On the Risk of Nearest Neighbor Rules

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen
Fakultät
der Christian-Albrechts-Universität
zu Kiel

vorgelegt von
Mohamed Mahmoud Rizk
Kiel 2004

Referent: Prof. Dr. Albrecht Irle
Korreferent: Prof. Dr. Uwe Rösler
Tag der mündlichen Prüfung: 6. Juli 2004
Zum Druck genehmigt: 8. Juli 2004

Der Dekan

Acknowledgements

I would like to express my deep thanks and gratitude to Prof. Dr. A. Irle, not only for his suggesting the point of the work but also for his instructive supervision, guidance, valuable advice and his valuable suggestions expending much time and effort through the progress of this thesis.

I would like to express my special thanks and gratitude to Prof. Dr. U. Rösler and the staff members for their kind cooperation.

At last, but not least, my special thanks go to my wife Hanan for her patience, understanding and continuous support during my work, without which I would not be able to carry out this work, and my children Merrihan and Mahmoud.

Mohamed Rizk

To my mother

and

to the memory of my father

Contents

Kurzfassung

Das Nächste-Nachbar-(NN) Verfahren ist eine der nichtparametrischen Klassifikationtechniken, wobei ein nichtparametrischer Klassifikator auf keinen Annahmen hinsichtlich der Struktur der zugrundeliegenden Verteilung beruht. Das NN-Verfahren wurde zuerst durch Fix und Hodges [9], [10] studiert.

Cover und Hart [2] bewiesen, daß unter bestimmten Bedingungen an die Verteilungen der erwartete Fehler R_m des NN-Verfahrens gegen einen Wert R_∞ konvergiert, der zwischen dem Bayes-Risiko $R^\star$ und dem doppelten Bayes-Risiko liegt. Cover [3] untersuchte die Eigenschaft des NN-Klassifikators für den ein-dimensionalen Fall mit beschränkten Träger und Mischungsdichte $f \geq c > 0$ und fand heraus, dass R_m asymptotisch durch $O\left(m^{-2}\right)$ beschränkt ist, wobei m der Umfang der Trainingsfolge ist. Psaltis, Snapp und Venkatesh [19] leiteten eine asymptotische Darstellung von R_m unter der euklidischen Metrik für ein Zweikategorienproblem ab. Dieses wurde ausgeweitet auf weitere Metriken von Snapp und Venkatesh [20]. Kulkarni und Posner [16] studierten die Rate der Konvergenz für nächste Nachbarschätzung mittels der Überdeckungzahlen total beschränkter Mengen und fanden obere Schranken der Konvergenzrate für Verteilungen mit Trägern auf total beschränkten Teilmengen eines separable metrischen Raumes, ausgedrückt durch deren Überdeckungzahlen.

Es gibt eine Fülle von Konvergenzresultaten anderer Ausrichtung für NN-Verfahren: siehe die Sammlung von Dasarathy [4] und die Monographie von Devroye, Györfi und Lugosi [6].

Der Hauptinhalt dieser These wird wie folgt zusammengefaßt: Begründet auf einem exakten Ausdruck für das Risiko, wird eine asymptotische Auswertung des bedingten Risikos $R_m(x)$ für unbeschränkten Träger gefunden. Dann werden die Probleme und die Moglichkeiten bei der Integration dieser asymptotischen Entwicklung behandelt. Anschließend wird eine alternative asymptotische Entwicklung mit der Methode von Laplace gegeben. Schließlich werden NN-Abstände für unbeschränkte Träger behandelt.

1 Introduction and Model

Pattern recognition is about inference on the unknown nature of an observation. More formally, an observation is a $d-$dimensional vector x, and the unknown nature of the observation is called a class. It is denoted by ϑ and takes values in a finite set $M = \{1, 2, ..., C\}$. Suppose that we have a function $\delta : R^d \to \{1, 2, ..., C\}$ where $\delta(x)$ represents one's guess of ϑ given x. This mapping is called a classifier. Our classifier errs on x if $\delta(x) \neq \vartheta$.

That is, pattern recognition considers the following basic situation: A random variable (X, θ) consists of an observed pattern $X \in R^d$ from which we wish to infer the unobservable class θ. This class belongs to the known finite set $M = \{1, 2, ..., C\}$. The probability of error for a classifier δ is $P(\delta(X) \neq \theta)$.

If the joint distribution of $(X\ ,\ \theta)$ is known, then we may compute the Bayes classifier $\delta^\star$ which is defined by

$$\delta^\star(x) = \arg\min_{i=1,...,C}\ P(\theta \neq i | X = x)$$

The problem of finding $\delta^\star$ is called the Bayes problem and the resulting probability of misclassification is usually called the Bayes risk.

In general the joint distribution of (X, θ) will be unknown, and we have a training sequence $Z_m = \left(\left(X^{(1)}, \theta^{(1)}\right), \left(X^{(2)}, \theta^{(2)}\right), ..., \left(X^{(m)}, \theta^{(m)}\right)\right)$ at our disposal, where patterns and corresponding classes are observed. We shall assume that $\left(X^{(1)}, \theta^{(1)}\right), \left(X^{(2)}, \theta^{(2)}\right), ..., \left(X^{(m)}, \theta^{(m)}\right)$, the data, stem from a sequence of independent identically distributed (iid) random pairs with the same distribution as (X, θ).

1.1 Nearest Neighbor Procedure

The nearest neighbor rule is one of the nonparametric classification techniques, where a nonparametric classifier does not rely on any assumptions concerning the structure of the underlying distribution.

Let $\left(X^{(1)}, \theta^{(1)}\right), \left(X^{(2)}, \theta^{(2)}\right), ..., \left(X^{(m)}, \theta^{(m)}\right)$ be independent identically distributed random variables taking values in $R^d \times \{1, 2, ..., C\}$. Let (X, θ) be another independent sample of the same distribution, such that X is an observed pattern and it is desired to estimate θ. The nearest neighbor rule assigns X to a class $\theta^{(i)}$ with the property

$$\left\|X - X^{(i)}\right\| \leq \left\|X - X^{(j)}\right\| \qquad \text{for all } i \neq j,$$

using suitable tie-breaking.

1.1.1 Definition

The nearest neighbor procedure assigns any input feature vector to the class given by the label θ' of the nearest reference vector.

1.2 Literature Review

The nearest neighbor rule was first studied by Fix and Hodges [9] and [10]. Cover and Hart [2] proved that under certain conditions on the distribution the expected error of the nearest neighbor rule converges, as the sample size tends to infinity, to a value R_∞ which lies between the Bayes error $R^\star$ (the minimum probability of error over all decision rules) and twice the Bayes error, i.e. $R^\star \leq R_\infty \leq 2R^\star(1 - R^\star)$. Cover [3] investigated the finite-sample performance of the nearest neighbor classifier for the one-dimensional case with bounded support and mixture density $f \geq c > 0$ and found under some additional conditions that the bias of the nearest neighbor error from its asymptotic value is bounded by $O\left(m^{-2}\right)$ where m is the sample size.

Fukunaga and Hummels [11] studied the rate of convergence of the above bias in d-dimensional feature space using a series of nonrigorous approximations based on a second-order Taylor series expansion, they obtained the heuristic estimate $R_m \sim R_\infty + B\frac{\Gamma(m+1)}{\Gamma\left(m+1+\frac{2}{d}\right)}$, where Γ is the gamma function and B is a distribution-dependent constant. This approximation indicates $m^{-2/d}$ as the rate of convergence of R_m to R_∞.

Psaltis, Snapp and Venkatesh [19] derived an asymptotic representation of the finite sample risk of a nearest neighbor classifier under the Euclidean metric for a two-class problem. They assume bounded support and that the class-conditional distributions are absolutely continuous with densities admitting uniform asymptotic expansions, that the mixture density satisfies $f \geq c > 0$

and that one of the class-conditional densities vanishes close to the boundary of the support. They proved that $R_m \sim R_\infty + \sum_{k=2}^{\infty} c_k m^{-k/d} \quad (m \longrightarrow \infty)$, where the coefficients c_k are distribution-dependent constants independent of the sample size m. This was extended to other metrics in Snapp and Venkatesh [20].

Kulkarni and Posner [16] studied the rate of convergence for nearest neighbor estimation in terms of the covering numbers of totally bounded sets. They found upper bounds on the convergence rate for distributions with support on a totally bounded subset of a separable metric space in terms of the covering numbers of this support.

There is a wealth of consistency results in different directions available for nearest neighbor rules; see the collection of Dasarathy [4] and the monograph by Devroye, Györfi and Lugosi [6].

1.3 Results of the work

The main contents of this thesis are summarized as follows: Based on an exact integral expression for the risk, we find an asymptotic evaluation of the conditional risk $R_m(x)$ for unbounded support. Then the problems and the applicability of integrating these asymptotic expansions are discussed. This is followed by an alternative asymptotic approach using Laplace's method. Finally nearest neighbor distances are treated, again for unbounded support.

In the next section we give the integral expressions for $R_m(x)$ and R_m in the form

$$R_m(x) = P(\theta' \neq \theta \,|X = x\,) = \tfrac{p_1 p_2 f_1(x)}{f(x)} I + \tfrac{p_1 p_2 f_2(x)}{f(x)} J,$$

where

$$I = I(x) = \textstyle\int_S f_2(x') m P(|X - x| > |x' - x|)^{m-1} dx',$$

$$J = J(x) = \textstyle\int_S f_1(x') m P(|X - x| > |x' - x|)^{m-1} dx',$$

hence

$$R_m = p_1 p_2 \textstyle\int_S \int_S m(P(|X - x| > |x' - x|))^{m-1}$$

$$\cdot \, (f_1(x) f_2(x') + f_1(x') f_2(x)) \, dx' dx,$$

where the densities f_l are those of the class-conditional distributions which are assumed absolutely continuous, for $l = 1, 2$, $f = p_1 f_1 + p_2 f_2$ denotes the mixture density, S being its support in R^d.

Chapter 2 evaluates the probability of error conditioned on the event that $X = x$ (m-sample conditional risk $R_m(x)$) for different supports S in R^1 by using partial integration and presents a general representation for $R_m(x)$ when X has support in R^d.

Chapter 3 discusses the problem of integrating $R_m(x)$ with respect to x to obtain R_m. We find that, in example like the normal and exponential distribution, the integrals diverge. This seems to be typical for the case of unbounded support. For the triangular distributions as an example for the case of bounded support we find that the integrals exist and the rate of convergence of R_m to R_∞ is $O(m^{-2})$, which is in accord with Cover's result [3].

Chapter 4 presents another method to evaluate $R_m(x)$ by using the asymptotic expansion by Laplace's method. We derive an exact integral expression for I and J in the form $\int_S g e^{-mh}$, where g and h are nonnegative functions. For large m, as in typical Laplace integrals, most of the contribution to the integral arises from a neighborhood of the point where h has a minimum. We represent g and h as asympototic power series in a neighborhood of this minimum, and then the integral itself may be represented as an asymptotic power series in reciprocal powers of m. We look at the error estimates for this case.

In chapter 5 we study the rates of convergence of nearest neighbor classification in terms of metric covering numbers of the underlying space, present an upper bound on the expected nearest neighbor distance for all distributions with support on a totally bounded subset of a separable metric space in terms of the covering numbers of the support (see [16]). We then give some contributions in the case of unbounded support for which we find upper and lower bounds for the normal and exponential distributions as typical.

1.4 The Finite Sample Risk

In this section we shall derive an exact integral expression for the finite-sample risk R_m.

1.4.1 Definition

The risk of the nearest neighbor procedure from a training sequence of size m is defined by

$$R(\delta_{1,m}) = P(\delta_{1,m}(X, Z_m) \neq \theta)\cdot$$

We can write this in the simple form $R_m = P\left(\theta' \neq \theta\right)$.

The finite-sample risk R_m can be written in integral form by taking the expectation of the probability of the event $\theta' \neq \theta$ conditioned on the training sequence and the test feature vector. Then the asymptotic risk is given by the following Lemma, compare [19].

For this Lemma, we suppose that the class-conditional distributions F_l are absolutely continuous with corresponding densities f_l, for each $l \in M$. Let $f = \sum_{l=1}^{C} p_l f_l$ denote the mixture density, and let S be its support in R^d. Introduce the notation $B(\rho, x) \equiv \left\{x' \in R^d : \|x - x'\| \leq \rho\right\}$ for the closed ball of radius ρ at x. We shall assume $C = 2$, i.e. $M = \{1, 2\}$, in the following.

1.4.2 Lemma

$$R_m = p_1 p_2 \int_S \int_S m(P(|X - x| > |x' - x|))^{m-1}$$

$$\cdot\left(f_1(x) f_2(x') + f_1(x') f_2(x)\right) dx' dx$$

Proof:
Let X' denote the nearest neighbor feature vector in the training sequence $Z_m = \left(\left(X^{(1)}, \theta^{(1)}\right), \left(X^{(2)}, \theta^{(2)}\right), ..., \left(X^{(m)}, \theta^{(m)}\right)\right)$ that is closest to the random test vector X, and let θ' be the class label associated with X'. Then from the definition (1.4.1)

$$R_m = P(\theta' \neq \theta) = \int_S P\left(\theta' \neq \theta \,|X = x\right) f(x) dx, \tag{1.4.1}$$

where $P\,(\theta' \neq \theta\,|x\,)$ denotes the probability of error conditioned on the event that $X = x$.

Taking expectation with respect to the value of the nearest neighbor of x, we hence obtain:

$$P(\theta' \neq \theta\,|X = x\,) = \int_S P\,(\theta' \neq \theta\,|X' = x'\,, X = x)\, f_m(x' \mid x)\, dx', \qquad (1.4.2)$$

where $f_m\,(x'\,|x\,)$ denotes the conditional density of X' given $X = x$. That is, the event $X' = x'$ occurs if one of the training sequence $X^{(j)}$ assumes the value x' and every other feature vector $X^{(k)}$, $k \neq j$, assumes a value outside $B(\rho, x)$ with $\rho = |x' - x|$. We thus obtain:

$$f_m\,(x'\,|x\,) = P(\,one\,of\,the\,X_j's \in B(\rho, x),\, all\,others \notin B(\rho, x))f(x')$$

$$= \sum_{j=1}^{m} \left(\prod_{k \neq j} P\left[X^{(k)} \notin B(|x' - x|\,, x)\right]\right) f(x')$$

$$= m(1 - P(X \in B(|x' - x|\,, x))^{m-1} f(x'),$$

where X is a feature vector in R^d drawn from the mixture distribution $F(x)$. Thus we can write $f_m\,(x'\,|x\,)$ in the form:

$$f_m\,(x'\,|x\,) = m(P(|X - x| > |x' - x|))^{m-1} f(x'). \qquad (1.4.3)$$

Furthermore:

$$P\,(\theta' \neq \theta\,|X' = x'\,, X = x)$$

$$= P\,(\theta = 1, \theta' = 2\,|X' = x', X = x\,) + P\,(\theta = 2, \theta' = 1\,|X' = x', X = x\,)$$

$$= P(\theta = 1\,|X = x)P(\theta' = 2\,|X' = x') + P(\theta = 2\,|X = x)\; P(\theta' = 1\,|X' = x')$$

$$= \tfrac{p_1 p_2}{f(x) f(x')}\,(f_1(x) f_2(x') + f_1(x') f_2(x)). \qquad (1.4.4)$$

Substituting (1.4.3) and (1.4.4) in (1.4.2) yields

$$P(\theta' \neq \theta \,|X = x\,) = \int_S m(P(|X - x| > |x' - x|))^{m-1}$$

$$\cdot \tfrac{p_1 p_2}{f(x)f(x')} \left(f_1(x)f_2(x') + f_1(x')f_2(x)\right) f(x')dx'$$

$$= \tfrac{p_1 p_2}{f(x)} \int_S m(P(|X - x| > |x' - x|))^{m-1}$$

$$\cdot \left(f_1(x)f_2(x') + f_1(x')f_2(x)\right) dx'. \qquad (1.4.5)$$

Then

$$R_m = p_1 p_2 \int_S \int_S \, m(P(|X - x| > |x' - x|))^{m-1}$$

$$\cdot \left(f_1(x)f_2(x') + f_1(x')f_2(x)\right) dx'dx. \qquad (1.4.6)$$

1.4.3 Definition

We denote the probability of error conditioned on the event that $X = x$ by $R_m(x)$, that is $R_m(x) = P(\theta' \neq \theta \,|X = x\,)$.

From equation (1.4.5)

$$R_m(x) = \tfrac{p_1 p_2 f_1(x)}{f(x)} \int_S f_2(x') m P(|X - x| > |x' - x|)^{m-1} dx'$$

$$+ \tfrac{p_1 p_2 f_2(x)}{f(x)} \int_S f_1(x') m P(|X - x| > |x' - x|)^{m-1} dx'$$

Put

$$I = I(x) = \int_S f_2(x') m P(|X - x| > |x' - x|)^{m-1} dx', \qquad (1.4.7)$$

$$J = J(x) = \int_S f_1(x') m P(|X - x| > |x' - x|)^{m-1} dx' \qquad (1.4.8)$$

Then

$$R_m(x) = P(\theta' \neq \theta \,|X = x\,) = \tfrac{p_1 p_2 f_1(x)}{f(x)} I + \tfrac{p_1 p_2 f_2(x)}{f(x)} J \qquad (1.4.9)$$

$$= \frac{p_2}{1 + \frac{p_2 f_2(x)}{p_1 f_1(x)}} \, I(x) + \frac{p_1}{1 + \frac{p_1 f_1(x)}{p_2 f_2(x)}} \, J(x). \qquad (1.4.10)$$

2 The Asymptotic Evaluation of $R_m(x)$

In this chapter we evaluate the probability of error conditioned on the event that $X = x$ for a two-class pattern recognition problem for different supports S in R^1.

2.1 Support $S = (-\infty, \infty)$:

Firstly, we evaluate the asymptotic expansions for I and J in (1.4.9).

2.1.1 Lemma

Let $x \in R^d$, $x \in S$. Assume that the densities f_i are $k-$times differentiable and $(f(x-\rho) + f(x+\rho)) > 0$ for all $\rho > 0$. Define

$$q_\circ(x,\rho) = \frac{f_2(x-\rho)+f_2(x+\rho)}{f(x-\rho)+f(x+\rho)} \quad \text{and} \quad q_k(x,\rho) = \frac{q'_{k-1}(x,\rho)}{f(x-\rho)+f(x+\rho)} \text{ for } k \geq 1,$$

and

$$\overline{q}_\circ(x,\rho) = \frac{f_1(x-\rho)+f_1(x+\rho)}{f(x-\rho)+f(x+\rho)} \quad \text{and} \quad \overline{q}_k(x,\rho) = \frac{\overline{q}'_{k-1}(x,\rho)}{f(x-\rho)+f(x+\rho)} \text{ for } k \geq 1.$$

Then

$$I = q_\circ(x,0) + \frac{1}{m+1}q_1(x,0) + \frac{1}{(m+1)(m+2)}q_2(x,0)$$
$$+ ... + \frac{1}{(m+1)(m+2)...(m+k)}q_k(x,0) + \frac{1}{(m+1)(m+2)...(m+k)}I_{k+1}$$

and

$$J = \overline{q}_\circ(x,0) + \frac{1}{m+1}\overline{q}_1(x,0) + \frac{1}{(m+1)(m+2)}\overline{q}_2(x,0)$$
$$+ ... + \frac{1}{(m+1)(m+2)...(m+k)}\overline{q}_k(x,0) + \frac{1}{(m+1)(m+2)...(m+k)}J_{k+1}$$

where

$$I_{k+1} = I_{k+1}(m,x) = \int_0^\infty q'_k(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+k}\,d\rho \qquad k = 1,2,3,...$$

and

$$J_{k+1} = J_{k+1}(m,x) = \int_0^\infty \overline{q}'_k(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+k}\,d\rho \qquad k = 1,2,3,...$$

Proof:
From equation (1.4.7)

$$
\begin{aligned}
I &= \int_{-\infty}^{\infty} f_2(x') m P(|X-x| > |x'-x|)^{m-1} dx' \\
&= m \int_{-\infty}^{x} f_2(x') P(|X-x| > |x'-x|)^{m-1} dx' \\
&\qquad + m \int_{x}^{\infty} f_2(x') P(|X-x| > |x'-x|)^{m-1} dx' \\
&= m \int_{-\infty}^{x} f_2(z) \left[P(X < z) + P(X > x + (x-z))\right]^{m-1} dz \\
&\qquad + m \int_{x}^{\infty} f_2(z) \left[P(X > z) + P(X < x - (z-x))\right]^{m-1} dz \\
&= m \int_{0}^{\infty} f_2(x-\rho) \left[P(X < x-\rho) + P(X > x+\rho)\right]^{m-1} d\rho \\
&\qquad + m \int_{0}^{\infty} f_2(x+\rho) \left[P(X < x-\rho) + P(X > x+\rho)\right]^{m-1} d\rho \\
&= m \int_{0}^{\infty} (f_2(x-\rho) + f_2(x+\rho)) \left[P(X < x-\rho) + P(X > x+\rho)\right]^{m-1} d\rho \\
&= - \int_{0}^{\infty} \tfrac{f_2(x-\rho)+f_2(x+\rho)}{f(x-\rho)+f(x+\rho)} \tfrac{d}{d\rho} \left[P(X < x-\rho) + P(X > x+\rho)\right]^{m} d\rho \\
&= - \int_{0}^{\infty} q_\circ(x,\rho) \tfrac{d}{d\rho} \left[P(X < x-\rho) + P(X > x+\rho)\right]^{m} d\rho,
\end{aligned}
$$

where $\quad q_\circ(x,\rho) = \frac{f_2(x-\rho)+f_2(x+\rho)}{f(x-\rho)+f(x+\rho)}.$ (2.1.1)

Let $\quad u = q_\circ(x,\rho), \qquad dv = \frac{d}{d\rho}\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m} d\rho,$

$\quad du = q'_\circ(x,\rho)\, d\rho, \qquad v = \left[P(X < x-\rho) + P(X > x+\rho)\right]^{m}.$

Then, by partial integration

$$
\begin{aligned}
I &= \int_{0}^{\infty} q'_\circ(x,\rho) \left[P(X < x-\rho) + P(X > x+\rho)\right]^{m} d\rho \\
&\qquad - \left[q_\circ(x,\rho)(P(X < x-\rho) + P(X > x+\rho))^{m}\right]_0^{\infty} \\
&= \int_{0}^{\infty} q'_\circ(x,\rho) \left[P(X < x-\rho) + P(X > x+\rho)\right]^{m} d\rho \\
&\qquad - \{(q_\circ(x,\infty) \left[(P(X < x-\infty) + P(X > x+\infty)\right]^{m}) \\
&\qquad\qquad - (q_\circ(x,0) \left[(P(X < x) + P(X > x)\right]^{m})\}
\end{aligned}
$$

$= \int_0^\infty q_\circ'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^m \,d\rho - (0 - q_\circ(x,0))$

$= q_\circ(x,0) + \int_0^\infty q_\circ'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^m \,d\rho$

$= q_\circ(x,0) + I_1,$ (2.1.2)

where $\quad I_1 = I_1(x) = \int_0^\infty q_\circ'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^m \,d\rho.$

Now, we evaluate I_1.

$I_1 = \int_0^\infty q_\circ'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^m \,d\rho$

$= \frac{-1}{m+1} \int_0^\infty \frac{q_\circ'(x,\rho)}{f(x-\rho)+f(x+\rho)} \frac{d}{d\rho}\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1} \,d\rho$

$= \frac{-1}{m+1} \int_0^\infty q_1(x,\rho) \frac{d}{d\rho}\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1} \,d\rho,$

where $\;q_1(x,\rho) = \frac{q_\circ'(x,\rho)}{f(x-\rho)+f(x+\rho)}.$ (2.1.3)

We integrate by parts with

$u = q_1(x,\rho), \;\; dv = \frac{d}{d\rho}\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1} \,d\rho$, then

$I_1 = \frac{1}{m+1} \int_0^\infty q_1'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1} \,d\rho$

$-\frac{1}{m+1}\,[q_1(x,\rho)(P(X < x-\rho) + P(X > x+\rho))^{m+1}]_0^\infty$

$= \frac{1}{m+1} \int_0^\infty q_1'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1} \,d\rho + \frac{1}{m+1} q_1(x,0)$

$= \frac{1}{m+1} q_1(x,0) + \frac{1}{m+1} I_2,$ (2.1.4)

where $\;I_2 = \int_0^\infty q_1'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1} \,d\rho.$

Similarly,

$I_2 = \frac{-1}{m+2} \int_0^\infty \frac{q_1'(x,\rho)}{f(x-\rho)+f(x+\rho)} \frac{d}{d\rho}\,[P(X < x-\rho) + P(X > x+\rho)]^{m+2} \,d\rho,$

where $\;q_2(x,\rho) = \frac{q_1'(x,\rho)}{f(x-\rho)+f(x+\rho)}.$ (2.1.5)

Now, we evaluate I_2.

$$I_2 = \frac{1}{m+2}\int_0^\infty q_2'(x,\rho)\left[P(X<x-\rho)+P(X>x+\rho)\right]^{m+2}d\rho$$

$$-\frac{1}{m+2}\left[q_2(x,\rho)(P(X<x-\rho)+P(X>x+\rho))^{m+2}\right]_0^\infty$$

$$= \frac{1}{m+2}\int_0^\infty q_2'(x,\rho)\left[P(X<x-\rho)+P(X>x+\rho)\right]^{m+2}d\rho + \frac{1}{m+2}q_2(x,0)$$

$$= \frac{1}{m+2}q_2(x,0) + \frac{1}{m+2}I_3,$$

where $I_3 = I_3(x) = \int_0^\infty q_2'(x,\rho)\left[P(X<x-\rho)+P(X>x+\rho)\right]^{m+2}d\rho.$

By repeating this procedure, we obtain an asymptotic expansion for $I(x)$ in the form:

$$I = q_\circ(x,0) + \frac{1}{m+1}q_1(x,0) + \frac{1}{(m+1)(m+2)}q_2(x,0) + ... + \frac{1}{(m+1)(m+2)...(m+k)}q_k(x,0)$$

$$+\frac{1}{(m+1)(m+2)...(m+k)}\int_0^\infty q_k'(x,\rho)\left[P(X<x-\rho)+P(X>x+\rho)\right]^{m+k}d\rho$$

$$= q_\circ(x,0) + \frac{1}{m+1}q_1(x,0) + \frac{1}{(m+1)(m+2)}q_2(x,0) + ... + \frac{1}{(m+1)(m+2)...(m+k)}q_k(x,0)$$

$$+\frac{1}{(m+1)(m+2)...(m+k)}I_{k+1},$$

where

$$I_{k+1} = I_{k+1}(m,x) = \int_0^\infty q_k'(x,\rho)\left[P(X<x-\rho)+P(X>x+\rho)\right]^{m+k}d\rho \quad k = 1,2,3,....$$

Similarly,

$$J = \overline{q}_\circ(x,0) + \frac{1}{m+1}\overline{q}_1(x,0) + \frac{1}{(m+1)(m+2)}\overline{q}_2(x,0) + ... + \frac{1}{(m+1)(m+2)...(m+k)}\overline{q}_k(x,0)$$

$$+\frac{1}{(m+1)(m+2)...(m+k)}J_{k+1},$$

where

$$J_{k+1} = J_{k+1}(m,x) = \int_0^\infty \overline{q}_k'(x,\rho)\left[P(X<x-\rho)+P(X>x+\rho)\right]^{m+k}d\rho \quad k = 1,2,3,....$$

Now we show that under suitable conditions $I_{k+1}(m) \to 0$ when $m \to \infty$ for all $k \geq 2$.

2.1.2 Lemma

Assume that there exist j, l such that the following conditions are satisfied

(i) $\left| q'_k(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^j \right|$ is bounded for ρ and

(ii) $[P(X < x-\rho) + P(X > x+\rho)]^l$ is integrable for ρ.

Then $\quad I_{k+1} \to 0 \quad$ when $m \to \infty$,

where $\quad I_{k+1} = \int_0^\infty q'_k(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+k}\,d\rho$.

Proof:

$$
\begin{aligned}
I_{k+1} = \int_0^\infty q'_k(x,\rho)\,&[P(X < x-\rho) + P(X > x+\rho)]^j \\
&\cdot [P(X < x-\rho) + P(X > x+\rho)]^l \\
&\cdot [P(X < x-\rho) + P(X > x+\rho)]^{m+k-j-l}\,d\rho
\end{aligned}
$$

$$
\begin{aligned}
|I_{k+1}| = \Big| \int_0^\infty q'_k(x,\rho)\,&[P(X < x-\rho) + P(X > x+\rho)]^j \\
&\cdot [P(X < x-\rho) + P(X > x+\rho)]^l \\
&\cdot [P(X < x-\rho) + P(X > x+\rho)]^{m+k-j-l}\,d\rho \Big| \\
\le \int_0^\infty \Big| q'_k(x,\rho)\,&[P(X < x-\rho) + P(X > x+\rho)]^j \Big| \\
&\cdot [P(X < x-\rho) + P(X > x+\rho)]^l \\
&\cdot [P(X < x-\rho) + P(X > x+\rho)]^{m+k-j-l}\,d\rho \\
\le \sup_{\rho'} \Big| q'_k(x,\rho')\,&[P(X < x-\rho') + P(X > x+\rho')]^j \Big| \\
&\cdot \int_0^\infty [P(X < x-\rho) + P(X > x+\rho)]^l \\
&\cdot [P(X < x-\rho) + P(X > x+\rho)]^{m+k-j-l}\,d\rho \\
\le C \int_0^\infty f(\rho,m)\,&d\rho,
\end{aligned}
$$

where C is a constant, and

$$
\begin{aligned}
f(\rho, m) = {} & [P(X < x-\rho) + P(X > x+\rho)]^{l} \\
& \cdot [P(X < x-\rho) + P(X > x+\rho)]^{m+k-j-l} \\
& \leq f(\rho) = [P(X < x-\rho) + P(X > x+\rho)]^{l}
\end{aligned}
$$

We have $f(\rho, m) \to 0$ for all ρ when $m \to \infty$, and from condition (ii) $\int_0^\infty f(\rho)\, d\rho < \infty$. This implies $\int_0^\infty f(\rho, m)\, d\rho \to 0$ by the dominated convergence theorem. Then $I_{k+1} \to 0$ when $m \to \infty$.

Similarly, we show that $J_{k+1} \to 0$ when $m \to \infty$.

2.1.3 Lemma

Assume that there exist j, l such that the following are satisfied

(i) $\left| \overline{q}'_k(x, \rho)\, [P(X < x-\rho) + P(X > x+\rho)]^{j} \right|$ is bounded for ρ and

(ii) $[P(X < x-\rho) + P(X > x+\rho)]^{l}$ is integrable for ρ.

Then $J_{k+1} \to 0$ when $m \to \infty$,

where $J_{k+1} = \int_0^\infty \overline{q}'_k(x, \rho)\, [P(X < x-\rho) + P(X > x+\rho)]^{m+k}\, d\rho$.

Proof:
As the proof of Lemma 2.1.2.

Now we give an example for normal distribution to show that the conditions (i) and (ii) in the above Lemmas are satisfied when the support is unbounded.

2.1.4 Example

Let $f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2}}$, $f_2(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-b)^2}{2}}$ be two densities for normal distributions with prior probabilities p_1, p_2 such that $p_1 + p_2 = 1$, and $f = p_1 f_1 + p_2 f_2$.

Firstly we show that there exist j such that

$$\left| q_2'(x,\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^j \right| \text{ is bounded for } \rho.$$

Since $q_2(x,\rho) = \frac{q_1'}{g}$ then

$$q_2'(x,\rho) = \left(\frac{q_1'}{g}\right)' = \frac{h'''g - 6h''g'g^2 - h'g''g^2 - 7hgg'g'' + 3h'(g')^2 g + 3h(g')^3}{g^6}$$

where $h(x,\rho) = f_2(x+\rho) + f_2(x-\rho)$, and $g(x,\rho) = f(x+\rho) + f(x-\rho)$.

Substituting this functions in above equation, then $q_2'(x,\rho)$ can be bounded in the following form where we assume $x > 0$:

$$q_2'(x,\rho) \leq \frac{(a_\circ + a_1(x+\rho) + a_2(x+\rho)^2 + a_3(x+\rho)^3)\left(e^{-\frac{((x+\rho)-a)^2}{2}}\right)^4}{\left(e^{-\frac{((x+\rho)-a)^2}{2}}\right)^6} \leq \frac{C}{\left(e^{-\frac{((x+\rho)-a)^2}{2}}\right)^3},$$

where $a_\circ$, a_1, a_2, a_3, and C are constants. Then

$$\begin{aligned}
&\left| q_2'(x,\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^j \right| \\
&\quad \leq C\, e^{\frac{3((x+\rho)-a)^2}{2}} \left[P(X < x-\rho) + P(X > x+\rho)\right]^j \\
&\quad \leq C_1\, e^{\frac{C_2\rho^2}{2}} \left[P(X < x-\rho) + P(X > x+\rho)\right]^j.
\end{aligned}$$

But

$$\begin{aligned}
\left[P(X < x-\rho) + P(X > x+\rho)\right]^j &\leq C_3 \left[\left(e^{-\frac{((x-\rho)+a)^2}{2}}\right) + \left(e^{-\frac{((x+\rho)+a)^2}{2}}\right)\right]^j \\
&\leq C_4 e^{-\frac{C_5 j\rho^2}{2}}
\end{aligned}$$

That is, we can find j such that

$$\left| q_2'(x,\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^j \right| \text{ is bounded for } \rho.$$

Now we show that there exists l such that $\left[P(X < x-\rho) + P(X > x+\rho)\right]^l$ is integrable for ρ.

Since $[P(X < x-\rho) + P(X > x+\rho)] = [P(|X-x| > \rho)]$

$$= P\left(e^{t|X-x|} > e^{t\rho}\right) \leq \frac{Ee^{t|X-x|}}{e^{t\rho}} \leq \frac{C_6}{e^{t\rho}},$$

where $t > 0$, then

$$\int_0^\infty \left([P(X < x-\rho) + P(X > x+\rho)]\right)^l d\rho \leq C_6^l \int_0^\infty e^{-lt\rho} d\rho = \frac{C_6^l}{lt} = \frac{C_7}{l}.$$

Thus there exists l such that

$[P(X < x-\rho) + P(X > x+\rho)]^l$ is integrable for ρ.

Here $C_1, C_2, ..., C_7$ are constants only depending on x, a, b.

2.1.5 Corollary

$$I = q_\circ(x,0) + \sum_{k=2}^\infty \frac{q_k(x,0)}{\prod_{n=1}^k (m+n)}, \tag{2.1.6}$$

$$J = \overline{q}_\circ(x,0) + \sum_{k=2}^\infty \frac{\overline{q}_k(x,0)}{\prod_{n=1}^k (m+n)} \tag{2.1.7}$$

under the condition that $\frac{I_{k+1}}{\prod_{n=1}^k (m+n)}$ and $\frac{J_{k+1}}{\prod_{n=1}^k (m+n)}$ tend to zero when $k \to \infty$,

where

$$I_{k+1} = \int_0^\infty q_k'(x,\rho) \left[P(X < x-\rho) + P(X > x+\rho)\right]^{m+k} d\rho,$$

$$J_{k+1} = \int_0^\infty \overline{q}_k'(x,\rho) \left[P(X < x-\rho) + P(X > x+\rho)\right]^{m+k} d\rho,$$

and $q_k(x,0), \overline{q}_k(x,0)$ are defined as in Lemma 2.1.1 when $\rho = 0$. We note that $q_1(x,0) = 0$ and $\overline{q}_1(x,0) = 0$.

Proof:
This is immediate from Lemmas 2.1.1 and 2.1.3.

2.1.6 Theorem

Let the conditions of Lemmas 2.1.1-2.1.3 and Corollary 2.1.5 be satisfied. Then

$$R_m(x) = P(\theta' \neq \theta \,|X = x\,) = \tfrac{2p_1p_2f_1(x)f_2(x)}{f^2(x)} + \textstyle\sum_{k=2}^{\infty} \frac{\eta_k(x,0)}{\prod_{n=1}^{k}(m+n)} \tag{2.1.8}$$

where $\eta_k(x,0) = \frac{p_1p_2\,f_1(x)}{f(x)} q_k(x,0) + \frac{p_1p_2\,f_2(x)}{f(x)} \overline{q}_k(x,0)$, and $q_k(x,0)$, $\overline{q}_k(x,0)$ are defined as in Lemma 2.1.1 when $\rho = 0$.

Proof:
This is immediate from the above results. By substituting (2.1.6) and (2.1.7) in to (1.4.9) we obtain

$$\begin{aligned} R_m(x) &= \tfrac{p_1p_2\,f_1(x)}{f(x)} \left\{ q_\circ(x,0) + \textstyle\sum_{k=2}^{\infty} \frac{q_k(x,0)}{\prod_{n=1}^{k}(m+n)} \right\} \\ &\quad + \tfrac{p_1p_2\,f_2(x)}{f(x)} \left\{ \overline{q}_\circ(x,0) + \textstyle\sum_{k=2}^{\infty} \frac{\overline{q}_k(x,0)}{\prod_{n=1}^{k}(m+n)} \right\} \\ &= \tfrac{2\,p_1p_2f_1(x)f_2(x)}{f^2(x)} + \textstyle\sum_{k=2}^{\infty} \frac{\eta_k(x,0)}{\prod_{n=1}^{k}(m+n)}. \end{aligned}$$

2.1.7 Multidimensional Case

Now we present a general representation for $R_m(x)$ when X has support in R^d. Since $R_m(x)$ take the following form:

$$\begin{aligned} R_m(x) &= \tfrac{p_1p_2f_1(x)}{f(x)} \textstyle\int_S f_2(x')mP(|X-x| > |x'-x|)^{m-1}dx' \\ &\quad + \tfrac{p_1p_2f_2(x)}{f(x)} \textstyle\int_S f_1(x')mP(|X-x| > |x'-x|)^{m-1}dx' \\ &= \tfrac{p_1p_2f_1(x)}{f(x)} I + \tfrac{p_1p_2f_2(x)}{f(x)} J, \end{aligned}$$

where

$$I = I(x) = \textstyle\int_S f_2(x')mP(|X-x| > |x'-x|)^{m-1}dx',$$

$$J = J(x) = \textstyle\int_S f_1(x')mP(|X-x| > |x'-x|)^{m-1}dx'.$$

Using the function $H(\rho) = mP(|X-x| > \rho)^{m-1}$, we have

$$\begin{aligned} I(x) &= E_2H\left(|X-x|\right) = \textstyle\int H\left(|X-x|\right)dP_2 = \int_0^{\infty} H\left(\rho\right) P_2^{|X-x|}\,d\rho \\ &= \textstyle\int_0^{\infty} H\left(\rho\right) f_2^{|X-x|}\,d\rho, \end{aligned}$$

similarly,
$J(x) = E_1 H\left(|X-x|\right) = \int H\left(|X-x|\right) dP_1 = \int_0^{\infty} H\left(\rho\right) P_1^{|X-x|}\, d\rho$

$$= \int_0^{\infty} H\left(\rho\right) f_1^{|X-x|}\, d\rho,$$

where E_1, E_2 denote the expectations with respect to the densities f_1, f_2 respectively. Then $R_m(x)$ take the following form :

$$R_m(x) = \frac{p_1 p_2 f_1(x)}{f(x)} \int_0^{\infty} H\left(\rho\right) f_2^{|X-x|}\, d\rho + \frac{p_1 p_2 f_2(x)}{f(x)} \int_0^{\infty} H\left(\rho\right) f_1^{|X-x|}\, d\rho$$

We note that we may expand as in Lemma 2.1.1 for any dimension d if X has unbounded support in R^d with f_i and f replaced by $f_i^{|X-x|}$ and $f^{|X-x|}$ the densities of $|X_i - x|$ and $|X-x|$ respectively, where $i = 1, 2$.

2.1.8 Example

Consider $X = (X_1, X_2)$ having a joint density $f(x_1, x_2) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x_1-\mu_1)^2+(x_2)^2}{2\sigma^2}}$.

Define $Z = |X-x| = \sqrt{\left(X_1 - x_1\right)^2 + \left(X_2 - x_2\right)^2} = \sqrt{X_1'^2 + X_2'^2}$

where $X_1' = X_1 - x_1$, and $X_2' = X_2 - x_2$.

Then the region $\triangle D_z$ of the plane such that $z < \sqrt{X_1'^2 + X_2'^2} < z + dz$ is a circular ring with inner radius z and thickness dz. With

$x_1' = z\cos\theta$, $x_2' = z\sin\theta$, we have $dx_1' dx_2' = z dz d\theta$, it follows that

$f_{|X-x|}(z) = f_Z(z) = \int\int_{\triangle D_z} f(x_1', x_2')\, dx_1' dx_2'$

$$= \frac{1}{2\pi\sigma^2} \int_0^{2\pi} e^{-\frac{(z\cos\theta-\mu)^2+(z\sin\theta)^2}{2\sigma^2}}\, z dz d\theta$$

Hence

$$f_{|X-x|}(z) = \frac{z}{2\pi\sigma^2} e^{-\frac{(z+\mu^2)}{2\sigma^2}} \int_0^{2\pi} e^{\frac{z\mu\cos\theta}{\sigma^2}} d\theta = \frac{z}{\sigma^2} I_\circ\left(\frac{z\mu}{\sigma^2}\right) e^{-\frac{(z+\mu^2)}{2\sigma^2}}, \qquad z > 0,$$

where $I_\circ\left(x\right) = \frac{1}{2\pi} \int_0^{2\pi} e^{x\cos\theta} d\theta = \sum_{n=0}^{\infty} \frac{x^{2n}}{2^{2n}(n!)^2}$ is the modified Bessel function of order zero, see [18].

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

Firstly, we evaluate the asymptotic expansions for I and J in (1.4.9). From equation (1.4.7)

$$I = \int_0^\infty f_2(x')m\left[P(|X-x| > |x'-x|)\right]^{m-1} dx'$$

$$= m\int_0^x f_2(x')\left[P(|X-x| > |x'-x|)\right]^{m-1} dx'$$

$$+m\int_x^\infty f_2(x')\left[P(|X-x| > |x'-x|)\right]^{m-1} dx'$$

$$= m\int_0^x f_2(z)\left[P(X < z) + P(X > x + (x-z))\right]^{m-1} dz$$

$$+m\int_x^\infty f_2(z)\left[P(X > z) + P(X < x - (z-x))\right]^{m-1} dz$$

$$= m\int_0^x f_2(x-\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m-1} d\rho$$

$$+m\int_0^\infty f_2(x+\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m-1} d\rho$$

$$= m\int_0^x (f_2(x-\rho) + f_2(x+\rho))\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m-1} d\rho$$

$$+m\int_x^\infty f_2(x+\rho)\left[P(X > x+\rho)\right]^{m-1} d\rho = I' + I'', \tag{2.2.1}$$

where

$$I' = m\int_0^x (f_2(x-\rho) + f_2(x+\rho))\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m-1} d\rho \tag{2.2.2}$$

and

$$I'' = m\int_x^\infty f_2(x+\rho)\left[P(X > x+\rho)\right]^{m-1} d\rho. \tag{2.2.3}$$

Now we estimate I' and I''.

2.2.1 Lemma

Let $x \in R^d$, $x \in S$. Assume that the densities f_i are $k-$times differentiable and $f(z) > 0$ for all $z \in S$. Define

$$q_\circ(x,\rho) = \frac{f_2(x-\rho)+f_2(x+\rho)}{f(x-\rho)+f(x+\rho)} \quad \text{and} \quad q_k(x,\rho) = \frac{q'_{k-1}(x,\rho)}{f(x-\rho)+f(x+\rho)} \text{ for } k \geq 1,$$

and

$\lambda_\circ(x,\rho) = \frac{f_2(x+\rho)}{f(x+\rho)}$ and $\lambda_k(x,\rho) = \frac{\lambda'_{k-1}(x,\rho)}{f(x+\rho)}$ for $k \geq 1$.

Then

$$I' = q_\circ(x,0) - q_\circ(x,x)\cdot(P(X>2x))^m + \frac{q_1(x,0)}{m+1} - \frac{q_1(x,x)}{m+1}(P(X>2x))^{m+1}$$

$$+\frac{q_2(x,0)}{(m+1)(m+2)} - \frac{q_2(x,x)}{(m+1)(m+2)}(P(X>2x))^{m+2} + ... + \frac{q_k(x,0)}{(m+1)(m+2)...(m+k)}$$

$$-\frac{q_k(x,x)}{(m+1)(m+2)...(m+k)}(P(X>2x))^{m+k} + \frac{1}{(m+1)(m+2)...(m+k)}I'_{k+1}$$

and

$$I'' = \lambda_\circ(x,x)\cdot(P(X>2x))^m + \frac{\lambda_1(x,x)}{m+1}(P(X>2x))^{m+1}$$

$$+\frac{\lambda_2(x,x)}{(m+1)(m+2)}(P(X>2x))^{m+2} + ... + \frac{\lambda_k(x,x)}{(m+1)(m+2)...(m+k)}(P(X>2x))^{m+k}$$

$$+ \frac{1}{(m+1)(m+2)...(m+k)}I''_{k+1}$$

where

$$I'_{k+1} = I'_{k+1}(x) = \int_0^x q'_k(x,\rho)\,[P(X<x-\rho)+P(X>x+\rho)]^{m+k}\,d\rho \qquad k=1,2,3,...$$

and

$$I''_{k+1} = I''_{k+1}(x) = \int_0^x \lambda'_k(x,\rho)\,[P(X>x+\rho)]^{m+k}\,d\rho \qquad k=1,2,3,...$$

Proof:
First we estimate I'.

$$I' = m\int_0^x (f_2(x-\rho)+f_2(x+\rho))\,[P(X<x-\rho)+P(X>x+\rho)]^{m-1}\,d\rho$$

$$= -\int_0^x \frac{f_2(x-\rho)+f_2(x+\rho)}{f(x-\rho)+f(x+\rho)}\frac{d}{d\rho}\,[P(X<x-\rho)+P(X>x+\rho)]^m\,d\rho$$

$$= -\int_0^x q_\circ(x,\rho)\frac{d}{d\rho}\,[P(X<x-\rho)+P(X>x+\rho)]^m\,d\rho,$$

where $q_\circ(x,\rho) = \frac{f_2(x-\rho)+f_2(x+\rho)}{f(x-\rho)+f(x+\rho)}$. (2.2.4)

Let $u = q_\circ(x,\rho)$, $dv = \frac{d}{d\rho}\,[P(X<x-\rho)+P(X>x+\rho)]^m\,d\rho$,

$du = q'_\circ(x,\rho)d\rho$, $v = [P(X<x-\rho)+P(X>x+\rho)]^m$.

Then, by partial integration

$$I' = \int_0^x q_\circ'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^m\,d\rho$$

$$- [q_\circ(x,\rho)(P(X < x-\rho) + P(X > x+\rho))^m]_0^x$$

$$= \int_0^x q_\circ'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^m\,d\rho$$

$$- [q_\circ(x,x)\cdot(P(X > 2x))^m - q_\circ(x,0)]$$

$$= q_\circ(x,0) - q_\circ(x,x)\cdot(P(X > 2x))^m + I_1', \tag{2.2.5}$$

where $I_1' = I_1'(x) = \int_0^x q_\circ'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^m\,d\rho$.

We evaluate I_1'.

$$I_1' = \int_0^x q_\circ'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^m\,d\rho$$

$$= \tfrac{-1}{m+1}\int_0^x \tfrac{q_\circ'(x,\rho)}{f(x-\rho)+f(x+\rho)}\tfrac{d}{d\rho}\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1}\,d\rho$$

$$= \tfrac{-1}{m+1}\int_0^x q_1(x,\rho)\tfrac{d}{d\rho}\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1}\,d\rho,$$

where $q_1(x,\rho) = \frac{q_\circ'(x,\rho)}{f(x-\rho)+f(x+\rho)}$. (2.2.6)

We integrate by parts with

$u = q_1(x,\rho)$, $dv = \frac{d}{d\rho}\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1}\,d\rho$, then

$$I_1' = \tfrac{1}{m+1}\int_0^x q_1'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1}\,d\rho$$

$$-\tfrac{1}{m+1}\,[q_1(x,\rho)(P(X < x-\rho) + P(X > x+\rho))^{m+1}]_0^x$$

$$= \tfrac{1}{m+1}\int_0^x q_1'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1}\,d\rho$$

$$-\tfrac{1}{m+1}\,[q_1(x,x)\cdot(P(X > 2x))^{m+1} - q_1(x,0)]$$

$$= \tfrac{1}{m+1}I_2' + \tfrac{1}{m+1}q_1(x,0) - \tfrac{1}{m+1}q_1(x,x)\cdot(P(X > 2x))^{m+1}, \tag{2.2.7}$$

where $I_2' = \int_0^x q_1'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1}\,d\rho$.

Evaluating I_2':

$$I_2' = \int_0^x q_1'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+1}\,d\rho$$

$$= \tfrac{-1}{m+2}\int_0^x \tfrac{q_1'(x,\rho)}{f(x-\rho)+f(x+\rho)}\tfrac{d}{d\rho}\,[P(X < x-\rho) + P(X > x+\rho)]^{m+2}\,d\rho$$

$$= \tfrac{-1}{m+2}\int_0^x q_2(x,\rho)\tfrac{d}{d\rho}\,[P(X < x-\rho) + P(X > x+\rho)]^{m+2}\,d\rho,$$

where $\quad q_2(x,\rho) = \frac{q_1'(x,\rho)}{f(x-\rho)+f(x+\rho)}$. (2.2.8)

$$I_2' = \tfrac{1}{m+2}\int_0^x q_2'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+2}\,d\rho$$

$$-\tfrac{1}{m+2}\,[q_2(x,\rho)(P(X < x-\rho) + P(X > x+\rho))^{m+2}]_0^x$$

$$= \tfrac{1}{m+2}\int_0^x q_2'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+2}\,d\rho$$

$$-\tfrac{1}{m+2}\,[q_2(x,x)\cdot(P(X > 2x))^{m+2} - q_2(x,0)]$$

$$= \tfrac{1}{m+2}q_2(x,0) - \tfrac{1}{m+2}q_2(x,x)\cdot(P(X > 2x))^{m+2} + \tfrac{1}{m+2}I_3', \tag{2.2.9}$$

where $I_3' = I_3'(x) = \int_0^x q_2'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+2}\,d\rho$.

By repeating this procedure, we can obtain an asymptotic expansion for $I'(x)$ in the form:

$$I' = q_\circ(x,0) - q_\circ(x,x)\cdot(P(X > 2x))^m + \tfrac{q_1(x,0)}{m+1} - \tfrac{q_1(x,x)}{m+1}(P(X > 2x))^{m+1}$$

$$+\tfrac{q_2(x,0)}{(m+1)(m+2)} - \tfrac{q_2(x,x)}{(m+1)(m+2)}(P(X > 2x))^{m+2} + \ldots + \tfrac{q_k(x,0)}{(m+1)(m+2)\ldots(m+k)}$$

$$-\tfrac{q_k(x,x)}{(m+1)(m+2)\ldots(m+k)}(P(X > 2x))^{m+k} + \tfrac{1}{(m+1)(m+2)\ldots(m+k)}I_{k+1}',$$

where $I_{k+1}' = I_{k+1}'(x) = \int_0^x q_k'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)]^{m+k}\,d\rho$
$k = 1, 2, 3, \ldots$

Now we evaluate I''.

$I'' = m \int_x^\infty f_2(x+\rho)\,[P(X > x+\rho)]^{m-1}\,d\rho$

$\quad = -\int_x^\infty \frac{f_2(x+\rho)}{f(x+\rho)} \frac{d}{d\rho}\,[P(X > x+\rho)]^m\,d\rho$

$\quad = -\int_x^\infty \lambda_\circ(x,\rho) \frac{d}{d\rho}\,[P(X > x+\rho)]^m\,d\rho,$

where $\lambda_\circ(x,\rho) = \frac{f_2(x+\rho)}{f(x+\rho)}.$ (2.2.11)

Let $u = \lambda_\circ(x,\rho), \qquad dv = \frac{d}{d\rho}\,[P(X > x+\rho)]^m\,d\rho,$

$du = \lambda'_\circ(x,\rho)d\rho, \qquad v = [P(X > x+\rho)]^m.$

Then

$I'' = \int_x^\infty \lambda'_\circ(x,\rho)\,[P(X > x+\rho)]^m - [\lambda_\circ(x,\rho)(P(X > x+\rho))^m]_x^\infty$

$\quad = \int_x^\infty \lambda'_\circ(x,\rho)\,[P(X > x+\rho)]^m\,d\rho + \lambda_\circ(x,x)\cdot(P(X > 2x))^m$

$\quad = \lambda_\circ(x,x)\cdot(P(X > 2x))^m + I''_1,$ (2.2.12)

where $I''_1 = I''_1(x) = \int_x^\infty \lambda'_\circ(x,\rho)\,[P(X > x+\rho)]^m\,d\rho.$

We estimate I''_1.

$I''_1 = \int_x^\infty \lambda'_\circ(x,\rho)\,[P(X > x+\rho)]^m\,d\rho$

$\quad = \frac{-1}{m+1} \int_x^\infty \frac{\lambda'_\circ(x,\rho)}{f(x+\rho)} \frac{d}{d\rho}\,[P(X > x+\rho)]^{m+1}\,d\rho$

$\quad = \frac{-1}{m+1} \int_x^\infty \lambda_1(x,\rho) \frac{d}{d\rho}\,[P(X > x+\rho)]^{m+1}\,d\rho,$

where $\lambda_1(x,\rho) = \frac{\lambda'_\circ(x,\rho)}{f(x+\rho)}.$ (2.2.13)

We integrate by parts with

$u = \lambda_1(x,\rho), \;\; dv = \frac{d}{d\rho}\,[P(X > x+\rho)]^{m+1}\,d\rho$, then

$$I_1'' = \frac{1}{m+1}\int_x^\infty \lambda_1'(x,\rho)\left[P(X > x+\rho)\right]^{m+1} d\rho$$

$$-\frac{1}{m+1}\left[\lambda_1(x,\rho)(P(X > x+\rho))^{m+1}\right]_x^\infty$$

$$= \frac{1}{m+1}\int_x^\infty \lambda_1'(x,\rho)\left[P(X > x+\rho)\right]^{m+1} d\rho + \frac{1}{m+1}\lambda_1(x,x)(P(X > 2x))^{m+1}$$

$$= \frac{1}{m+1}\lambda_1(x,x)(P(X > 2x))^{m+1} + \frac{1}{m+1} I_2'', \qquad (2.2.14)$$

where $I_2'' = I_2''(x) = \int_x^\infty \lambda_1'(x,\rho)\left[P(X > x+\rho)\right]^{m+1} d\rho$.

We estimate I_2''.

$$I_2'' = \int_x^\infty \lambda_1'(x,\rho)\left[P(X > x+\rho)\right]^{m+1} d\rho$$

$$= \frac{-1}{m+2}\int_x^\infty \frac{\lambda_1'(x,\rho)}{f(x+\rho)}\frac{d}{d\rho}\left[P(X > x+\rho)\right]^{m+2} d\rho$$

$$= \frac{-1}{m+2}\int_x^\infty \lambda_2(x,\rho)\frac{d}{d\rho}\left[P(X > x+\rho)\right]^{m+2} d\rho,$$

where $\lambda_2(x,\rho) = \frac{\lambda_1'(x,\rho)}{f(x+\rho)}$. (2.2.15)

We integrate by parts with

$u = \lambda_2(x,\rho)$, $dv = \frac{d}{d\rho}\left[P(X > x+\rho)\right]^{m+2} d\rho$, then

$$I_2'' = \frac{1}{m+2}\int_x^\infty \lambda_2'(x,\rho)\left[P(X > x+\rho)\right]^{m+2} d\rho$$

$$-\frac{1}{m+2}\left[\lambda_2(x,\rho)(P(X > x+\rho))^{m+2}\right]_x^\infty$$

$$= \frac{1}{m+2}\int_x^\infty \lambda_2'(x,\rho)\left[P(X > x+\rho)\right]^{m+2} d\rho + \frac{1}{m+2}\lambda_2(x,x)(P(X > 2x))^{m+2}$$

$$= \frac{1}{m+2}\lambda_2(x,x)\,(P(X > 2x))^{m+2} + \frac{1}{m+2} I_3'', \qquad (2.2.17)$$

where $I_3'' = I_3''(x) = \int_x^\infty \lambda_2'(x,\rho)\left[P(X > x+\rho)\right]^{m+2} d\rho$.

By repeating this procedure, we can obtain an asymptotic expansion for $I''(x)$ in the form:

$$I'' = \lambda_\circ(x,x)\cdot(P(X>2x))^m + \tfrac{\lambda_1(x,x)}{m+1}(P(X>2x))^{m+1}$$

$$+\tfrac{\lambda_2(x,x)}{(m+1)(m+2)}(P(X>2x))^{m+2} + ... + \tfrac{\lambda_k(x,x)}{(m+1)(m+2)...(m+k)}(P(X>2x))^{m+k}$$

$$+\ \tfrac{1}{(m+1)(m+2)...(m+k)} I''_{k+1},$$

where $I''_{k+1} = I''_{k+1}(x) = \int_x^\infty \lambda'_k(x,\rho)\,[P(X>x+\rho)]^{m+k}\,d\rho \qquad k=1,2,3,...$

Similarly, we can show that $I'_{k+1}(m), I''_{k+1}(m) \to 0$ when $m\to\infty$ for all $k\geq 2$ under suitable conditions as in part (2.1).

2.2.2 Lemma

Assume that there exist j,l such that the following conditions are satisfied

(i) $\left|q'_k(x,\rho)\,[P(X<x-\rho)+P(X>x+\rho)\ ]^j\right|$ is bounded for ρ and

(ii) $[P(X<x-\rho)+P(X>x+\rho)\ ]^l$ is integrable for ρ.

Then $\quad I'_{k+1}\to 0 \quad$ when $m\to\infty$,

where $\quad I'_{k+1} = \int_0^x q'_k(x,\rho)\,[P(X<x-\rho)+P(X>x+\rho)\ ]^{m+k}\,d\rho.$

Proof:
As in part (2.1).

2.2.3 Lemma

Assume that there exist j,l such that the following conditions are satisfied

(i) $\left|\lambda'_k(x,\rho)\,[P(X>x+\rho)]^j\right|$ is bounded for ρ and

(ii) $[P(X>x+\rho)]^l$ is integrable for ρ.

Then $\quad I''_{k+1}\to 0 \quad$ when $m\to\infty$,

where $\quad I''_{k+1} = \int_x^\infty \lambda'_k(x,\rho)\,[P(X>x+\rho)]^{m+k}\,d\rho.$

Proof:
As in part (2.1)

Now we give an example for exponential distribution to show that the conditions (i) and (ii) in the above Lemmas are satisfied in the case of support $(0, \infty)$.

2.2.4 Example

Let $f_1(x) = ae^{-ax}$, $f_2(x) = be^{-bx}$ be two densities for exponential distributions with prior probabilities p_1, p_2 such that $p_1+p_2 = 1$, and $f = p_1 f_1+p_2 f_2$.

Fix $x > 0$. Firstly we show that there exist j such that

$$\left| q_2'(x, \rho) \; [P(X < x - \rho) + P(X > x + \rho) \;]^j \right| \text{ is bounded for } \rho.$$

Since $q_2(x, \rho) = \frac{q_1'}{g}$ then

$$q_2'(x, \rho) = \left(\frac{q_1'}{g}\right)' = \frac{h'''g - 6h''g'g^2 - h'g''g^2 - 7hgg'g'' + 3h'(g')^2 g + 3h(g')^3}{g^6},$$

where $h(x, \rho) = f_2(x + \rho) + f_2(x - \rho)$, and $g(x, \rho) = f(x + \rho) + f(x - \rho)$.

Substituting these functions in the above equation, $q_2'(x, \rho)$ can be bounded in the form:

$$q_2'(x, \rho) \leq \frac{C_1}{\left(e^{-a(x+\rho)}\right)^2} \leq C_2 e^{C_3 \rho}.$$

Since $[P(X < x - \rho) + P(X > x + \rho) \;] = [P\left(|X - x| > \rho\right)]$

$$= P\left(e^{t|X-x|} > e^{t\rho}\right) \leq \frac{Ee^{t|X-x|}}{e^{t\rho}} \leq \frac{C_4}{e^{t\rho}}, \text{ then}$$

$$[P(X < x - \rho) + P(X > x + \rho) \;]^j \leq C_5 e^{-jt\rho},$$

where $0 < t < \min\{a, b\}$. Then

$$\left| q_2'(x, \rho) \; [P(X < x - \rho) + P(X > x + \rho) \;]^j \right| \leq C_2 e^{C_3 \rho} C_5 e^{-jt\rho} \leq C_6 e^{(C_3 - jt)\rho}$$

That is, we can find j such that

$$\left|q_2'(x,\rho)\ [P(X<x-\rho)+P(X>x+\rho)\ ]^j\right| \text{ is bounded for } \rho.$$

Now we show that there exists l such that $[P(X<x-\rho)+P(X>x+\rho)\ ]^l$ is integrable for ρ.

Above we showed that $\Big[P(X<x-\rho)+P(X>x+\rho)\ \Big]^l \leq C_7 e^{-lt\rho}$ which immediatedly shows integrability.

2.2.5 Corollary

$$I' = q_\circ(x,0) - q_\circ(x,x)\cdot(P(X>2x))^m + \sum_{k=1}^{\infty} \frac{q_k(x,0)}{\prod_{n=1}^{k}(m+n)} - \sum_{k=1}^{\infty} \frac{q_k(x,x)\cdot(P(X>2x))^{m+k}}{\prod_{n=1}^{k}(m+n)} \tag{2.2.18}$$

and

$$I'' = \lambda_\circ(x,x)\cdot(P(X>2x))^m + \sum_{k=1}^{\infty} \frac{\lambda_k(x,x)\cdot(P(X>2x))^{m+k}}{\prod_{n=1}^{k}(m+n)} \tag{2.2.19}$$

under the condition that $\frac{I'_{k+1}}{\prod_{n=1}^{k}(m+n)}$ and $\frac{I''_{k+1}}{\prod_{n=1}^{k}(m+n)}$ tend to zero when $k\to\infty$,

where

$$I'_{k+1} = \int_0^x q_k'(x,\rho)\,[P(X<x-\rho)+P(X>x+\rho)\ ]^{m+k}\,d\rho,$$

$$I''_{k+1} = \int_x^\infty \lambda_k'(x,\rho)\,[P(X>x+\rho)]^{m+k}\,d\rho,$$

$q_k(x,0)$, $q_k(x,x)$ and $\lambda_k(x,x)$ are defined as in Lemma 2.2.1.

2.2.6 Lemma

Let the conditions of Lemmas 2.2.1-2.2.3 and Corollary 2.2.5 be satisfied. Then

$$I = q_\circ(x,0) + (\lambda_\circ(x,x)-q_\circ(x,x))\cdot(P(X>2x))^m + \sum_{k=1}^{\infty} \frac{q_k(x,0)}{\prod_{n=1}^{k}(m+n)} + \sum_{k=1}^{\infty} \frac{(\lambda_k(x,x)-q_k(x,x))\cdot(P(X>2x))^{m+k}}{\prod_{n=1}^{k}(m+n)} \tag{2.2.20}$$

where $q_k(x,0)$, $q_k(x,x)$ and $\lambda_k(x,x)$ are defined as in Lemma 2.2.1.

Proof:
Substituting (2.2.18) and (2.2.19) into (2.2.1).

Similarly, we can obtain an asymptotic expansion for J where

$$J = \int_0^\infty f_1(x')mP(|X-x| > |x'-x|)^{m-1}dx'$$

$$= m\int_0^x (f_1(x-\rho)+f_1(x+\rho))\,[P(X < x-\rho)+P(X > x+\rho)\,]^{m-1}\,d\rho$$

$$+m\int_x^\infty f_1(x+\rho)\,[P(X > x+\rho)]^{m-1}\,d\rho.$$

2.2.7 Lemma

Under conditions as in Lemma 2.2.6, then

$$J = \overline{q}_\circ(x,0) + \left(\overline{\lambda}_\circ(x,x) - \overline{q}_\circ(x,x)\right)\cdot(P(X > 2x))^m + \sum_{k=1}^\infty \frac{\overline{q}_k(x,0)}{\prod_{n=1}^k (m+n)}$$

$$+\sum_{k=1}^\infty \frac{\left(\overline{\lambda}_k(x,x)-\overline{q}_k(x,x)\right)\cdot(P(X>2x))^{m+k}}{\prod_{n=1}^k (m+n)} \qquad (2.2.21)$$

where

$$\overline{q}_\circ(x,\rho) = \frac{f_1(x-\rho)+f_1(x+\rho)}{f(x-\rho)+f(x+\rho)} \quad \text{and} \quad \overline{q}_k(x,\rho) = \frac{\overline{q}'_{k-1}(x,\rho)}{f(x-\rho)+f(x+\rho)} \ \text{ for } k \geq 1,$$

and

$$\overline{\lambda}_\circ(x,\rho) = \frac{f_1(x+\rho)}{f(x+\rho)} \quad \text{and} \quad \overline{\lambda}_k(x,\rho) = \frac{\overline{\lambda}'_{k-1}(x,\rho)}{f(x+\rho)} \ \text{ for } k \geq 1.$$

2.2.8 Theorem

Let the conditions of Lemmas 2.2.6 and 2.2.7 be satisfied. Then

$$R_m(x) = P(\theta' \neq \theta\,|X = x\,)$$

$$= \frac{p_1p_2f_1(x)f_2(x)}{f^2(x)} + \sum_{k=1}^\infty \frac{\alpha_k}{\prod_{n=1}^k (m+n)} + \beta_\circ(P(X > 2x))^m + \sum_{k=1}^\infty \frac{\beta_k(P(X>2x))^{m+k}}{\prod_{n=1}^k (m+n)} \qquad (2.2.22)$$

where $\alpha_k = \frac{p_1 p_2}{f(x)}(f_1(x) q_k(x,0) + f_2(x)\overline{q}_k(x,0)), \qquad k = 1,2,3,...$

$$\beta_k = \frac{p_1 p_2}{f(x)}\left(f_1(x)\left(\lambda_k(x,x) - q_k(x,x)\right) + f_2(x)\left(\overline{\lambda}_k(x,x) - \overline{q}_k(x,x)\right)\right), \qquad k = 0,1,2,...$$

Proof:
Substituting (2.2.20) and (2.2.21) in to (1.4.9) we obtain

$$R_m(x) = \frac{p_1 p_2\, f_1(x)}{f(x)} \{q_\circ(x,0) + (\lambda_\circ(x,x) - q_\circ(x,x)) \cdot (P(X > 2x))^m$$
$$+ \sum_{k=1}^{\infty} \frac{q_k(x,0)}{\prod_{n=1}^{k}(m+n)} + \sum_{k=1}^{\infty} \frac{(\lambda_k(x,x) - q_k(x,x)) \cdot (P(X>2x))^{m+k}}{\prod_{n=1}^{k}(m+n)}\Bigg\}$$
$$+ \frac{p_1 p_2\, f_2(x)}{f(x)} \left\{\overline{q}_\circ(x,0) + \left(\overline{\lambda}_\circ(x,x) - \overline{q}_\circ(x,x)\right) \cdot (P(X > 2x))^m\right.$$
$$\left. + \sum_{k=1}^{\infty} \frac{\overline{q}_k(x,0)}{\prod_{n=1}^{k}(m+n)} + \sum_{k=1}^{\infty} \frac{\left(\overline{\lambda}_k(x,x) - \overline{q}_k(x,x)\right) \cdot (P(X>2x))^{m+k}}{\prod_{n=1}^{k}(m+n)}\right\}$$

$$= \frac{p_1 p_2 f_1(x) f_2(x)}{f^2(x)} + \frac{p_1 p_2}{f(x)} \sum_{k=1}^{\infty} \frac{f_1(x) q_k(x,0) + f_2(x)\overline{q}_k(x,0)}{\prod_{n=1}^{k}(m+n)}$$
$$+ \frac{p_1 p_2 (P(X>2x))^m}{f(x)} \left(f_1(x)\left(\lambda_\circ(x,x) - q_\circ(x,x)\right) + f_2(x)\left(\overline{\lambda}_\circ(x,x) - \overline{q}_\circ(x,x)\right)\right)$$
$$+ \frac{p_1 p_2}{f(x)} \sum_{k=1}^{\infty} \frac{f_1(x)(\lambda_k(x,x) - q_k(x,x)) + f_2(x)\left(\overline{\lambda}_k(x,x) - \overline{q}_k(x,x)\right)(P(X>2x))^{m+k}}{\prod_{n=1}^{k}(m+n)}$$

$$= \frac{p_1 p_2 f_1(x) f_2(x)}{f^2(x)} + \sum_{k=1}^{\infty} \frac{\alpha_k}{\prod_{n=1}^{k}(m+n)} + \beta_\circ (P(X > 2x))^m + \sum_{k=1}^{\infty} \frac{\beta_k (P(X>2x))^{m+k}}{\prod_{n=1}^{k}(m+n)},$$

where α_k and β_k are defined as above.

2.3 Support $S = (a, b)$:

Firstly, we evaluate the asymptotic expansions for I and J in (1.4.9). From equation (1.4.7)

$$I = \int_a^b f_2(x')m\left[P(|X-x| > |x'-x|)\right]^{m-1} dx'$$

$$= m\int_a^x f_2(x')\left[P(|X-x| > |x'-x|)\right]^{m-1} dx'$$

$$+ m\int_x^b f_2(x')\left[P(|X-x| > |x'-x|)\right]^{m-1} dx'$$

$$= m\int_a^x f_2(z)\left[P(X < z) + P(X > x + (x-z))\right]^{m-1} dz$$

$$+ m\int_x^b f_2(z)\left[P(X > z) + P(X < x - (z-x))\right]^{m-1} dz$$

$$= m\int_0^{x-a} f_2(x-\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m-1} d\rho$$

$$+ m\int_0^{b-x} f_2(x+\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m-1} d\rho$$

For the rest of 2.3 we only treat $x - a \leq b - x$ i.e. $x \leq \frac{a+b}{2}$. The case $x \geq \frac{a+b}{2}$ is treated similarly. Then

$$I = m\int_0^{x-a}(f_2(x-\rho) + f_2(x+\rho))\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m-1} d\rho$$

$$+ m\int_{x-a}^{b-x} f_2(x+\rho)\left[P(X > x+\rho)\right]^{m-1} d\rho = I' + I'', \tag{2.3.1}$$

where

$$I' = m\int_0^{x-a}(f_2(x-\rho) + f_2(x+\rho))\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m-1} d\rho, \tag{2.3.2}$$

$$I'' = m\int_{x-a}^{b-x} f_2(x+\rho)\left[P(X > x+\rho)\right]^{m-1} d\rho. \tag{2.3.3}$$

2.3.1 Lemma

Let $x \in S$, $x \leq \frac{a+b}{2}$. Assume that the densities f_i are $k-$times differentiable and $f(z) > 0$ for all $z \in S$. Define

$q_\circ(x,\rho) = \frac{f_2(x-\rho)+f_2(x+\rho)}{f(x-\rho)+f(x+\rho)}$ and $q_k(x,\rho) = \frac{q'_{k-1}(x,\rho)}{f(x-\rho)+f(x+\rho)}$ for $k \geq 1$,
and
$\lambda_\circ(x,\rho) = \frac{f_2(x+\rho)}{f(x+\rho)}$ and $\lambda_k(x,\rho) = \frac{\lambda'_{k-1}(x,\rho)}{f(x+\rho)}$ for $k \geq 1$.

Then

$$I' = q_\circ(x,0)+\frac{q_1(x,0)}{m+1} + \frac{q_2(x,0)}{(m+1)(m+2)} - q_\circ(x,x-a)\cdot(P(X>2x-a))^m$$
$$-\frac{q_1(x,x-a)}{m+1}(P(X>2x-a))^{m+1} - \frac{q_2(x,x-a)}{(m+1)(m+2)}(P(X>2x-a))^{m+2}$$
$$+... + \frac{q_k(x,0)}{(m+1)(m+2)...(m+k)} - \frac{q_k(x,x-a)}{(m+1)(m+2)...(m+k)}(P(X>2x-a))^{m+k}$$
$$+\frac{1}{(m+1)(m+2)...(m+k)}I'_{k+1}$$

and

$$I'' = \lambda_\circ(x,x-a)\cdot(P(X>2x-a))^m + \frac{\lambda_1(x,x-a)}{m+1}(P(X>2x-a))^{m+1}$$
$$+\frac{\lambda_2(x,x-a)}{(m+1)(m+2)}(P(X>2x-a))^{m+2}$$
$$+... + \frac{\lambda_k(x,x-a)}{(m+1)(m+2)...(m+k)}(P(X>2x-a))^{m+k} + \frac{1}{(m+1)(m+2)...(m+k)}I''_{k+1}$$

where
$$I'_{k+1} = I'_{k+1}(x) = \int_0^{x-a} q'_k(x,\rho)\,[P(X<x-\rho)+P(X>x+\rho)]^{m+k}\,d\rho \qquad k=1,2,3,...$$

and
$$I''_{k+1} = I''_{k+1}(x) = \int_{x-a}^{b-x} \lambda'_k(x,\rho)\,[P(X>x+\rho)]^{m+k}\,d\rho \qquad k=1,2,3,...$$

Proof:
First we estimate I'.

$$I' = -\int_0^{x-a} \frac{f_2(x-\rho)+f_2(x+\rho)}{f(x-\rho)+f(x+\rho)}\frac{d}{d\rho}[P(X<x-\rho)+P(X>x+\rho)]^m\,d\rho$$
$$= -\int_0^{x-a} q_\circ(x,\rho)\frac{d}{d\rho}[P(X<x-\rho)+P(X>x+\rho)]^m\,d\rho,$$

where $q_\circ(x,\rho) = \frac{f_2(x-\rho)+f_2(x+\rho)}{f(x-\rho)+f(x+\rho)}.$ (2.3.4)

Let $\quad u = q_\circ(x,\rho), \qquad dv = \frac{d}{d\rho}\left[P(X < x-\rho) + P(X > x+\rho)\right]^m d\rho,$

$$du = q'_\circ(x,\rho)d\rho, \qquad v = \left[P(X < x-\rho) + P(X > x+\rho)\right]^m.$$

Then
$I' = \int_0^{x-a} q'_\circ(x,\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^m d\rho$

$$- \left[q_\circ(x,\rho)(P(X < x-\rho) + P(X > x+\rho))^m\right]_0^{x-a}$$

$$= \int_0^{x-a} q'_\circ(x,\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^m d\rho$$

$$- \left[\ q_\circ(x,x-a)\cdot(P(X > 2x-a))^m - q_\circ(x,0)\ \right]$$

$$= q_\circ(x,0) - q_\circ(x,x-a)\cdot(P(X > 2x-a))^m + I'_1, \tag{2.3.5}$$

where $\ I'_1 = I'_1(x) = \int_0^{x-a}\ q'_\circ(x,\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^m d\rho.$

We evaluate I'_1.

$$I'_1 = \int_0^{x-a}\ q'_\circ(x,\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^m d\rho$$

$$= \frac{-1}{m+1}\int_0^{x-a} \frac{q'_\circ(x,\rho)}{f(x-\rho)+f(x+\rho)} \frac{d}{d\rho}\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m+1} d\rho$$

$$= \frac{-1}{m+1}\int_0^{x-a} q_1(x,\rho)\, \frac{d}{d\rho}\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m+1} d\rho,$$

where $\ q_1(x,\rho) = \frac{q'_\circ(x,\rho)}{f(x-\rho)+f(x+\rho)}.$ (2.3.6)

We integrate by parts with

$u = q_1(x,\rho)$, $dv = \frac{d}{d\rho}\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m+1} d\rho$, then

$$I'_1 = \frac{1}{m+1}\int_0^{x-a} q'_1(x,\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m+1} d\rho$$

$$-\frac{1}{m+1}\left[q_1(x,\rho)(P(X < x-\rho) + P(X > x+\rho))^{m+1}\right]_0^{x-a}$$

$$= \frac{1}{m+1}\int_0^{x-a} q'_1(x,\rho)\left[P(X < x-\rho) + P(X > x+\rho)\right]^{m+1} d\rho$$

$$-\frac{1}{m+1}\left[\ q_1(x,x-a)\cdot(P(X > 2x))^{m+1} - q_1(x,0)\ \right]$$

$$= \tfrac{1}{m+1} q_1(x,0) - \tfrac{1}{m+1} q_1(x, x-a)\cdot (P(X>2x-a))^{m+1} + \tfrac{1}{m+1} I_2', \qquad (2.3.7)$$

where $I_2' = \int_0^{x-a} q_1'(x,\rho)\,[P(X<x-\rho)+P(X>x+\rho)]^{m+1}\,d\rho.$

Evaluating I_2':

$$I_2' = \int_0^{x-a} q_1'(x,\rho)\,[P(X<x-\rho)+P(X>x+\rho)\,]^{m+1}\;d\rho$$

$$= \tfrac{-1}{m+2}\int_0^{x-a} \tfrac{q_1'(x,\rho)}{f(x-\rho)+f(x+\rho)}\tfrac{d}{d\rho}\,[P(X<x-\rho)+P(X>x+\rho)\,]^{m+2}\,d\rho$$

$$= \tfrac{-1}{m+2}\int_0^{x-a} q_2(x,\rho)\tfrac{d}{d\rho}\,[P(X<x-\rho)+P(X>x+\rho)\,]^{m+2}\,d\rho,$$

where $q_2(x,\rho) = \frac{q_1'(x,\rho)}{f(x-\rho)+f(x+\rho)}.$ (2.3.8)

$$I_2' = \tfrac{1}{m+2}\int_0^{x-a} q_2'(x,\rho)\,[P(X<x-\rho)+P(X>x+\rho)\,]^{m+2}\,d\rho$$

$$-\tfrac{1}{m+1}\,[q_1(x,\rho)(P(X<x-\rho)+P(X>x+\rho))^{m+2}]_0^{x-a}$$

$$= \tfrac{1}{m+2}\int_0^{x-a} q_2'(x,\rho)\,[P(X<x-\rho)+P(X>x+\rho)]^{m+2}\,d\rho$$

$$-\tfrac{1}{m+2}\,[q_2(x,x-a)\cdot(P(X>2x-a))^{m+2} - q_2(x,0)]$$

$$= \tfrac{1}{m+2} q_2(x,0) - \tfrac{1}{m+2} q_2(x,x-a)\cdot(P(X>2x-a))^{m+2} + \tfrac{1}{m+2} I_3', \qquad (2.3.9)$$

where $I_3' = I_3'(x) = \int_0^{x-a} q_2'(x,\rho)\,[P(X<x-\rho)+P(X>x+\rho)\,]^{m+2}\,d\rho.$

By repeating this procedure, we can obtain an asymptotic expansion for $I'(x)$ in the form:

$$I' = q_\circ(x,0) + \tfrac{q_1(x,0)}{m+1} + \tfrac{q_2(x,0)}{(m+1)(m+2)} - q_\circ(x,x-a)\cdot(P(X>2x-a))^m$$

$$-\tfrac{q_1(x,x-a)}{m+1}(P(X>2x-a))^{m+1} - \tfrac{q_2(x,x-a)}{(m+1)(m+2)}(P(X>2x-a))^{m+2}$$

$$+ ... + \tfrac{q_k(x,0)}{(m+1)(m+2)...(m+k)} - \tfrac{q_k(x,x-a)}{(m+1)(m+2)...(m+k)}(P(X>2x-a))^{m+k}$$

$$+\tfrac{1}{(m+1)(m+2)...(m+k)} I_{k+1}',$$

where

$$I'_{k+1} = I'_{k+1}(x) = \int_0^{x-a} q'_k(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)\,]^{m+k}\,d\rho$$
$$k = 1, 2, 3, ...$$

Now we evaluate I''.

$$I'' = m\int_{x-a}^{b-x} f_2(x+\rho)\,[P(X > x+\rho)]^{m-1}\,d\rho$$

$$= -\int_{x-a}^{b-x} \frac{f_2(x+\rho)}{f(x+\rho)}\frac{d}{d\rho}\,[P(X > x+\rho)]^m\,d\rho$$

$$= -\int_{x-a}^{b-x} \lambda_\circ(x,\rho)\frac{d}{d\rho}\,[P(X > x+\rho)]^m\,d\rho,$$

where $\lambda_\circ(x,\rho) = \frac{f_2(x+\rho)}{f(x+\rho)}$. (2.3.10)

Let $u = \lambda_\circ(x,\rho)$, $dv = \frac{d}{d\rho}\,[P(X > x+\rho)]^m\,d\rho$,

$du = \lambda'_\circ(x,\rho)d\rho$, $v = [P(X > x+\rho)]^m$.

Then

$$I'' = \int_{x-a}^{b-x} \lambda'_\circ(x,\rho)\,[P(X > x+\rho)]^m\,d\rho - [\lambda_\circ(x,\rho)(P(X > x+\rho))^m]_{x-a}^{b-x}$$

$$= \int_{x-a}^{b-x} \lambda'_\circ(x,\rho)\,[P(X > x+\rho)]^m\,d\rho + \lambda_\circ(x, x-a)\cdot(P(X > 2x-a))^m$$

$$= \lambda_\circ(x, x-a)\cdot(P(X > 2x-a))^m + I''_1, \quad (2.3.11)$$

where $I''_1 = I''_1(x) = \int_{x-a}^{b-x} \lambda'_\circ(x,\rho)\,[P(X > x+\rho)]^m\,d\rho$.

We estimate I''_1.

$$I''_1 = \int_{x-a}^{b-x} \lambda'_\circ(x,\rho)\,[P(X > x+\rho)]^m\,d\rho$$

$$= \frac{-1}{m+1}\int_{x-a}^{b-x} \frac{\lambda'_\circ(x,\rho)}{f(x+\rho)}\frac{d}{d\rho}\,[P(X > x+\rho)]^{m+1}\,d\rho$$

$$= \frac{-1}{m+1}\int_{x-a}^{b-x} \lambda_1(x,\rho)\frac{d}{d\rho}\,[P(X > x+\rho)]^{m+1}\,d\rho,$$

where $\lambda_1(x,\rho) = \frac{\lambda'_\circ(x,\rho)}{f(x+\rho)}$. (2.3.12)

We integrate by parts with

$u = \lambda_1(x,\rho)$, $dv = \frac{d}{d\rho}\left[P(X > x+\rho)\right]^{m+1} d\rho$, then

$$I_1'' = \tfrac{1}{m+1}\int_{x-a}^{b-x} \lambda_1'(x,\rho)\left[P(X > x+\rho)\right]^{m+1} d\rho$$

$$-\tfrac{1}{m+1}\left[\lambda_1(x,\rho)(P(X > x+\rho))^{m+1}\right]_{x-a}^{b-x}$$

$$= \tfrac{1}{m+1}\int_{x-a}^{b-x} \lambda_1'(x,\rho)\left[P(X > x+\rho)\right]^{m+1} d\rho$$

$$+\tfrac{1}{m+1}\lambda_1(x,x-a)\,(P(X > 2x-a))^{m+1}$$

$$= \tfrac{1}{m+1}\lambda_1(x,x-a)(P(X > 2x-a))^{m+1} + \tfrac{1}{m+1}I_2'', \tag{2.3.13}$$

where $I_2'' = I_2''(x) = \int_{x-a}^{b-x} \lambda_1'(x,\rho)\left[P(X > x+\rho)\right]^{m+1} d\rho.$

We estimate I_2''.

$$I_2'' = \int_{x-a}^{b-x} \lambda_1'(x,\rho)\left[P(X > x+\rho)\right]^{m+1} d\rho$$

$$= \tfrac{-1}{m+2}\int_{x-a}^{b-x} \tfrac{\lambda_1'(x,\rho)}{f(x+\rho)}\tfrac{d}{d\rho}\left[P(X > x+\rho)\right]^{m+2} d\rho$$

$$= \tfrac{-1}{m+2}\int_{x-a}^{b-x} \lambda_2(x,\rho)\tfrac{d}{d\rho}\left[P(X > x+\rho)\right]^{m+2} d\rho,$$

where $\lambda_2(x,\rho) = \frac{\lambda_1'(x,\rho)}{f(x+\rho)}$. (2.3.14)

We integrate by parts with

$u = \lambda_2(x,\rho)$, $dv = \frac{d}{d\rho}\left[P(X > x+\rho)\right]^{m+2} d\rho$, then

$$I_2'' = \tfrac{1}{m+2}\int_{x-a}^{b-x} \lambda_2'(x,\rho)\left[P(X > x+\rho)\right]^{m+2} d\rho$$

$$-\tfrac{1}{m+2}\left[\lambda_2(x,\rho)(P(X > x+\rho))^{m+2}\right]_{x-a}^{b-x}$$

$$= \tfrac{1}{m+2}\int_{x-a}^{b-x} \lambda_2'(x,\rho)\left[P(X > x+\rho)\right]^{m+2} d\rho$$

$$+\tfrac{1}{m+2}\lambda_2(x,x-a)(P(X > 2x-a))^{m+2}$$

$$= \frac{1}{m+2}\lambda_2(x, x-a)(P(X > 2x-a))^{m+2} + \frac{1}{m+2} I_3'', \qquad (2.3.15)$$

where $I_3'' = I_3''(x) = \int_{x-a}^{b-x} \lambda_2'(x,\rho)\,[P(X > x+\rho)]^{m+2}\, d\rho.$

By repeating this procedure, we can obtain an asymptotic expansion for $I''(x)$ in the form:

$$I'' = \lambda_\circ(x, x-a)\cdot(P(X > 2x-a))^m + \frac{\lambda_1(x,x-a)}{m+1}(P(X > 2x-a))^{m+1}$$

$$+\frac{\lambda_2(x,x-a)}{(m+1)(m+2)}(P(X > 2x-a))^{m+2}$$

$$+ ... + \frac{\lambda_k(x,x-a)}{(m+1)(m+2)...(m+k)}(P(X > 2x-a))^{m+k} + \frac{1}{(m+1)(m+2)...(m+k)} I_{k+1}''$$

where $I_{k+1}'' = I_{k+1}''(x) = \int_{x-a}^{b-x} \lambda_k'(x,\rho)\,[P(X > x+\rho)]^{m+k}\, d\rho \qquad k = 1, 2, 3, ...$

Similarly, we can show that $I_{k+1}'(m)$, $I_{k+1}''(m) \to 0$ when $m \to \infty$ for all $k \geq 2$ under suitable conditions as in part (2.1).

2.3.2 Lemma

Assume that there exist j, l such that the following conditions are satisfied

(i) $\left| q_k'(x,\rho)\ [P(X < x-\rho) + P(X > x+\rho)\ ]^j \right|$ is bounded for ρ and

(ii) $[P(X < x-\rho) + P(X > x+\rho)\ ]^l$ is integrable for ρ.

Then $\quad I_{k+1}' \to 0 \quad$ when $m \to \infty$,

where $\quad I_{k+1}' = \int_0^{x-a} q_k'(x,\rho)\,[P(X < x-\rho) + P(X > x+\rho)\ ]^{m+k}\, d\rho.$

Proof:
As in part (2.1).

2.3.3 Lemma

Assume that there exist j, l such that the following conditions are satisfied

(i) $\left|\lambda_k'(x,\rho)\,[P(X > x+\rho)]^j\right|$ is bounded for ρ and

(ii) $[P(X > x+\rho)]^l$ is integrable for ρ.

Then $\quad I''_{k+1} \to 0 \quad$ when $m \to \infty$,

where $\quad I''_{k+1} = \int_{x-a}^{b-x} \lambda_k'(x,\rho)\,[P(X > x+\rho)]^{m+k}\,d\rho.$

Proof:
As in part (2.1).

2.3.4 Corollary

$$I' = q_\circ(x,0) - q_\circ(x,x-a)(P(X > 2x-a))^m$$

$$+\sum_{k=1}^{\infty} \frac{q_k(x,0)}{\prod_{n=1}^{k}(m+n)} - \sum_{k=1}^{\infty} \frac{q_k(x,x-a)(P(X>2x-a))^{m+k}}{\prod_{n=1}^{k}(m+n)} \quad (2.3.16)$$

and

$$I'' = \lambda_\circ(x,x-a)(P(X > 2x-a))^m + \sum_{k=1}^{\infty} \frac{\lambda_k(x,x-a)(P(X>2x-a))^{m+k}}{\prod_{n=1}^{k}(m+n)} \quad (2.3.17)$$

under the condition $\frac{I'_{k+1}}{\prod_{n=1}^{k}(m+n)}$ and $\frac{I''_{k+1}}{\prod_{n=1}^{k}(m+n)}$ that tend to zero when $k \to \infty$,

where $q_k(x,0)$, $q_k(x,x-a)$ and $\lambda_k(x,x-a)$ are defined as in Lemma 2.3.1.

2.3.5 Lemma

Let the conditions of Lemmas 2.3.1-2.3.3 and Corollary 2.3.4 be satisfied. Then

$$I = q_\circ(x,0) + (\lambda_\circ(x,x-a) - q_\circ(x,x-a))\,(P(X > 2x-a))^m$$

$$+\sum_{k=1}^{\infty} \frac{q_k(x,0)}{\prod_{n=1}^{k}(m+n)} + \sum_{k=1}^{\infty} \frac{(\lambda_k(x,x-a)-q_k(x,x-a))(P(X>2x-a))^{m+k}}{\prod_{n=1}^{k}(m+n)}, \quad (2.3.18)$$

where $q_k(x,0)$, $q_k(x,x-a)$ and $\lambda_k(x,x-a)$ are defined as above in Lemma 2.3.1.

Proof:
Substituting (2.3.16) and (2.3.17) in to (2.3.1).

Similarly, we can obtain an asymptotic expansion for J where

$$J = \int_a^b f_1(x') m P(|X - x| > |x' - x|)^{m-1} dx'$$

$$= m \int_0^{x-a} (f_1(x-\rho) + f_1(x+\rho)) \left[P(X < x - \rho) + P(X > x + \rho) \right]^{m-1} d\rho$$

$$+ m \int_{x-a}^{b-x} f_1(x+\rho) \left[P(X > x + \rho) \right]^{m-1} d\rho$$

2.3.6 Lemma

Under suitable conditions as in Lemma 2.3.5

$$J = \overline{q}_\circ(x, 0) + \left(\overline{\lambda}_\circ(x, x-a) - \overline{q}_\circ(x, x-a)\right)(P(X > 2x - a))^m$$

$$+ \sum_{k=1}^{\infty} \frac{\overline{q}_k(x,0)}{\prod_{n=1}^{k}(m+n)} + \sum_{k=1}^{\infty} \frac{\left(\overline{\lambda}_k(x,x-a) - \overline{q}_k(x,x-a)\right)(P(X > 2x-a))^{m+k}}{\prod_{n=1}^{k}(m+n)} \qquad (2.3.19)$$

where

$\overline{q}_\circ(x, \rho) = \frac{f_1(x-\rho) + f_1(x+\rho)}{f(x-\rho) + f(x+\rho)}$ and $\overline{q}_k(x, \rho) = \frac{\overline{q}'_{k-1}(x,\rho)}{f(x-\rho) + f(x+\rho)}$ for $k \geq 1$,

and

$\overline{\lambda}_\circ(x, \rho) = \frac{f_1(x+\rho)}{f(x+\rho)}$ and $\overline{\lambda}_k(x, \rho) = \frac{\overline{\lambda}'_{k-1}(x,\rho)}{f(x+\rho)}$ for $k \geq 1$.

2.3.7 Theorem

Let the conditions of Lemmas 2.3.5 and 2.3.6 be satisfied. Then

$$R_m(x) = P(\theta' \neq \theta \,|\, X = x) = \frac{p_1 p_2 f_1(x) f_2(x)}{f^2(x)} + \sum_{k=1}^{\infty} \frac{\zeta_k}{\prod_{n=1}^{k}(m+n)}$$

$$+ \xi_\circ (P(X > 2x - a))^m + \sum_{k=1}^{\infty} \frac{\xi_k (P(X > 2x-a))^{m+k}}{\prod_{n=1}^{k}(m+n)} \qquad (2.3.20)$$

where

$$\zeta_k = \frac{p_1 p_2}{f(x)} \left[f_1(x)\, q_k(x, 0) + f_2(x)\, \overline{q}_k(x, 0) \right], \qquad k = 1, 2, 3, ...$$

$$\xi_k = \frac{p_1 p_2}{f(x)} \left[f_1(x) \left(\lambda_k(x, x-a) - q_k(x, x-a) \right) \right.$$

$$\left. + f_2(x) \left(\overline{\lambda}_k(x, x-a) - \overline{q}_k(x, x-a) \right) \right], \qquad k = 0, 1, 2, ...$$

Proof:
Substituting (2.3.18) and (2.3.19) in to (1.4.9). We hence obtain

$$R_m(x) = \tfrac{p_1p_2f_1(x)}{f(x)} \left\{q_\circ(x,0) + (\lambda_\circ(x,x-a) - q_\circ(x,x-a))\,(P(X>2x-a))^m \right.$$
$$\left. + \textstyle\sum_{k=1}^{\infty} \tfrac{q_k(x,0)}{\prod_{n=1}^{k}(m+n)} + \sum_{k=1}^{\infty} \tfrac{(\lambda_k(x,x-a)-q_k(x,x-a))(P(X>2x-a))^{m+k}}{\prod_{n=1}^{k}(m+n)} \right\}$$

$$+\tfrac{p_1p_2f_2(x)}{f(x)} \left\{\overline{q}_\circ(x,0) + \left(\overline{\lambda}_\circ(x,x-a) - \overline{q}_\circ(x,x-a)\right)(P(X>2x-a))^m \right.$$
$$\left. + \textstyle\sum_{k=1}^{\infty} \tfrac{\overline{q}_k(x,0)}{\prod_{n=1}^{k}(m+n)} + \sum_{k=1}^{\infty} \tfrac{\left(\overline{\lambda}_k(x,x-a)-\overline{q}_k(x,x-a)\right)(P(X>2x-a))^{m+k}}{\prod_{n=1}^{k}(m+n)} \right\}$$

$$= \tfrac{p_1p_2f_1(x)f_2(x)}{f^2(x)} + \tfrac{p_1p_2}{f(x)} \textstyle\sum_{k=1}^{\infty} \tfrac{f_1(x)q_k(x,0)+f_2(x)\overline{q}_k(x,0)}{\prod_{n=1}^{k}(m+n)}$$

$$+\tfrac{p_1p_2(P(X>2x-a))^m}{f(x)} \left[f_1(x)(\lambda_\circ(x,x-a) - q_\circ(x,x-a)) \right.$$
$$\left. +f_2(x)(\overline{\lambda}_\circ(x,x-a) - \overline{q}_\circ(x,x-a)\right]$$

$$+\tfrac{p_1p_2}{f(x)} \textstyle\sum_{k=1}^{\infty} \tfrac{f_1(x)(\lambda_k(x,x-a)-q_k(x,x-a))+f_2(x)\left(\overline{\lambda}_k(x,x-a)-\overline{q}_k(x,x-a)\right)(P(X>2x-a))^{m+k}}{\prod_{n=1}^{k}(m+n)}$$

$$= \tfrac{p_1p_2f_1(x)f_2(x)}{f^2(x)} + \textstyle\sum_{k=1}^{\infty} \tfrac{\zeta_k}{\prod_{n=1}^{k}(m+n)} + \xi_\circ(P(X>2x-a))^m$$
$$+ \textstyle\sum_{k=1}^{\infty} \tfrac{\xi_k(P(X>2x-a))^{m+k}}{\prod_{n=1}^{k}(m+n)}.$$

2.4 Special Case: $S = (0,1)$

We can obtain an expansion for $P(\theta' \neq \theta \,|X = x)$ when the support is $S = (0,1)$ by using $a = 0$, $b = 1$ in the previous part. Then we obtain the following form for $0 \leq x \leq \frac{1}{2}$:

$$R_m(x) = \tfrac{p_1 p_2 f_1(x)}{f(x)} \{q_\circ(x,0) + (\lambda_\circ(x,x) - q_\circ(x,x))\,(P(X > 2x))^m + \textstyle\sum_{k=1}^{\infty} \tfrac{q_k(x,0)}{\prod_{n=1}^{k}(m+n)} + \textstyle\sum_{k=1}^{\infty} \tfrac{(\lambda_k(x,x) - q_k(x,x))(P(X>2x))^{m+k}}{\prod_{n=1}^{k}(m+n)}\Big\}$$

$$+\tfrac{p_1 p_2\, f_2(x)}{f(x)} \Big\{\overline{q}_\circ(x,0) + \Big(\overline{\lambda}_\circ(x,x) - \overline{q}_\circ(x,x)\Big)\,(P(X > 2x))^m + \textstyle\sum_{k=1}^{\infty} \tfrac{\overline{q}_k(x,0)}{\prod_{n=1}^{k}(m+n)} + \textstyle\sum_{k=1}^{\infty} \tfrac{\left(\overline{\lambda}_k(x,x) - \overline{q}_k(x,x)\right)(P(X>2x))^{m+k}}{\prod_{n=1}^{k}(m+n)}\Big\}$$

$$= \tfrac{p_1 p_2 f_1(x) f_2(x)}{f^2(x)} + \textstyle\sum_{k=1}^{\infty} \tfrac{\zeta_k}{\prod_{n=1}^{k}(m+n)} + \xi_\circ (P(X > 2x))^m + \textstyle\sum_{k=1}^{\infty} \tfrac{\xi_k\,(P(X>2x))^{m+k}}{\prod_{n=1}^{k}(m+n)}, \tag{2.4.1}$$

where

$$\zeta_k = \tfrac{p_1 p_2}{f(x)} (f_1(x) q_k(x,0) + f_2(x) \overline{q}_k(x,0))\ , \qquad k = 1, 2, 3, ...$$

$$\xi_k = \tfrac{p_1 p_2}{f(x)} \Big[f_1(x)(\lambda_k(x,x) - q_k(x,x)) + f_2(x)(\overline{\lambda}_k(x,x) - \overline{q}_k(x,x))\Big], \qquad k = 0, 1, 2, ...$$

3 Integrating the Asymptotic Expansion for $R_m(x)$

In the previous chapter we evaluated the probability of error conditioned on the event $\{X = x\}$ by averaging $P(\theta' \neq \theta \,|X' = x', X = x\,)$ over the value of the nearest neighbor of x to obtain an asymptotic expansion for $R_m(x) = P\left(\theta' \neq \theta \,|X = x\right)$.

Now we shall attempt to average this result over x in order to obtain an asymptotic expansion for R_m. Hence R_m take the form as (1.4.1):

$$R_m = \textstyle\int_S \, P(\theta' \neq \theta \,|X = x\,)\, f(x)\, dx \tag{3.0.1}$$

There we find essentially different situations for bounded and unbounded support.

Integrating $P\left(\theta' \neq \theta \,|X = x\right)$ with respect to x we find that, in examples like the normal and exponential distribution, the integrals diverge. This seems to be typical for the case of unbounded support.

3.1 The Case of Unbounded Support

In this section we shall present an example for normal distributions where $S = (-\infty, \infty)$, and an example for exponential distributions where $S = (0, \infty)$.

3.1.1 The Example of Normal Distributions

Let $f_1(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma_1^2}}$, $f_2(x) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(x-b)^2}{2\sigma_2^2}}$ be two densities for normal distributions with prior probabilities p_1, p_2 such that $p_1 + p_2 = 1$, and let $\sigma_1 = \sigma_2 = 1$.

Let us look at the expansion up to the second order. Then

$$R_m(x) = \tfrac{2\,p_1p_2f_1(x)f_2(x)}{f^2(x)} + \tfrac{1}{(m+1)(m+2)}\left(\tfrac{p_1p_2\,f_1(x)}{f(x)} q_2(x,0) + \tfrac{p_1p_2\,f_2(x)}{f(x)} \overline{q}_2(x,0)\right)$$
$$+ \tfrac{1}{(m+1)(m+2)}(I_3 + J_3),$$

where I_3, and J_3 as in Lemma 2.1.1.

Substituting this equation in to (3.0.1) we would obtain formally

$$R_m = \int_{-\infty}^{\infty} \frac{2\,p_1p_2\,f_1(x)f_2(x)}{f(x)}\,dx$$

$$+ \int_{-\infty}^{\infty} \frac{1}{(m+1)(m+2)}\left(p_1p_2f_1(x)q_2(x,0) + p_1p_2f_2(x)\overline{q}_2(x,0)\right)dx$$

$$+ \int_{-\infty}^{\infty} \frac{f(x)(I_3+J_3)}{(m+1)(m+2)}dx,$$

where

$$q_2(x,0) = \frac{h''(x,0)g(x,0)-h(x,0)g''(x,0)}{g^4(x,0)}, \qquad \overline{q}_2(x,0) = \frac{\overline{h}''(x,0)g(x,0)-\overline{h}(x,0)g''(x,0)}{g^4(x,0)},$$

$$h(x,\rho) = f_2(x+\rho)+f_2(x-\rho), \qquad \overline{h}(x,\rho) = f_1(x+\rho)+f_1(x-\rho),$$

$$g(x,\rho) = f(x+\rho)+f(x-\rho), \text{ and } \qquad f(x) = p_1f_1(x)+p_2f_2(x).$$

Then

$$h(x,0) = 2f_2(x) = \frac{2}{\sqrt{2\pi}}e^{-\frac{(x-b)^2}{2}}, \qquad \overline{h}(x,0) = 2f_1(x) = \frac{2}{\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2}},$$

$$h''(x,0) = \frac{2((x-b)^2-1)}{\sqrt{2\pi}}e^{-\frac{(x-b)^2}{2}}, \qquad \overline{h}''(x,0) = \frac{2((x-a)^2-1)}{\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2}},$$

$$g(x,0) = 2f(x) = \frac{2p_1}{\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2}} + \frac{2p_2}{\sqrt{2\pi}}e^{-\frac{(x-b)^2}{2}}, \text{ and}$$

$$g''(x,0) = \frac{2p_1((x-a)^2-1)}{\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2}} + \frac{2p_2((x-b)^2-1)}{\sqrt{2\pi}}e^{-\frac{(x-b)^2}{2}}.$$

Now, by using these functions to evaluate the second integral we find that

$$L = \int_{-\infty}^{\infty} \frac{1}{(m+1)(m+2)}\left(p_1p_2f_1(x)q_2(x,0) + p_1p_2f_2(x)\overline{q}_2(x,0)\right)dx$$

$$= \frac{p_1p_2}{(m+1)(m+2)}\left(\int_{-\infty}^{\infty} \frac{f_1(x)h''(x,0)g(x,0)-f_1(x)h(x,0)g''(x,0)}{g^4(x,0)}dx\right.$$

$$\left. + \int_{-\infty}^{\infty} \frac{f_2(x)\overline{h}''(x,0)g(x,0)-f_2(x)\overline{h}(x,0)g''(x,0)}{g^4(x,0)}dx\right)$$

These integrals can be written in the form

$$L = \frac{p_1p_2}{(m+1)(m+2)} \left(\int_{-\infty}^{\infty} C_1(x)e^{\frac{(x-a)^2}{2}}dx + \int_{-\infty}^{\infty} C_2(x)e^{\frac{(x-b)^2}{2}}dx \right)$$

Looking at $C_1(x)$ and $C_2(x)$, it is easily seen that these integrals are divergent, that is, the integrals in the expansion of $R_m(x)$ with respect to x in the case of normal distribution diverge.

3.1.2 The Example of Exponential Distributions

Let $f_1(x) = ae^{-ax}$, $f_2(x) = be^{-bx}$ be two densities for exponential distributions with prior probabilities p_1, p_2 such that $p_1 + p_2 = 1$, and $a, b > 0$.

Let us look at the expansion up to the second order. Then

$$\begin{aligned}
R_m(x) = {} & \frac{2p_1p_2f_1(x)f_2(x)}{f^2(x)} + \frac{P_1P_2}{(m+1)(m+2)f(x)} \left(f_1(x)q_2(x,0) + f_2(x)\overline{q}_2(x,0)\right) \\
& + \frac{p_1p_2}{f(x)} \left(f_1(x) \left(\lambda_\circ(x,x) - q_\circ(x,x)\right)\right. \\
& \qquad \left. + f_2(x) \left(\overline{\lambda}_\circ(x,x) - \overline{q}_\circ(x,x)\right)\right) (P(X > 2x))^m \\
& + \frac{p_1p_2}{(m+1)f(x)} \left(f_1(x) \left(\lambda_1(x,x) - q_1(x,x)\right)\right. \\
& \qquad \left. + f_2(x) \left(\overline{\lambda}_1(x,x) - \overline{q}_1(x,x)\right)\right) (P(X > 2x))^{m+1} \\
& + \frac{p_1p_2}{(m+1)(m+2)f(x)} \left(f_1(x) \left(\lambda_2(x,x) - q_2(x,x)\right)\right. \\
& \qquad \left. + f_2(x) \left(\overline{\lambda}_2(x,x) - \overline{q}_2(x,x)\right)\right) (P(X > 2x))^{m+2} + \frac{(I_3+J_3)}{(m+1)(m+2)}
\end{aligned}$$

where $I_3 = I_3' + I_3''$, and $J_3 = J_3' + J_3''$ as in part 2.2.

Substituting this equation in to (3.0.1) we would obtain formally

$$\begin{aligned}
R_m = {} & \int_0^\infty \frac{2p_1p_2f_1(x)f_2(x)}{f(x)}dx + \int_0^\infty \frac{p_1p_2}{(m+1)(m+2)} \left(f_1(x)q_2(x,0) + f_2(x)\overline{q}_2(x,0)\right) dx \\
& + \int_0^\infty \beta_\circ(P(X > 2x))^m f(x)dx + \int_0^\infty \frac{p_1p_2}{(m+1)} \left(f_1(x) \left(\lambda_1(x,x) - q_1(x,x)\right)\right. \\
& \qquad \left. + f_2(x) \left(\overline{\lambda}_1(x,x) - \overline{q}_1(x,x)\right)\right) (P(X > 2x))^{m+1}dx
\end{aligned}$$

$$+\int_0^\infty \tfrac{p_1 p_2}{(m+1)(m+2)} \left(f_1(x)\left(\lambda_2(x,x) - q_2(x,x)\right)\right.$$

$$\left.+f_2(x)\left(\overline{\lambda}_2(x,x) - \overline{q}_2(x,x)\right)\right)(P(X > 2x))^{m+2}dx$$

$$+\int_0^\infty \tfrac{f(x)(I_3+J_3)}{(m+1)(m+2)}dx,$$

where

$$q_2(x,0) = \tfrac{h''(x,0)g(x,0)-h(x,0)g''(x,0)}{g^4(x,0)}, \qquad \overline{q}_2(x,0) = \tfrac{\overline{h}''(x,0)g(x,0)-\overline{h}(x,0)g''(x,0)}{g^4(x,0)},$$

$$h(x,\rho) = f_2(x+\rho) + f_2(x-\rho), \qquad \overline{h}(x,\rho) = f_1(x+\rho) + f_1(x-\rho),$$

$$g(x,\rho) = f(x+\rho) + f(x-\rho), \text{ and } \qquad f(x) = p_1 f_1(x) + p_2 f_2(x).$$

Then

$$h(x,0) = 2f_2(x) = 2b\,e^{-bx}, \qquad \overline{h}(x,0) = 2f_1(x) = 2a\,e^{-ax},$$

$$h''(x,0) = 2b^3e^{-bx}, \qquad \overline{h}''(x,0) = 2a^3e^{-ax},$$

$$g(x,0) = 2f(x) = 2p_1 a e^{-ax} + 2p_2 b e^{-bx}, \text{ and}$$

$$g''(x,0) = 2p_1 a^3 e^{-ax} + 2p_2 b^3 e^{-bx}.$$

Now, by using these functions to evaluate the second integral we find that

$$\overline{L} = \int_0^\infty \tfrac{p_1 p_2}{(m+1)(m+2)} \left(f_1(x)q_2(x,0) + f_2(x)\overline{q}_2(x,0)\right)dx$$

$$= \tfrac{p_1 p_2}{(m+1)(m+2)} \left(\int_0^\infty \tfrac{f_1(x)h''(x,0)g(x,0)-f_1(x)h(x,0)g''(x,0)}{g^4(x,0)}dx\right.$$

$$\left.+\int_0^\infty \tfrac{f_2(x)\overline{h}''(x,0)g(x,0)-f_2(x)\overline{h}(x,0)g''(x,0)}{g^4(x,0)}dx\right)$$

These integrals can be written in the form

$$\overline{L} = \tfrac{p_1 p_2}{(m+1)(m+2)} \left(\int_0^\infty C_1 e^{ax}dx + \int_0^\infty C_2 e^{bx}dx\right)$$

Looking at $C_1(x)$ and $C_2(x)$, it is clear that these integrals are divergent, that is, the integrals in the expansion of $R_m(x)$ with respect to x in the case of exponential distribution diverge.

3.2 The Case of Bounded Support

In this case we present the following condition, and an asymptotic expansion for R_m in the case of triangular distributions.

3.2.1 A General Condition

Assume that the denominator in our expansions is bounded away from zero. Then we may integrate the asymptotic expansion.

We present the example of triangular distributions such that $S = (0,1)$.

3.2.2 The case of triangular distributions

Consider the one-dimensional triangular distribution over the unit interval, $f_1(x) = 2 - 2x$, and $f_2(x) = 2x$ with prior probabilities $p_1 = p_2 = \frac{1}{2}$. Then the density $f = p_1 f_1 + p_2 f_2$ is uniform on $[0,1]$. We may use the special case (2.4) when $S = (0,1)$. Thus using the expansion up to second order we obtain

$$
\begin{aligned}
R_m = \textstyle\int_0^1 p_1 p_2 f_1(x) \{ & q_\circ(x,0) + (\lambda_\circ(x,x) - q_\circ(x,x)) \cdot (P(X > 2x))^m \\
& + \tfrac{1}{m+1} \lambda_1(x,x)(P(X > 2x))^{m+1} \Big\} \, dx \\
+ \textstyle\int_0^1 p_1 p_2 f_2(x) & \Big\{ \overline{q}_\circ(x,0) + \Big(\overline{\lambda}_\circ(x,x) - \overline{q}_\circ(x,x)\Big) \cdot (P(X > 2x))^m \\
& + \tfrac{1}{m+1} \overline{\lambda}_1(x,x)(P(X > 2x))^{m+1} \Big\} \, dx + \ldots
\end{aligned}
$$

Since $q_1 = \overline{q}_1 = 0$ and all second derivatives of f_1 and f_2 are identically zero the remainder term is equal to zero.

Using symmetry with respect to the point 1/2 we get

$$
\begin{aligned}
R_m &= \tfrac{1}{2} \textstyle\int_0^{\frac{1}{2}} (2 - 2x) \cdot \Big\{ 2x + 2x\,(1 - 2x)^m + \tfrac{2}{m+1} (1 - 2x)^{m+1} \Big\} \, dx \\
&\quad + \tfrac{1}{2} \textstyle\int_0^{\frac{1}{2}} 2x \cdot \Big\{ (2 - 2x) - 2x\,(1 - 2x)^m - \tfrac{2}{m+1} (1 - 2x)^{m+1} \Big\} \, dx \\
&= 2 \Big\{ \textstyle\int_0^{\frac{1}{2}} (x - x^2) dx + \int_0^{\frac{1}{2}} (x - x^2)(1 - 2x)^m dx +
\end{aligned}
$$

$$+\tfrac{1}{m+1}\int_0^{\frac{1}{2}}(1-x)(1-2x)^{m+1}dx+\int_0^{\frac{1}{2}}(x-x^2)dx$$

$$-\int_0^{\frac{1}{2}}x^2(1-2x)^m dx-\tfrac{1}{m+1}\int_0^{\frac{1}{2}}x(1-2x)^{m+1}dx\Big\}$$

$$=4\int_0^{\frac{1}{2}}(x-x^2)dx+2\int_0^{\frac{1}{2}}(x-2x^2)(1-2x)^m dx$$

$$+\tfrac{2}{m+1}\int_0^{\frac{1}{2}}(1-2x)(1-2x)^{m+1}dx$$

$$=\tfrac{1}{3}+\tfrac{1}{2(m+1)(m+2)}-\tfrac{1}{(m+1)(m+2)(m+3)}+\tfrac{1}{(m+1)(m+3)}$$

$$=\tfrac{1}{3}+\tfrac{3m+5}{2(m+1)(m+2)(m+3)}$$

We remark that in this case the infinite sample risk $R_\infty=\frac{1}{3}$ and the rate of convergence of R_m to R_∞ is $O(m^{-2})$. This example was treated by differnt methods in [2].

4 The Asymptotic Evaluation of $R_m(x)$ by Laplace's Method

In this chapter we present another method to evaluate $R_m(x)$. In chapter one, we showed that

$$R_m(x) = P(\theta^{'} \neq \theta \,| X = x\,) = \frac{p_1 p_2 f_1(x)}{f(x)}\, I + \frac{p_1 p_2 f_2(x)}{f(x)}\, J$$

The method of evaluation $R_m(x)$ proceeds in several stages:

First, we derive an exact integral expression for I and J in the form $\int_S g\, e^{-mh}$ where g and h are nonnegative functions. For large m, this integral appears to be in a form amenable to Laplace's asymptotic method. This asserts that for large m the dominant contribution to the integral arises from a neighborhood of the point where h has a minimum. If h has more than one minimum, then the domain of integration can be partitioned so that each subdomain contains only one minimum.

Second, we represent g and h as asymptotic power series in a neighborhood of this minimum, and then the integral itself may be represented as an asymptotic power series in reciprocal powers of m, compare [19], [20].

4.1 A General Result

From chapter one, the probability of error conditioned on the event $\{X = x\}$ can be written in the form

$$R_{m+1}(x) = P(\theta^{'} \neq \theta \,| X = x\,) = \tfrac{p_1 p_2 f_1(x)}{f(x)} I + \tfrac{p_1 p_2 f_2(x)}{f(x)} J \tag{4.1.1}$$

where
$I = I(x) = (m+1) \int_{-\infty}^{\infty} f_2(x')(P(|X-x| > |x'-x|))^m dx'$

$= (m+1) \int_0^{\infty} (f_2(x-\rho) + f_2(x+\rho))\, [P(X < x-\rho) + P(X > x+\rho)\ ]^m\, d\rho$

$$= (m+1) \int_0^{\infty} (f_2(x-\rho) + f_2(x+\rho))\ (1 - (F(x+\rho) - F(x-\rho)))^m\, d\rho \tag{4.1.2}$$

and
$J = J(x) = (m+1) \int_{-\infty}^{\infty} f_1(x')(P(|X-x| > |x'-x|))^m\, dx'$

$= (m+1) \int_0^\infty (f_1(x-\rho) + f_1(x+\rho)) \, [P(X < x-\rho) + P(X > x+\rho) \,]^m \, d\rho$

$$= (m+1) \int_0^\infty (f_1(x-\rho)+f_1(x+\rho)) \; (1-(F(x+\rho)-F(x-\rho)))^m \, d\rho \quad (4.1.3)$$

Putting $P(x,\rho) = -\log\,(1-(F(x+\rho)-F(x-\rho)))$ (4.1.4)

$\implies e^{-mP(x,\rho)} = (1-(F(x+\rho)-F(x-\rho)))^m$

Then the integrals of I and J take the forms:

$$I = (m+1) \int_0^\infty (f_2(x-\rho)+f_2(x+\rho)) e^{-mP(x,\rho)} d\rho = (m+1) I_1 \quad (4.1.5)$$

$$J = (m+1) \int_0^\infty (f_1(x-\rho)+f_1(x+\rho)) e^{-mP(x,\rho)} d\rho = (m+1) J_1 \quad (4.1.6)$$

where

$I_1 = \int_0^\infty (f_2(x-\rho) + f_2(x+\rho)) e^{-mP(x,\rho)} d\rho,$

and

$J_1 = \int_0^\infty (f_1(x-\rho) + f_1(x+\rho)) e^{-mP(x,\rho)} d\rho.$

The above integrals have the desired form, so that we can apply Laplace's method for fixed x.

The asymptotic expansions for I_1 and J_1 are given by the following Lemma. For this Lemma, we set $Q(x,\rho) = f_2(x-\rho) + f_2(x+\rho)$, and $\overline{Q}(x,\rho) = f_1(x-\rho) + f_1(x+\rho)$. We proceed as in [17].

4.1.1 Lemma

Let the functions $P(x,\rho)$, $Q(x,\rho)$, and $\overline{Q}(x,\rho)$ be defined as above. Assume that the following conditions are satisfied:

(i) f_1, f_2 have a power series expansion.

(ii) $I_1(m) \equiv \int_0^\infty e^{-mP(x,\rho)} \, Q(x,\rho) \, d\rho, \quad J_1(m) \equiv \int_0^\infty e^{-mP(x,\rho)} \, \overline{Q}(x,\rho) \, d\rho$

converge absolutely throughout its range for all sufficiently large m.

Then

$$I_1 = \int_0^\infty e^{-mP(x,\rho)}\, Q(x,\rho)\, d\rho \sim e^{-mP(x,0)} \sum_{s=0}^{N-1} \Gamma(s+1)\, \frac{a_s}{m^{s+1}} + O(m^{-(N+1)}),$$

and

$$J_1 = \int_0^\infty e^{-mP(x,\rho)}\, \overline{Q}(x,\rho)\, d\rho \sim e^{-mP(x,0)} \sum_{s=0}^{N-1} \Gamma(s+1)\, \frac{a'_s}{m^{s+1}} + O(m^{-(N+1)})$$

where the a_s and a'_s are defined through the proof, and Γ denotes the Gamma function s.t. $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$.

Proof:
We start by expanding I_1:

(i) We compute an asymptotic expansion for the function $Q(x,\rho)$.

By using a Taylor expansion for the functions $f_2(x+\rho)$ and $f_2(x-\rho)$, we obtain

$$f_2(x+\rho) = f_2(x) + f_2'(x)\frac{\rho}{1!} + f_2''(x)\frac{\rho^2}{2!} + f_2'''(x)\frac{\rho^3}{3!} + ...,$$

$$f_2(x-\rho) = f_2(x) - f_2'(x)\frac{\rho}{1!} + f_2''(x)\frac{\rho^2}{2!} - f_2'''(x)\frac{\rho^3}{3!} + ...$$

Then

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

that is,

$$Q(x,\rho) = \sum_{s=0}^{\infty} \frac{2}{(2s)!} f_2^{(2s)}(x)\, \rho^{2s} = \sum_{s=0}^{\infty} \alpha_s\,(x)\, \rho^s \qquad (\rho \longrightarrow 0) \tag{4.1.7}$$

where

$$\alpha_\circ = 2f_2(x),\ \alpha_1 = 0,\ \alpha_2 = f_2''(x),\ \alpha_3 = 0,\ \alpha_4 = \frac{f_2^{(4)}(x)}{12},\ \alpha_5 = 0, ...$$

(ii) We now compute an asymptotic expansion for the function $P(x,\rho)$.

Firstly, we need good asymptotic estimates, as $(\rho \longrightarrow 0)$, for $F(x+\rho) - F(x-\rho)$. By using the Taylor expansion for the functions $F(x+\rho)$ and $F(x-\rho)$ we obtain

$$F(x+\rho) = F(x) + \tfrac{f(x)\rho}{1!} + \tfrac{f'(x)\rho^2}{2!} + \tfrac{f''(x)\rho^3}{3!} + \tfrac{f'''(x)\rho^4}{4!} + \tfrac{f^{(4)}(x)\rho^5}{5!} + \dots ,$$

$$F(x-\rho) = F(x) - \tfrac{f(x)\rho}{1!} + \tfrac{f'(x)\rho^2}{2!} - \tfrac{f''(x)\rho^3}{3!} + \tfrac{f'''(x)\rho^4}{4!} - \tfrac{f^{(4)}(x)\rho^5}{5!} + \dots$$

Then

$$\begin{aligned} F(x+\rho) - F(x-\rho) &= \tfrac{2f(x)\rho}{1!} + \tfrac{2f''(x)\rho^3}{3!} + \tfrac{2f^{(4)}(x)\rho^5}{5!} + \dots \\ &= \textstyle\sum_{s=0}^{\infty} \widetilde{P_s}(x)\, \rho^{s+1}, \end{aligned} \tag{4.1.8}$$

where $\widetilde{P}_{2n}(x) = \frac{2f^{(2n)}(x)}{(2n+1)!}$ for $n = 0, 1, 2, \dots$, and $\widetilde{P_1}(x) = \widetilde{P_3}(x) = \dots = 0$.

Use (4.1.8) and the series representation

$$\textstyle\sum_{s=1}^{\infty} \frac{y^s}{s} = -\log(1-y), \qquad\qquad y \in (-1, 1).$$

After substituting in (4.1.4) we obtain

$$\begin{aligned} P(x,\rho) &= \textstyle\sum_{s=1}^{\infty} \frac{\left(\sum_{k=0}^{\infty} \widetilde{P}_k(x)\, \rho^{k+1}\right)^s}{s} \\ &= \rho \widetilde{P}_\circ(x) + \rho^2 \left(\tfrac{\widetilde{P}_\circ^2}{2} + \widetilde{P}_1\right) + \rho^3 \left(\tfrac{\widetilde{P}_\circ^3}{3} + \widetilde{P}_\circ \widetilde{P}_1 + \widetilde{P}_2\right) + \dots \\ &= \textstyle\sum_{s=0}^{\infty} P_s(x) \rho^{s+1} \qquad\qquad (\rho \longrightarrow 0) \end{aligned} \tag{4.1.9}$$

where $P_\circ(x) = 2f(x)$, $P_1(x) = 2f^2(x)$, $P_2(x) = \frac{8}{3} f^3(x) + \frac{1}{3} f''(x)$, ...

We note that $P(x, 0) = 0$.

Equation (4.1.9) can be differentiated, that is

$$P'(x,\rho) = \textstyle\sum_{s=0}^{\infty} (s+1)\, P_s(x)\, \rho^s \qquad\qquad (\rho \longrightarrow 0) \tag{4.1.10}$$

(iii) Now we change the variable of integration.

We can find a number z which is close enough to 0 to ensure that in $(0, z]$, $P'(x,\rho)$ is continuous and positive and $Q(x,\rho)$ is continuous. Since $P(x,\rho)$ increases in $(0, z)$ we may take

$$v = P(x,\rho) - P(x,0).$$

as new integration variable in this interval. Then v and ρ are continuous functions of each other and

$$e^{-mP(x,0)} \int_0^z Q(x,\rho)\, e^{-mP(x,\rho)}\, d\rho = \int_0^Z e^{-mv}\, f(v)\, dv$$

where $\quad Z = P(x,z) - P(x,0)\ , \qquad f(v) = Q(x,\rho)\frac{d\rho}{dv} = \frac{Q(x,\rho)}{P'(x,\rho)}$

Although $P(x,0) = 0$ we insert this term to show that the argument is also valid for $P(x,0) > 0$.

(iv) We now need an asymptotic expansion for $f(v)$.

Since $v = P(x,\rho) - P(x,0) = \sum_{s=0}^{\infty} P_s(x)\rho^{s+1}$, we obtain, compare [5], by using Lagrange inversion formula an expansion of the form

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

where the coefficients C_s are

$$C_s = \frac{1}{s!}\left[\left(\frac{d}{d\rho}\right)^{k-1} (h(\rho))^k\right]_{\rho=0} ,$$

$h(\rho) = \left(\sum_{s=0}^{\infty} P_s(x)\rho^s\right)^{-1}$, and $v = \frac{\rho}{h(\rho)} = \frac{\rho}{\left(\sum_{s=0}^{\infty} P_s(x)\rho^s\right)^{-1}}$.

The first three coefficients are

$$C_1 = \frac{1}{P_\circ},\ \ C_2 = -\frac{P_1}{P_\circ^3},\ \ C_3 = \frac{4P_1^2 - 2P_\circ P_2}{2P_\circ^5}.$$

We substitute this result in (4.1.7) and (4.1.10), and use the equation $f(v) = \frac{Q(x,\rho)}{P'(x,\rho)}$. Then

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

where $b_\circ = \frac{\alpha_\circ}{P_\circ}$, $b_1 = \frac{\alpha_1 P_\circ - 2\alpha_\circ P_1}{P_\circ^2}$,

$$b_2 = \frac{P_\circ^2\alpha_2 - 3P_\circ P_2\alpha_\circ - 2P_\circ P_1\alpha_1 + 2P_1^2\alpha_\circ}{P_\circ^3}, \ ...$$

Since $\rho = \sum_{s=1}^{\infty} C_s\, v^s$ when $(v \longrightarrow 0)$, then

$$f(v) = \sum_{s=0}^{\infty} b_s\, (\sum_{k=1}^{\infty} C_k v^k)^s = \sum_{s=0}^{\infty} b_s\, (C_1 v^1 + C_2 v^2 + C_3 v^3 + ...)^s$$

$$= b_\circ + b_1 C_1 v + (b_1 C_2 + b_2 C_1^2)\, v^2 + ... = \sum_{s=0}^{\infty} a_s\, v^s \qquad (v \longrightarrow 0)$$

where $a_\circ = b_\circ$, $a_1 = b_1 C_1$, $a_2 = b_1 C_2 + b_2 C_1^2, ...$, hence

$a_\circ = \frac{f_2}{f}$, $a_1 = \frac{-f_2}{f}$, $a_2 = \frac{f_2}{2f} + \frac{f_2'' f - f'' f_2}{8f^4} = \frac{4f_2 f^3 + f_2'' f - f'' f_2}{8f^4}$, ...

(v) Asymptotic evaluation on the range of integration $(0, z)$.

For each positive integer N, let the remainder term $f_N(v)$ be defined by $f_N(0) = a_N$ and

$$f(v) = \sum_{s=0}^{N-1} a_s v^s + v^N f_N(v) \qquad (v > 0). \qquad (4.1.11)$$

Then

$$\int_0^Z e^{-mv} f(v) dv = \int_0^Z e^{-mv} \left(\sum_{s=0}^{N-1} a_s v^s + v^N f_N(v) \right) dv$$

$$= \int_0^Z e^{-mv} \sum_{s=0}^{N-1} a_s v^s dv + \int_0^Z e^{-mv} v^N f_N(v) dv$$

$$= \sum_{s=0}^{N-1} a_s\, (\int_0^\infty e^{-mv} v^s dv - \int_Z^\infty e^{-mv} v^s dv) + \int_0^Z e^{-mv} v^N f_N(v) dv$$

$$= \sum_{s=0}^{N-1} \Gamma(s+1) \frac{a_s}{m^{s+1}} - \sum_{s=0}^{N-1} \Gamma(s+1, Zm) \frac{a_s}{m^{s+1}} + \int_0^Z e^{-mv} v^N f_N(v) dv$$

$$= \sum_{s=0}^{N-1} \Gamma(s+1) \frac{a_s}{m^{s+1}} - \varepsilon_{N,1}(m) + \varepsilon_{N,2}(m),$$

where $\varepsilon_{N,1}(m) = \sum_{s=0}^{N-1} \Gamma(s+1, Zm) \frac{a_s}{m^{s+1}}$, (4.1.12)

$$\varepsilon_{N,2}(m) = \int_0^Z e^{-mv} v^N f_N(v) dv \qquad (4.1.13)$$

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

$$\varepsilon_{N,1}(m) = O\left(\frac{e^{-Zm}}{m}\right). \qquad (4.1.14)$$

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

$$\varepsilon_{N,2}(m) = \int_0^Z e^{-mv}\, v^N O(1)\, dv = O\left(\frac{1}{m^{N+1}}\right). \qquad (4.1.15)$$

(vi) Asymptotic evaluation on the range (z, ∞).

For the remaining range (z, ∞), let M be a value of m for which $I_1(m)$ is absolutely convergent and write

$$\zeta \equiv inf_{[k,\infty)}\{P(x, \rho) - P(x, 0)\}$$

Since $P(x, 0) = 0$, and $P(x, \rho)$ strictly increasing in ρ from (4.1.9), hence $(P(x, \rho) - P(x, 0)) > 0$ for all $\rho > 0$. Then ζ is positive. Restricting $m \geq M$, we have

$$mP(x, \rho) - mP(x, 0) = (m - M)\{P(x, \rho) - P(x, 0)\} + M\{P(x, \rho) - P(x, 0)\}$$

$$\geq (m - M)\zeta + MP(x, \rho) - MP(x, 0),$$

then we obtain

$$\left|e^{mP(x,0)} \int_z^{\infty} Q(x, \rho) e^{-mP(x,\rho)} d\rho\right| \leq$$

$$e^{-(m-M)\zeta + MP(x,0)} \int_z^{\infty} |Q(x, \rho)| e^{-MP(x,\rho)} d\rho \qquad (4.1.16)$$

Thus
$I_1 = \int_0^{\infty} Q(x, \rho) e^{-mP(x,\rho)}\, d\rho \sim e^{-mP(x,0)} \sum_{s=0}^{\infty} \Gamma(s+1) \frac{a_s}{m^{(s+1)}}$

$$= e^{-mP(x,0)} \sum_{s=0}^{N-1} \Gamma(s+1) \frac{a_s}{m^{s+1}} + O(m^{-(N+1)}) \qquad (4.1.17)$$

Similarly,
$J_1 = \int_0^{\infty} \overline{Q}(x, \rho)\, e^{-mP(x,\rho)}\, d\rho \sim e^{-mP(x,0)} \sum_{s=0}^{\infty} \Gamma(s+1) \frac{a'_s}{m^{(s+1)}}$

$$= e^{-mP(x,0)} \sum_{s=0}^{N-1} \Gamma(s+1) \frac{a'_s}{m^{s+1}} + O(m^{-(N+1)}) \qquad (4.1.18)$$

where $a'_\circ = \frac{f_1}{f}$, $a'_1 = \frac{-f_1}{f}$, $a'_2 = \frac{f_1}{2f} + \frac{f''_1 f - f'' f_1}{8f^4} = \frac{4f_1 f^3 + f''_1 f - f'' f_1}{8f^4}$, ...

4.1.2 Theorem

Let the conditions of Lemma 4.1.1 be satisfied for all N so that the expansions (4.1.7), (4.1.9), and (4.1.10) hold. Then

$$R_{m+1}(x) = e^{-mP(x,0)} \left(\frac{2p_1p_2 f_1(x) f_2(x)}{f^2(x)} + \frac{p_1p_2}{m^2} \left(\frac{f_1 f''_2 f - 2f_1 f_2 f'' + f''_1 f_2 f}{4f^5} \right) \right.$$

$$+ \frac{p_1p_2}{m^3} \left(\frac{2f_1(x)f_2(x)}{f^2} + \frac{f_1 f''_2 f - 2f_1 f_2 f'' + f''_1 f_2 f}{4f^5} \right)$$

$$\left. + (m+1) \sum_{s=3}^{N-1} \frac{A_s}{m^{s+1}} + O\left(m^{-N}\right) \right)$$

where $A_s = p_1p_2\Gamma(s+1)\left(\frac{f_1 a_s + f_2 a'_s}{f}\right)$, and a_s, a'_s as in Lemma 4.1.1.

Proof:
Substituting (4.1.17) in to (4.1.5) we obtain

$$I = (m+1)I_1 = (m+1)(e^{-mP(x,0)} \sum_{s=0}^{N-1} \Gamma(s+1) \frac{a_s}{m^{s+1}} + O(m^{-(N+1)}))$$

$$= (m+1)e^{-mP(x,0)} \left(\frac{a_\circ}{m} + \frac{a_1}{m^2} + \frac{2a_2}{m^3} + \sum_{s=3}^{N-1} \Gamma(s+1) \frac{a_s}{m^{s+1}} + O(m^{-(N+1)})) \right)$$

$$= e^{-mP(x,0)} \left(\frac{m+1}{m} \frac{f_2}{f} - \frac{m+1}{m^2} \frac{f_2}{f} + \frac{2(m+1)}{m^3} \left(\frac{f_2}{2f} + \frac{f''_2 f - f'' f_2}{8f^4} \right) + ... \right)$$

$$= e^{-mP(x,0)} \left(\frac{f_2}{f} + \frac{1}{m} \frac{f_2}{f} - \frac{1}{m} \frac{f_2}{f} - \frac{1}{m^2} \frac{f_2}{f} + \frac{1}{m^2} \frac{f_2}{f} + \frac{1}{m^3} \frac{f_2}{f} \right.$$

$$\left. + \frac{2(m+1)}{m^3} \left(\frac{f''_2 f - f'' f_2}{8f^4} \right) + ... \right)$$

$$= e^{-mP(x,0)} \left(\frac{f_2}{f} + \frac{1}{m^2} \left(\frac{f''_2 f - f'' f_2}{4f^4} \right) + \frac{1}{m^3} \left(\frac{f_2}{f} + \frac{f''_2 f - f'' f_2}{4f^4} \right) + ... \right)$$

$$= e^{-mP(x,0)} \left(\frac{f_2}{f} + \frac{1}{m^2} \left(\frac{f''_2 f - f'' f_2}{4f^4} \right) + \frac{1}{m^3} \left(\frac{f_2}{f} + \frac{f''_2 f - f'' f_2}{4f^4} \right) \right.$$

$$\left. + (m+1) \sum_{s=3}^{N-1} \Gamma(s+1) \frac{a_s}{m^{s+1}} + O(m^{-(N+1)}) \right) \qquad (4.1.19)$$

Similarly, substituting (4.1.18) in to (4.1.6) we obtain

$$J=(m+1)J_1=(m+1)(e^{-mP(x,0)}\ \textstyle\sum_{s=1}^{N-1}\ \Gamma(s+1)\ \frac{a'_s}{m^{s+1}}+O(m^{-(N+1)}))$$

$$J=e^{-mP(x,0)}\left(\tfrac{f_1}{f}+\tfrac{1}{m^2}\left(\tfrac{f''_1 f-f'' f_1}{4f^4}\right)+\tfrac{1}{m^3}\left(\tfrac{f_1}{f}+\tfrac{f''_1 f-f'' f_1}{4f^4}\right)\right.$$
$$\left.+(m+1)\textstyle\sum_{s=3}^{N-1}\Gamma(s+1)\frac{a'_s}{m^{s+1}}+O(m^{-(N+1)})\right) \qquad (4.1.20)$$

Substituting (4.1.19) and (4.1.20) in to (4.1.1), then we obtain

$$R_{m+1}(x)=\tfrac{p_1p_2f_1(x)}{f(x)}e^{-mP(x,0)}\left(\tfrac{f_2}{f}+\tfrac{1}{m^2}\left(\tfrac{f''_2 f-f'' f_2}{4f^4}\right)+\tfrac{1}{m^3}\left(\tfrac{f_2}{f}+\tfrac{f''_2 f-f'' f_2}{4f^4}\right)\right.$$
$$\left.+\ (m+1)\textstyle\sum_{s=3}^{N-1}\Gamma(s+1)\frac{a_s}{m^{s+1}}+O(m^{-(N+1)})\right)$$
$$+\tfrac{p_1p_2f_2(x)}{f(x)}e^{-mP(x,0)}\left(\tfrac{f_1}{f}+\tfrac{1}{m^2}\left(\tfrac{f''_1 f-f'' f_1}{4f^4}\right)+\tfrac{1}{m^3}\left(\tfrac{f_1}{f}+\tfrac{f''_1 f-f'' f_1}{4f^4}\right)\right.$$
$$\left.+(m+1)\textstyle\sum_{s=3}^{N-1}\Gamma(s+1)\frac{a'_s}{m^{s+1}}+O(m^{-(N+1)})\right)$$
$$=e^{-mP(x,0)}\left(\tfrac{2P_1P_2f_1(x)f_2(x)}{f^2(x)}+\tfrac{p_1p_2}{m^2}\left(\tfrac{f_1f''_2f-2f_1f_2f''+f''_1f_2f}{4f^5}\right)\right.$$
$$+\tfrac{p_1p_2}{m^3}\left(\tfrac{2f_1(x)f_2(x)}{f^2}+\tfrac{f_1f''_2f-2f_1f_2f''+f''_1f_2f}{4f^5}\right)$$
$$\left.+(m+1)\textstyle\sum_{s=3}^{N-1}\ \frac{A_s}{m^{s+1}}+O\left(m^{-N}\right)\right),$$

where $A_s=p_1p_2\Gamma(s+1)\left(\frac{f_1a_s+f_2a'_s}{f}\right)$.

The following Corollary adresses $P(x,0)=0$, that is, $e^{-mP(x,0)}=1$.

4.1.3 Corollary

Under the conditions in Theorem 4.1.2, and if $P(x,0)=0$ we have

$$R_{m+1}(x)=\tfrac{2p_1p_2f_1(x)f_2(x)}{f^2(x)}+\tfrac{p_1p_2}{m^2}\left(\tfrac{f_1f''_2f-2f_1f_2f''+f''_1f_2f}{4f^5}\right)$$
$$+\tfrac{p_1p_2}{m^3}\left(\tfrac{2f_1(x)f_2(x)}{f^2}+\tfrac{f_1f''_2f-2f_1f_2f''+f''_1f_2f}{4f^5}\right)$$
$$+(m+1)\textstyle\sum_{s=3}^{N-1}\ \frac{A_s}{m^{s+1}}+O\left(m^{-N}\right),$$

where $A_s=p_1p_2\Gamma(s+1)\ \left(\frac{f_1a_s+f_2a'_s}{f}\right)$.

4.2 Error Estimates

It is well-known, compare [17], and can be seen from the proof of the Lemma 4.1.1 that the $N\,th$ truncation error of the expansion (4.1.17) can be expressed as

$$\int_0^\infty Q(x,\rho)e^{-mP(x,\rho)}\,d\rho - e^{-mP(x,0)}\sum_{s=1}^{N-1}\Gamma(s+1)\,\frac{a_s}{m^{s+1}}$$

$$= -e^{-mP(x,0)}\varepsilon_{N,1}(m)+e^{-mP(x,0)}\varepsilon_{N,2}(m)+\int_z^\infty Q(x,\rho)e^{-mP(x,\rho)}d\rho, \quad (4.2.1)$$

where z is a number in $(0,\infty]$, and $\varepsilon_{N,1}(m)$ and $\varepsilon_{N,2}(m)$ are defined by (4.1.12) and (4.1.13).

If the requirement in the proof of the Lemma 4.1.1 that z and Z be finite does not apply in (4.2.1), that is, if we take $z=\infty$, and $P(x,\infty)=\infty$, we obtain $Z=P(x,\infty)-P(x,0)=\infty$. Then the first error $\varepsilon_{N,1}(m)$ in (4.2.1) is absent.

In other cases, since the asymptotic expansion for the complemently incomplete Gamma funcation $\Gamma(\alpha,\, m)$ can be writen in the form:

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

$$+\varepsilon_N(m),$$

where N is an arbitrary nonnegative integer, and

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

We can show that

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

For the particular case $N=0$, we have

$$\Gamma(\alpha,m)\le\frac{e^{-m}\,m^\alpha}{m-\alpha+1},\qquad(\alpha>1,\ m>\alpha-1)$$

and for the special case $\alpha=1$, $m>0$ we have $\Gamma(1,m)\le e^{-m}$. Then

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

Substituting in (4.1.12) by means of this inequality, we obtain

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

where $Z = P(x,z) - P(x,0)$ as before, and $\alpha_N = \max\{(N-1)\,, 0\}$.

The second error term $e^{-mP(x,0)}\, \varepsilon_{N,2}(m)$, can be bounded by the following method: We introduce a number σ_N such that the function $v^N f_N(v)$ is majorized by

$$\left|v^N f_N(v)\right| \leq |a_N|\, v^N e^{\sigma_N\, v}. \tag{4.2.2}$$

Then

$$\left|\int_0^Z e^{-mv}\, v^N\, f_N(v)\, dv\right| \leq \left|\int_0^Z |a_N|\, e^{-(m-\sigma_N)v}\, v^N dv\right|$$

$$\leq \Gamma(N+1) \frac{|a_N|}{(m-\sigma_N)^{N+1}} \qquad (m > \sigma_N) \tag{4.2.3}$$

The best value of σ_N is given by

$$\sigma_N = \sup_{(0,\infty)} \{\psi_N(v)\}, \tag{4.2.4}$$

where $\psi_N(v) = \frac{1}{v} \ln \left|\frac{v^N\, f_N(v)}{a_N\, v^N}\right|$.

The bounded (4.2.3) has the property of being asymptotic to the absolute value of the actual error when $m \longrightarrow \infty$. But the preceding 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 $a_N \neq 0$, then the failure occurs when $\psi_N(v)$ tends to $+\infty$ as $v \longrightarrow 0^+$. But for small v, we have from (4.1.11)

$$v^N\, f_N(v) = a_N\, v^{N+1} + a_{N+1}\, v^{N+2} + ...$$

Therfore

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

For the tail, the inequality (4.1.16) can be used, the integral on the right-hand side being found numerically for a suitably chosen value of M.

5 Risk and Nearest Neighbor Distances

In this chapter we study the rates of convergence of nearest neighbor classification in terms of metric covering numbers of the underlying space, present an upper bound on the expected nearest neighbor distance for distributions having support on a totally bounded set and give some contributions in the case of unbounded support for special distributions.

5.1 Introduction

Recall that the problem to be considered is the classification of a random variable θ taking values in $M = \{1, 2\}$ given a sample X in χ. Somewhat more generally we consider X taking values in some general separable metric space χ equipped with metric ρ which we denote as the pair (χ, ρ). That is, a random variable (X, θ) consists of an observed pattern $X \in \chi$ from which we wish to infer the unobservable class θ, such that $\theta \in \{1, 2\}$. The probability of error for a classifier δ is $P(\delta(X) \neq \theta)$.

For a given x, a classifier δ yields a conditional risk $P(\theta \neq \delta(x) | X = x)$. If the joint distribution of (X, θ) is known then the best classifier is the Bayes classifier, see Section 1. The Bayes classifier $\delta^\star$ minimizes this risk resulting in the conditional Bayes risk

$$r^\star(x) = P(\theta \neq \delta^\star(x) | X = x) \leq P(\theta \neq \delta(x) | X = x) \quad \text{for all classifier } \delta.$$

The Bayes risk is given by $R^\star = Er^\star(x) = \int r^\star(x)\, P^X(dx)$.

Define the conditional mean of θ given $X = x$ as

$$m(x) = P(\theta = 1 | X = x) = E(\theta | X = x)$$

and the conditional variance as

$$\begin{aligned} \sigma^2(x) &= P(\theta = 1 | X = x) - [P(\theta = 1 | X = x)]^2 = m(x) - (m(x))^2 \\ &= P(\theta = 1 | X = x)\, P(\theta = 0 | X = x) \end{aligned}$$

5.2 Nearest Neighbor Classification

As in section 1 we have a training sequence $Z_m = \left(\left(X^{(1)}, \theta^{(1)}\right), \left(X^{(2)}, \theta^{(2)}\right)\right.$ $, ..., \left.\left(X^{(m)}, \theta^{(m)}\right)\right)$ at our disposal, where patterns and corresponding classes are observed. Recall from chapter one that the nearest neighbor procedure assigns any input feature vector to the class given by the label θ' of the nearest reference vector.

The conditional probability of error for the nearest neighbor rule is defined as the probability of error in classification θ by θ' given X and its nearest neighbor X' and denoted by $R_m(X, X')$, that is $R_m(X, X') = P(\theta \neq \theta'|X, X')$. By averaging $P(\theta \neq \theta'|X, X')$ over X', we obtain the m-sample conditional average probability of error

$$R_m(X) = P(\theta \neq \theta'|X) = \int P(\theta' \neq \theta\,|X'\ , X)\, f_m(x'|X)dx',$$

and by averaging $P(\theta \neq \theta'|X)$ with respect to X, we obtain the unconditional probability of error

$$R_m = P(\theta \neq \theta') = \int P(\theta' \neq \theta|X)\, f(x)dx$$

$$= \int\int P(\theta' \neq \theta\,|X' = x'\ , X = x)\, f_m(x'|x)f(x)dx'dx.$$

Define the nearest distance at time m as $d_m = \rho(X, X')$.

5.2.1 Lemma

$$R_m(X, X') = \sigma^2(X) + \sigma^2(X') + (m(X) - m(X'))^2\,.$$

Proof:
We have

$$R_m(X, X') = P(\theta \neq \theta'|X, X')$$

$$= P(\theta = 1, \theta' = 0|X, X') + (\theta = 0, \theta' = 1|X, X')$$

$$= P(\theta = 1|X)\, P(\theta' = 0|X') + P(\theta = 0|X)\, P(\theta' = 1|X')$$

$= m(X)((1-m(X'))+m(X')((1-m(X)) = m(X)+m(X')-2m(X)m(X')$

$= m(X)+m(X')-2m(X)m(X')+(m(X))^2-(m(X))^2+(m(X'))^2-(m(X'))^2$

$= [m(X)-(m(X))^2]+[m(X')-(m(X'))^2]+[(m(X))^2-m(X)m(X')]$

$\quad +[(m(X'))^2-m(X)m(X')]$

$= [m(X)-(m(X))^2]+[m(X')-(m(X'))^2]+m(X)\,[m(X)-m(X')]$

$\quad -m(X')\,[m(X)-m(X')]$

$= \sigma^2(X)+\sigma^2(X')+(m(X)-m(X'))^2.$

The following Corollary provides an upper bound on $R_m(X, X')$ in terms of d_m.

5.2.2 Corollary

If, for some $K_1 > 0$ and $\alpha > 0$ we have $|m(x) - m(x')| \leq K_1 \rho(x, x')^\alpha$ for all $x, x' \in \chi$, then, for some suitable $K > 0$ independent of m,

$$R_m(X, X') \leq 2\sigma^2(X) + K\left(d_m^\alpha + d_m^{2\alpha}\right).$$

Proof:
From Lemma 5.2.1

$$R_m(X, X') = \sigma^2(X)+\sigma^2(X')+(m(X)-m(X'))^2$$

$$= 2\sigma^2(X)+[\sigma^2(X')-\sigma^2(X)]+(m(X)-m(X'))^2.$$

Note that

$$|\sigma^2(X')-\sigma^2(X)| = |m(X')((1-m(X'))-m(X)((1-m(X))|$$

$$\leq |m(X')-m(X)|.$$

Then

$$
\begin{aligned}
R_m(X,X') &\leq 2\sigma^2(X) + |m(X') - m(X)| + (m(X) - m(X'))^2 \\
&\leq 2\sigma^2(X) + K_1\rho(x,x')^\alpha + K_1^2\rho(x,x')^{2\alpha} \\
&\leq 2\sigma^2(X) + K\,(d_m^\alpha + d_m^{2\alpha}),
\end{aligned}
$$

where $K = \max\{K_1, K_1^2\}$.

5.3 Covering Numbers and Supports

Define the open ball of radius ϵ about a point $x \in \chi$ as

$$B(x,\epsilon) = \{y \in \chi\,|\rho(x,y) < \epsilon\}.$$

5.3.1 Definition

Let A be a subset of the metric space (χ, ρ). The metric covering number $\mathcal{N}(\epsilon)$ is defined as the smallest number of open balls of radius ϵ that cover the set A. That is

$$\mathcal{N}(\epsilon) = inf\left\{k : \exists x_1, ..., x_k \in \chi\, s.t.\, A \subset \bigcup_{i=1}^{k} B(x_i,\epsilon)\right\}.$$

The logarithm of the metric covering number is often referred to as the metric entropy or $\epsilon-$*entropy*. A set A is said to be totally bounded if $\mathcal{N}(\epsilon, A, \rho) < \infty$ for all $\epsilon > 0$.

5.3.2 Definition

The metric covering radius $\mathcal{N}^{-1}(k)$ is defined as the smallest radius such that there exist k balls of this radius which cover the set A, that is

$$\mathcal{N}^{-1}(k) = inf\left\{\epsilon : \exists x_1, ..., x_k \in \chi\, s.t.\, A \subset \bigcup_{i=1}^{k} B(x_i,\epsilon)\right\}.$$

Note that $\mathcal{N}^{-1}(.)$ is a nonincreasing discrete function of k. In particular, $\mathcal{N}^{-1}(1)$ is the radius of the smallest ball to cover A and is referred to as the radius of A.

5.3.3 Example:

For any bounded set A in Euclidean $r-$space, the covering number of A satisfies $\mathcal{N}(\epsilon, A) \leq (\beta/\epsilon)^r$ for all $\epsilon \leq \beta = \mathcal{N}^{-1}(1, A)$ (the radius of the set) and the covering radius satisfies $\mathcal{N}^{-1}(m, A) \leq \beta/m^{1/r}$. In addition, if A contains an interior point in R^r then $\mathcal{N}(\epsilon, A) \geq (\beta_1/\epsilon)^r$ for some $\beta_1 > 0$, and $\mathcal{N}^{-1}(m, A) \geq \beta_1/m^{1/r}$, see [15], [16].

5.3.4 Lemma

Let A be a totally bounded subset of (χ, ρ), then

$$\lim_{m\to\infty} \mathcal{N}^{-1}(m, A, \rho) = 0$$

Proof:
Assume the statement is false. Then since $\mathcal{N}^{-1}(m)$ is nonincreasing, there exists $\varepsilon > 0$ such that $\mathcal{N}^{-1}(m) \geq \epsilon$ for all m. But this implies that $\mathcal{N}(\epsilon) \geq m$ for all m, i.e., $\mathcal{N}(\epsilon) = \infty$, which contradicts the fact A is totally bounded.

We next define the standard notion of the support of a measure

5.3.5 Definition

The support of a probability measure μ defined on (χ, ρ) is defined as

$$\kappa(\mu) = \{x \in \chi : \mu(B(x, \epsilon)) > 0 \;\; \forall \epsilon > 0\} .$$

For a probability measure μ on a separable metric space χ it is well known that $\mu(support(\mu)) = \mu(\kappa(\mu)) = 1$.

5.4 A Bound for the Risk

In this section we find an upper bound on the finite sample performance in terms of the nearest neighbor distance.

5.4.1 Lemma

Under the assumptions of 5.2.2 with $\alpha \leq 1$

$$R_m \leq R_\infty + K\left[(Ed_m)^\alpha + \left(Ed_m^2\right)^\alpha\right]$$

Proof:
From Corollary 5.2.2

$$R_m(X, X') \le 2\sigma^2(X) + K\,(d_m^\alpha + d_m^{2\alpha})$$

By taking expected values on this conclusion, we obtain with $R_\infty = 2E[\sigma^2(X)] = 2R^\star$

$$R_m \le R_\infty + K\,[E\,(d_m^\alpha) + E\,(d_m^{2\alpha})].$$

Using Jensen's inequality since $h(t) = t^\alpha$ is concave for $0 < \alpha \le 1$

$$R_m \le R_\infty + K\left[(Ed_m)^\alpha + (Ed_m^2)^\alpha\right].$$

5.5 The case of bounded support

We consider the case of totally bounded support of μ. The following theorem, compare [16], provides a bound on Ed_m and Ed_m^2.

5.5.1 Theorem

Let $X, X_1, X_2, ...$ be i.i.d. according to a probability measure μ with $\kappa(\mu)$ a totally bounded subset of (χ, ρ). Then

$$Ed_m \le \tfrac{3}{m}\textstyle\sum_{i=1}^m \mathcal{N}^{-1}(i, \kappa(\mu))$$

and

$$Ed_m^2 \le \tfrac{8}{m}\textstyle\sum_{i=1}^m \left[\mathcal{N}^{-1}(i, \kappa(\mu))\right]^2.$$

Proof:
Note that $P[d_m > \epsilon | X] = (1 - \mu(B(X, \epsilon)))^m$ and $P[X \in \kappa(\mu)] = 1$.

Fix $\epsilon > 0$. Now take an $\epsilon/2$-covering of $\kappa(\mu)$, $B_1, B_2, ...B_{\mathcal{N}(\epsilon/2)}$. Then for $X \in \kappa(\mu)$, there exists an i such that $B_i \subset B(X, \epsilon)$. Let $N \equiv \mathcal{N}(\epsilon/2)$. Now define an $\epsilon/2$-partition as follows. For each $i = 1, ..., N$, let $P_i = B_i - \bigcup_{k=1}^{i-1} B_k$. Then

$$P_i \subset B_i$$

$$\textstyle\bigcup_{i=1}^N B_i = \bigcup_{i=1}^N P_i$$

and $P_i \bigcap P_j = 0$. Also $\sum_{i=1}^{N} \mu(P_i) = 1$.

Then for $X \in \kappa(\mu)$ there exists an i such that $P_i \subset B_i \subset B(X, \epsilon)$ and in turn $p_i \equiv \mu(P_i) \leq \mu(B(X, \epsilon))$. Hence

$$P[d_m > \epsilon | X \in P_i] \leq (1 - p_i)^m$$

and

$$P[d_m > \epsilon] = \sum_{i=1}^{N} P[d_m > \epsilon | X \in P_i] \mu(P_i) \leq \sum_{i=1}^{N} p_i (1 - p_i)^m.$$

As $d_m \geq 0$, then $Ed_m = \int_0^\infty P[d_m > \epsilon] d\epsilon$.

We now prove that

$$\sum_{i=1}^{N} p_i (1 - p_i)^m \leq \begin{cases} 1 & m \leq N \\ \frac{N}{2m} & m > N \end{cases} .$$

The case $m \leq N$ follows from

$$\sum_{i=1}^{N} p_i (1 - p_i)^m \leq \sum_{i=1}^{N} max_{p_i}\, p_i (1 - p_i)^m = \frac{N}{N}(1 - \frac{1}{N})^m \leq 1,$$

the case $m > N$ from

$$\sum_{i=1}^{N} p_i (1 - p_i)^m \leq \sum_{i=1}^{N} max_{p_i}\, p_i (1 - p_i)^m = \sum_{i=1}^{N} \frac{1}{m}(1 - \frac{1}{m})^m \leq \frac{N}{2m}.$$

Hence we have that

$$P[d_m > \epsilon] \leq \sum_{i=1}^{\mathcal{N}(\epsilon/2)} p_i (1 - p_i)^m \leq \begin{cases} 1 & m \leq \mathcal{N}(\epsilon/2) \\ \frac{\mathcal{N}(\epsilon/2)}{2m} & m > \mathcal{N}(\epsilon/2) \end{cases}$$

That is $P[d_m > \epsilon] \leq \begin{cases} 1 & \epsilon \leq 2\mathcal{N}^{-1}(m) \\ \frac{\mathcal{N}(\epsilon/2)}{2m} & \epsilon > 2\mathcal{N}^{-1}(m) \end{cases}$

Since $P[d_m > \epsilon] = 0$ for $\epsilon > 2\mathcal{N}^{-1}(1)$, we have

$$Ed_m = \int_0^\infty P[d_m > \epsilon] d\epsilon \leq \int_0^{2\mathcal{N}^{-1}(m)} d\epsilon$$

$$+ \int_{2\mathcal{N}^{-1}(m)}^{2\mathcal{N}^{-1}(1)} \frac{\mathcal{N}(\epsilon/2)}{2m} d\epsilon = 2\mathcal{N}^{-1}(m) + \frac{1}{m} \int_{\mathcal{N}^{-1}(m)}^{\mathcal{N}^{-1}(1)} \mathcal{N}(\epsilon) d\epsilon.$$

Since $\mathcal{N}(\epsilon) = i$ for $\mathcal{N}^{-1}(i) \leq \epsilon < \mathcal{N}^{-1}(i-1)$ we get

$\int_{\mathcal{N}^{-1}(m)}^{\mathcal{N}^{-1}(1)} \mathcal{N}(\epsilon)d\epsilon = \sum_{i=2}^{m} \int_{\mathcal{N}^{-1}(i)}^{\mathcal{N}^{-1}(i-1)} \mathcal{N}(\epsilon)d\epsilon$

$= \sum_{i=2}^{m} i\,[\mathcal{N}^{-1}(i-1) - \mathcal{N}^{-1}(i)]$

$= 2\mathcal{N}^{-1}(1) + \mathcal{N}^{-1}(2) + ... + \mathcal{N}^{-1}(m-1) - m\mathcal{N}^{-1}(m)$

$= 2\mathcal{N}^{-1}(1) - m\mathcal{N}^{-1}(m) + \sum_{i=2}^{m-1} \mathcal{N}^{-1}(i) = \mathcal{N}^{-1}(1) - m\mathcal{N}^{-1}(m) + \sum_{i=1}^{m-1} \mathcal{N}^{-1}(i)$

Hence

$$Ed_m \leq \mathcal{N}^{-1}(m) + \tfrac{\mathcal{N}^{-1}(1)}{m} + \tfrac{1}{m} \sum_{i=1}^{m-1} \mathcal{N}^{-1}(i) \leq \tfrac{3}{m} \sum_{i=1}^{m} \mathcal{N}^{-1}(i). \qquad (5.5.1)$$

Similarly

$$Ed_m^2 = \int_0^\infty P[d_m^2 > \epsilon]d\epsilon \leq \int_0^\infty P[d_m > \sqrt{\epsilon}]d\epsilon$$

$$= 4(\mathcal{N}^{-1}(m))^2 + \tfrac{4}{m} \int_{(\mathcal{N}^{-1}(m))^2}^{(\mathcal{N}^{-1}(1))^2} \mathcal{N}(\sqrt{\epsilon})d\epsilon$$

and since

$$\int_{(\mathcal{N}^{-1}(m))^2}^{(\mathcal{N}^{-1}(1))^2} \mathcal{N}(\sqrt{\epsilon})d\epsilon = [\mathcal{N}^{-1}(1)]^2 - m[\mathcal{N}^{-1}(m)]^2 + \sum_{i=1}^{m-1} [\mathcal{N}^{-1}(i)]^2$$

Then

$$Ed_m^2 \leq \tfrac{4[\mathcal{N}^{-1}(1)]^2}{m} + \tfrac{4}{m} \sum_{i=1}^{m-1} [\mathcal{N}^{-1}(i)]^2. \qquad (5.5.2)$$

As an example, take $\kappa(\mu)$ a bounded subset of R^r for some integer $r > 1$. Then

$$E[d_m] \leq \tfrac{3\beta r}{r-1} m^{-1/r},$$

where β is the radius of $\kappa(\mu)$.

5.6 Specific bounds for nearest neighbor distances

5.2 shows that the risk of nearest neighbor procedures may be bounded using the distance to the nearest neighbor. For bounded support, covering numbers may be used to estimate the expected distance. It is not clear how to use covering numbers for unbounded support. Here we give some contributions to these questions for special distributions. We look at real-valued observations and remark on the multidimensional case.

5.6.1 Deriving a lower bound

We know, letting S denote the support of our random variables

$$Ed_m = \int_0^\infty P\left(d_m > \epsilon\right) d\epsilon$$

$$= \int_S \int_0^\infty P\left(d_m > \epsilon | X = x\right) d\epsilon\, P^X(dx)$$

$$= \int_S \int_0^\infty P\left(|X - x| > \epsilon\right)^m d\epsilon\, P^X(dx)$$

$$= \int_0^\infty \int_S P\left(|X - x| > \epsilon\right)^m P^X(dx)\, d\epsilon$$

$$\geq \int_0^\infty \left(\int_S P\left(|X - x| > \epsilon\right) P^X(dx)\right)^m d\epsilon,$$

using Jensen's inequality.

This shows

$$Ed_m \geq \int_0^\infty P\left(\left|X - \widetilde{X}\right| > \epsilon\right)^m d\epsilon = 2\int_0^\infty P\left(Z > \epsilon\right)^m d\epsilon, \text{ say.}$$

Here $\widetilde{X}$ is an independent copy of X, $Z = \left|X - \widetilde{X}\right|$ and $X - \widetilde{X}$ has a symmetric distribution on $S - S = \{x - y : x, y \in S\}$.

Let ψ denote the density of Z and assume smoothness to apply partial integration. Then in the case of unbounded support and positive density

$$Ed_m = -2\int_0^\infty \frac{1}{(m+1)\psi(\epsilon)} \frac{d}{d\epsilon}\left[P\left(Z > \epsilon\right)^{m+1}\right] d\epsilon$$

$$= \frac{-2}{(m+1)}\left[\frac{1}{\psi(\epsilon)} P\left(Z > \epsilon\right)^{m+1}\right]_0^\infty - \frac{2}{(m+1)}\int_0^\infty \frac{\psi'(\epsilon)}{(\psi(\epsilon))^2} P\left(Z > \epsilon\right)^{m+1} d\epsilon.$$

If we have $\frac{1}{\psi(\epsilon)} P\left(Z > \epsilon\right)^{m+1} \longrightarrow 0 \qquad$ for $\epsilon \longrightarrow \infty$ and

$$\int_0^\infty \frac{\psi'(\epsilon)}{(\psi(\epsilon))^2} P\left(Z > \epsilon\right)^{m+1} d\epsilon \longrightarrow 0 \qquad \text{for } m \longrightarrow \infty,$$

as is e.g. the case for exponential and normal distributions, then

$$Ed_m \geq \frac{2}{(m+1)} \frac{1}{\psi(0)}$$

An analogous lower bound also holds for the multidimensional case, using

$Ed_m = \int_0^\infty P\left(Z > \epsilon\right)^{m+1} d\epsilon$ with Z denoting Euclidean norm of $X - \widetilde{X}$.

5.6.2 A lower bound for the exponential distribution

Applying a different method, we look at d_m in the case of an exponential distribution with density e^{-x}, $x > 0$. Then

$$\begin{aligned} Ed_m &= \int_0^\infty \int_0^\infty P\left(|X - x| > \epsilon\right)^m d\epsilon\, e^{-x} dx \\ &\geq \int_0^\infty \int_0^\infty P\left(X < x - \epsilon\right)^m d\epsilon\, e^{-x} dx \\ &= \int_0^\infty \int_0^x P\left(X < z\right)^m dz\, e^{-x} dx \\ &= \int_0^\infty \int_0^x \left(1 - e^{-z}\right)^m dz\, e^{-x} dx \\ &= \int_0^\infty \int_z^\infty e^{-x} dx \left(1 - e^{-z}\right)^m dz \\ &= \int_0^\infty e^{-z} \left(1 - e^{-z}\right)^m dz. \end{aligned}$$

To evaluate this integral note that

$$e^{-z}\left(1 - e^{-z}\right)^m = \sum_{k=0}^m (-1)^k \begin{pmatrix} m \\ k \end{pmatrix} e^{-z(k+1)} \ ,$$

hence

$$\int_0^\infty e^{-z}\left(1 - e^{-z}\right)^m dz = \sum_{k=0}^m (-1)^k \begin{pmatrix} m \\ k \end{pmatrix} \frac{1}{k+1} \ .$$

This sum may be computed as

$$\frac{1}{m+1}\sum_{k=0}^{m}(-1)^k \begin{pmatrix} m+1 \\ k+1 \end{pmatrix}$$

$$= -\frac{1}{m+1}\left[\sum_{k=0}^{m+1}(-1)^k \begin{pmatrix} m+1 \\ k \end{pmatrix} - \begin{pmatrix} m+1 \\ 0 \end{pmatrix}\right]$$

$$= -\frac{1}{m+1}\left[(1-1)^{m+1}-1\right] = \frac{1}{m+1}.$$

We obtain the lower bound $Ed_m \geq \frac{1}{m+1}$.

We may also use the integral at hand using Stieltjes integration:

$$\int_0^\infty e^{-z}\left(1-e^{-z}\right)^m dz = \int_0^\infty \left(1-e^{-z}\right)^m d\left(1-e^{-z}\right)$$

$$= \int_0^1 y^m dy = \frac{1}{m+1}.$$

5.6.3 A lower bound for the normal distribution

We look at d_m in the case of a normal distribution with density $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, distribution function $\Phi(x)$ and let $H(x) = 1-\Phi(x)$, using the method of 5.6.2.

Then as in 5.6.2

$$Ed_m = \int_{-\infty}^\infty \int_0^\infty P\left(|X-x| > \epsilon\right)^m d\epsilon\, \phi(x)dx$$

$$\geq \int_{-\infty}^\infty \int_0^\infty P\left(X < x-\epsilon\right)^m d\epsilon\, \phi(x)dx$$

$$= \int_{-\infty}^\infty \int_{-\infty}^x P\left(X < z\right)^m dz\, \phi(x)dx$$

$$= \int_{-\infty}^\infty \int_z^\infty \phi(x)dx P\left(X < z\right)^m dz$$

$$= \int_{-\infty}^\infty \left(1-\Phi(z)\right)\left(\Phi(z)\right)^m dz$$

$$\geq \int_0^\infty H(z)\left(1-H(z)\right)^m dz.$$

For a first estimate use Stieltjes integration to obtain

$$Ed_m \geq \int_{-\infty}^{\infty} (1-\Phi(z))\,(\Phi(z))^m\,dz$$

$$\geq \int_{-\infty}^{\infty} \phi(z)\,(1-\Phi(z))\,(\Phi(z))^m\,dz = \int_{-\infty}^{\infty} (1-\Phi(z))\,(\Phi(z))^m\,d\Phi(z)$$

$$= \int_0^1 (1-y)y^m dy = \frac{\Gamma(2)\Gamma(m+1)}{\Gamma(m+3)} = \frac{1}{(m+1)(m+2)}.$$

To proceed, note first that for any $c > 0$

$$\int_0^c H(z)\,(1-H(z))^m\,dz \longrightarrow 0 \text{ exponentially fast as } m \longrightarrow \infty.$$

So we only have to look at $\int_c^{\infty} H(z)\,(1-H(z))^m\,dz$, fixing some $c > 1$.

Now use the well known inequalities

$$\frac{1}{x}\phi(x) \geq H(x) \geq \frac{1}{x}\left(1-\frac{1}{x^2}\right)\phi(x), \qquad x > 0,$$

which imply

$$\phi(x) \geq H(x) \geq \frac{1}{x}\left(1-\frac{1}{c^2}\right)\phi(x), \qquad x \geq c > 1.$$

Hence

$$\int_c^{\infty} H(z)\,(1-H(z))^m\,dz \geq \int_c^{\infty} \left(1-\frac{1}{c^2}\right)\frac{1}{z}\phi(z)\,(1-\phi(z))^m\,dz,$$

and we look at

$$\int_c^{\infty} \frac{1}{z}\phi(z)\,(1-\phi(z))^m\,dz = \int_c^{\infty} \frac{1}{z}\gamma e^{\frac{-z^2}{2}}\left(1-\gamma e^{\frac{-z^2}{2}}\right)^m dz$$

$$= \int_a^{\infty} \frac{1}{2y}\gamma e^{-y}\,(1-\gamma e^{-y})^m\,dy, \qquad y = \frac{z^2}{2},\ a = \frac{c^2}{2},\ \gamma = \frac{1}{\sqrt{2\pi}}.$$

Thus we have to treat, as $m \longrightarrow \infty$,

$$\int_a^{\infty} \frac{1}{y}\gamma e^{-y}\,(1-\gamma e^{-y})^m\,dy.$$

If we could show that, for $a > 0$,

$$\int_a^\infty \frac{1}{y}\gamma e^{-y}\left(1-\gamma e^{-y}\right)^m dy \geq \frac{c}{m},$$

then this would imply $Ed_m \geq \frac{\widetilde{c}}{m}$ for some suitable $\widetilde{c} > 0$. But as we see now, this is not the case:

Assume that for some $c > 0$, $a > 1$ we have

$$\int_a^\infty \frac{1}{y}\gamma e^{-y}\left(1-\gamma e^{-y}\right)^m dy \geq \frac{c}{m+1} \text{ for all } m.$$

This implies by partial integration for all m

$$c \leq \int_a^\infty \frac{1}{y}\frac{d}{dy}\left(1-\gamma e^{-y}\right)^{m+1} dy$$

$$= \left[\frac{1}{y}\left(1-\gamma e^{-y}\right)^{m+1}\right]_a^\infty + \int_a^\infty \frac{1}{y^2}\left(1-\gamma e^{-y}\right)^{m+1} dy.$$

Obviously $\left[\frac{1}{y}\left(1-\gamma e^{-y}\right)^{m+1}\right]_a^\infty \longrightarrow 0$ as $m \longrightarrow \infty$ and by dominated convergence also $\int_a^\infty \frac{1}{y^2}\left(1-\gamma e^{-y}\right)^{m+1} dy \longrightarrow 0$ as $m \longrightarrow \infty$.

This gives a contradiction. Our arguments thus do not provide the bound of 5.6.1. Here it is interesting to note that the bound $P\left(|X-x| > \epsilon\right) \geq P\left(X \leq x-\epsilon\right)$ does not yield the right asymptotic bound for normal distributions.

We can came arbitrarily close to a bound of the type $\frac{c}{m}$ by the following method:

For any $\beta > 1$ choose $a\left(\beta\right) > 1$ such that $\frac{1}{y} \geq e^{-\beta y}$ for $y \geq a\left(\beta\right)$. Then

$$\int_{a(\beta)}^\infty \frac{1}{y}e^{-y}\left(1-e^{-y}\right)^m dy \geq \int_{a(\beta)}^\infty e^{-(\beta+1)y}\left(1-e^{-y}\right)^m dy,$$

and

$$\int_0^\infty e^{-(\beta+1)y}\left(1-e^{-y}\right)^m dy = \int_0^1 y^{-\beta}\left(1-y\right)^m dy = \frac{\Gamma(\beta+1)\Gamma(m+1)}{\Gamma(\beta+m+2)}.$$

5.6.4 An upper bound

We use constants, $-\infty < K_1(m) \le 0 \le K_2(m) < \infty$ depending on m, to write

$$
\begin{aligned}
Ed_m &= \textstyle\int_{-\infty}^{\infty} \int_0^{\infty} P\left(|X-x| > \epsilon\right)^m d\epsilon\, P^X(dx) \\
&= \textstyle\int_{-\infty}^{K_1(m)} \int_0^{\infty} P\left(|X-x| > \epsilon\right)^m d\epsilon\, P^X(dx) \\
&\quad + \textstyle\int_{K_2(m)}^{\infty} \int_0^{\infty} P\left(|X-x| > \epsilon\right)^m d\epsilon\, P^X(dx) \\
&\quad + \textstyle\int_{K_1(m)}^{K_2(m)} \int_0^{\infty} P\left(|X-x| > \epsilon\right)^m d\epsilon\, P^X(dx) \\
&= L_1(m) + L_2(m) + L_3(m), \text{ say.}
\end{aligned}
$$

5.6.4.1 Bounding $L_1(m)$, $L_2(m)$

We assume for the following that $|X-x|$ has a finite moment generating function

$$\psi(t,x) = Ee^{t|X-x|}, \quad x \in R,\, 0 < t < 1.$$

By Markov's inequality for any $0 < t < 1$

$$
\begin{aligned}
\textstyle\int_0^{\infty} P\left(|X-x| > \epsilon\right)^m d\epsilon &= \textstyle\int_0^{\infty} P\left(e^{t|X-x|} > e^{t\epsilon}\right)^m d\epsilon \\
&\le \textstyle\int_0^{\infty} \psi(t,x)^m e^{-mt\epsilon} d\epsilon = \frac{1}{mt}\psi(t,x)^m,
\end{aligned}
$$

hence for $t = \frac{1}{\alpha m}$, $\alpha \ge 1$

$$\textstyle\int_0^{\infty} P\left(|X-x| > \epsilon\right)^m d\epsilon \le \alpha\psi(\frac{1}{\alpha m}, x)^m.$$

It follows

$$L_1(m) \le \alpha \textstyle\int_{-\infty}^{K_1(m)} \psi(\frac{1}{m}, x)^m\, P^X(dx),$$

$$L_2(m) \le \alpha \textstyle\int_{K_2(m)}^{\infty} \psi(\frac{1}{m}, x)^m\, P^X(dx).$$

5.6.4.2 The case of the exponential distribution

Let X have density e^{-x}, $x > 0$.

Then we take $K_1(m) = 0$ and $L_1(m)$ vanishes.

We have

$$\psi(t,x) = Ee^{t|X-x|} \le Ee^{tX+tx} = e^{tx} \int_0^\infty e^{ty}e^{-y}dy = e^{tx}\tfrac{1}{1-t},$$

hence for $t = \frac{1}{2m}$

$$\psi(\tfrac{1}{2m},x)^m \le e^{\frac{x}{2}} \left(\tfrac{1}{1-\frac{1}{2m}}\right)^m = e^{\frac{x}{2}} \left(1 + \tfrac{1}{2m-1}\right)^m .$$

It follows, using 5.6.4.1

$$\begin{aligned} L_2(m) \le 2\int_{K_2(m)}^\infty e^{\frac{x}{2}} \left(1+\tfrac{1}{2m-1}\right)^m e^{-x}dx &= 2\left(1+\tfrac{1}{2m-1}\right)^m \int_{K_2(m)}^\infty e^{\frac{-x}{2}}dx \\ &= 4\left(1+\tfrac{1}{2m-1}\right)^m e^{-K_2(m)}. \end{aligned}$$

For $K_2(m) = \log m$, it follows

$$L_2(m) \le 2\left(1+\tfrac{1}{2m-1}\right)^m \tfrac{1}{m} = O\left(\tfrac{1}{m}\right),$$

since $\left(1+\frac{1}{2m-1}\right)^m \longrightarrow e^{\frac{1}{2}} \qquad (m \longrightarrow \infty)$.

5.6.4.3 The case of the normal distribution

Let X have density $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, $x \in R$.

We take $-K_1(m) = K_2(m) > 0$.

For $x \in R$, $t > 0$

$$\psi(t,x) = Ee^{t|X-x|} \le Ee^{t|X|+t|x|} = e^{t|x|}Ee^{t|X|},$$

hence for $t = \frac{1}{m}$, using Jensen's inequality

$$\psi(\tfrac{1}{m}, x)^m \le e^{|x|}\left(Ee^{\frac{1}{m}|X|}\right)^m \le e^{|x|}E\left(e^{\frac{1}{m}|X|}\right)^m = e^{|x|}Ee^{|X|}$$

$$= 2e^{|x|}\int_0^\infty \tfrac{1}{\sqrt{2\pi}}e^y e^{-\frac{y^2}{2}}dy = 2e^{|x|}e^{\frac{1}{2}}\int_0^\infty \tfrac{1}{\sqrt{2\pi}}e^{-\frac{(y-1)^2}{2}}dy \le 2e^{|x|}e^{\frac{1}{2}}.$$

It follows for $K(m) > 1$

$$L_1(m) + L_2(m) \le 4e^{\frac{1}{2}}\int_{K_2(m)}^\infty e^x\phi(x)dx$$

$$= 4e^{\frac{1}{2}+\frac{1}{2}}\int_{K_2(m)}^\infty \tfrac{1}{\sqrt{2\pi}}e^{-\frac{(x-1)^2}{2}}dx = 4\,e\,P\left(X+1 \ge K_2(m)\right)$$

$$\le \tfrac{12}{K_2(m)-1}\,e^{-\frac{(K_2(m)-1)^2}{2}}.$$

For $K_2(m) = \sqrt{2\log m} + 1$, it follows

$$L_1(m) + L_2(m) = o\left(\tfrac{1}{m}\right).$$

5.6.4.4 Bounding $L_3(m)$

We write, for X with density f,

$$L_3(m) = \int_{K_1(m)}^{K_2(m)}\int_0^\infty P\left(|X-x| > \epsilon\right)^m d\epsilon\, P^X(dx)$$

$$= \int_{K_1(m)}^{K_2(m)}\int_0^\infty e^{-mG(x,\epsilon)}f(x)d\epsilon\, dx,$$

where $G(x,\epsilon) = -\log P\left(|X-x| > \epsilon\right)$.

Assume that the following inequality holds:

There exists $c > 0$ such that for all x in the support of X and for all $\epsilon > 0$

$$G(x,\epsilon) \ge c\epsilon f(x).$$

From this inequality we obtain

$$\int_{K_1(m)}^{K_2(m)}\int_0^\infty e^{-mG(x,\epsilon)}f(x)d\epsilon\, dx \le \int_{K_1(m)}^{K_2(m)}\int_0^\infty e^{-mc\epsilon f(x)}f(x)d\epsilon\, dx$$

$$= \int_{K_1(m)}^{K_2(m)}\tfrac{1}{cm}dx = \left(K_2(m) - K_1(m)\right)\tfrac{1}{cm}.$$

So if $-K_1(m)$, $K_2(m)$ have logarithmic growth as in 5.6.4.2, 5.6.4.3, we obtain

$$L_3(m) = o\left(\frac{1}{m^\beta}\right) \qquad \text{for all } \beta < 1.$$

We have to investigate validity of the inequality

$$-\log P\left(|X-x| > \epsilon\right) = -\log\left(1 - P\left(|X-x| \leq \epsilon\right)\right) \geq c\epsilon f(x).$$

Noting

$$-\log\left(1-y\right) \geq y \quad \text{for all } 0 \leq y \leq 1,$$

we see that

$$-\log P\left(|X-x| > \epsilon\right) \geq P\left(|X-x| \leq \epsilon\right).$$

Hence a sufficient condition for

$$G(x,\epsilon) \geq c\epsilon f(x)$$

is given by

$$P\left(|X-x| \leq \epsilon\right) \geq c\epsilon f(x).$$

Note that this second condition will always be violated for unbounded support letting ϵ tend to ∞.

Furthermore we remark:

If $[x, x+\epsilon]$ is contained in the support of X and f is increasing on $[x, x+\epsilon]$ then

$$P\left(|X-x| \leq \epsilon\right) \geq \epsilon f(x).$$

This also holds if $[x-\epsilon, x]$ is contained in the support of X and f is decreasing on $[x-\epsilon, x]$.

If $[x-\epsilon, x+\epsilon]$ is contained in the support of X and f is convex on $[x-\epsilon, x+\epsilon]$ then

$$P\left(|X-x| \leq \epsilon\right) \geq 2\epsilon f(x).$$

5.6.4.5 The case of the exponential distribution

We look at the validity of the inequality in 5.6.4.4.

If $[x-\epsilon, x+\epsilon] \subset [0, \infty)$, then by convexity

$$P\left(|X-x| \leq \epsilon\right) \geq 2\epsilon f(x).$$

If $x - \epsilon \leq 0$, then

$$\begin{aligned} -\log P\left(|X-x| > \epsilon\right) &= -\log P\left(X > x+\epsilon\right) \\ &= -\log e^{-(x+\epsilon)} = x+\epsilon \geq \epsilon f(x). \end{aligned}$$

This shows validity with $c = 1$.

Using $K_1(m) = 0$ and $K_2(m) = \log m$ we obtain with 5.6.4.2

$$Ed_m \leq 2\tfrac{1}{m} + \tfrac{\log m}{m}.$$

5.6.4.6 The case of the normal distribution

Again we look at the validity of the inequality in 5.6.4.4.

Due to symmetry it is enough to treat $x > 0$. Let $\epsilon > 0$.

(a) If $x - \epsilon \geq -x$, then

$$P\left(|X-x| \leq \epsilon\right) \geq P\left(x-\epsilon \leq X \leq x\right) \geq \epsilon\phi(x).$$

(b) So assume $x - \epsilon < -x$, i.e. $\epsilon > 2x$. We can show

$$
\begin{aligned}
-\log P\left(|X-x|>\epsilon\right) &= -\log\left(P\left(X<-(\epsilon-x)\right)+P\left(X>x+\epsilon\right)\right)\\
&\geq -\log 2P\left(X<-(\epsilon-x)\right) = -\log 2-\log P\left(X>\epsilon-x\right)\\
&\geq -\log 2-\log\left(\tfrac{1}{\epsilon-x}\tfrac{1}{\sqrt{2\pi}}e^{-\frac{(\epsilon-x)^2}{2}}\right)\\
&= \log\left(\tfrac{\sqrt{2\pi}}{2}\right)+\log(\epsilon-x)+(\epsilon-x)^2.
\end{aligned}
$$

(c) Firstly, let $x \geq 1$. Then $\log(\epsilon-x)\geq 0$ and from $\epsilon>2x$

$$(\epsilon-x)^2 \geq \tfrac{\epsilon^2}{4} \geq \tfrac{1}{2}\epsilon\phi(x).$$

(d) Finally, let $x<1$. If $\epsilon \geq 2$ then

$\log(\epsilon-x)\geq 0$ and we proceed as in (c).

So it remains to consider $x<1$, $\epsilon<2$.

But then $P\left(|X-x|\leq\epsilon\right)\geq c\epsilon$,

where $c=\inf_{|y|<3}\,\phi(y)$, hence

$$P\left(|X-x|\leq\epsilon\right)\geq c\epsilon\phi(x)$$

(e) Conclusion.

Retracting (a)-(d) we find a constant $c^\star>0$ such that

$$-\log P\left(|X-x|>\epsilon\right)\geq c^\star\epsilon\phi(x), \quad \text{for all } x,\epsilon>0.$$

Using $-K_1(m)=K_2(m)=\sqrt{2\log m}+1$ we obtain with 5.6.4.3.

$$Ed_m \leq o\left(\tfrac{1}{m}\right)+c^\star\tfrac{\sqrt{2\log m}+1}{m}.$$

References

[1] Billingsley, P. (1986). Probability and Measure, 2nd ed., Wiley & Sons, New York.

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

[3] Cover, T. M. (1968b). Rates of convergence for nearest neighbor procedures. In Proceedings of the Hawaii International Conference on Systems Sciences, pages 413-415. Honolulu, HI.

[4] Dasarathy, B. (1991). Nearest Neighbor Classification Techniques. IEEE, Los Alamitos, CA.

[5] De Bruijn, N. G. (1958). Asymptotic Methods in Analysis. Wiley, New York.

[6] Devroye, L., Györfi, L., and Lugosi, G. (1996). Probabilistic Theory of Pattern Recognition. Springer-Verlag, New York.

[7] Duda, R. O. and Hart, P. E. (1973). Pattern Classification and Scene Analysis. Wiley & Sons, New York.

[8] Durrett, R. (1991). Probability: Theory and Examples. Wadsworth and Brooks/Cole, Pacific Grove, CA.

[9] Fix, E., Hodges, J. (1991). Discriminatory analysis, nonparametric discrimination, consistency properties. In Nearest Neighbor Classification Techniques, Dasarathy, B., editor, pages 32-39, IEEE Computer Society Press, Los Alamitos, CA.

[10] Fix, E., Hodges, J. (1991). Discriminatory analysis, small sample performance. In Nearest Neighbor Classification Techniques, Dasarathy, B., editor, pages 40-56, IEEE Computer Society Press, Los Alamitos, CA.

[11] Fukunaga, K. and Hummels, D. M. (1987). Bias of nearest neighbor estimates. IEEE Trans. Pattern Anal. Mach. Intell., 9: 103-112.

[12] Györfi, L. (1978). On the rate of convergence of nearest neighbor rules. IEEE Transactions on Information Theory, 5: 449-457.

[13] Györfi, L. (1981). The rate of convergence of $k_n - NN$ regression estimates and classification rules. IEEE Transactions on Information Theory, 27: 362-364.

[14] Irle, A. (2001). Wahrscheinlichkeitstheorie und Statistik. Teubner, Stuttgart.

[15] Kolmogorov, A. N. and Tihomirov, V. M. (1961). $\epsilon-$entropy and $\epsilon-$capacity of sets in functional spaces. Amer. Math. Soc. Translations. 17: 277-364.

[16] Kulkarni, S. R. and Posner, S. E. (1995). Rates of convergence of nearest neighbor estimation under arbitrary sampling. IEEE Transactions on Information Theory, 41: 1028-1039.

[17] Olver, F. W. J. (1974). Asymptotics and Special Functions. Academic Press, New York.

[18] Papoulis, A. (1991). Probability Random Variables and Stochastic Processes, 3nd ed., McGraw-Hill.

[19] Psaltis, D., Snapp, R, R. and Venkatesh, S. S. (1994). On the finite sample performance of the nearest neighbor classifier. IEEE Transactions on Information Theory, 40: 820-837.

[20] Snapp, R. R. and Venkatesh, S. S. (1998). Asymptotic expansion of the k nearest neighbor risk. Ann. Statist. 26: 850-878.

[21] Stone, C. (1977). Consistent nonparametric regression. Annals of Statistics, 5: 595-645.