_0^Z e^{-m\upsilon}f(\upsilon)d\upsilon,$$

where $Z=P(x,u)-P(x,0)$, $f(\upsilon)=Q(x,\rho)\frac{d\rho}{d\upsilon}=\frac{Q(x,\rho)}{P'(x,\rho)}$.

In the following, we do not discard the term $P(x,0)$ (recall that $P(x,0)=0$) in order to show that the argument is also valid for $P(x,0)>0$.

By Lagrange inversion formula [17], and since $\upsilon = P(x,\rho)-P(x,0) = \sum_{s=0}^{\infty}P_s(x)\rho^{s+1}$, we have an expansion of the form:

$$\rho = \sum_{s=1}^{\infty}C_s\upsilon^s \quad (\upsilon\longrightarrow 0),$$

where $C_s = \frac{1}{s!}\left[\left(\frac{d}{d\rho}\right)^{s-1}(h(\rho))^s\right]_{\rho=0}$, $h(\rho) = \left(\sum_{s=0}^{\infty}P_s(x)\rho^s\right)^{-1}$, and $\upsilon = \frac{\rho}{h(\rho)} = \frac{\rho}{\left(\sum_{s=0}^{\infty}P_s(x)\rho^s\right)^{-1}}$.
Note that

$$C_1 = \frac{1}{P_0},\quad C_2 = -\frac{P_1}{P_0^3},\quad C_3 = \frac{4P_1^2-2P_0P_2}{2P_0^5}.$$

After substituting this result in (21) and (24), and use $f(\upsilon)=\frac{Q(x,\rho)}{P'(x,\rho)}$ we obtain

$$f(\upsilon) = \sum_{s=0}^{\infty}b_s\rho^s \quad (\rho\longrightarrow 0),$$

where

$$b_0 = \frac{l_0}{P_0},\quad b_1 = \frac{l_1P_0-2l_0P_1}{P_0^2},$$
$$b_2 = \frac{P_0^2l_2-3P_0P_2l_0-2P_0P_1l_1+4P_1^2l_0}{P_0^3},\cdots.$$

Then

$$\begin{aligned} f(\upsilon) &= \sum_{s=0}^{\infty}b_s\left(\sum_{i=1}^{\infty}C_i\upsilon^i\right)^s \\ &= \sum_{s=0}^{\infty}b_s\left(C_1\upsilon+C_2\upsilon^2+C_3\upsilon^3+\cdots\right)^s \\ &= b_0+b_1C_1\upsilon+\left(b_1C_2+b_2C_1^2\right)\upsilon^2+\ldots \\ &= \sum_{s=0}^{\infty}a_s\upsilon^s\,(\upsilon\longrightarrow 0), \end{aligned}$$

where $a_0=b_0$, $a_1=b_1C_1$, $a_2=b_1C_2+b_2C_1^2,\cdots$, hence

$$a_0 = \frac{f_2}{f},\ a_1 = -\frac{f_2}{f},$$
$$a_2 = \frac{f_2}{2f}+\frac{f_2''f-f_2f''}{8f^4} = \frac{4f_2f^3+f_2''f-f_2f''}{8f^4},$$
$$a_3 = \frac{-3(4f_2f^3+f_2''f-f_2f'')}{8f^4},\cdots.$$

Third, we compute the asymptotic evaluation on the range of integration $(0,x-\beta_2)$. We start to compute the asymptotic evaluation on the range $(0,u)$.

Let the remainder term $f_N(\upsilon)$ be defined by $f_N(0)=a_N$, for each positive integer N, and

$$f(\upsilon)=\sum_{s=0}^{N-1} a_s\upsilon^s+\upsilon^N f_N(\upsilon)\ \ (\upsilon>0). \tag{25}$$

Then

$$\begin{aligned}
&\int_0^Z e^{-m\upsilon}f(\upsilon)d\upsilon\\
&=\int_0^Z e^{-m\upsilon}\left(\sum_{s=0}^{N-1} a_s\upsilon^s+\upsilon^N f_N(\upsilon)\right)d\upsilon\\
&=\int_0^Z e^{-m\upsilon}\sum_{s=0}^{N-1} a_s\upsilon^s d\upsilon+\int_0^Z e^{-m\upsilon}\upsilon^N f_N(\upsilon)d\upsilon\\
&=\sum_{s=0}^{N-1} a_s\left(\int_0^\infty e^{-m\upsilon}\upsilon^s d\upsilon-\int_Z^\infty e^{-m\upsilon}\upsilon^s d\upsilon\right)\\
&\quad+\int_0^Z e^{-m\upsilon}\upsilon^N f_N(\upsilon)d\upsilon\\
&=\sum_{s=0}^{N-1}\Gamma(s+1)\frac{a_s}{m^{s+1}}-\sum_{s=0}^{N-1}\Gamma(s+1,Zm)\frac{a_s}{m^{s+1}}\\
&\quad+\int_0^Z e^{-m\upsilon}\upsilon^N f_N(\upsilon)d\upsilon\\
&=\sum_{s=0}^{N-1}\frac{\Gamma(s+1)a_s}{m^{s+1}}-A_{N,1}(m)+A_{N,2}(m),
\end{aligned}$$

where

$$A_{N,1}(m)=\sum_{s=0}^{N-1}\Gamma(s+1,Zm)\frac{a_s}{m^{s+1}}, \tag{26}$$

$$A_{N,2}(m)=\int_0^Z e^{-m\upsilon}\upsilon^N f_N(\upsilon)d\upsilon. \tag{27}$$

Since $\Gamma(\alpha,m)\sim e^{-m}m^{\alpha-1}\sum_{s=0}^{\infty}\frac{(\alpha-1)(\alpha-2)\cdots(\alpha-s)}{m^s}$ for fixed α and large m, then

$$A_{N,1}(m)=O\left(\frac{e^{-Zm}}{m}\right). \tag{28}$$

Also, since Z is finite and $f_N(\upsilon)$ is continuous in $[0,Z]$, then $|f_N|$ is bounded, as $\upsilon\longrightarrow 0$, that is $f_N(\upsilon)=O(1)$, it follows that

$$A_{N,2}(m)=\int_0^Z e^{-m\upsilon}\upsilon^N O(1)d\upsilon=O\left(\frac{1}{m^{N+1}}\right). \tag{29}$$

We now derive the asymptotic evaluation on the range $(u,x-\beta_2)$.

For the range $(u,x-\beta_2)$, Suppose M is a value of m such that $I_1'(m)$ is absolutely convergent and define $\eta\equiv inf_{[u,x-\beta_2)}\{P(x,\rho)-P(x,0)\}$.

Since $P(x,0)=0$, and $P(x,\rho)$ strictly increasing in ρ from (23), we thus obtain $(P(x,\rho)-P(x,0))>0$ for all $\rho>0$, so η is positive. For restricting $m\geq M$, we have

$$\begin{aligned}
mP(x,\rho)-mP(x,0)&=(m-M)(P(x,\rho)-P(x,0))\\
&\quad+M(P(x,\rho)-P(x,0))\\
&\geq(m-M)\eta+MP(x,\rho)-MP(x,0),
\end{aligned}$$

then, we obtain

$$\left|e^{mP(x,0)}\int_u^{x-\beta_2}Q(x,\rho)e^{-mP(x,\rho)}d\rho\right|\leq$$
$$e^{-(m-M)\eta+MP(x,0)}\int_u^{x-\beta_2}|Q(x,\rho)|e^{-MP(x,\rho)}d\rho. \tag{30}$$

The proof of lemma 3.1.1 is completed, if we take m large enough to guarantee that the right-hand sides of (28) and (30) are both bounded εm^{-1} for an arbitrary number $\varepsilon>0$ and this is always possible since Z and η are positive. Thus

$$\begin{aligned}
I_1'&\equiv\int_0^{x-\beta_2}Q(x,\rho)e^{-mP(x,\rho)}d\rho\\
&\sim e^{-mP(x,0)}\sum_{s=0}^{N-1}\frac{\Gamma(s+1)a_s}{m^{s+1}}+O\left(m^{-(N+1)}\right),
\end{aligned} \tag{31}$$

similarly

$$\begin{aligned}
I_2'&\equiv\int_0^{x-\beta_2}\bar{Q}(x,\rho)e^{-mP(x,\rho)}d\rho\\
&\sim e^{-mP(x,0)}\sum_{s=0}^{N-1}\frac{\Gamma(s+1)a_s'}{m^{s+1}}+O\left(m^{-(N+1)}\right),
\end{aligned} \tag{32}$$

where $a_0'=\frac{f_1}{f},\ a_1'=-\frac{f_1}{f}$,
$a_2'=\frac{f_1}{2f}+\frac{f_1''f-f_1f''}{8f^4}=\frac{4f_1f^3+f_1''f-f_1f''}{8f^4}$,
$a_3'=-3a_2'=\frac{-3(4f_1f^3+f_1''f-f_1f'')}{8f^4},\cdots$.

Now we evaluate J, from (8)

$$\begin{aligned}
J&=(m+1)p_1p_2\int_k^\infty\int_{\beta_2}^\infty\left(f_1(x)f_2(x')+f_1(x')f_2(x)\right)\times\\
&\quad\left(P(|X-x|>|x'-x|)\right)^m dx'dx\\
&\leq(m+1)p_1p_2\int_k^\infty f_1(x)\left[\int_{\beta_2}^\infty f_2(x')dx'\right]dx\\
&\quad+(m+1)p_1p_2\int_k^\infty f_2(x)\left[\int_{\beta_2}^\infty f_1(x')dx'\right]dx\\
&\leq(m+1)p_1p_2\int_k^\infty f_1(x)dx+(m+1)p_1p_2\int_k^\infty f_2(x)dx\\
&=(m+1)p_1p_2\int_k^\infty(f_1(x)+f_2(x))dx,
\end{aligned} \tag{33}$$

since $P(|X-x|>|x'-x|)\leq 1$, and $f_1(x')$ and $f_2(x')$ are density functions.

Theorem 3.1.2 Assume that the conditions for lemma 3.1.1 are satisfied for all N, that the expansions (21), (23) and (24) are held. Then

$$\begin{aligned}
R_{m+1}\leq p_1p_2\int_{\beta_1}^k&\left[e^{-mP(x,0)}\left(\frac{2f_1f_2}{f}\right.\right.\\
&+\frac{1}{m^2}\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right)\\
&+\frac{1}{m^3}\left(\frac{2f_1f_2}{f}+\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right)\right)\\
&\left.\left.+(m+1)\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\right)\right]dx
\end{aligned}$$

$$+\int_{\beta_1}^{k} f(x)e^{-(m+1)F(2x-\beta_2)}dx$$
$$+(m+1)p_1p_2\int_{k}^{\infty}(f_1(x)+f_2(x))\,dx,$$

where $L_s = p_1p_2\Gamma(s+1)\left(\frac{f_1a_s+f_2a_s'}{f}\right)$, and a_s, a_s' as in lemma 3.1.1.

Proof: From (31) we have

$$\begin{aligned}(m+1)I_1' &= (m+1)\left[e^{-mP(x,0)}\sum_{s=0}^{N-1}\frac{\Gamma(s+1)a_s}{m^{s+1}} + O\left(m^{-(N+1)}\right)\right]\\
&= (m+1)e^{-mP(x,0)}\left[\frac{a_0}{m}+\frac{a_1}{m^2}+\frac{a_2}{m^3} + \sum_{s=3}^{N-1}\Gamma(s+1)\frac{a_s}{m^{s+1}} + O\left(m^{-(N+1)}\right)\right]\\
&= e^{-mP(x,0)}\left[\frac{(m+1)}{m}\frac{f_2}{f} - \frac{(m+1)}{m^2}\frac{f_2}{f} + \frac{2(m+1)}{m^3}\left(\frac{f_2}{2f}+\frac{f_2''f-f_2f''}{8f^4}\right) + (m+1)\sum_{s=3}^{N-1}\Gamma(s+1)\frac{a_s}{m^{s+1}} + O\left(m^{-N}\right)\right]\\
&= e^{-mP(x,0)}\left[\frac{f_2}{f}+\frac{1}{m^2}\left(\frac{f_2''f-f_2f''}{4f^4}\right) + \frac{1}{m^3}\left(\frac{f_2}{f}+\frac{f_2''f-f_2f''}{4f^4}\right) + (m+1)\sum_{s=3}^{N-1}\frac{\Gamma(s+1)a_s}{m^{s+1}} + O\left(m^{-N}\right)\right], \quad (34)\end{aligned}$$

similarly, from (32) we have

$$(m+1)I_2' = e^{-mP(x,0)}\left[\frac{f_1}{f}+\frac{1}{m^2}\left(\frac{f_1''f-f_1f''}{4f^4}\right) + \frac{1}{m^3}\left(\frac{f_1}{f}+\frac{f_1''f-f_1f''}{4f^4}\right) + (m+1)\sum_{s=3}^{N-1}\frac{\Gamma(s+1)a_s'}{m^{s+1}} + O\left(m^{-N}\right)\right]. \quad (35)$$

Using (34) and (16) with (12) we obtain

$$(m+1)I_1 \le e^{-mP(x,0)}\left[\frac{f_2}{f}+\frac{1}{m^2}\left(\frac{f_2''f-f_2f''}{4f^4}\right) + \frac{1}{m^3}\left(\frac{f_2}{f}+\frac{f_2''f-f_2f''}{4f^4}\right) + (m+1)\sum_{s=3}^{N-1}\frac{\Gamma(s+1)a_s}{m^{s+1}} + O\left(m^{-N}\right)\right] + \frac{1}{p_2}e^{-(m+1)F(2x-\beta_2)}, \quad (36)$$

similarly, using (35) and (17) with (14) we obtain

$$(m+1)I_2 \le e^{-mP(x,0)}\left[\frac{f_1}{f}+\frac{1}{m^2}\left(\frac{f_1''f-f_1f''}{4f^4}\right) + \frac{1}{m^3}\left(\frac{f_1}{f}+\frac{f_1''f-f_1f''}{4f^4}\right) + (m+1)\sum_{s=3}^{N-1}\Gamma(s+1)\frac{a_s'}{m^{s+1}} + O\left(m^{-N}\right)\right] + \frac{1}{p_1}e^{-(m+1)F(2x-\beta_2)}. \quad (37)$$

Substituting (36) and (37) in (9) we obtain

$$\begin{aligned}I \le\ & p_1p_2\int_{\beta_1}^{k} f_1(x)\left[e^{-mP(x,0)}\left(\frac{f_2}{f}+\frac{1}{m^2}\left(\frac{f_2''f-f_2f''}{4f^4}\right) + \frac{1}{m^3}\left(\frac{f_2}{f}+\frac{f_2''f-f_2f''}{4f^4}\right) + (m+1)\sum_{s=3}^{N-1}\Gamma(s+1)\frac{a_s}{m^{s+1}} + O\left(m^{-N}\right)\right) + \frac{1}{p_2}e^{-(m+1)F(2x-\beta_2)}\right]dx\\
&+p_1p_2\int_{\beta_1}^{k} f_2(x)\left[e^{-mP(x,0)}\left(\frac{f_1}{f}+\frac{1}{m^2}\left(\frac{f_1''f-f_1f''}{4f^4}\right) + \frac{1}{m^3}\left(\frac{f_1}{f}+\frac{f_1''f-f_1f''}{4f^4}\right) + (m+1)\sum_{s=3}^{N-1}\Gamma(s+1)\frac{a_s'}{m^{s+1}} + O\left(m^{-N}\right)\right) + \frac{1}{p_1}e^{-(m+1)F(2x-\beta_2)}\right]dx. \quad (38)\end{aligned}$$

Then, substituting (38) and (33) in (6) we obtain

$$\begin{aligned}R_{m+1} \le\ & p_1p_2\int_{\beta_1}^{k}\left[e^{-mP(x,0)}\left(\frac{2f_1f_2}{f} + \frac{1}{m^2}\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right) + \frac{1}{m^3}\left(\frac{2f_1f_2}{f}+\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right)\right) + (m+1)\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}} + O\left(m^{-N}\right)\right)\right]dx\\
&+\int_{\beta_1}^{k} f(x)e^{-(m+1)F(2x-\beta_2)}dx\\
&+(m+1)p_1p_2\int_{k}^{\infty}(f_1(x)+f_2(x))\,dx, \quad (39)\end{aligned}$$

where $L_s = p_1p_2\Gamma(s+1)\left(\frac{f_1a_s+f_2a_s'}{f}\right)$.

In the following corollary, the upper bounds of the risk are given in the case $P(x,0)=0$.

Corollary 3.1.3 Under the conditions in Theorem 3.1.2 and for $P(x,0)=0$, i.e., $e^{-mP(x,0)}=1$, we have

$$R_{m+1} \le p_1p_2\int_{\beta_1}^{k}\left[\frac{2f_1f_2}{f}+\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4m^2f^4} + \frac{8f_1f_2f^3+f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4m^3f^4}\right.$$

$$+(m+1)\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\Bigg]dx$$
$$+\int_{\beta_1}^{k} f(x)e^{-(m+1)F(2x-\beta_2)}dx$$
$$+(m+1)p_1p_2\int_{k}^{\infty}(f_1(x)+f_2(x))\,dx, \qquad (40)$$

where $L_s = p_1p_2\Gamma(s+1)\left(\frac{f_1a_s+f_2a_s'}{f}\right)$.

Since $R_\infty = \int_{\beta_1}^{\infty}\frac{2p_1p_2f_1f_2}{f}dx$, then the upper bound of R_{m+1} can be written in the form

$$R_{m+1}\le R_\infty+\frac{p_1p_2}{m^2}\int_{\beta_1}^{k}\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right)dx$$
$$+\frac{p_1p_2}{m^3}\int_{\beta_1}^{k}\left(\frac{8f_1f_2f^3+f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right)dx$$
$$+p_1p_2\int_{\beta_1}^{k}\left((m+1)\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\right)dx$$
$$+\int_{\beta_1}^{k} f(x)e^{-(m+1)F(2x-\beta_2)}dx$$
$$+(m+1)p_1p_2\int_{k}^{\infty}(f_1(x)+f_2(x))\,dx.$$

Error Estimates 3.1.4

We can see from the proof of the lemma 3.1.1, compare Irle and Rizk [10], and Olver [18], that the Nth truncation error of the expansion (31) can be expressed as

$$\int_0^{x-\beta_2}Q(x,\rho)e^{-mP(x,\rho)}d\rho-e^{-mP(x,0)}\sum_{s=1}^{N-1}\Gamma(s+1)\frac{a_s}{m^{s+1}}$$
$$=e^{-mP(x,0)}A_{N,1}(m)+e^{-mP(x,0)}A_{N,2}(m)$$
$$+\int_u^{x-\beta_2}Q(x,\rho)e^{-mP(x,\rho)}d\rho, \qquad (41)$$

where u is a number in $(0,x-\beta_2]$ that satisfies the criteria of lemma 3.1.1, and the terms $A_{N,1}(m)$ and $A_{N,2}(m)$ are defined by (26) and (27), with $Z=P(x,u)-P(x,0)$. Then if $u=x-\beta_2$ and $P(x,x-\beta_2)=\infty$, we obtain $Z=\infty$. Therefore the first error term $e^{-mP(x,0)}A_{N,1}(m)$ in (41) is absent (If the requirement in the proof of the lemma 3.1.1 that u and Z be finite does not apply in (41)). In other cases, we have

$$\Gamma(\alpha,m)\le\frac{e^{-m}m^{\alpha}}{m-\max(\alpha-1,0)},\quad (m>\max(\alpha-1,0)), \qquad (42)$$

where $\Gamma(\alpha,m)$ refers to the complementally incomplete gamma function, and can be taken the form:

$$\Gamma(\alpha,m)=e^{-m}m^{\alpha-1}\left(1+\frac{(\alpha-1)}{m}+\frac{(\alpha-1)(\alpha-2)}{m^2}+\cdots+\frac{(\alpha-1)(\alpha-2)\cdots(\alpha-N+1)}{m^{N-1}}\right)$$
$$+A_N(m),$$

where N is an arbitrary nonnegative integer, and

$$A_N(m)=(\alpha-1)(\alpha-2)\cdots(\alpha-N)\int_m^{\infty}e^{-t}t^{\alpha-N+1}dt.$$

Then

$$|A_N(m)|\le\frac{(\alpha-1)(\alpha-2)\cdots(\alpha-N)e^{-m}m^{\alpha-N}}{m-\alpha+N+1},$$
$$(m>\alpha-N-1>0).$$

For the case $N=0$, we have $\Gamma(\alpha,m)\le\frac{e^{-m}m^{\alpha}}{m-\alpha+1},(\alpha>1,m>\alpha-1)$, and therefore $\Gamma(1,m)\le e^{-m}$ for the special case $\alpha=1$, $m>0$. Then we obtain (42).

Substituting (42) in (26), we obtain

$$\left|e^{-mP(x,0)}A_{N,1}(m)\right|\le\frac{e^{-mP(x,u)}}{Zm-\alpha_N}\sum_{s=1}^{N-1}|a_s|Z^{s+1},\ \left(m>\frac{\alpha_N}{Z}\right),$$

where $Z=P(x,u)-P(x,0)$, and $\alpha_N=\max\{(N-1),0\}$. By the following method, we show that the second error term $e^{-mP(x,0)}A_{N,2}(m)$ is bounded. Let σ_N is a number such that the function $\upsilon^N f_N(\upsilon)$ is majorized by $\left|\upsilon^N f_N(\upsilon)\right|\le|a_N|\upsilon^N e^{\sigma_N\upsilon}$. Then

$$\left|\int_0^Z e^{-m\upsilon}\upsilon^N f_N(\upsilon)d\upsilon\right|\le\left|\int_0^Z|a_N|e^{-(m-\sigma_N)\upsilon}\upsilon^N d\upsilon\right|$$
$$\le\Gamma(N+1)\frac{|a_N|}{(m-\sigma_N)^{N+1}},\quad (m>\sigma_N). \qquad (43)$$

Note that, the best value of σ_N is given by $\sigma_N=\sup_{(0,\infty)}\{\psi_N(\upsilon)\}$, where $\psi_N(\upsilon)=\frac{1}{\upsilon}\ln\left|\frac{\upsilon^N f_N(\upsilon)}{\sigma_N\upsilon^N}\right|$, and the bounded (43) has the property of being asymptotic to the absolute value of the actual error when $m\to\infty$. But the previous approach fails when σ_N is infinite. This happens when $a_N=0$, so we would proceed to a higher value of N at this case. If $a_N\ne0$, then the failure occurs when $\psi_N(\upsilon)$ tends to $+\infty$ as $\upsilon\to0^+$.

But for small υ, we have from (25) that:
$\upsilon^N f_N(\upsilon)=a_N\upsilon^{N+1}+a_{N+1}\upsilon^{N+2}+\cdots$.

Therefore $\psi_N(\upsilon)\sim\frac{a_{N+1}}{a_N}+\left(\frac{a_{N+2}}{a_N}-\frac{a_{N+1}^2}{2a_N^2}\right)\upsilon+\cdots$.

For the tail, the inequality (30) can be used, the integral on the right-hand side being found numerically for a suitably chosen value of M.

Now we evaluate the finite sample risk for support $S=(0,\infty)$.

3.2 Support $S=(0,\infty)$:

Since

$$R_{m+1}=(m+1)p_1p_2\int_S\int_S\left(f_1(x)f_2(x')+f_1(x')f_2(x)\right)\times$$
$$\left(P\left(|X-x|>|x'-x|\right)\right)^m dx'dx.$$

Then R_{m+1} for support $S=(0,\infty)$ takes the form:

$$R_{m+1}=(m+1)p_1p_2\int_0^{\infty}\int_0^{\infty}\left(f_1(x)f_2(x')+f_1(x')f_2(x)\right)\times$$
$$\left(P\left(|X-x|>|x'-x|\right)\right)^m dx'dx$$

$$= (m+1)p_1p_2\int_0^k\int_0^\infty \left(f_1(x)f_2(x')+f_1(x')f_2(x)\right)\times \left(P\left(|X-x|>|x'-x|\right)\right)^m dx'dx + (m+1)p_1p_2\int_k^\infty\int_0^\infty \left(f_1(x)f_2(x')+f_1(x')f_2(x)\right)\times \left(P\left(|X-x|>|x'-x|\right)\right)^m dx'dx.$$

Then, by using the same method in subsection 3.1 with $\beta_1=\beta_2=0$ we obtain the following theorem:

Theorem 3.2.1

Let the conditions of lemma 3.1.1 be satisfied for all N, that the expansions (21), (23) and (24) are held. Then

$$R_{m+1}\le p_1p_2\int_0^k\Bigg[e^{-mP(x,0)}\left(\frac{2f_1f_2}{f} + \frac{1}{m^2}\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right) + \frac{1}{m^3}\left(\frac{2f_1f_2}{f}+\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right)\right)\right) + (m+1)\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\Bigg]dx + \int_0^k f(x)e^{-(m+1)F(2x)}dx + (m+1)p_1p_2\int_k^\infty \left(f_1(x)+f_2(x)\right)dx,$$

where $L_s = p_1p_2\Gamma(s+1)\left(\frac{f_1a_s+f_2a'_s}{f}\right)$, $k=k(m)$ is a constant depending on m and a_s, a'_s are defined as in lemma 3.1.1.

Proof: This is immediate from theorem 3.1.2 with $\beta_1=\beta_2=0$.

Corollary 3.2.2 Under the conditions in Theorem 3.2.1 and for $P(x,0)=0$, we have

$$R_{m+1}\le p_1p_2\int_0^k\Bigg[\left(\frac{2f_1f_2}{f} + \frac{1}{m^2}\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right) + \frac{1}{m^3}\left(\frac{2f_1f_2}{f}+\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right)\right)\right) + (m+1)\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\Bigg]dx + \int_0^k f(x)e^{-(m+1)F(2x)}dx + (m+1)p_1p_2\int_k^\infty \left(f_1(x)+f_2(x)\right)dx, \quad (44)$$

where $L_s = p_1p_2\Gamma(s+1)\left(\frac{f_1a_s+f_2a'_s}{f}\right)$, $k=k(m)$ is a constant depending on m and a_s, a'_s are defined as in lemma 3.1.1.

4 Results and discussion

In preceding section, an asymptotic evaluation of the finite sample risk R_m for distributions having unbounded supports $S=(\beta,0),\beta>0$ and $S=(0,\infty)$ are derived ((40) and (44), respectively) by using the Laplace's method, such that we found an exact integral expression for I'_1 and I'_2 in the form $\int_S Qe^{-mP}$, where Q and P are nonnegative functions, we then applied Laplace's method for large m. We discussed the error bounds for this case. Some of the problems that came to us are related to find an upper bounded of the risk R_m for distributions having unbounded, such that the integration at the boundary of the domain of the integration (at infinity in our cases) is not easy to handle, so we divided the integrals I_1 and I_2 into two parts I'_1, I''_1 and I'_2, I''_2, respectively, and appropriate values of $k(m)$ were chosen to obtain a good approximation of the upper bounds.

In the following examples, we find upper bounds of the risk R_m of the Pareto and exponential distributions as typical for distributions having unbounded supports $S=(\beta,0),\beta>0$ and $S=(0,\infty)$, respectively to explain the theoretical results for our work, and evaluate some values of R_m numerically for those in tables 1-3, for which we use the first three terms in the expansion of the risk R_m to calculate. For the Pareto distribution, we choose $k(m)=m^{\frac{2}{\alpha}}$, where $\alpha>0$ denotes a shape parameter, also, choose $k(m)=m$ for the exponential distribution.

Example 4.1: (Pareto distribution)

We will derive the upper bounds on the finite sample risk of the Pareto distribution as typical, i.e., the case $S=(\beta,\infty),\beta>0$, so that it can often be used as a modal for heavy-tailed data.

Assume that $f_i(x)=\frac{\alpha_i\beta_i^{\alpha_i}}{x^{\alpha_i+1}}$ for $x\ge\beta_i, i=1,2$ respectively, such that the shape parameters $\alpha_i>0$ and the scale parameters $\beta_i>0$ for $i=1,2$. Define $k=k(m)$ be a constant depending on m and without loss of generality $f_1(x)=\max\{f_1(x),f_2(x)\}$.

We evaluate the risk R_m of Pareto distribution for: (i) Same shapes and different scales, (ii) Same shapes and scales, (iii) Different shapes and scales, and (iv) Same scales and different shapes.

(i) Same shapes and different scales.

Taking $\alpha_1=\alpha_2=\alpha$ say, in equation (40), we get $f_1f_2''f-2f_1f_2f''+f_1''f_2f=0$, then

$$R_{m+1}\le 2p_1p_2\int_{\beta_1}^k\frac{f_1f_2}{f}dx+\frac{p_1p_2}{m^3}\int_{\beta_1}^k\frac{2f_1f_2}{f}dx + \int_{\beta_1}^k f(x)e^{-(m+1)F(2x-\beta_2)}dx + (m+1)p_1p_2\int_k^\infty\left(f_1(x)+f_2(x)\right)dx + (m+1)p_1p_2\int_{\beta_1}^k\left(\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\right)dx$$

$$\leq \frac{p_1p_2\alpha\beta_1^\alpha\beta_2^\alpha}{(p_1\beta_1^\alpha+p_2\beta_2^\alpha)}\left(2+\frac{1}{m^3}\right)\left[\frac{1}{\beta_1^\alpha}-\frac{1}{k^\alpha}\right] + \frac{\left[e^{-(m+1)F(\beta_1)}-e^{-(m+1)F(k)}\right]}{(m+1)} + \frac{(m+1)p_1p_2\left[\beta_1^\alpha+\beta_2^\alpha\right]}{k^\alpha} + (m+1)p_1p_2\int_{\beta_1}^k\left(\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\right)dx,$$

where $\int_{\beta_1}^k \frac{f_1f_2}{f}dx = \frac{\alpha\beta_1^\alpha\beta_2^\alpha}{(p_1\beta_1^\alpha+p_2\beta_2^\alpha)}\left[\frac{1}{\beta_1^\alpha}-\frac{1}{k^\alpha}\right]$,

$$\int_{\beta_1}^k f(x)e^{-(m+1)F(2x-\beta_2)}dx \leq \int_{\beta_1}^k f(x)e^{-(m+1)F(x)}dx = \frac{-1}{(m+1)}\left[e^{-(m+1)F(x)}\right]_{\beta_1}^k = \frac{\left[e^{-(m+1)F(\beta_1)}-e^{-(m+1)F(k)}\right]}{(m+1)},$$

(Since $f_1=\max\{f_1,f_2\} \Longrightarrow \beta_1>\beta_2$, hence we have $F(2x-\beta_2)\geq F(x) \Longrightarrow e^{-F(x-\beta_2)}\leq e^{-F(x)}$)
$\int_k^\infty (f_1(x)+f_2(x))\,dx = \frac{[\beta_1^\alpha+\beta_2^\alpha]}{k^\alpha}$.

Here, we can choose a suitable $k=k(m)=m^{2/\alpha}$.
(ii) Same shapes and scales.
The result of this case is immediate from (i) by substituting $\beta_1=\beta_2=\beta$, say.
(iii) Different shapes and scales.
Let $\alpha_2>\alpha_1$, we obtain

$$R_{m+1} \leq 2p_1p_2\int_{\beta_1}^k\frac{f_1f_2}{f}dx + \frac{p_1p_2}{4m^2}\int_{\beta_1}^k\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{f^4}\right)dx + \frac{p_1p_2}{m^3}\int_{\beta_1}^k\left(\frac{2f_1f_2}{f}+\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right)dx + \int_{\beta_1}^k f(x)e^{-(m+1)F(2x-\beta_2)}dx + (m+1)p_1p_2\int_k^\infty(f_1(x)+f_2(x))\,dx + (m+1)p_1p_2\int_{\beta_1}^k\left(\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\right)dx$$

$$\leq p_2\beta_2^{\alpha_2}\left(2+\frac{1}{m^3}\right)\left[\frac{1}{\beta_1^{\alpha_2}}-\frac{1}{k^{\alpha_2}}\right] + \frac{p_1p_2(M_1-M_2)}{4m^2(2\alpha_1-\alpha_2)}\left(1+\frac{1}{m}\right)\left[k^{(2\alpha_1-\alpha_2)}-\beta_1^{(2\alpha_1-\alpha_2)}\right] + \frac{p_1p_2M_3}{4m^2(3\alpha_1-2\alpha_2)}\left(1+\frac{1}{m}\right)\left[k^{(3\alpha_1-2\alpha_2)}-\beta_1^{(3\alpha_1-2\alpha_2)}\right] + \frac{\left[e^{-(m+1)F(\beta_1)}-e^{-(m+1)F(k)}\right]}{(m+1)} + (m+1)p_1p_2\left[\frac{\beta_1^{\alpha_1}}{k^{\alpha_1}}+\frac{\beta_2^{\alpha_2}}{k^{\alpha_2}}\right] + (m+1)p_1p_2\int_{\beta_1}^k\left(\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\right)dx,$$

where

$$\int_{\beta_1}^k\frac{f_1f_2}{f}dx \leq \int_{\beta_1}^k\frac{f_1f_2}{p_1f_1}dx = \frac{\beta_2^{\alpha_2}}{p_1}\left[\frac{1}{\beta_1^{\alpha_2}}-\frac{1}{k^{\alpha_2}}\right],$$

$$\int_k^\infty(f_1(x)+f_2(x))\,dx = \left(\frac{\beta_1}{k}\right)^{\alpha_1}+\left(\frac{\beta_2}{k}\right)^{\alpha_2},$$

$\int_{\beta_1}^k f(x)e^{-(m+1)F(2x-\beta_2)}dx$ evaluate as in part (i), and

$$\int_{\beta_1}^k\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{f^4}dx = \int_{\beta_1}^k\frac{f_1f_2''}{f^3}dx+\int_{\beta_1}^k\frac{f_1''f_2}{f^3}dx-2\int_{\beta_1}^k\frac{f_1f_2f''}{f^4}dx \leq \int_{\beta_1}^k\frac{f_2''}{f_1^2}dx+\int_{\beta_1}^k\frac{f_1''f_2}{f_1^3}dx -2p_1\int_{\beta_1}^k\frac{f_2f_1''}{f_1^3}dx-2p_2\int_{\beta_1}^k\frac{f_2f_2''}{f_1^3}dx$$
$$= \frac{M_1\left(k^{(2\alpha_1-\alpha_2)}-\beta_1^{(2\alpha_1-\alpha_2)}\right)}{(2\alpha_1-\alpha_2)} - \frac{M_2\left(k^{(2\alpha_1-\alpha_2)}-\beta_1^{(2\alpha_1-\alpha_2)}\right)}{(2\alpha_1-\alpha_2)} - \frac{M_3\left(k^{(3\alpha_1-2\alpha_2)}-\beta_1^{(3\alpha_1-2\alpha_2)}\right)}{(3\alpha_1-2\alpha_2)},$$

where $f=p_1f_1+p_2f_2\geq p_1f_1 \Longrightarrow \frac{1}{f}\leq\frac{1}{p_1f_1}$, $f\leq f_1 \Longrightarrow \frac{-1}{f}\leq\frac{-1}{f_1}$, $F(2x-\beta_2)\geq F(x) \Longrightarrow e^{-F(2x-\beta_2)}\leq e^{-F(x)}$ such that $x-\beta_2>0$ from (12), and $M_1=\frac{\alpha_2(\alpha_2+1)(\alpha_2+2)\beta_2^{\alpha_2}}{\alpha_1^2\beta_1^{2\alpha_1}}$, $M_2=\frac{(1-2p_1)\alpha_2(\alpha_1+1)(\alpha_1+2)\beta_2^{\alpha_2}}{\alpha_1^2\beta_1^{2\alpha_1}}$, and $M_3=\frac{2p_2\alpha_2^2(\alpha_2+1)(\alpha_2+2)\beta_2^{2\alpha_2}}{\alpha_1^3\beta_1^{3\alpha_1}}$ Here we can choose a suitable $k=k(m)=m^{\frac{3}{2\alpha_1}}$.
(iv) Same scales and different shapes.
The result of this case is immediate from (iii) by substituting $\beta_1=\beta_2=\beta$, say.

Example 4.2. (Exponential distribution)
Assume that $f_i(x)=\lambda_ie^{-\lambda_ix}$ for $x,\lambda_i>0, i=1,2$, and let $k=k(m)$ be a constant depending on m and without loss of generality $f_1(x)=\max\{f_1(x),f_2(x)\}$ for every point x such that $\lambda_2>\lambda_1$. From (44) R_{m+1} takes the form:

$$R_{m+1} \leq 2p_1p_2\int_0^k\frac{f_1f_2}{f}dx + \frac{p_1p_2}{4m^2}\int_0^k\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{f^4}\right)dx + \frac{p_1p_2}{m^3}\int_0^k\left(\frac{2f_1f_2}{f}+\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{4f^4}\right)dx + \int_0^k f(x)e^{-(m+1)F(2x)}dx$$

$$+(m+1)p_1p_2\int_k^{\infty}(f_1(x)+f_2(x))\,dx$$
$$+(m+1)p_1p_2\int_0^k\left(\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\right)dx$$
$$\le p_2\left(2+\frac{1}{m^3}\right)\left[1-e^{-\lambda_2k}\right]$$
$$+\frac{p_1p_2\lambda_2}{4m^2p_1^3(\lambda_2-2\lambda_1)}\left(1+\frac{1}{m}\right)\left(\frac{1}{\lambda_1^2}-\frac{(1-2p_1)}{m}\right)\times$$
$$\left[1-e^{-(2\lambda_2-3\lambda_1)k}\right]$$
$$+\frac{2p_1p_2^2\lambda_2^4}{4m^2p_1^3\lambda_1^3(2\lambda_2-3\lambda_1)}\left(1+\frac{1}{m}\right)\left[1-e^{-(2\lambda_2-3\lambda_1)k}\right]$$
$$+\frac{\left[1-e^{-(m+1)F(k)}\right]}{(m+1)}+(m+1)p_1p_2\left[e^{-\lambda_1k}+e^{-\lambda_2k}\right]$$
$$+(m+1)p_1p_2\int_0^k\left(\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\right)dx,$$

where

$$\int_0^k\frac{f_1f_2}{f}dx\le\int_0^k\frac{f_1f_2}{p_1f_1}dx$$
$$=\frac{1}{p_1}\int_0^k\lambda_2e^{-\lambda_2x}dx=\frac{1}{p_1}\left(1-e^{-\lambda_2k}\right),$$

$$\int_0^k\left(\frac{f_1f_2''f-2f_1f_2f''+f_1''f_2f}{f^4}\right)dx$$
$$=\int_0^k\frac{f_1f_2''}{f^3}dx+\int_0^k\frac{f_1''f_2}{f^3}dx-2\int_0^k\frac{f_1f_2f''}{f^4}dx$$
$$\le\frac{1}{p_1^3}\left(\int_0^k\frac{f_2''}{f_1^2}dx+\int_0^k\frac{f_1''f_2}{f_1^3}dx-2p_1\int_0^k\frac{f_2f_1''}{f_1^3}dx\right.$$
$$\left.-2p_2\int_0^k\frac{f_2f_2''}{f^3}dx\right)$$
$$=\frac{1}{p_1^3}\left(\left[\frac{\lambda_2^3}{\lambda_1^2}e^{-(\lambda_2-2\lambda_1)x}\right]_0^k-(1-2p_1)\left[\lambda_2e^{-(\lambda_2-2\lambda_1)x}\right]_0^k\right.$$
$$\left.-2p_2\left[\frac{\lambda_2^4}{\lambda_1^3}e^{-(2\lambda_2-3\lambda_1)x}\right]_0^k\right)$$
$$=\frac{\lambda_2^3\left[1-e^{-(\lambda_2-2\lambda_1)k}\right]}{p_1^3\lambda_1^2(\lambda_2-2\lambda_1)}$$
$$-\frac{(1-2p_1)\lambda_2\left[1-e^{-(\lambda_2-2\lambda_1)k}\right]}{p_1^3(\lambda_2-2\lambda_1)}$$
$$-\frac{2p_2\lambda_2^4\left[1-e^{-(2\lambda_2-3\lambda_1)k}\right]}{p_1^3\lambda_1^3(2\lambda_2-3\lambda_1)},$$

$$\int_0^k f(x)e^{-(m+1)F(2x)}dx\le\int_0^k f(x)e^{-(m+1)F(x)}dx$$
$$=\frac{-1}{(m+1)}\left[e^{-(m+1)F(x)}\right]_0^k=\frac{\left[1-e^{-(m+1)F(k)}\right]}{(m+1)},$$
$$\int_k^{\infty}(f_1(x)+f_2(x))\,dx=e^{-\lambda_1k}+e^{-\lambda_2k}.$$

Since $f=p_1f_1+p_2f_2\ge p_1f_1\Longrightarrow\frac{1}{f}\le\frac{1}{p_1f_1},f\le f_1\Longrightarrow\frac{-1}{f}\le\frac{-1}{f_1}$ (such that $f_1(x)=\max\{f_1(x),f_2(x)\}$), and $F(2x)\ge F(x)\Longrightarrow e^{-F(2x)}\le e^{-F(x)}$.

Here we can choose a suitable $k=k(m)=\log\left(m^{\frac{3}{2\lambda_1}}\right)$.

Note that, in the case $\lambda_1=\lambda_2=\lambda$ say, we have $f_1f_2''f-2f_1f_2f''+f_1''f_2f=0$, then equation (44) takes the form:

$$R_{m+1}\le 2p_1p_2\int_0^k\frac{f_1f_2}{f}dx+\frac{p_1p_2}{m^3}\int_0^k\frac{2f_1f_2}{f}dx$$
$$+\int_0^k f(x)e^{-(m+1)F(2x)}dx$$
$$+(m+1)p_1p_2\int_k^{\infty}(f_1(x)+f_2(x))\,dx$$
$$+(m+1)p_1p_2\int_0^k\left(\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\right)dx$$
$$\le 2p_1p_2\left(1+\frac{1}{m^3}\right)\left[1-e^{-\lambda k}\right]$$
$$+\frac{\left[1-e^{-(m+1)F(k)}\right]}{(m+1)}+2(m+1)p_1p_2e^{-\lambda k}$$
$$+(m+1)p_1p_2\int_0^k\left(\sum_{s=3}^{N-1}\frac{L_s}{m^{s+1}}+O\left(m^{-N}\right)\right)dx.$$

We can choose a suitable $k=k(m)=m$.

5 Conclusions

We derived an asymptotic evaluation of the finite sample risk R_m for distributions having unbounded supports by using Laplace methods, and evaluated upper bounds for the risk R_m of nearest neighbor of Pareto and exponential distributions as typical.

The examples 4.1 and 4.2 presented in section 4 (see also tables 1-3) indicate that the finite sample of the nearest neighbor risk R_m approaches its infinite-sample limit R_∞, and thus, the asymptotic series expansions presented in the theorems 3.1.2 and 3.2.1, and corollaries 3.1.3 and 3.2.2 helps us to use the finite sample (for large m) instead of the infinite sample.

References

[1] E. Fix and J. Hodges, Discriminatory analysis, nonparametric discrimination, consistency properties, In nearest neighbor classification techniques, B. Dasarathy, editor, IEEE Computer Society Press, Los Alamitos, CA, 32-39 (1991).

[2] E. Fix and J. Hodges, Discriminatory analysis, small samples editor, performance, In nearest neighbor classification techniques, B. Dasarathy, editor, IEEE Computer Society Press, Los Alamitos, CA, 40-56 (1991).

[3] T. M. Cover and P. E. Hart, Nearest neighbor pattern classification, IEEE Transactions on Information Theory, 13, 21-27 (1967).

[4] T. M. Cover, Rates of convergence for neatest neighbor procedures, In Proceedings of the Hawaii International Conference on System Sciences, Honolulu, HI, 413-415 (1968).
[5] T. J. Wagner, Convergence of the nearest neighbor rule, IEEE Trans. Inform. Theory, 17, 566-571 (1971).
[6] J. Fritz, Distribution-free exponential error bound for nearest neighbor pattern classification, IEEE Trans. Inform. Theory, 21, 552-557 (1975).
[7] K. Fukunaga and D. M. Hummels, Bias of nearest neighbor estimates, IEEE Trans. Pattern Anal. Mach. Intel., 9, 103-112 (1987).
[8] D. Psaltis, R. Snapp and S. S. Venkatesh, On the finite sample performance ofthe nearest neighbor classifier, IEEE Transactions on Information Theory, 40, 820-837 (1994).
[9] R. R. Snapp and S. S. Venkatesh, Asymptotic expansion of the *k* nearest neighbor risk, Ann. Statist, 26, 850-878 (1998).
[10] A. Irle and M. Rizk, On the risk of nearest neighbor rules, Ph.D. thesis, Kiel University, Germany, (2004).
[11] B. Dasarathy, Nearest neighbor classification techniques, IEEE Computer Society Press, Los Alamitos, CA (1991).
[12] L. Devroye, L. Györfi and G. Lugosi, A probabilistic theory of pattern recognition, Springer-Verlag, New York (1996).
[13] K. Fukunaga, Introduction to statistical pattern recognition, Second Edition, Academic Press, (1990).
[14] G. Biau and L. Devroye, Lectures on the nearest neighbor method, Springer, (2015).
[15] M. Döring, L. Györfi and H. Walk, Rate of convergence of k-nearest-neighbor classification rule, journal of machine learning research, 18, 1-16 (2018).
[16] P. Zhao and L. Lai, Analysis of KNN information estimators for smooth distributions, IEEE Trans. Inform. Theory, 66 (6), 3798-3826 (2019).
[17] N. G. De Bruijn, Asymptotic methods in analysis, Wiley, New York (1958).
[18] F. Olver, Asymptotics and special functions, Academic Press, INC (1974).

Tables

Table 1: Numerical values of $R_{(m+1)}$ and R_∞ for some different values of β_2, when $\alpha = 1, \beta_1 = 1$ and $p_1 = 0.2$ for Pareto distribution in the cases (i) and (ii).

m	$\beta_2 = 0.2$	$\beta_2 = 0.5$	$\beta_2 = 1$
1.0×10^2	0.1796992889	0.2690641333	0.32320116
1.0×10^3	0.1779681930	0.2669068801	0.32032000
1.0×10^4	0.1777969779	0.2666906664	0.32003200
1.0×10^6	0.1777779700	0.2666669067	0.32000032
1.0×10^9	0.1777777782	0.2666666668	0.32000000
R_∞	0.1777777778	0.2666666667	0.32

Table 2: Numerical values of $R_{(m+1)}$ and R_∞ for some different values of β_2, when $\alpha = 1, \beta_1 = 1$ and $p_1 = 0.5$ for Pareto distribution in the cases (i) and (ii).

m	$\beta_2 = 0.2$	$\beta_2 = 0.5$	$\beta_2 = 1$
1.0×10^2	0.169680083	0.3370876667	0.505000250
1.0×10^3	0.166966801	0.3337083752	0.500500001
1.0×10^4	0.166696668	0.3333708338	0.500050000
1.0×10^6	0.166666967	0.3333337083	0.500000500
1.0×10^9	0.166666667	0.3333333337	0.500000500
R_∞	0.166666667	0.3333333333	0.5

Table 3: Numerical values of $R_{(m+1)}$ and R_∞ for some different values of p_1, when $\lambda_1 = \lambda_2 = \lambda = 1$ for exponential distribution.

m	$p_1 = 0.2$	$p_1 = 0.3$	$p_1 = 0.5$
1.0×10^1	0.411372837	0.511468293	0.5915446569
1.0×10^2	0.329901310	0.429901410	0.5099014901
1.0×10^3	0.320999002	0.420999002	0.5009990015
1.0×10^4	0.320099990	0.420099990	0.5000999900
1.0×10^6	0.320001000	0.420001000	0.5000010000
1.0×10^9	0.320000001	0.420000001	0.5000000010
R_∞	0.32	0.42	0.5

Mohamed Mahmoud Rizk received his B.Sc. and M.Sc. in pure mathematics from faculty of science, Menoufia University, Egypt, and his the Ph.D. degree in mathematics from institute of mathematics and natural science, Christian-Albrechts University of Kiel, Germany. His current research interests are in mathematical statistics, probability theory, pattern recognition.