Title	極値統計量と順序統計量の理論的研究
Author(s)	髙橋, 倫也
Citation	大阪大学, 1986, 博士論文
Version Type	VoR
URL	https://hdl.handle.net/11094/157
rights	
Note	

CONTRIBUTION TO THE THEORY

OF

EXTREME STATISTIC AND ORDER STATISTICS

AUGUST 1986

BY

RINYA TAKAHASHI

SUMMARY

This thesis is concerned with the theory of extreme statistic and order statistics, including applications of the latter in some fields of statistical inference.

In Chapter 1, an asymptotic theory of univariate extreme statistic was treated. Properties of tail equivalence and extreme value distributions were investigated. Simple sufficient conditions for the domain of attraction of the double exponential distribution and analytic expressions of the normalizing constants were given. A number of examples to illustrate the results were shown.

In Chapter 2, an asymptotic theory of multivariate extreme statistic was treated. Basic properties of multivariate extreme value distributions were proved. Characterizations of multivariate extreme value distribution by its marginal independence were shown. Also, multivariate extensions of the results on univariate tail equivalence and extreme value distributions were established.

In Chapter 3, an application of the distribution theory of order statistics was considered. It was proved that the family of Weibull distributions with varying shape parameter is outlier-prone completely. Properties of tails of outlier-prone and outlier-resistant distributions were established. Relations between the properties of outliers and the types of extreme value distributions were shown.

Finally in Chapter 4, an application of the order statistics to statistical inference was dealt with. In the time truncated life testing problem, a sequential procedure was proposed and its properties were investigated. In the non-regular estimation problem, the second order efficiency of the maximum probability estimators was introduced and discussed.

CONTENTS

0 INTRODUCTION 1

1 ASYMPTOTIC THEORY OF EXTREME STATISTIC - UNIVARIATE CASE

1.1 Introduction 7

1.2 The domain of attraction of Ψ_α 12

1.3 The domain of attraction of Λ 15

1.4 Examples 24

2 ASYMPTOTIC THEORY OF EXTREME STATISTIC - MULTIVARIATE CASE

2.1 Introduction 28

2.2 Multivariate extreme value distributions 32

2.3 Multivariate tail equivalence 38

2.4 Examples 43

Appendix I, II 46

3 OUTLIER-PRONE AND OUTLIER-RESISTANT DISTRIBUTIONS

3.1 Introduction 47

3.2 Outlier properties of the family of Weibull distributions 51

3.3 Properties of outlier-prone and outlier-resistant distributions 53

3.4 Relations betweem the outlier properties and the types of extreme value distributions 57

4 ORDER STATISTICS IN SEQUENTIAL LIFE TESTING AND NON-REGULAR ESTIMATION

4.1 Introduction 59

4.2 A sequential test in time truncated life testing, I 62

4.3 A sequential test in time truncated life testing, II 65

4.4 Non-regular estimation problem 68

Appendix 71

REFERENCES 74

CHAPTER 0

INTRODUCTION

Extreme statistic and order statistics have been studied by many authors, because they play an important role in statistical theory. In this thesis, properties of univariate and multivariate extreme statistic and applications of order statistics were studied.

In Chapter 1 we dealt with an univariate extreme statistic. Let X_1, X_2, $\cdots$ be a sequence of independent random variables with common distribution function F and then define the extreme statistic $Z_n = \max(X_1, X_2, \cdots, X_n)$. If there exist sequences of constants $a_n > 0$ and b_n for which $(Z_n - b_n)/a_n$ has a nondegenerate limit distribution H, then F is said to be in the domain of attraction of H and H is said to be an extreme value distribution.

Fisher and Tippett [5] have established that a distribution function with non empty domain of attraction is of one of the following three types:

$$\Phi_\alpha(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ \exp(-x^{-\alpha}) & \text{if } x > 0, \end{cases}$$

$$\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^{\alpha}) & \text{if } x < 0, \\ 1 & \text{if } x \geq 0, \end{cases}$$

$$\Lambda(x) = \exp(-e^{-x}), \qquad -\infty < x < \infty,$$

where α is a positive constant. Sufficient conditions for the domain of attraction of the limit distributions were given by Mises [15]. Gnedenko [8] has established necessary and sufficient conditions for the domain of attraction of Φ_α, Ψ_α and Λ. Gnedenko points out, however, that he has not found satisfactory results in the case of Λ. Haan [12] has established other necessary and sufficient conditions for the domain of attraction of Λ. Gnedenko's results were generalized by Smirnov [22]. Resnick [19] has established the results on tail equivalence and asymptotic distribution of extreme statistic.

In Section 1.2 we considered the normalizing constants of distribution functions which belong to the domain of attraction of Ψ_α, and extended the results of Resnick [19]. This section is based on Takahashi [23]. In Section 1.3 we considered the domain of attraction of Λ. We showed some properties of normalizing constants. Simple sufficient conditions that a distribution function belongs to the domain of attraction of Λ, and expressions of the normalizing constants were given. In Section 1.4, a number of examples to illustrate the results of Section 1.3 were shown. Sections 1.3 and 1.4 are based on Takahashi [24, 28].

In Chapter 2 we dealt with a multivariate extreme statistic. Let $\mathbf{X}^{(1)}$, $\mathbf{X}^{(2)}$, $\cdots$ be a sequence of independent k-dimensional random vectors with common distribution function F and let

$$Z_i^{(n)} = \max_{1 \leq j \leq n} X_i^{(j)}, \quad i = 1, \cdots, k.$$

If there exist vectors $\mathbf{a}^{(n)}$ and $\mathbf{b}^{(n)}$ in $\mathbf{R}^k$ for which $(\mathbf{Z}^{(n)} - \mathbf{b}^{(n)})/\ \mathbf{a}^{(n)}$ has a nondegenerate limit distribution H, then F is said to be in the domain of attraction of H and H is said to be a multivariate extreme value distribution. Results for the multivariate extreme value distributions, obtained by Sibuya [20], Nair [16], Galambos [6] and others, have been summarized by Galambos [7]. Recently, Marshall and Olkin [14] have established a multivariate analogs of Gnedenko's necessary and sufficient conditions for the domain of attraction.

In Section 2.2 we proved the basic properties of multivariate extreme value distributions. Next we characterized the multivariate extreme value distribution by its marginal independence. The content in this section is due to Takahashi [27, 30]. In Section 2.3 we established multivariate extensions of the results of Resnick [19] on tail equivalence and asymptotic distributions of extremes. This section is based on Takahashi [27]. Some examples were shown in Section 2.4.

In Chapter 3 we dealt with an application of order statistics. Using the order statistics, Neyman and Scott [17] considered the ideas of outlier-proneness and outlier-resistance of a family of distributions and showed that the family of gamma distributions and the family of lognormal distributions are outlier-prone completely. On the other hand, Green [10] considered the ideas of outlier-proneness and outlier-resistance of individual distributions and gave theorems concerning necessary and sufficient conditions for outlier-proneness and outlier-resistance of distributions.

In Section 3.2, using the same device of Neyman and Scott [17], we showed the family of Weibull distributions with varying shape parameter is outlier-prone completly. In Sections 3.3 and 3.4, we considered the outlier-prone and outlier-resistant distributions introduced by Green [10] proved some properties of tails of outlier-prone and outlier-resistant distributions. Next we showed the relations between the properties of outlier and the types of extreme value distributions. The contents of this chapter is due to Takahashi [29].

Finally in Chapter 4 we dealt with the problems of statistical inference, using the order statistics. Padgett and Wai [18] considered the following life test. Let n independent items with life distribution F_θ be put on life test at the outset, and suppose items are not replaced upon failure. It is assumed that F_θ is stachastically increasing in θ. They considered the hypotheses $H_0: \theta \geq \theta_0$ versus $H_1: \theta \leq \theta_1$ $(\theta_1 < \theta_0)$ and proposed a one-sided sequential test based on $X_n(t)$, the number of failed items before or at time $t \leq t_0$, where t_0 is a specified truncation time for the test. They derived the average sampling time, and developed an interval estimation procedure for θ after acceptance of H_0.

In Section 4.2 we improved and extended the results on the sequential test proposed by Padgett and Wei [18], and derived the average sampling time (see Takahashi [25]). In Section 4.3, another sequential test procedure based on $X_n(t)$ was proposed and analyzed its properties. This procedure allows a quick acceptance of H_0 when H_0 is true (see Takahashi [26]).

Weiss and Wolfowitz [31] developed a theory of maximum probability estimators and applied it to some non-regular estimation problems in which a function of extreme statistic is the muximum probability estimator.

In Section 4.4 we introduced the notion of the second order efficiency which we apply to some non-regular estimation problems (see Takahashi [23]).

In this thesis we have used the following results of analysis and probability theory. Proofs and further details can be found in Haan [12].

DEFINITION 0.1. *Let* $\mathbf{R}^+$ *be the set of positive real numbers. A function* $R : \mathbf{R}^+ \to \mathbf{R}^+$ *varies regularly at infinity if there exists a* $\rho \in \mathbf{R}$ *such that for all* $x \in \mathbf{R}^+$

$$\lim_{t\to\infty} R(tx) / R(t) = x^\rho .$$

This number ρ *is called the exponent of regular variation. In the particular case* $\rho = 0$, *R is often called slowly varying at infinity. For brevity we also say that R is* ρ*-varying at infinity.*

THEOREM 0.1. (a) *If a function* $R : \mathbf{R}^+ \to \mathbf{R}^+$ *is Lebesgue-integrable on finite intervals and regularly varying with exponent* ρ, *then there exist functions* $a : \mathbf{R}^+ \to \mathbf{R}$ *and* $c : \mathbf{R}^+ \to \mathbf{R}^+$ *with*

$$\lim_{x\to\infty} c(x) = c_0 \quad (0 < c_0 < \infty), \quad \lim_{x\to\infty} a(x) = \rho,$$

such that $$R(x) = c(x)\exp\{ \int_1^x \frac{a(t)}{t}\, dt\} \text{ for all positive } x.$$

(b) *If R is ρ-varying at infinity ($-\infty < \rho < \infty$), then for all sequences $\{a_n\}$ and $\{a'_n\}$ of positive numbers with*

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} a'_n = \infty \quad \text{and}$$

$$\lim_{n\to\infty} a_n / a'_n = c \qquad (0 < c < \infty),$$

we have

$$\lim_{n\to\infty} R(a_n) / R(a'_n) = c^{\rho}.$$

(c) *Suppose R_1 and R_2 are non-increasing and ρ-varying at infinity ($-\infty < \rho < 0$). For $0 \leq c \leq \infty$ we have*

$$\lim_{x\to\infty} R_1(x) / R_2(x) = c$$

if and only if

$$\lim_{x\to\infty} R_1^*(x) / R_2^*(x) = c^{-1/\rho}.$$

where $R_i^*(x) = \inf\{y : R_i(y) \leq 1/x\}$, $i = 1, 2$.

LEMMA 0.1. *Suppose that for positive functions f and g on $\mathbf{R}$ both*

$$\int_0^\infty f(t)dt \qquad \text{and} \qquad \int_0^\infty g(t)dt$$

are finite and

$$\lim_{t\to\infty} f(t) / g(t) = c \quad \text{with} \quad 0 \leq c \leq \infty .$$

Then

$$\lim_{t\to\infty} \frac{\int_t^\infty f(s)ds}{\int_t^\infty g(s)ds} = c.$$

THEOREM 0.2. *Suppose that the distribution function H is continuous on the whole real line and strictly increasing on $\{x : 0 < H(x) < 1\}$. Suppose for a sequence $\{F_n\}$ of distribution functions there exist $a_n > 0$ and b_n, $n = 1, 2, \cdots$ such that*

$$F_n(a_n x + b_n) \to H(x).$$

Let $\{\alpha_n\}$ and $\{\beta_n\}$ be sequences of positive numbers tending for $n \to \infty$ to limits α and β respectively, where $0 < \alpha < \beta < 1$. If we define

$$B_n = \xi^{(n)}_{\alpha_n} , \qquad B = H^{-1}(\alpha) ,$$

$$A_n = \xi^{(n)}_{\beta_n} - B_n , \qquad A = H^{-1}(\beta) - B,$$

where $\xi^{(n)}_{\gamma} = \inf\{x : F_n(x) \geq \gamma\}$, then we have

$$F_n(A_n x + B_n) \to H(Ax + B).$$

LEMMA 0.2. *Let U and V be two distribution functions neither of which is concentrated at one point. If for a sequence $\{F_n\}$ of distribution functions and constants $a_n > 0$, $\alpha_n > 0$, b_n and β_n*

$$F_n(a_n x + b_n) \to U(x), \qquad F_n(\alpha_n x + \beta_n) \to V(x) \tag{0.1}$$

then

$$\lim_{n\to\infty} \alpha_n/a_n = A > 0, \qquad \lim_{n\to\infty} (\beta_n - b_n) / a_n = B, \tag{0.2}$$

and

$$V(x) = U(Ax + B) \tag{0.3}$$

is true. Conversely, if (0.2) holds then each of the two relations (0.1) implies the other and (0.3).

CHAPTER 1

ASYMPTOTIC THEORY OF EXTREME STATISTIC - UNIVARIATE CASE

1.1. Introduction

In this chapter we deal with the univariate extreme statistic. We extend the results of Resnick [19] (see Theorem 1.7), and give simple sufficient conditions for a distribution function to belong to the domain of attraction of Λ and a method to determine the normalizing constants.

Let $X_1, X_2, \cdots, X_n$ be a sequence of independent random variables with common distribution function F and let

$$Z_n = \max_{1 \leq i \leq n} X_i .$$

If there exist sequences of constants $a_n > 0$ and b_n such that $(Z_n - b_n)/a_n$ has a nondegenerate limit distribution H, or $F^n(a_n x + b_n) \to H(x)$, then F is said to be in the domain of attraction of H (notation $F \in \mathbf{D}(H)$), H is said to be an extreme value distribution, and a_n and b_n are called normalizing constants.

In the sequel we use the following notations

$$x_F^0 = \sup\{x: F(x) < 1\}$$

and

$$\overline{F}^{-1}(p) = \inf\{x: 1 - F(x) \leq p\}, \quad 0 < p < 1.$$

Note that $x_F^0 \leq \infty$. The number x_F^0 is called the endpoint of the distribution function F.

Now we summarize some univariate results which are used in this thesis.

Gnedenko [8] showed the following classical results of necessary and sufficient conditions for a distribution function to belong to the domain of attraction of extreme value distributions.

LEMMA 1.1. *Suppose* $\{a_n\}$ *and* $\{b_n\}$ *are sequences of real numbers and* $a_n > 0$ *for* $n = 1, 2, \cdots$. *For distribution functions* F *and* H *we have for a fixed real* x *with* $0 < H(x) < 1$

$$\lim_{n\to\infty} F^n(a_n x + b_n) = H(x)$$

if and only if

$$\lim_{n\to\infty} n\{1 - F(a_n x + b_n)\} = -\log H(x).$$

THEOREM 1.1. *A distribution function* F *belongs to* $\mathbf{D}(\Phi_\alpha)$ *if and only if* $1 - F$ *is* $-\alpha$*-varying at infinity. Moreover*

$$F^n(a_n x) \to \Phi_\alpha(x)$$

holds with $a_n = \overline{F}^{-1}(1/n)$, $n = 1, 2, \cdots$.

THEOREM 1.2. *A distribution function* F *belongs to* $\mathbf{D}(\Psi_\alpha)$ *if and only if* F *has a finite endpoint* x^0 *and the function* R *defined by* $R(x) = 1 - F(x^0 - 1/x)$ *for all* $x > 0$ *is* $-\alpha$*-varying at infinity. Moreover*

$$F^n(a_n x + x^0) \to \Psi_\alpha(x)$$

holds with $a_n = \{x^0 - \overline{F}^{-1}(1/n)\}^{-1}$, $n = 1, 2, \cdots$.

Now we quote the following theorems concerning the domain of attraction of Λ.

THEOREM 1.3. (Necessary and sufficient conditions.) *For a distribution function* F, *let*

$$b(s) = \overline{F}^{-1}(1/s),$$

$$a(s) = b(es) - b(s)$$

and

$$x^0 = x^0_F .$$

Then the following assertions are equivalent.

(a) $F \in \mathbf{D}(\Lambda)$.

(b) (Gnedenko [8] and Haan [12]) *There exist* $a_n > 0$, b_n, $n = 1, 2, \cdots$ *such that*

$$F^n(a_n x + b_n) \to \Lambda(x).$$

Moreover, we can take $b_n = b(n)$, $a_n = a(n)$ *or* $f_a(b_n)$, *where* f_a *is the auxiliary function of* F (see (c)).

(c) (Gnedenko [8] and Haan [12]) *There exists a function* $f_a\colon \mathbf{R} \to \mathbf{R}^+$ *such that*

$$\lim_{t \uparrow x^0} \frac{1 - F(t + x \cdot f_a(t))}{1 - F(t)} = e^{-x} \qquad \textit{for all } x \in \mathbf{R}.$$

Moreover, we can take

$$f_a(t) = \frac{\int_t^{x^0} \{1 - F(s)\}ds}{1 - F(t)} \qquad \textit{for all } t < x^0.$$

Such a f_a *is said to be an auxiliary function of* F.

(d) (Mejzler, see Haan [12]) *For every positive* x *and* y $(y \neq 1)$

$$\lim_{s \to \infty} \frac{b(sx) - b(s)}{b(sy) - b(s)} = \frac{\log x}{\log y}.$$

(e) (Haan [12]) *There exist functions* $\alpha : \mathbf{R} \to \mathbf{R}^+$ *and* $\beta : \mathbf{R} \to \mathbf{R}$, *such that*

$$\lim_{s \to \infty} s\{1 - F(\alpha(s)x + \beta(s))\} = e^{-x} \qquad \textit{for all } x \in \mathbf{R}.$$

Moreover, we can take $\alpha(s) = a(s)$ *and* $\beta(s) = b(s)$.

THEOREM 1.4. (Sufficient conditions.) *If one of the following conditions holds ,then we have* $F \in \mathbf{D}(\Lambda)$.

(a) (Mises [15]) *The endpoint of F is infinite, F is twice differentiable at least for all x greater than some value x_0, and is such that*

$$\lim_{x\to\infty} \frac{d}{dx}\left[\frac{1 - F(x)}{F'(x)}\right] = 0,$$

where the prime denoting the derivative.

(b) (Haan [12]) *There exists a distribution function $G \in \mathbf{D}(\Lambda)$ such that the function $R : \mathbf{R}^+ \to \mathbf{R}^+$ defined by $R(x) = [1 - F(x)]/[1 - G(x)]$ is regularly varying at infinity.*

(c) (Haan [12]) *There exist a distribution function $G \in \mathbf{D}(\Lambda)$ and a function r which has a ρ-varying derivative $(-1 < \rho < \infty)$ such that $F(x) = G(r(x))$ for $x > 0$.*

(d) (Haan [12]) *There exists a distribution function $G \in \mathbf{D}(\Lambda)$ such that $F(x) = G(e^x)$.*

Note that we can take

$$f_a(t) = g_a(r(t))t/(\rho+1)r(t) \qquad \text{in (c)}$$

and

$$f_a(t) = g_a(e^t)e^{-t} \qquad \text{in (d)},$$

where f_a and g_a are auxiliary functions of F and G, respectively.

THEOREM 1.5. (Necessary conditions.) *For a distribution function F, let*

$$b(s) = \overline{F}^{-1}(1/s), \qquad a(s) = b(es) - b(s).$$

If $F \in \mathbf{D}(\Lambda)$, then the following results hold.

(a) (Marcus and Pinsky [13]) *The function $a(s)$ is slowly varying at infinity.*

(b) (Haan [12]) *If* $b(\infty) = \infty$, *then* $b(s)$ *is slowly varying at infinity. If* $b(\infty) < \infty$, *then* $b(\infty) - b(s)$ *is slowly varying at infinity.*

THEOREM 1.6. (Gnedenko [8]) *There exist* $a > 0$ *and* b_n, $n = 1, 2, \cdots$ *such that*

$$F^n(ax + b_n) \to \Lambda(x)$$

if and only if $F(\log x)$ *is* $-\alpha$*-varying at infinity and* $\alpha a = 1$.

Resnick [19] proved the following theorem of tail equivalence and asymptotic distributions of extreme statistics.

THEOREM 1.7. *Let* F, G *be distribution functions and let* H *be an extreme value distribution. Suppose* $F \in \mathbf{D}(H)$ *and that* $F^n(a_n x + b_n) \to H(x)$ *for normalizing constants* $a_n > 0$, b_n, $n = 1, 2, \cdots$. *Then, for some* $A > 0$, B: $G^n(a_n x + b_n) \to H(Ax + B)$ *if and only if*

$$x_F^0 = x_G^0 = x^0,$$

$$\lim_{x \uparrow x^0} \frac{1 - F(x)}{1 - G(x)} \quad \textit{exists.}$$

Moreover,

(a) *if* $H = \Phi_\alpha$, *then* $B = 0$ *and* $\lim_{x \to \infty} [1 - F(x)]/[1 - G(x)] = A^\alpha$;

(b) *if* $H = \Psi_\alpha$, *then* $B = 0$ *and* $\lim_{x \uparrow x^0} [1 - F(x)]/[1 - G(x)] = A^{-\alpha}$;

(c) *if* $H = \Lambda$, *then* $A = 1$ *and* $\lim_{x \uparrow x^0} [1 - F(x)]/[1 - G(x)] = e^B$.

1.2. The domain of attraction of Ψ_α

We show some properties of normalizing constants of the extreme statistic from a distribution which belongs to $D(\Psi_\alpha)$.

THEOREM 2.1. *Let F and G be distribution functions. Suppose $F \in D(\Psi_\alpha)$ or there exist $a_n > 0$ and b_n such that*

$$F^n(a_n x + b_n) \to \Psi_\alpha(x) \tag{2.1}$$

and there exist $\alpha_n > 0$ and β_n such that

$$G^n(\alpha_n x + \beta_n) \to H(x), \quad H \text{ nondegenerate.} \tag{2.2}$$

Then there exists c_0 $(0 < c_0 < \infty)$ such that

$$\lim_{n\to\infty} a_n/\alpha_n = c_0 \quad \text{and} \quad x_F^0 = x_G^0 \ (= x^0,\ say) < \infty, \tag{2.3}$$

if and only if there exists $c_1 (0 < c_1 < \infty)$ such that

$$\lim_{x\uparrow x_1} \frac{1 - F(x)}{1 - G(x)} = c_1, \tag{2.4}$$

where $x_1 = x_F^0$ or x_G^0.

Moreover, if (2.3) or (2.4) holds, then there exist $A > 0$ and B such that $H(x) = \Psi_\alpha(Ax+B)$.

PROOF. For $s > 0$, define

$$a(s) = \{x_F^0 - \bar{F}^{-1}(1/s)\}^{-1},$$

$$\alpha(s) = \{x_G^0 - \bar{G}^{-1}(1/s)\}^{-1},$$

$$\beta(s) = \bar{G}^{-1}(1/s)$$

and

$$\alpha'(s) = \beta(es) - \beta(s).$$

Sufficiency. By (2.4) it is easily seen that $x_1^0 = x_F^0 = x_G^0 = x^0 < \infty$.

From Theorem 1.2 we get that $1 - F(x^0 - 1/x)$ is $-\alpha$-varying. Then by (2.4) we have that $1 - G(x^0 - 1/x)$ is also $-\alpha$-varying, too. Hence we have $G \in \mathbf{D}(\Psi_\alpha)$ and $H(x) = \Psi_\alpha(Ax+B)$. By Lemma 0.2 in Chapter 0 and Theorem 1.2 we can suppose without loss of generality that $a_n = a(n)$ and $\alpha_n = A\alpha(n)$. By (2.4) and Theorem 0.1 (c) in Chapter 0 we have

$$\lim_{n\to\infty} a_n/\alpha_n = c_1^{-1/\alpha}/A .$$

Necessity. By (2.2), $x_G^0 < \infty$ and Theorems 1.1, 1.2 and 1.3 it holds that $H(x) = \Lambda(Ax+B)$ or $H(x) = \Psi_\beta(Ax+B)$, where A, $\beta > 0$. If $H(x) = \Lambda(Ax+B)$, then we can take $\alpha_n = A\alpha'(n)$. By Theorem 1.5 (a) $\alpha'(s)$ is slowly varying. On the other hand, $a(s)$ is $-\alpha^{-1}$-varying. Therefore, (2.3) does not hold. Thus we have $H(x) = \Psi_\beta(Ax+B)$ and we can suppose that $\alpha_n = A\alpha(n)$. Since $a(s)$ and $\alpha(s)$ are non-increasing and $a(s)$ is $-\alpha^{-1}$-varying, we have

$$\lim_{s\to\infty} a(s)/\,A\alpha(s) = c_0 .$$

Hence we have that $\alpha(s)$ is also $-\alpha^{-1}$-varying, too. By Theorem 0.1 (c) in Chapter 0 it holds that

$$\lim_{x\uparrow x^0} \frac{1 - F(x)}{1 - G(x)} = (Ac_0)^{-\alpha}.$$

So we have $\beta = \alpha$. Q.E.D.

COROLLARY 2.1. *Let F be a distribution function with endpoint $x_F^0 = 0$ and $0 < c < \infty$. Then*

$$\lim_{x\uparrow 0} \frac{1 - F(x)}{1 - \Psi_\alpha(x)} = c$$

if and only if

$$F^n(x/n^{1/\alpha}) \to \Psi_\alpha(c^{1/\alpha}x) .$$

COROLLARY 2.2. *Let F be a distribution function with density f and $0 < c < \infty$. Then*

$$F^n(x/n) \to \Psi_1(cx)$$

if and only if

$$\lim_{x \downarrow 0} f(x) = c \quad \text{and} \quad f(x) = 0 \quad \text{for } x > 0.$$

REMARK 2.1. Theorem 2.1 is an extension of Theorem 1.7 (b) with a simple proof.

REMARK 2.2. Corollary 2.2 is closely related to the order of efficient estimators dealt with in Section 4.4.

1.3. The domain of attraction of Λ

In this section we consider the distribution functions which belong to $\mathbf{D}(\Lambda)$.

First we prove four theorems useful in choosing normalizing constants.

THEOREM 3.1. *Let F be a distribution function.*

(a) *Suppose $F \in \mathbf{D}(\Lambda)$ or there exist $a_n > 0$ and b_n such that*

$$F^n(a_n x + b_n) \to \Lambda(x). \tag{3.1}$$

Then

$$\lim_{n\to\infty} n\{1 - F(b_n)\} = 1. \tag{3.2}$$

(b) *Conversely, if $F \in \mathbf{D}(\Lambda)$, then* (3.1) *holds with b_n which satisfies* (3.2) *and $a_n = f_a(b_n)$, where f_a is defined in* Theorem 1.3.

PROOF. By Lemma 1.1 and Theorem 1.3 (b), (c), the proof is trivial.

THEOREM 3.2. *Let F and G be distribution functions. Suppose there exist $a_n > 0$, b_n and $\alpha_n > 0$ such that* (3.1) *holds, and it holds*

$$G^n(\alpha_n x + b_n) \to \Lambda(Ax+B), \tag{3.3}$$

where $A > 0$. Then we have

$$\lim_{n\to\infty} \alpha_n/a_n = A. \tag{3.4}$$

PROOF. For $s > 0$, let $b(s) = \overline{F}^{-1}(1/s)$ and $\beta(s) = \overline{G}^{-1}(1/s)$. By Lemma 0.2 and Theorem 0.2 in Chapter 0 we can suppose without loss of generality that $a_n = b(en) - b(n)$, $b_n = b(n) = \beta(e^B n)$ and $\alpha_n = \beta(e^A e^B n) - \beta(e^B n)$. It holds that

$$\beta(e^B[en]) \leq \beta(ee^B n) \leq \beta(e^B([en]+1)) \quad \text{for all } n,$$

where $[en]$ is the greatest integer less than or equal to en. Thus we have

$$(3.5) \qquad \frac{b([en]) - b(n)}{b(en) - b(n)} \le \frac{\beta(ee^{B}n) - \beta(e^{B}n)}{b(en) - b(n)} \le \frac{b([en]+1) - b(n)}{b(en) - b(n)} .$$

As b is non-decreasing, we have for fixed c $(1 < c < e)$ and sufficiently large n

$$1 \ge \frac{b([en]) - b(n)}{b(en) - b(n)} \ge \frac{b(cn) - b(n)}{b(en) - b(n)}.$$

By Theorem 1.3 (d) we have

$$\lim_{n\to\infty} \frac{b(cn) - b(n)}{b(en) - b(n)} = \log c .$$

Hence we have

$$(3.6) \qquad \lim_{n\to\infty} \frac{b([en]) - b(n)}{b(en) - b(n)} = 1 .$$

Similarly

$$(3.7) \qquad \lim_{n\to\infty} \frac{b([en]+1) - b(n)}{b(en) - b(n)} = 1 .$$

By (3.5), (3.6) and (3.7), it holds that

$$(3.8) \qquad \lim_{n\to\infty} \frac{\beta(ee^{B}n) - \beta(e^{B}n)}{b(en) - b(n)} = 1 .$$

On the other hand, by Theorem 1.3 (d) it holds that

$$(3.9) \qquad \lim_{n\to\infty} \frac{\beta(e^{A}e^{B}n) - \beta(e^{B}n)}{\beta(ee^{B}n) - \beta(e^{B}n)} = A .$$

By (3.8) and (3.9) we have (3.4). Q.E.D.

The converse of Theorem 3.2 becomes the following theorem.

THEOREM 3.3. *Let F and G be distribution functions. Suppose there exist $a_n > 0$, b_n, $\alpha_n > 0$ and β_n such that* (3.1) *holds, and it holds that*

$$(3.10) \qquad G^{n}(\alpha_n x + \beta_n) \to \Lambda(x).$$

(a) *If* $x_F^0 = x_G^0 = \infty$ *and for* $0 \le c \le \infty$,

(3.11) $$\lim_{n\to\infty} a_n/\alpha_n = c \ ,$$

then

$$\lim_{n\to\infty} b_n/\beta_n = c \ .$$

(b) *If* $x_F^0 = x_G^0 = x^0 < \infty$ *and for* $0 \le c \le \infty$, (3.11) *holds, then*

$$\lim_{n\to\infty} \frac{x^0 - b_n}{x^0 - \beta_n} = c \ .$$

PROOF. For $s > 0$, define

$$a(s) = b(es) - b(s)$$

and

$$\alpha(s) = \beta(es) - \beta(s)$$

where $b(s) = \overline{F}^{-1}(1/s)$ and $\beta(s) = \overline{G}^{-1}(1/s)$. By Lemma 0.2 in Chapter 0 we can suppose without loss of generality that $a_n = a(n)$, $b_n = b(n)$, $\alpha_n = \alpha(n)$ and $\beta_n = \beta(n)$.

(a) By Theorem 1.5 (b) we have

$$b(s) \sim \int_1^s \frac{a(t)}{t}\, dt$$

and

$$\beta(s) \sim \int_1^s \frac{\alpha(t)}{t}\, dt$$

Suppose (3.11) holds, then by Theorem 1.5 (a) it holds that

$$\lim_{s\to\infty} a(s)/\alpha(s) = c \ .$$

Thus we have

$$\lim_{s\to\infty} \frac{b(s)}{\beta(s)} = \lim_{s\to\infty} \frac{\int_1^s \frac{a(t)}{t}\, dt}{\int_1^s \frac{\alpha(t)}{t}\, dt} = c \ .$$

(b) By Theorem 1.5 (b) we have

$$x^0 - b(s) \sim \int_1^s \frac{a(t)}{t}\,dt$$

and

$$x^0 - \beta(s) \sim \int_1^s \frac{\alpha(t)}{t}\,dt ,$$

so that a similar argument as in (a) proves (b). Q.E.D.

COROLLARY 3.1. *Let F be a distribution function with infinite endpoint. Then $1 - F(\log x)$ is $-\alpha$-varying if and only if there exists b_n such that*

$$F^n(x/\alpha + b_n) \to \Lambda(x)$$

and $b_n \sim \log n/\alpha$, where $\alpha > 0$.

PROOF. It is clear from Theorems 1.6 and 3.3.

THEOREM 3.4. *Let F and G be distribution functions. Suppose $F \in \mathbf{D}(\Lambda)$ or there exist a_n and b_n such that (3.1) holds. Then there exists $\alpha_n > 0$ such that*

$$G^n(\alpha_n x + b_n) \to \Lambda(Ax+B) \quad \text{and} \quad A > 0 \tag{3.12}$$

if and only if $x_F^0 = x_G^0$ ($= x^0$, say), $\alpha_n/a_n \to A$, and

$$\lim_{x \uparrow x^0} \frac{1 - G(x)}{1 - F(x)} = e^{-B}. \tag{3.13}$$

PROOF. Sufficiency. By Lemma 1.1 we have

$$\lim_{n\to\infty} n\{1 - F(a_n x + b_n)\} = e^{-x}. \tag{3.14}$$

So it holds that $a_n x + b_n \to x^0$ as $n \to \infty$. Hence we have

$$\lim_{n\to\infty} \frac{1 - G(a_n x + b_n)}{1 - F(a_n x + b_n)} = e^{-B}.$$

Therefore it holds that

$$(3.15) \qquad \lim_{n\to\infty} n\{1 - G(a_n x + b_n)\} = e^{-(x+B)}.$$

By $\lim_{n\to\infty} \alpha_n/a_n = A > 0$, we have (3.12).

Necessity. By Theorem 3.2 we have $\lim_{n\to\infty} \alpha_n/a_n = A$. By (3.1), (3.12) and Lemma 1.1 we have (3.14) and (3.15), hence we get $b_n \to x_F^0$ and $b_n \to x_G^0$. So it holds that $x_F^0 = x_G^0 = x^0$. By Lemma 0.2 in Chapter 0 we can suppose without loss of generarity that b_n is defined by $\overline{F}^{-1}(1/n)$, and hence $b_n \uparrow x^0$ as $n \to \infty$. For any x sufficiently near $x^0 (x < x^0)$, there exists an n such that $b_n \le x \le b_{n+1}$. Then we have

$$\frac{1 - G(b_{n+1})}{1 - F(b_n)} \le \frac{1 - G(x)}{1 - F(x)} \le \frac{1 - G(b_n)}{1 - F(b_{n+1})} .$$

By (3.14) and (3.15), taking the limits of above inequalities, we have (3.13). Q.E.D.

REMARK 3.1. Theorem 3.4 is an extension of Theorem 1.7 (c) with a simpler proof.

REMARK 3.2. Given a distribution function $G \in \mathbf{D}(\Lambda)$, from Theorems 3.1 and 3.2 we can present the following procedure to find normalizing constants of G. We first choose an appropriate simple distribution function F such that (3.13) holds for some B, seek b_n which satisfy (3.2) and then define $a_n = f_a(b_n)$. The explicit form of the function f_a is given by Theorem 1.3 (c).

The following two theorems determine the transformations of the normilizing constants induced by a transformations of the underlying random variable.

THEOREM 3.5. *Let a_n and b_n be normalizing constants of a distribution function $F \in \mathbf{D}(\Lambda)$ with infinite endpoint. For $x > 0$, let $G(x) = F(r(x))$*

where r is a function which has a ρ-varying derivative $(-1 < \rho < \infty)$. *Then* $G \in \mathbf{D}(\Lambda)$, *and*

$$G^n(\alpha_n x + \beta_n) \to \Lambda(x)$$

where $\alpha_n = a_n r^{-1}(b_n)/(\rho+1)b_n$ *and* $\beta_n = r^{-1}(b_n)$.

PROOF. By Theorem 1.4 (c) it holds that $G \in \mathbf{D}(\Lambda)$ and $g_a(t) = f_a(r(t))t/(\rho+1)r(t)$, where f_a and g_a are auxiliary functions of F and G, respectively. Then we can suppose without loss of generality that $\beta_n = \bar{G}^{-1}(1/n)$ and $\alpha_n = g_a(\beta_n)$ (see Theorem 1.3 (b)). The inverse function r^{-1} of r is well determined, because r is continuous and strictly increasing. Thus we have $\beta_n = r^{-1}(b_n)$ and $\alpha_n = a_n r^{-1}(b_n)/(\rho+1)b_n$. Q.E.D.

COROLLARY 3.2. *Let F be a distribution function with infinite endpoint. For $\alpha > 0$ and $\rho > 0$, we have*

$$\frac{1 - F(t + xt^{1-\alpha}/\alpha\rho)}{1 - F(t)} \to e^{-x} \qquad \text{as } t \to \infty$$

if and only if $1 - F((\log x)^{1/\alpha})$ *is* $-\rho$*-varying.*

PROOF. It is clear from Theorems 1.6 and 3.5.

THEOREM 3.6. *Let α_n and β_n be normalizing constants of a distribution function $G \in \mathbf{D}(\Lambda)$ with infinite endpoint. Let $F(x) = G(e^x)$. Then $F \in \mathbf{D}(\Lambda)$ and*

$$F^n(a_n x + b_n) \to \Lambda(x),$$

where $a_n = \alpha_n/\beta_n$ *and* $b_n = \log \beta_n$.

PROOF. By Theorem 1.4 (d) it holds that $F \in \mathbf{D}(\Lambda)$ and $f_a(t) = g_a(e^t)e^{-t}$, where f_a and g_a are auxiliary functions of F and G, respectively. We can suppose without loss of generality that $\beta_n = \bar{G}^{-1}(1/n)$ and $\alpha_n = g_a(\beta_n)$ (see

Theorem 1.3 (b)). Then we have $b_n = \overline{F}^{-1}(1/n) = \log \beta_n$ and $a_n = f_a(b_n)$ $= g_a(\beta_n)\exp(-b_n) = \alpha_n/\beta_n$. Q.E.D.

Now we show some simple sufficient conditions that a distribution function belongs to $\mathbf{D}(\Lambda)$ and determine the normalizing constants.

THEOREM 3.7. *Let F be a distribution function. Suppose that there exist $\alpha > 0$, $a > 0$, $c > 0$ and $b \in \mathbf{R}$ such that*

$$\lim_{x\to\infty} \frac{1 - F(x)}{ax^b e^{-cx^\alpha}} = 1.$$

For $\mu \in \mathbf{R}$ and $\sigma > 0$, let $F_*(x) = F((x-\mu)/\sigma)$. Then $F_* \in \mathbf{D}(\Lambda)$ and

$$F_*^n(a_n^* x + b_n^*) \to \Lambda(x),$$

where $a_n^* = \sigma a_n$, $b_n^* = \sigma b_n + \mu$, $a_n = ((\log n)/c)^{1/\alpha-1}/\alpha c$ and $b_n = ((\log n)/c)^{1/\alpha}$ $+ \{(b/\alpha)(\log(\log n) - \log c) + \log a\}/\{((\log n)/c)^{1-1/\alpha}\alpha c\}$.

PROOF. Consider the following distribution functions

$$G_0(x) = 1 - e^{-x}, \quad x > 0$$

and

$$G(x) = 1 - ax^b e^{-cx^\alpha}, \quad x > x_1,$$

where x_1 is a constant such that $G(x)$, $G'(x) > 0$ for all $x > x_1$. Then $G_0 \in \mathbf{D}(\Lambda)$. By Theorem 3.5 and Theorem 1.4 (b) we have $G \in \mathbf{D}(\Lambda)$. It is easy to show that

$$n\{1 - G(b_n)\} \to 1 \quad \text{as } n \to \infty.$$

By Corollary 3.2 , Lemma 0.2 in Chapter 0 and $a_n \sim b_n^{1-\alpha}/\alpha c$, we have

$$G^n(a_n x + b_n) \to \Lambda(x).$$

Thus, by Theorem 1.7 (c), we have

$$F^n(a_n x + b_n) \to \Lambda(x)$$

Some obvious calculation gives the statement of the theorem. Q.E.D.

THEOREM 3.8. *Suppose a distribution function F has a density f such that*

$$\lim_{x\to\infty} \frac{f(x)}{\alpha a c x^{\alpha+b-1} e^{-cx^{\alpha}}} = 1 ,$$

where $\alpha > 0$, $a > 0$, $c > 0$ and $b \in \mathbf{R}$. For $\mu \in \mathbf{R}$ and $\sigma > 0$, let $F_*(x) = F((x-\mu)/\sigma)$. Then the same conclusion as that of Theorem 3.7 *holds.*

PROOF. Consider the same distribution function G as in the proof of Theorem 3.7. Let $g(x) = G'(x)$, then by the assumption it holds that $\lim_{x\to\infty} f(x)/g(x) = 1$. Thus, by Lemma 0.1 in Chapter 0, we have

$$\lim_{x\to\infty} \frac{1 - F(x)}{ax^b e^{-cx^{\alpha}}} = \lim_{x\to\infty} \frac{\int_x^\infty f(s)ds}{\int_x^\infty g(s)ds} = 1 .$$

By Theorem 3.7 we have the conclusion. Q.E.D.

THEOREM 3.9. *Let G and G_* be distribution functions such that $G(x) = F(\log x)$ and $G_*(x) = F_*(\log x)$ for $x > 0$, where F and F_* are distribution functions satisfying the conditions in* Theorem 3.7. *If $\alpha > 1$, then $G_* \in \mathbf{D}(\Lambda)$ and*

$$G_*^n(\alpha_n^* x + \beta_n^*) \to \Lambda(x),$$

where $\alpha_n^* = a_n^* \beta_n^*$, $\beta_n^* = \exp(b_n^*)$ and a_n^*, b_n^* are the same normalizing constants as that of Theorem 3.7.

PROOF. Let $H(x) = 1 - \exp\{-c(\log x)^{\alpha}\}$, for $x > 1$. By Theorem 1.4 (a) we have $H \in \mathbf{D}(\Lambda)$. Thus by Theorem 1.4 (b) we have $G \in \mathbf{D}(\Lambda)$. By the relation $G_*(x) = G(e^{-\mu/\sigma} x^{1/\sigma})$ and Theorem 1.4 (c) we have $G_* \in \mathbf{D}(\Lambda)$.

Therefore, by Theorems 3.6 and 3.7 we have the conclusion. Q.E.D.

The following theorem can be proved similarly as Theorem 3.8.

THEOREM 3.10. *Let G and G_* be distribution functions. Suppose $G(x) = F(\log x)$ and $G_*(x) = F_*(\log x)$ for $x > 0$, where F and F_* are distribution functions satisfying the conditions in* Theorem 3.8. *If $\alpha > 1$, then $G_* \in \mathbf{D}(\Lambda)$ and the same conclusion as that of* Theorem 3.9 *holds.*

REMARK 3.3. (a) Theorems 3.7–3.10 are very simple and the conditions are convenient to verify. In the next section we will see that many distribution functions which are used in applied statistics satisfy these conditions.
(b) Note that $\alpha > 0$ in Theorems 3.7 and 3.8, whereas $\alpha > 1$ in Theorems 3.9 and 3.10.
(c) It is well-known that b_n is given by $\bar{F}^{-1}(1/n)$ (see Theorem 1.3 (b)). But with some distributions its exact values can not be expressed by sample size n in simple forms as in Theorems 3.7–3.10.

1.4. Examples

We shall apply the results of Section 1.3 to some specific distribution functions which are often used in applied statistics.

EXAMPLE 4.1 (Inverse Gaussian distribution). Suppose a distribution function F has a density function

$$f(x) = \begin{cases} \left(\dfrac{\lambda}{2\pi x^3}\right)^{1/2} \exp - \left\{\dfrac{\lambda(x-\mu)^2}{2\mu^2 x}\right\}, & x > 0 \\ 0, & \text{otherwise,} \end{cases}$$

where $\mu > 0$ and $\lambda > 0$. Then by Theorem 3.8 we have $F \in \mathbf{D}(\Lambda)$ and

$$F^n(a_n x + b_n) \to \Lambda(x),$$

where $a_n = 2\mu^2/\lambda$ and

$b_n = \mu^2\{ 2\log n - 3\log(\log n) + \log(\lambda^2/(4\pi\mu^2)) + 2\lambda/\mu\}/\lambda$.

EXAMPLE 4.2 (Gamma distribution and Mixture of gamma distributions).

(I) Suppose a distribution function G has a density function

$$g(x) = x^{p-1}e^{-x}/\ \Gamma(p),$$

where $p > 0$ and $\Gamma(\cdot)$ is the gamma function. For $\mu \in \mathbf{R}$ and $\sigma > 0$, let $G_*(x) = G((x-\mu)/\sigma)$. Then by Theorem 3.8 we have $G_* \in \mathbf{D}(\Lambda)$ and

$$G_*^n(a_n^* x + b_n^*) \to \Lambda(x),$$

where $a_n^* = \sigma a_n = \sigma$, $b_n^* = \sigma b_n + \mu$, $a_n = 1$ and

$b_n = \log [n/\Gamma(p)] + (p-1)\log (\log n)$.

Note that Gumbel's approximation

$$u_2 = \log [n/\Gamma(p)] + (p-1)\log \log [n/\Gamma(p)]$$

of $\inf\{x: 1 - G(x) \leq 1/n\}$ (see Gumbel [11], p.145) is not an exact normalizing constant.

(II) Suppose F is a distribution function such that

$$F(x) = pG((x-\mu_1)/\sigma_1) + qG((x-\mu_2)/\sigma_2),$$

where μ_1, μ_2 ε $\mathbf{R}$, σ_1, $\sigma_2 > 0$, p, $q > 0$ and $p + q = 1$. The density function of F is

$$f(x) = pg((x-\mu_1)/\sigma_1)/\sigma_1 + qg((x-\mu_2)/\sigma_2)/\sigma_2.$$

(1) If $\sigma_1 = \sigma_2 = \sigma$, then

$$f(x)/\{g(x/\sigma)/\sigma\} \to pe^{\mu_1/\sigma} + qe^{\mu_2/\sigma} \qquad \text{as } x \to \infty.$$

Hence, by Theorem 3.8 we have $F \,\varepsilon\, \mathbf{D}(\Lambda)$ and

$$F^n(a_n^1 x + b_n^1) \to \Lambda(x),$$

where $a_n^1 = \sigma$ and $b_n^1 = \sigma(b_n + \log(pe^{\mu_1/\sigma} + qe^{\mu_2/\sigma}))$.

(2) If $\sigma_1 < \sigma_2$, then

$$f(x)/\{qg((x-\mu_2)/\sigma_2)/\sigma_2\} \to 1 \qquad \text{as } x \to \infty.$$

Hence, by Theorem 3.8 we have $F \,\varepsilon\, \mathbf{D}(\Lambda)$ and

$$F^n(a_n^2 x + b_n^2) \to \Lambda(x),$$

where $a_n^2 = \sigma_2$ and $b_n^2 = \sigma_2(b_n + \log q) + \mu_2$.

EXAMPLE 4.3 (Normal distribution, Lognormal distribution and Mixture of normal distributions).

(I) Suppose G is a distribution function such that

$$G(x) = \Phi(\log x) \quad \text{for } x > 0,$$

where Φ is a standard normal distribution function. For $\mu \,\varepsilon\, \mathbf{R}$ and $\sigma > 0$ let $\Phi_*(x) = \Phi((x-\mu)/\sigma)$ and $G_*(x) = \Phi_*(\log x)$ for $x > 0$. Then by Theorem 3.8 we have $\Phi_* \,\varepsilon\, \mathbf{D}(\Lambda)$ and

$$\Phi_*^n(a_n^* x + b_n^*) \to \Lambda(x),$$

where $a_n^* = \sigma a_n$, $b_n^* = \sigma b_n + \mu$, $a_n = (2 \log n)^{-1/2}$ and $b_n = (2 \log n)^{1/2} - (\log(\log n) + \log 4\pi)/\ 2(2 \log n)^{1/2}$.

On the other hand, by Theorem 3.9 we have $G_* \in \mathbf{D}(\Lambda)$ and

$$G_*^n(\alpha_n^* x + \beta_n^*) \to \Lambda(x)$$

where $\alpha_n^* = a_n^* \beta_n^*$ and $\beta_n^* = \exp(b_n^*)$.

Approaches presented in Galambos [7] (p.65 and p.67) and Singpurwalla [21] to evaluate normalizing constants seem to be more complicated than ours. It is hard to evaluate $\bar{\Phi}^{-1}(1/n)$ and sometimes inadequate sequences $(2 \log n)^{1/2}$ are claimed to be b_n (for example see David [2], p.209).

(II) Suppose F is a distribution function such that

$$F(x) = p\Phi((x-\mu_1)/\sigma_1) + q\Phi((x-\mu_2)/\sigma_2),$$

where $\mu_1, \mu_2 \in \mathbf{R}$, $\sigma_1, \sigma_2 > 0$, $p, q > 0$ and $p + q = 1$. The density of F is

$$f(x) = p\phi((x-\mu_1)/\sigma_1)/\sigma_1 + q\phi((x-\mu_2)/\sigma_2)/\sigma_2,$$

where $\phi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$.

(1) If $\sigma_1 = \sigma_2 = \sigma$ and $\mu_1 < \mu_2$, then

$$f(x)/\{q\phi((x-\mu_2)/\sigma)/\sigma\} \to 1 \quad \text{as } x \to \infty.$$

Hence by Theorem 3.8 we have $F \in \mathbf{D}(\Lambda)$ and

$$F^n(a_n^1 x + b_n^1) \to \Lambda(x),$$

where $a_n^1 = \sigma a_n$ and $b_n^1 = \sigma(b_n + a_n \log q) + \mu_2$.

(2) If $\sigma_1 < \sigma_2$, then

$$f(x)/\{q\phi((x-\mu_2)/\sigma_2)/\sigma_2\} \to 1 \quad \text{as } x \to \infty.$$

Hence by Theorem 3.8 we have $F \in \mathbf{D}(\Lambda)$ and

$$F^n(a_n^2 x + b_n^2) \to \Lambda(x),$$

where $a_n^2 = \sigma_2 a_n$ and $b_n^2 = \sigma_2(b_n + a_n \log q) + \mu_2$.

CHAPTER 2

ASYMPTOTIC THEORY OF EXTREME STATISTIC - MULTIVARIATE CASE

2.1. Introduction

In this chapter, we establish some properties of multivariate extreme value distributions, and by using the results of Marshall and Olkin [14] (see Propositions 1.1 - 1.3) we extend some of the results given by Resnick [19] (see Theorem 1.7 in Section 1.1) to the multivariate case. We may use the same notations as in Marshall and Olkin [14].

For $\mathbf{a}, \mathbf{b}, \mathbf{x} \in \mathbf{R}^k$, write $\mathbf{ax} + \mathbf{b}$ to denote the vector

$$(a_1x_1 + b_1, \cdots, a_kx_k + b_k).$$

Basic arithmetical operations are always meant componentwise. Let $\mathbf{X}^{(1)}$, $\mathbf{X}^{(2)}$, $\cdots$ be a sequence of independent k-dimensional random vectors with common distribution function F and let

$$Z_i^{(n)} = \max_{1\le j\le n} X_i^{(j)}, \quad i = 1,\cdots, k.$$

If there exist $\mathbf{a}^{(n)} > \mathbf{0}$, $\mathbf{b}^{(n)} \in \mathbf{R}^k$, $n = 1, 2, \cdots$ ($\mathbf{a}^{(n)} > \mathbf{0}$ means $a_i^{(n)} > 0$, $i = 1,\cdots, k$) such that $(\mathbf{Z}^{(n)} - \mathbf{b}^{(n)})/\ \mathbf{a}^{(n)}$ converges in distribution to a random vector $\mathbf{U}$ with nondegenerate distribution function H (i.e., all univariate marginals of H are nondegenerate), then F is said to be in the domain of attraction of H with the notation $F \in \mathbf{D}(H)$ and H is said to be a multivariate extreme value distribution. The convergence in distribution is equivalent to the condition

$$(1.1) \qquad \lim_{n\to\infty} F^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)}) = H(\mathbf{x}) \qquad \text{for all } \mathbf{x},$$

because multivariate extreme value distributions are continuous (see Theorem 1.1).

Note that if $(\mathbf{Z}^{(n)} - \mathbf{b}^{(n)})/\ \mathbf{a}^{(n)}$ converges in distribution to $\mathbf{U}$, then the ith component of $(\mathbf{Z}^{(n)} - \mathbf{b}^{(n)})/\ \mathbf{a}^{(n)}$ converges to the ith component of $\mathbf{U}$ and thus normalizing constants $\{a_i^{(n)}\}$, $\{b_i^{(n)}\}$ can be determined from well-known univariate considerations, $i = 1,\cdots, k$.

For $k > 1$, let G be the joint distribution of $(Y_1,Y_2, \cdots ,Y_k)$ then $G_{i_1 i_2\cdots i_m}$ denotes the joint distribution of $(Y_{i_1},Y_{i_2}, \cdots ,Y_{i_m})$, and $\bar{G}_{i_1 i_2\cdots i_m}$ denotes the survival function of $(Y_{i_1},Y_{i_2}, \cdots ,Y_{i_m})$, i.e.,

$$\bar{G}_{i_1 i_2\cdots i_m}(y_{i_1},\cdots,y_{i_m}) = P(Y_{i_1} > y_{i_1}, \cdots ,Y_{i_m} > y_{i_m}),$$

where $1 \leq i_1 < i_2 < \cdots < i_m \leq k$. Let

$$\mathbf{x}_G^0 = (x_{G_1}^0, \cdots ,x_{G_k}^0),$$

where $x_{G_i}^0$ is the endpoint of G_i, $i = 1, 2,\cdots, k$.

Now we summarize some multivariate results which are used in later sections.

We make extensive use of the following result (see Marshall and Olkin [14] and Theorem 5.3.1 of Galambos [7]).

LEMMA 1.1. *Equation* (1.1) *is equivalent to*

$$\lim_{n\to\infty} n\{1 - F(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)})\} = - \log H(\mathbf{x})$$

for all $\mathbf{x}$ *such that* $0 < H(\mathbf{x}) < 1$.

It is well-known that the weak convergence of distributions implies the weak convergence of any finite-dimensional marginal distribution (for example see Billingsley [1], p.30). Thus Lemma 1.1 holds for any marginal distribution of F.

Following results are seen in Galambos [7] (Chapter 5).

THEOREM 1.1. *Let H be an extreme value distribution. Then H is continuous, and its univariate marginals H_i (i = 1, 2,$\cdots$, k) belongs to one of the types Φ_α, Ψ_α and Λ.*

THEOREM 1.2. *Equation* (1.1) *holds if and only if for each fixed* $\{i_1, i_2, \cdots, i_m\}$ *and for all* $\mathbf{x}$ *such that* $H(\mathbf{x}) > 0$,

$$(1.2) \qquad \lim_{n\to\infty} n\bar{F}_{i_1 i_2 \cdots i_m}(a_{i_1}^{(n)} x_{i_1} + b_{i_1}^{(n)}, \cdots, a_{i_m}^{(n)} x_{i_m} + b_{i_m}^{(n)})$$

$$= h_{i_1 i_2 \cdots i_m}(x_{i_1}, \cdots, x_{i_m})$$

are finite and the function

$$(1.3) \qquad H(\mathbf{x}) = \exp\{ \sum_{m=1}^{k} (-1)^m \sum_{1 \le i_1 < \cdots < i_m \le k} h_{i_1 i_2 \cdots i_m}(x_{i_1}, \cdots, x_{i_m}) \}$$

is a nondegenerate distribution function. The actual limit distribution function of $F^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)})$ *is the one given in* (1.3).

COROLLARY 1.1. *Assume that* $(\mathbf{Z}^{(n)} - \mathbf{b}^{(n)})/\mathbf{a}^{(n)}$ *has a nondegenerate asymptotic distribution* $H(\mathbf{x})$. *Then the components of* $(\mathbf{Z}^{(n)} - \mathbf{b}^{(n)})/\mathbf{a}^{(n)}$ *are asymptotically independent if and only if the limits in* (1.2) *are identically zero for* $m = 2$.

The following results were proved by Marshall and Olkin [14].

PROPOSITION 1.1. *Let* H *be an extreme value distribution such that* $H_i = \Phi_{\alpha_i}$, $\alpha_i > 0$, $i = 1, \cdots, k$. *Then* $F \in \mathrm{D}(H)$ *if and only if*

$$\lim_{t\to\infty} \frac{1 - F(tx_1, \phi_2(t)x_2, \cdots, \phi_k(t)x_k)}{1 - F_1(t)} = -\log H(\mathbf{x})$$

for all $\mathbf{x}$ *such that* $H(\mathbf{x}) > 0$, *where* $\phi_i(t) = \bar{F}_i^{-1}\bar{F}_1(t)$, $i = 2, \cdots, k$.

PROPOSITION 1.2. *Let* H *be an extreme value distribution such that* $H_i = \Psi_{\alpha_i}$, $\alpha_i > 0$, $i = 1, \cdots, k$. *Then* $F \in \mathrm{D}(H)$ *if and only if* $\mathbf{x}_F^0 \in \mathbf{R}^k$ *and*

$$\lim_{t\downarrow 0} \frac{1 - F((tx_1, \phi_2(t)x_2, \cdots, \phi_k(t)x_k) + \mathbf{x}^0)}{1 - F_1(x_{F_1}^0 - t)} = -\log H(\mathbf{x})$$

for all $\mathbf{x}$ such that $H(\mathbf{x}) > 0$, where $\phi_i(t) = x^0_{F_i} - \bar{F}_i^{-1}(\bar{F}_1(x^0_{F_1} - t))$, $i = 2, \cdots, k$.

PROPOSITION 1.3. *Let H be an extreme value distribution such that $H_i = \Lambda$, $i = 1, \cdots, k$. Then $F \in \mathbf{D}(H)$ if and only if*

$$\lim_{t \uparrow x^0_{F_1}} \frac{1 - F(\mathbf{a}(t)\mathbf{x} + \mathbf{b}(t))}{1 - F_1(t)} = -\log H(\mathbf{x})$$

for all $\mathbf{x}$ such that $H(\mathbf{x}) > 0$, where $a_i(t) = \bar{F}_i^{-1}(\bar{F}_1(t)/e) - \bar{F}_i^{-1}\bar{F}_1(t)$, $b_i(t) = \bar{F}_i^{-1}\bar{F}_1(t)$, $i = 1, \cdots, k$.

2.2. Multivariate extreme value distributions

In this section we establish some properties of multivariate extreme value distributions.

THEOREM 2.1. *A nondegenerate k-dimensional distribution function H is an extreme value distribution if and only if for all* $s > 0$ *there exist vectors* $\mathbf{A}^{(s)} > \mathbf{0}$ *and* $\mathbf{B}^{(s)}$ *such that*

$$H^s(\mathbf{A}^{(s)}\mathbf{x} + \mathbf{B}^{(s)}) = H(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbf{R}^k. \tag{2.1}$$

PROOF. Sufficiency is obvious so that we shall prove necessity. If H is an extreme value distribution, then there exist a distribution function F and vectors $\mathbf{a}^{(n)}$ and $\mathbf{b}^{(n)}$ such that

$$\lim_{n\to\infty} F^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)}) = H(\mathbf{x}). \tag{2.2}$$

It follows from Lemma 1.1 that

$$\lim_{n\to\infty} n\{1 - F(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)})\} = -\log H(\mathbf{x}).$$

Hence for all $s > 0$

$$\lim_{n\to\infty} [ns]\{1 - F(\mathbf{a}^{([ns])}\mathbf{x} + \mathbf{b}^{([ns])})\} = -\log H(\mathbf{x}),$$

where $[ns]$ is the greatest integer less than or equal to ns. Then by Lemma 1.1

$$\lim_{n\to\infty} F^n(\mathbf{a}^{([ns])}\mathbf{x} + \mathbf{b}^{([ns])}) = H^{1/s}(\mathbf{x}). \tag{2.3}$$

Hence by (2.2), (2.3) and Lemma 0.2 in Chapter 0 (which can easily be extended to the multivariate case, see also the proof of Theorem 5.2.1 of Galambos [7]), there exist vectors $\mathbf{A}^{(s)} > \mathbf{0}$ and $\mathbf{B}^{(s)}$ such that

$$H^s(\mathbf{A}^{(s)}\mathbf{x} + \mathbf{B}^{(s)}) = H(\mathbf{x})$$

for all $\mathbf{x} \in \mathbf{R}^k$. Q.E.D.

COROLLARY 2.1. *Let H be an extreme value distribution. Then for any $t > 0$, H^t is an extreme value distribution.*

COROLLARY 2.2. *Let H be an extreme value distribution. Then for all $s > 0$, there exists vectors $\mathbf{A}^{(s)} > \mathbf{0}$ and $\mathbf{B}^{(s)}$ such that* (2.1) *holds, and if*

(a) $H_i = \Phi_{\alpha_i}$, $i = 1,\cdots, k$, *then* $\mathbf{A}^{(s)} = (s^{1/\alpha_1},\cdots,s^{1/\alpha_k})$ *and* $\mathbf{B}^{(s)} = \mathbf{0}$;

(b) $H_i = \Psi_{\alpha_i}$, $i = 1,\cdots, k$, *then* $\mathbf{A}^{(s)} = (s^{-1/\alpha_1},\cdots,s^{-1/\alpha_k})$ *and* $\mathbf{B}^{(s)} = \mathbf{0}$;

(c) $H_i = \Lambda$, $i=1,\cdots,k$, *then* $\mathbf{A}^{(s)} = \mathbf{1} = (1,\cdots,1)$ *and* $\mathbf{B}^{(s)} = (\log s,\cdots,\log s)$,

where $\alpha_i > 0$, $i = 1,\cdots, k$.

EXAMPLE. (See Galambos [7], p.254.) The distribution function

$$H(x_1,x_2,\cdots,x_k) = \exp\{-\exp[-\min(x_1,x_2,\cdots,x_k)]\}$$

is an extreme value distribution (with $H_i = \Lambda$, $i = 1,\cdots, k$), since for all $s > 0$

$$H^s(x_1 + \log s, \cdots , x_k + \log s) = H(x_1, \cdots ,x_k).$$

On the other hand, the distribution function

$$H(x_1,x_2) = \Lambda(x_1)\Lambda(x_2)[1 + (1-\Lambda(x_1))(1-\Lambda(x_2))/2]$$

is not an extreme value distribution, since

$$\begin{aligned} H^s(x_1+ \log s,x_2+ \log s) &= \Lambda(x_1)\Lambda(x_2)[1+(1-\Lambda^{1/s}(x_1))(1-\Lambda^{1/s}(x_2))/2]^s \\ &\neq H(x_1,x_2) \end{aligned}$$

COROLLARY 2.3. (Lemma 5.4.1 of Galambos [7].) *Let H be an extreme value distribution and denote by $D_H(\mathbf{y}) = H(H_1^{-1}(y_1),\cdots,H_k^{-1}(y_k))$, $\mathbf{y} \in (0,1)^k$, its dependence function. Then D_H satisfies for all $s > 0$*

$$D_H^s(\mathbf{y}^{1/s}) = D_H(\mathbf{y}).$$

LEMMA 2.1. *Let H be an extreme value distribution and D_H be the dependence function of H. If there exists a real number $c \in (0,1)$ such that*

(2.4) $$D_H(\mathbf{y}) = y_1 y_2 \cdots y_k \quad \textit{for all } \mathbf{y} \in (c,1)^k,$$

then

$$D_H(\mathbf{y}) = y_1 y_2 \cdots y_k \quad \textit{for all } \mathbf{y} \in (0,1)^k.$$

PROOF. For any $\mathbf{y} \in (0,1)^k$, there exists an $s > 0$ such that $\mathbf{y}^{1/s} \in (c,1)^k$. Hence by Corollary 2.3 and (2.4)

$$D_H(\mathbf{y}) = (D_H(\mathbf{y}^{1/s}))^s = (y_1^{1/s} y_2^{1/s} \cdots y_k^{1/s})^s$$
$$= y_1 y_2 \cdots y_k$$

for all $\mathbf{y} \in (0,1)^k$. Q.E.D.

LEMMA 2.2. *Let H be an extreme value distribution. Let $1 \le i_1 < i_2 < \cdots < i_m \le k$, $1 \le j_1 < j_2 < \cdots < j_{k-m} \le k$ and $\{i_1, i_2, \cdots, i_m\} \cap \{j_1, j_2, \cdots, j_{k-m}\} = \phi$. Then it holds that*

(2.5) $$H(x_1, x_2, \cdots, x_k) \ge H_{i_1 i_2 \cdots i_m}(x_{i_1}, x_{i_2}, \cdots, x_{i_m}) \times H_{j_1 j_2 \cdots j_{k-m}}(x_{j_1}, x_{j_2}, \cdots, x_{j_{k-m}})$$

for all $\mathbf{x} \in \mathbf{R}^k$.

PROOF. Since H is an extreme value distribution, there exist $\mathbf{a}^{(n)} > \mathbf{0}$ and $\mathbf{b}^{(n)}$ such that

$$H^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)}) = H(\mathbf{x})$$

(see Theorem 2.1). It is easily seen that

$$1 - H(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)})$$
$$\le \{ 1 - H_{i_1 i_2 \cdots i_m}(a_{i_1}^{(n)} x_{i_1} + b_{i_1}^{(n)}, \cdots, a_{i_m}^{(n)} x_{i_m} + b_{i_m}^{(n)}) \}$$
$$+ \{ 1 - H_{j_1 j_2 \cdots j_{k-m}}(a_{j_1}^{(n)} x_{j_1} + b_{j_1}^{(n)}, \cdots, a_{j_{k-m}}^{(n)} x_{j_{k-m}} + b_{j_{k-m}}^{(n)}) \} .$$

Thus, by Lemma 1.1 it holds that

$$- \log H(\mathbf{x}) \leq - \log H_{i_1 i_2 \cdots i_m}(x_{i_1}, \cdots, x_{i_m})$$

$$- \log H_{j_1 j_2 \cdots j_{k-m}}(x_{j_1}, \cdots, x_{j_{k-m}}).$$

Hence we have (2.5). Q.E.D.

COROLLARY 2.4. (Theorem 5.4.1 of Galambos [7]). *Let H be an extreme value distribution, then we have*

$$H(\mathbf{x}) \geq H_1(x_1)H_2(x_2)\cdots H_k(x_k)$$

for all $\mathbf{x} = (x_1, x_2, \cdots, x_k) \in \mathbf{R}^k$.

Now we have the following results which characterize the multivariate extreme value distributions.

THEOREM 2.2. *Let H be an extreme value distribution. Then it holds either*

$$H(\mathbf{x}) = H_1(x_1)H_2(x_2)\cdots H_k(x_k) \qquad \text{for all } \mathbf{x} \in \mathbf{R}^k$$

or

$$H(\mathbf{x}) > H_1(x_1)H_2(x_2)\cdots H_k(x_k) \qquad \text{for all } \mathbf{x}$$

such that $0 < H_i(x_i) < 1$, $i = 1, 2, \cdots, k$.

This theorem is trivial from the next one. We only prove the latter.

THEOREM 2.3. *Let H be an extreme value distribution. Then*

(2.6) $$H(\mathbf{x}) = H_1(x_1)H_2(x_2)\cdots H_k(x_k) \qquad \text{for all } \mathbf{x} \in \mathbf{R}^k$$

if and only if there exists $\mathbf{p} = (p_1, p_2, \cdots, p_k) \in \mathbf{R}^k$ *such that* $0 < H_i(p_i) < 1$, $i = 1, 2, \cdots, k$ *and*

(2.7) $$H(\mathbf{p}) = H_1(p_1)H_2(p_2)\cdots H_k(p_k).$$

PROOF. Necessity is obvious so that we shall prove sufficiency. First we show that

$$H_{ij}(p_i,p_j) = H_i(p_i)H_j(p_j) \qquad \text{for any } i < j. \tag{2.8}$$

By Corollary 2.4 it holds that

$$H_{ij}(p_i,p_j) \geq H_i(p_i)H_j(p_j) \qquad \text{for any } i < j.$$

Suppose (2.8) does not hold, i.e., if for example for $i = 1$ and $j = 2$

$$H_{12}(p_1,p_2) > H_1(p_1)H_2(p_2),$$

then by Lemma 2.2 and Corollary 2.4 we have

$$H(p_1,p_2,\cdots,p_k) \geq H_{12}(p_1,p_2)H_{34\cdots k}(p_3,\cdots,p_k)$$

$$> H_1(p_1)H_2(p_2)H_3(p_3) \cdots H_k(p_k).$$

This contradicts (2.7), thus we have (2.8). By Lemma 1.1 and (2.8) we have

$$\lim_{n\to\infty} n\{ 1 - H_{ij}(a_i^{(n)}p_i + b_i^{(n)},\ a_j^{(n)}p_j + b_j^{(n)}) \} \tag{2.9}$$

$$= -\log H_{ij}(p_i,p_j) = -\log H_i(p_i) - \log H_j(p_j),$$

where $a_l^{(n)}$ and $b_l^{(n)}$ are normalizing constants of H_l, $l = i, j$. It holds that

$$\bar{H}_{ij}(x_i,x_j) = \{1 - H_i(x_i)\} + \{1 - H_j(x_j)\}$$

$$- \{1 - H_{ij}(x_i,x_j)\}.$$

Thus, by Lemma 1.1 and (2.9), we have

$$\lim_{n\to\infty} n\bar{H}_{ij}(a_i^{(n)}p_i + b_i^{(n)},\ a_j^{(n)}p_j + b_j^{(n)}) = 0.$$

From the definition of $\bar{H}_{ij}$ and the inequalities $a_i^{(n)}, a_j^{(n)} > 0$ it holds that

$$\bar{H}_{ij}(a_i^{(n)}p_i + b_i^{(n)}, a_j^{(n)}p_j + b_j^{(n)}) \geq \bar{H}_{ij}(a_i^{(n)}x_i + b_i^{(n)}, a_j^{(n)}x_j + b_j^{(n)})$$

for all $(x_i,x_j) \geq (p_i,p_j)$. Thus we have

$$\lim_{n\to\infty} n\bar{H}_{ij}(a_i^{(n)}x_i + b_i^{(n)},\ a_j^{(n)}x_j + b_j^{(n)}) = 0$$

for any $i < j$. So, by Theorem 1.2 we have

$$H(x_1,x_2, \cdots ,x_k) = H_1(x_1)H_2(x_2)\cdots H_k(x_k) \quad \text{for all } \mathbf{x} \geq \mathbf{p}.$$

Let $c = \max_{1\leq i\leq k} H_i(p_i)$, then by Lemma 2.1, (2.6) holds for all $\mathbf{x} \in \mathbf{R}^k$. Q.E.D.

REMARK. The proof of this theorem is based on the idea of Corollary 1.1.

COROLLARY 2.5. (Theorems 2.2, 2.3 and 2.4 of Takahashi [27].) *Let H be an extreme value distribution.*

(a) *Suppose $H_i = \Phi_{\alpha_i}$, $\alpha_i > 0$, $i = 1, 2, \cdots , k$. Then*

$$H(\mathbf{x}) = \Phi_{\alpha_1}(x_1)\Phi_{\alpha_2}(x_2)\cdots\Phi_{\alpha_k}(x_k) \quad \textit{for all } \mathbf{x} \in \mathbf{R}^k$$

if and only if $H(\mathbf{1}) = \Phi_{\alpha_1}(1)\Phi_{\alpha_2}(1)\cdots\Phi_{\alpha_k}(1)$.

(b) *Suppose $H_i = \Psi_{\alpha_i}$, $\alpha_i > 0$, $i = 1, 2, \cdots , k$. Then*

$$H(\mathbf{x}) = \Psi_{\alpha_1}(x_1)\Psi_{\alpha_2}(x_2)\cdots\Psi_{\alpha_k}(x_k) \quad \textit{for all } \mathbf{x} \in \mathbf{R}^k$$

if and only if $H(\mathbf{-1}) = \Psi_{\alpha_1}(-1)\Psi_{\alpha_2}(-1)\cdots\Psi_{\alpha_k}(-1)$.

(c) *Suppose $H_i = \Lambda$, $i = 1, 2, \cdots , k$. Then*

$$H(\mathbf{x}) = \Lambda(x_1)\Lambda(x_2)\cdots\Lambda(x_k) \quad \textit{for all } \mathbf{x} \in \mathbf{R}^k$$

if and only if $H(\mathbf{0}) = \Lambda(0)^k$.

2.3. Multivariate tail equivalence

In this section, by using Propositions 1.1 - 1.3 we extend Theorem 1.7 in Section 1.1 to the multivariate case.

The following theorem is a k-dimensional version of Theorem 1.7 (a) in Section 1.1.

THEOREM 3.1. *Let F and G be k-dimensional distribution functions. Suppose for normalizing vectors $\mathbf{a}^{(n)} > \mathbf{0}$, $\mathbf{b}^{(n)}$, $n \geq 1$, $F^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)}) \to H(\mathbf{x})$, where $H_i = \Phi_{\alpha_i}$, $\alpha_i > 0$, $i = 1,\cdots, k$. Then*

$$G^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)}) \to H(\mathbf{Ax}+\mathbf{B}), \quad and \quad \mathbf{A} = (c^{1/\alpha_1},\cdots,c^{1/\alpha_k}),\ c > 0$$

if and only if $\mathbf{B} = \mathbf{0}$ *and*

$$\lim_{t\to\infty} \frac{1 - F(tx_1,\phi_2(t)x_2,\cdots,\phi_k(t)x_k)}{1 - G(tx_1,\phi_2(t)x_2,\cdots,\phi_k(t)x_k)} = c \tag{3.1}$$

for all $\mathbf{x} = (x_1,x_2,\cdots,x_k)$ such that $0 < H(\mathbf{x}) < 1$, where $\phi_i(t) = \overline{F}_i^{-1}\overline{F}_1(t)$, $i = 2,\cdots, k$.

PROOF. If $F \in \mathbf{D}(H)$, then it is known that we can take $\mathbf{b}^{(n)} = \mathbf{0}$, $a_1^{(n)} = \overline{F}_1^{-1}(1/n)$ and $a_i^{(n)} = \overline{F}_i^{-1}\overline{F}_1(a_1^{(n)}) = \phi_i(a_1^{(n)})$, $i = 2,\cdots, k$, $n \geq 1$ (see Appendix I).

Sufficiency. Since $a_1^{(n)} \to \infty$ as $n \to \infty$ and (3.1), we have that for all $\mathbf{x}$ such that $0 < H(\mathbf{x}) < 1$,

$$c = \lim_{n\to\infty} \frac{1 - F(a_1^{(n)}x_1,\phi_2(a_1^{(n)})x_2,\cdots,\phi_k(a_1^{(n)})x_k)}{1 - G(a_1^{(n)}x_1,\phi_2(a_1^{(n)})x_2,\cdots,\phi_k(a_1^{(n)})x_k)}$$

$$= \lim_{n\to\infty} \frac{n\{1 - F(\mathbf{a}^{(n)}\mathbf{x})\}}{n\{1 - G(\mathbf{a}^{(n)}\mathbf{x})\}} .$$

Then by Lemma 1.1 and Corollary 2.2 we have

$$G^n(\mathbf{a}^{(n)}\mathbf{x}) \to H^{1/c}(\mathbf{x}) = H(c^{1/\alpha_1}x_1,\cdots,c^{1/\alpha_k}x_k) = H(\mathbf{Ax}).$$

Necessity. From the univariate result (Theorem 1.7 (a) in Section 1.1), it holds that $B_i = 0$, $i = 1,\cdots,k$. So we have $\mathbf{B} = \mathbf{0}$. Since $a_1^{(n)} \leq a_1^{(n+1)} \to \infty$, for any sufficiently large t there exists an $n \in \mathbf{N}$ such that $a_1^{(n)} \leq t \leq a_1^{(n+1)}$. For any $\mathbf{x} = (x_1,\cdots,x_k)$ such that $0 < H(\mathbf{x}) < 1$, we have $\mathbf{0} < \mathbf{x} \neq \boldsymbol{\infty}$. Moreover, ϕ_i is non-decreasing, $i = 2,\cdots, k$. Therefore we have

$$\frac{1 - F(\mathbf{a}^{(n+1)}\mathbf{x})}{1 - G(\mathbf{a}^{(n)}\mathbf{x})} \leq \frac{1 - F(tx_1,\phi_2(t)x_2,\cdots,\phi_k(t)x_k)}{1 - G(tx_1,\phi_2(t)x_2,\cdots,\phi_k(t)x_k)}$$

$$\leq \frac{1 - F(\mathbf{a}^{(n)}\mathbf{x})}{1 - G(\mathbf{a}^{(n+1)}\mathbf{x})}$$

for all $\mathbf{x}$ such that $0 < H(\mathbf{x}) < 1$. Taking the limits of above inequalities, we have (3.1). Q.E.D.

If we consider the particular case $F_1 = \cdots = F_k$, than we have the following handy result.

COROLLARY 3.1. *Let F and G be k-dimensional distribution functions. Suppose $F_1 = \cdots = F_k$ and that there exist $a^{(n)} > 0$, $b^{(n)}$, $n \geq 1$ such that $F^n(a^{(n)}\mathbf{x} + b^{(n)}\mathbf{1}) \to H(\mathbf{x})$, where $H_i = \Phi_\alpha$, $i = 1,\cdots, k$, $\alpha > 0$. Then*

$$G^n(a^{(n)}\mathbf{x} + b^{(n)}\mathbf{1}) \to H(\mathbf{Ax+B}) \quad \text{and} \quad \mathbf{A} = c^{1/\alpha}\mathbf{1},\ c > 0$$

if and only if $\mathbf{B} = \mathbf{0}$ *and*

$$\lim_{t\to\infty} \frac{1 - F(t\mathbf{x})}{1 - G(t\mathbf{x})} = c \quad \textit{for all } \mathbf{x}$$

such that $0 < H(\mathbf{x}) < 1$.

Next we establish a k-dimensional version of Theorem 1.7 (b) in Section 1.1 which can be proved similarly to Theorem 3.1.

THEOREM 3.2. *Let F and G be k-dimensional distribution functions and $\mathbf{a}^{(n)} > \mathbf{0}$, $\mathbf{b}^{(n)}$, $n \geq 1$ are normalizing vectors such that $F^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)}) \to H(\mathbf{x})$, where $H_i = \Psi_{\alpha_i}$, $\alpha_i > 0$, $i = 1,\cdots, k$. Then*

$$G^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)}) \to H(\mathbf{Ax+B}) \quad \text{and} \quad \mathbf{A} = (c^{-1/\alpha_1},\cdots,c^{-1/\alpha_k}),\ c > 0$$

if and only if $\mathbf{B} = \mathbf{0}$, $\mathbf{x}_F^0 = \mathbf{x}_G^0 = \mathbf{x}^0 \in \mathbf{R}^k$ *and*

$$\lim_{t \downarrow 0} \frac{1 - F((tx_1,\phi_2(t)x_2,\cdots,\phi_k(t)x_k) + \mathbf{x}^0)}{1 - G((tx_1,\phi_2(t)x_2,\cdots,\phi_k(t)x_k) + \mathbf{x}^0)} = c$$

for all $\mathbf{x} = (x_1,\cdots,x_k)$ *such that* $0 < H(\mathbf{x}) < 1$, *where* $x_i^0 = x_{F_i}^0$, $i = 1,\cdots, k$. *and* $\phi_i(t) = x_i^0 - \bar{F}_i^{-1}(\bar{F}_1(x_1^0-t))$, $i = 2,\cdots, k$.

COROLLARY 3.2. *Let F and G be k-dimensional distribution functions. Suppose $F_1 = \cdots = F_k$ and that there exist $a^{(n)} > 0$, $b^{(n)}$, $n \geq 1$ such that $F^n(a^{(n)}\mathbf{x} + b^{(n)}\mathbf{1}) \to H(\mathbf{x})$, where $H_i = \Psi_\alpha$, $i = 1,\cdots, k$, $\alpha > 0$. Then*

$$G^n(a^{(n)}\mathbf{x} + b^{(n)}\mathbf{1}) \to H(\mathbf{Ax+B}) \quad \text{and} \quad \mathbf{A} = c^{-1/\alpha}\mathbf{1},\ c > 0$$

if and only if $\mathbf{B} = \mathbf{0}$, $x_{F_i}^0 = x_{G_i}^0 = x^0 \in \mathbf{R}$, $i = 1,\cdots, k$, *and*

$$\lim_{t \downarrow 0} \frac{1 - F(t\mathbf{x} + x^0\mathbf{1})}{1 - G(t\mathbf{x} + x^0\mathbf{1})} = c \quad \text{for all } \mathbf{x}$$

such that $0 < H(\mathbf{x}) < 1$.

Finally, we establish a k-dimensional version of Theorem 1.7 (c) in Section 1.1.

THEOREM 3.3. *Let F and G be k-dimensional distribution functions, and* $\mathbf{a}^{(n)} > \mathbf{0}$, $\mathbf{b}^{(n)}$, $n \geq 1$ *are normalizing vectors such that* $F^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)}) \to H(\mathbf{x})$, *where* $H_i = \Lambda$, $i = 1,\cdots, k$. *Then*

$$(3.2) \qquad G^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)}) \to H(\mathbf{Ax+B}) \quad \text{and} \quad \mathbf{A} > \mathbf{0},\ \mathbf{B} = b\mathbf{1}$$

if and only if $\mathbf{A} = \mathbf{1}$, $\mathbf{x}_F^0 = \mathbf{x}_G^0 = \mathbf{x}^0$ *and*

$$(3.3) \qquad \lim_{t \uparrow x_1^0} \frac{1 - F(\mathbf{a}(t)\mathbf{x} + \mathbf{b}(t))}{1 - G(\mathbf{a}(t)\mathbf{x} + \mathbf{b}(t))} = e^b \quad \text{for all } \mathbf{x}$$

such that $0 < H(\mathbf{x}) < 1$, *where* $x_1^0 = x_{F_1}^0$, $a_i(t) = \bar{F}_i^{-1}(\bar{F}_1(t)/e) - \bar{F}_i^{-1}\bar{F}_1(t)$ *and* $b_i(t) = \bar{F}_i^{-1}\bar{F}_1(t)$, $i = 1,\cdots, k$.

PROOF. If $F \in \mathbf{D}(H)$, then we can suppose without loss of generality that

$$b_i^{(n)} = \bar{F}_i^{-1}\bar{F}_1(b_1^{(n)})$$

and

$$a_i^{(n)} = \bar{F}_i^{-1}(\bar{F}_1(b_1^{(n)})/e) - \bar{F}_i^{-1}\bar{F}_1(b_1^{(n)}),\ i = 1,\cdots, k.$$

(See Appendix II.)

Sufficiency. Since $\lim_{n\to\infty} b_1^{(n)} = x_1^0$ and $\{b_1^{(n)}\}$ is an increasing sequence, by (3.3) we have

$$e^b = \lim_{n\to\infty} \frac{1 - F(\mathbf{a}(b_1^{(n)})\mathbf{x} + \mathbf{b}(b_1^{(n)}))}{1 - G(\mathbf{a}(b_1^{(n)})\mathbf{x} + \mathbf{b}(b_1^{(n)}))}$$

$$= \lim_{n\to\infty} \frac{n\{1 - F(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)})\}}{n\{1 - G(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)})\}} .$$

Hence by Lemma 1.1 and Corollary 2.2 we have

$$G^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)}) \to H^{1/e^b}(\mathbf{x}) = H(\mathbf{x}+\mathbf{B}), \quad \text{where } \mathbf{B} = b\mathbf{1}.$$

Necessity. Consider the marginal distributions F_i, G_i and $H_i = \Lambda$, $i = 1, \cdots, k$, then by Theorem 1.7 (c) in Section 1.1, we have $\mathbf{A} = \mathbf{1}$ and $\mathbf{x}_F^0 = \mathbf{x}_G^0 = \mathbf{x}^0$. Proposition 1.3 implies

$$\lim_{t \uparrow x_1^0} \frac{1 - F(\mathbf{a}(t)\mathbf{x} + \mathbf{b}(t))}{1 - F_1(t)} = -\log H(\mathbf{x}) \quad \text{for all } \mathbf{x} \text{ such that} \tag{3.4}$$

$0 < H(\mathbf{x}) < 1$. Now we shall prove that for all $\mathbf{x}$ such that $0 < H(\mathbf{x}) < 1$

$$\lim_{t \uparrow x_1^0} \frac{1 - G(\mathbf{a}(t)\mathbf{x} + \mathbf{b}(t))}{1 - F_1(t)} = -e^{-b}\log H(\mathbf{x}). \tag{3.5}$$

From (3.2) and $\mathbf{A} = \mathbf{1}$, we have $G^n(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)}) \to H^{e^{-b}}(\mathbf{x})$. So it holds

$$\lim_{s \to \infty} s\{1 - G(\boldsymbol{\alpha}(s)\mathbf{x} + \boldsymbol{\beta}(s))\} = -e^{-b}\log H(\mathbf{x}),$$

where $\alpha_i(s) = \bar{F}_i^{-1}(1/(es)) - \bar{F}_i^{-1}(1/s)$ and $\beta_i(s) = \bar{F}_i^{-1}(1/s)$, $i = 1, \cdots, k$. (This result can be proved similarly to Theorem 1.3 (e) in Section 1.1.) Now, let $s(t) = 1/(1 - F_1(t))$, then $\boldsymbol{\alpha}(s(t)) = \mathbf{a}(t)$ and $\boldsymbol{\beta}(s(t)) = \mathbf{b}(t)$, thus we have (3.5). The relations (3.4) and (3.5) imply (3.3). Q.E.D.

COROLLARY 3.3. *Let F and G be k-dimensional distribution functions. Suppose $F_1 = \cdots = F_k$ and that there exist $a^{(n)} > 0$, $b^{(n)}$, $n \geq 1$ such that $F^n(a^{(n)}\mathbf{x} + b^{(n)}\mathbf{1}) \to H(\mathbf{x})$, where $H_i = \Lambda$, $i = 1, \cdots, k$. Then*

$$G^n(a^{(n)}\mathbf{x} + b^{(n)}\mathbf{1}) \to H(\mathbf{Ax+B}) \quad \textit{and } \mathbf{A} > \mathbf{0},\ \mathbf{B} = b\mathbf{1}$$

if and only if $\mathbf{A} = \mathbf{1}$, $x_{F_i}^0 = x_{G_i}^0 = x^0$, $i = 1, \cdots, k$ *and*

$$\lim_{t \uparrow x^0} \frac{1 - F(a(t)\mathbf{x} + t\mathbf{1})}{1 - G(a(t)\mathbf{x} + t\mathbf{1})} = e^b \quad \textit{for all } \mathbf{x}$$

such that $0 < H(\mathbf{x}) < 1$, *where* $a(t) = \bar{F}_1^{-1}(\bar{F}_1(t)/e) - t$.

2.4. Examples

In this section, some examples concerning the multivariate tail equivalence are shown.

First we consider an application of the multivariate tail equivalence to the limit distribution of an extreme statistic from a bivariate exponential distribution.

Let (X_1,X_2) be two-dimensional random vector, where X_1 and X_2 are unit exponential variate. Let $F(x_1,x_2)$ be their bivariate distribution function. We put

$$\bar{F}(x_1,x_2) = P(X_1 > x_1, X_2 > x_2).$$

Then

$$F(x_1,x_2) = 1 - e^{-x_1} - e^{-x_2} + \bar{F}(x_1,x_2). \tag{4.1}$$

Versions of the following theorem have been obtaind by Galambos [7] and Marshall and Olkin [14]; it has a simple direct proof.

THEOREM 4.1. *Let F be a distribution function defined in* (4.1). *Then*

$$F \in \mathbf{D}(H) \quad \textit{and} \quad H(x_1,x_2) = \Lambda(x_1)\Lambda(x_2)$$

if and only if

$$\lim_{s\to\infty} s\bar{F}(x_1 + \log s,\ x_2 + \log s) = 0. \tag{4.2}$$

PROOF. Let

$$F_0(x_1,x_2) = 1 - e^{-x_1} - e^{-x_2} + e^{-x_1}e^{-x_2} = (1 - e^{-x_1})(1 - e^{-x_2}),$$

Then $F_0 \in \mathbf{D}(H)$. By Corollary 3.3, $F \in \mathbf{D}(H)$ if and only if

$$\lim_{t\to\infty} \frac{1 - F(x_1 + t,\ x_2 + t)}{1 - F_0(x_1 + t,\ x_2 + t)} = 1 . \tag{4.3}$$

Then it is easy to see that (4.3) holds if and only if (4.2) holds. Q.E.D.

EXAMPLE 4.1. (See Galambos [7], p.247.) The Morgenstern distribution

$$\bar{F}(x_1,x_2) = e^{-x_1-x_2}[1 + \alpha(1 - e^{-x_1})(1 - e^{-x_2})],$$

where $-1 \leq \alpha \leq 1$, satisfies (4.2). The Marshall-Olkin distribution

$$\bar{F}(x_1,x_2) = \exp[- x_1 - x_2 - \lambda \max(x_1,x_2)], \text{ where } \lambda > 0,$$

satisfies (4.2). On the other hand, the Mardia's distribution

$$\bar{F}(x_1,x_2) = (e^{x_1} + e^{x_2} - 1)^{-1}$$

does not satisfy (4.2). Therefore, the Morgenstern and the Marshall-Olkin distribution belongs to $\mathbf{D}(H)$, but the Mardia's distribution does not belong to $\mathbf{D}(H)$, where $H(x_1,x_2) = \Lambda(x_1)\Lambda(x_2)$ for all $(x_1,x_2) \in \mathbf{R}^2$.

Note that in general univariate tail equivalence does not imply the multivariate tail equivalence as the following counter-example shows.

EXAMPLE 4.2. Consider the following two distribution functions.

$$F(x_1,x_2) = H(x_1)H(x_2) \quad \text{and} \quad G(x_1,x_2) = H(\min(x_1,x_2)),$$

where H is an univariate extreme value distribution.

(a) If $H = \Phi_\alpha$, then

$$F^n(n^{1/\alpha}x_1,n^{1/\alpha}x_2) = F(x_1,x_2) \quad \text{and} \quad G^n(n^{1/\alpha}x_1,n^{1/\alpha}x_2) = G(x_1,x_2).$$

But it does not hold that

$$\lim_{t\to\infty} \frac{1 - F(x_1,x_2)}{1 - G(x_1,x_2)} = 1 \quad \text{for all } x_1,\ x_2 > 0.$$

(b) If $H = \Psi_\alpha$, then

$$F^n(n^{-1/\alpha}x_1,n^{-1/\alpha}x_2) = F(x_1,x_2) \text{ and } G^n(n^{-1/\alpha}x_1,n^{-1/\alpha}x_2) = G(x_1,x_2).$$

But it does not hold that

$$\lim_{t\downarrow 0} \frac{1 - F(tx_1,tx_2)}{1 - F(tx_1,tx_2)} = 1 \quad \text{for all } x_1,\ x_2 < 0 .$$

(c) If $H = \Lambda$, then

$$F^n(x_1 + \log n,\ x_2 + \log n) = F(x_1,x_2)$$

and

$$G^n(x_1 + \log n,\ x_2 + \log n) = G(x_1,x_2)\ .$$

But it does not hold that

$$\lim_{t\to\infty} \frac{1 - F(x_1 + t, x_2 + t)}{1 - G(x_1 + t, x_2 + t)} = 1 \quad \text{for all } (x_1,x_2) \in \mathbf{R}^2.$$

Appendix I

By Lemma 0.2 in Chapter 0 it is sufficient to prove

$$\text{(A.1)} \qquad \lim_{n\to\infty} \frac{\bar{F}_i^{-1}(1/n)}{\bar{F}_i^{-1}(\bar{F}_1(a_1^{(n)}))} = 1, \quad i = 2,\cdots, k.$$

The relation

$$\lim_{n\to\infty} \frac{n}{1/\bar{F}_1(a_1^{(n)})} = \lim_{n\to\infty} n\{1 - F_1(a_1^{(n)})\} = 1$$

holds. Hence by Theorem 1.1 in Section 1.1 and Theorem 0.1 (b), (c) in Chapter 0 we have (A.1).

Appendix II

It is sufficient to prove

$$\text{(A.2)} \qquad F_i^n(a_i^{(n)}x + b_i^{(n)}) \to \Lambda(x),$$

where $a_i^{(n)} = \bar{F}_i^{-1}(\bar{F}_1(b_1^{(n)})/e) - \bar{F}_i^{-1}\bar{F}_1(b_1^{(n)})$ and $b_i^{(n)} = \bar{F}_i^{-1}\bar{F}_1(b_1^{(n)})$, $i=2,\cdots,k$.

Since the relation

$$\lim_{n\to\infty} n\bar{F}_1(b_1^{(n)}) = 1$$

holds, we obtain

$$\lim_{n\to\infty} \{1 - \bar{F}_1(b_1^{(n)})\}^n = e^{-1},$$

$$\lim_{n\to\infty} \{1 - \bar{F}_1(b_1^{(n)})/e\}^n = e^{-e^{-1}}.$$

Hence by Theorem 0.2 in Chapter 0 we have (A.2).

CHAPTER 3

OUTLIER-PRONE AND OUTLIER-RESISTANT DISTRIBUTIONS

3.1. Introduction

In this chapter we consider an application of the distribution theory of order statistics to a problem of outliers.

Let S_n be a sample of size n of independent observations of random variables with common distribution function F. In the following we only consider distribution functions with endpoint at infinity. Let the variable values be denoted by $X_1, X_2, \cdots, X_n$ and let the ordered values be $X_{n1}, X_{n2}, \cdots, X_{nn}$. For a positive number k, Neyman and Scott [17] called X_{nn} a k-outlier if $X_{nn} - X_{n\,n-1} > k(X_{n\,n-1} - X_{n1})$. Let $P(k,n|F)$ denote the probability that a sample S_n of observations from a distribution F will contain a k-outlier. Let $\boldsymbol{F}$ be a family of distributions and let $\Pi(k,n|\boldsymbol{F})$ stand for the least upper bound of probabilities $P(k,n|F)$ for $F \in \boldsymbol{F}$. Then Neyman and Scott [17] offered the following definitions of outlier-proneness and outlier-resistance of a family of distributions.

DEFINITION 1.1. *If* $\Pi(k,n|\boldsymbol{F}) < 1$ *then we shall say that the family* $\boldsymbol{F}$ *is* (k,n)*-outlier-resistant. Otherwise, that is, if* $\Pi(k,n|\boldsymbol{F}) = 1$*, we shall say that the family* $\boldsymbol{F}$ *is* (k,n)*-outlier-prone.*

DEFINITION 1.2. *If a family of distributions* $\boldsymbol{F}$ *is* (k,n)*-outlier-prone for all* $k > 0$ *and all* $n > 2$*, we shall say that* $\boldsymbol{F}$ *is outlier-prone completely.*

REMARK 1.1. Green [9] showed that the family of distributions $\boldsymbol{F}$ is outlier-prone completely if and only if it is (k,n)-outlier-prone for some $k > 0$, $n > 2$.

Neyman and Scott [17] showed the following lemmas.

LEMMA 1.1. *If a family* $\boldsymbol{F}$ *is composed of distributions* $F(x/\sigma)$ *that differ only in the scale parameter* $\sigma > 0$, *then the probability* $P(k,n|\sigma) = P(k,n|1)$ *and for any* $k > 0$ *and* $n > 2$, *the family* $\boldsymbol{F}$ *is outlier-resistant.*

LEMMA 1.2. *If a family* $\boldsymbol{F}$ *is composed of distributions* $F(x - \xi)$ *that differ only in their location parameter,* ξ, *then* $P(k,n|\xi) = P(k,n|0)$ *and for any* $k > 0$ *and* $n > 2$, *the family* $\boldsymbol{F}$ *is outlier-resistant.*

Neyman and Scott [17] mentioned the following:

The two lemmas imply that the family $N(\xi,\sigma^2)$ of normal distributions (with mean ξ and variance σ^2) is outlier-resistant. Also, perhaps unexpectedly, the family $C(\xi,\sigma)$ of Cauchy distributions (centered at ξ and having scale parameter σ) is outlier-resistant.

Neyman and Scott [17] proved the following two theorems.

THEOREM 1.1. *The family of Gamma distributions is outlier-prone completely.*

THEOREM 1.2. *The family of lognormal distributions is outlier-prone completely.*

In Section 3.2 we show that the family of Weibull distributions is outlier-prone completely.

On the other hand, Green [10] considered the ideas of outlier-proneness and outlier-resistance of individual distributions, introducing the following definitions.

DEFINITION 1.3. *A distribution F will be said to be absolutely outlier-resistant (notation* $F \in \{\mathrm{AOR}\}$*) if, for all* $\varepsilon > 0$*, we have*

$$P(X_{nn} - X_{n\ n-1} > \varepsilon) \to 0 \quad \text{as } n \to \infty.$$

DEFINITION 1.4. *A distribution F will be said to be relatively outlier-resistant (notation* $F \in \{\mathrm{ROR}\}$*) if, for all* $k > 1$*, we have*

$$P(X_{nn} \,/\, X_{n\ n-1} > k) \to 0 \quad \text{as } n \to \infty.$$

DEFINITION 1.5. *A distribution F will be said to be absolutely outlier-prone (notation* $F \in \{\mathrm{AOP}\}$*) if there exist constants* $\varepsilon > 0$*,* $\delta > 0$ *and an integer* n_0 *such that*

$$P(X_{nn} - X_{n\ n-1} > \varepsilon) \geq \delta \qquad \text{for all integers} \quad n \geq n_0.$$

DEFINITION 1.6. *A distribution F will be said to be relatively outlier-prone (notation* $F \in \{\mathrm{ROP}\}$*) if there exist constants* $k > 1$*,* $\delta > 0$ *and an integer* n_0 *such that*

$$P(X_{nn} \,/\, X_{n\ n-1} > k) \geq \delta \qquad \text{for all integers} \quad n \geq n_0.$$

Green [10] showed the following theorem connecting the definitions of outlier-proneness and outlier-resistance to the classical laws of large numbers for maxima and relative stability for maxima given by Gnedenko [8].

THEOREM 1.3. *Let F be a distribution function, then*

(a) $F \in \{\mathrm{AOR}\}$ *iff* *for all* $\varepsilon > 0$

$$\lim_{x\to\infty} [1 - F(x + \varepsilon)] \,/\, [1 - F(x)] = 0;$$

(b) $F \in \{\mathrm{ROR}\}$ *iff* *for all* $k > 1$

$$\lim_{x\to\infty} [1 - F(kx)] \,/\, [1 - F(x)] = 0;$$

(c) $F \varepsilon$ {AOP} *iff* *there exist constants* $\alpha > 0$ *and* $\beta > 0$ *such that*

$$[1 - F(x + \beta)] \,/\, [1 - F(x)] \geq \alpha \qquad \text{for all finite } x \;;$$

(d) $F \varepsilon$ {ROP} *iff* *there exist constants* $k > 1$ *and* $\delta > 0$ *such that*

$$[1 - F(kx)] \,/\, [1 - F(x)] \geq \delta \qquad \text{for all finite } x.$$

Green [10] also showed the following theorem.

THEOREM 1.4. *Suppose the distribution function F has a density f.*

(a) *If*

(A) $$f(x + \varepsilon)/f(x) \to 0 \quad \text{as } x \to \infty \quad \text{for all } \varepsilon > 0,$$

then $F \varepsilon$ {AOR}.

If f has a monotone right tail and $F \varepsilon$ {AOR}, *then* (A) *holds.*

(b) *If*

(B) $$f(kx)/f(x) \to 0 \quad \text{as } x \to \infty \quad \text{for all } k > 1,$$

then $F \varepsilon$ {ROR}.

If f has a monotone right tail and $F \varepsilon$ {ROR}, *then* (B) *holds.*

(c) *If there exist constants* $\varepsilon > 0$, $\delta > 0$, x_0 *such that*

$$f(x + \varepsilon)/f(x) \geq \delta \quad \text{for all } x \geq x_0,$$

then $F \varepsilon$ {AOP}.

(d) *If there exist constants* $k > 1$, $\delta > 0$, x_0 *such that*

$$f(kx)/f(x) \geq \delta \quad \text{for all } x \geq x_0,$$

then $F \varepsilon$ {ROP}.

In Section 3.3 some properties of outlier-proneness and outlier-resistance are established. In Section 3.4 examples of outlier-prone and outlier-resistant distributions are given.

3.2. Outlier properties of the family of Weibull distributions

For $\sigma > 0$, $\alpha > 0$ and ξ, we say that a random variable X follows a Weibull distribution if its distribution function is given by

$$W(x) = \begin{cases} 1 - \exp\{-((x-\xi)/\sigma)^{\alpha}\}, & \text{for } x > \xi \\ 0, & \text{otherwise.} \end{cases}$$

Because of Lemmas 1.1 and 1.2, in investigating the outlier properties of the Weibull distributions, we may restrict ourselves to the case $\xi = 0$ and $\sigma = 1$.

The weibull distribution function, depending upon a shape parameter $\alpha > 0$, will be denoted by $W(x|\alpha)$. The corresponding density will be

$$w(x|\alpha) = \begin{cases} \alpha\, x^{\alpha-1}\exp(-x^{\alpha}), & \text{for } x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

THEOREM 2.1. *The family of Weibull distributions is outlier-prone completely.*

PROOF. The proof of this theorem is based on the same device as that of Theorem 1.1 (see Neyman and Scott [17] (pp. 416-417)).

Let $Q(k,n|\alpha)$ denote the probability that the largest of the n sample menbers will exceed the next largest by more than a factor k+1. Then the probability $P(k,n|\alpha)$ of a k-outlier in a sample of n will satisfy

$$1 > P(k,n|\alpha) > Q(k,n|\alpha),$$

where

$$Q(k,n|\alpha) = n\int_0^{\infty} W^{n-1}(x/(k+1)|\alpha)w(x|\alpha)dx.$$

For any $k > 0$ and any $n > 2$, it holds that

$$Q(k,n|\alpha) = n\int_0^\infty (1 - e^{-(x/(k+1))^\alpha})^{n-1} \alpha\, x^{\alpha-1} e^{-x^\alpha} dx$$

$$= n\int_0^\infty (1 - e^{-y/(k+1)^\alpha})^{n-1} e^{-y} dy$$

$$> n\int_0^T (1 - e^{-y/(k+1)^\alpha})^{n-1} e^{-y} dy \quad \text{for any } T > 0$$

$$\rightarrow \quad (1 - e^{-T})^n \quad \text{as } \alpha \rightarrow 0.$$

Thus we have that

$$P(k,n|\alpha) > Q(k,n|\alpha) \rightarrow 1 \quad \text{as } \alpha \rightarrow 0.$$

Q.E.D.

3.3. Properties of outlier-prone and outlier-resistant distributions

In this section we establish some properties of outlier-prone and outlier-resistant distributions defined by Green [10] (see Definitions 1.3-1.6).

Following results are seen in Green [10].

PROPOSITION 3.1. *Let F be a distribution function.*

(a) *If* $F \in \{AOR\}$, *then* $F \in \{ROR\}$.

(b) *If* $F \in \{ROP\}$, *then* $F \in \{AOP\}$.

(c) $\{ROR\} \cap \{ROP\} = \phi$.

(d) $\{AOR\} \cap \{AOP\} = \phi$.

PROPOSITION 3.2. *Let F be a distribution function with infinite endpoint. Then F belongs to one of the following disjoint classes:*

Class I $= \{AOR\}$, Class II $= \{ROR\} \cap \{AOR\}^c \cap \{AOP\}^c$,

Class III $= \{AOP\} \cap \{ROR\}$, Class IV $= \{AOP\} \cap \{ROR\}^c \cap \{ROP\}^c$,

Class V $= \{ROP\}$ Class VI $= \{ROR\}^c \cap \{AOP\}^c$.

Now we have the following theorem.

THEOREM 3.1. *Let F and G be distribution functions and the function $R : \mathbf{R}^+ \to \mathbf{R}^+$ defined by*

$$R(x) = [1 - F(x)] \,/\, [1 - G(x)]$$

is ρ-varying at infinity, where $\rho \in \mathbf{R}$ and $\mathbf{R}^+ = (0,\infty)$. Then

(a) $F \in \{AOR\}$ *iff* $G \in \{AOR\}$;

(b) $F \in \{ROR\}$ *iff* $G \in \{ROR\}$;

(c) $F \in \{AOP\}$ *iff* $G \in \{AOP\}$;

(d) $F \in \{ROP\}$ *iff* $G \in \{ROP\}$.

PROOF. By assumption of $R(x)$, we have

(3.1) $$\lim_{x\to\infty} R(x+\varepsilon)/R(x) = 1 \quad \text{for all } \varepsilon > 0,$$

and

(3.2) $$\lim_{x\to\infty} R(kx)/R(x) = k^{\rho} \quad \text{for all } k > 1.$$

(By Theorem 0.1 (a) in Chapter 0, it is easy to show (3.1).)

(a) Suppose $G \in \{\text{AOR}\}$. By Theorem 1.3 (a), for all $\varepsilon > 0$

$$[1 - G(x+\varepsilon)]/[1 - G(x)] \to 0 \quad \text{as } x \to \infty.$$

Thus, by (3.1), for all $\varepsilon > 0$

$$\frac{1 - F(x+\varepsilon)}{1 - F(x)} = \frac{R(x+\varepsilon)}{R(x)} \cdot \frac{1 - G(x+\varepsilon)}{1 - G(x)} \to 0 \quad \text{as } x \to \infty.$$

Therefore, by Theorem 1.3 (a) it holds that $F \in \{\text{AOR}\}$.

Similarly, we can prove the converse.

(b) The proof is analogous to that of part (a).

(c) Suppose $G \in \{\text{AOP}\}$. By Theorem 1.3 (c) there exist $\alpha > 0$ and $\beta > 0$ such that

$$[1 - G(x+\beta)]/[1 - G(x)] \geq \alpha \quad \text{for all finite } x.$$

On the other hand,

$$\frac{1 - F(x+\beta)}{1 - F(x)} = \frac{R(x+\beta)}{R(x)} \cdot \frac{1 - G(x+\beta)}{1 - G(x)} .$$

Thus, by (3.1) there exist $x_0 \in \mathbf{R}$ such that

$$[1 - F(x+\beta)]/[1 - F(x)] \geq \alpha/2 \quad \text{for all finite } x \geq x_0.$$

Furthermore

$$\frac{1 - F(x+\beta)}{1 - F(x)} \geq 1 - F(x+\beta) \geq 1 - F(x_0+\beta) > 0$$

for all $x < x_0$.

Let $\alpha' = \min\{\alpha/2,\ 1 - F(x_0+\beta)\} > 0$, then it holds that

$$[1 - F(x + \beta)] / [1 - F(x)] \geq \alpha' \qquad \text{for all finite } x.$$

Therefore, by Theorem 1.3 (c), it holds that $F \in \{\mathrm{AOP}\}$.

Similarly, we can prove the converse.

(d) The proof is analogous to that of part (c). Q.E.D.

COROLLARY 3.1. *Suppose F and G are distribution functions as in* Theorem 3.1, *then F and G belong to the same class defined in* Proposition 3.2.

REMARK 3.1. Let F and G be distribution functions such that $[1 - F(x)]/[1 - G(x)] \to c$ as $x \to \infty$, where $0 < c < \infty$. Then F and G belong to the same class.

Now let us consider the heaviness of the tails of distributions, then we have the following theorem.

THEOREM 3.2. *Let F and G be distribution functions.*

(a) *If* $F \in \{\mathrm{AOR}\}$ *and* $G \in \{\mathrm{AOP}\}$, *then* $\lim_{x\to\infty} [1 - F(x)]/[1 - G(x)] = 0$.

(b) *If* $F \in \{\mathrm{ROR}\}$ *and* $G \in \{\mathrm{ROP}\}$, *then* $\lim_{x\to\infty} [1 - F(x)]/[1 - G(x)] = 0$.

(c) *If* $F \in \{\mathrm{AOR}\}$ *and* $G \in \{\mathrm{ROP}\}$, *then* $\lim_{x\to\infty} [1 - F(x)]/[1 - G(x)] = 0$.

PROOF. (a) Suppose $F \in \{\mathrm{AOR}\}$ and $G \in \{\mathrm{AOP}\}$, then by Theorem 1.3 (a) and (c), it holds that

$$\lim_{x\to\infty} [1 - F(x + \varepsilon)]/[1 - F(x)] = 0 \quad \text{for all } \varepsilon > 0,$$

and there exist $\alpha > 0$ and $\beta > 0$ such that

$$[1 - G(x + \beta)]/[1 - G(x)] \geq \alpha \quad \text{for all finite } x.$$

Thus there exists $x_0 \in \mathbf{R}$ such that

$$[1 - F(x + \beta)]/[1 - F(x)] \leq \alpha/2 \qquad \text{for all } x \geq x_0.$$

For sufficiently large x, there exists an $n \in \mathbf{N}$ such that

$$x_0 + n\beta \leq x < x_0 + (n+1)\beta.$$

Hence we have

$$\frac{1 - F(x)}{1 - G(x)} \leq \frac{1 - F(x_0 + n\beta)}{1 - G(x_0+(n+1)\beta)} \leq \frac{(\alpha/2)^n(1 - F(x_0))}{\alpha^{n+1}\ (1 - G(x_0))}$$

$$= (1/2)^n \frac{1 - F(x_0)}{\alpha(1 - G(x_0))} \to 0 \quad \text{as } x \to \infty,$$

because $n \to \infty$ as $x \to \infty$.

(b) The proof is analogous to that of part (a).

(c) Using the relation {AOR}$\subset${ROR} (see Proposition 3.1 (a)), the proof is trivial from (b). Q.E.D.

REMARK 3.2. This theorem shows a connection between heaviness of tails and outlier properties.

3.4. RELATIONS BETWEEN THE OUTLIER PROPERTIES AND THE TYPES OF EXTREME VALUE DISTRIBUTIONS

In this section we show some relations between the outlier properties and the types of extreme value distributions.

THEOREM 4.1. *If* $F \in \mathbf{D}(\Phi_\alpha)$, *then* $F \in$ {ROP}, *i.e.*, *F belongs to the* class V.

PROOF. Suppose $F \in \mathbf{D}(\Phi_\alpha)$, then $1 - F(x)$ is $-\alpha$ -varying at infinity (see Theorem 1.1 in Section 1.1). Thus, for any $k > 1$

$$\lim_{x\to\infty} [1 - F(kx)]/[1 - F(x)] = k^{-\alpha} > 0.$$

Hence there exists $x_0 \in \mathbf{R}$ such that

$$[1 - F(kx)]/[1 - F(x)] > k^{-\alpha}/2 > 0 \qquad \text{for all finite } x \geq x_0.$$

Moreover

$$[1 - F(kx)]/[1 - F(x)] \geq 1 - F(kx)$$

$$\geq 1 - F(kx_0) > 0 \quad \text{for all } x < x_0.$$

Thus, let $\delta = \min\{k^{-\alpha}/2,\ 1 - F(kx_0)\} > 0$, then

$$[1 - F(kx)]/[1 - F(x)] \geq \delta \qquad \text{for all finite } x.$$

Hence, by Theorem 1.3 (d), we have $F \in$ {ROP}. Q.E.D.

EXAMPLE 4.1. For $\alpha > 0$ and $c > 0$, let F be a distribution function such that $1 - F(x) \sim cx^{-\alpha}$ for $x > 1$. Then, by Theorem 4.1 we have $F \in$ {ROP}.

THEOREM 4.2. *If* $F \in \mathbf{D}(\Lambda)$, *then* $F \in$ {ROR}.

PROOF. Suppose $F \in \mathbf{D}(\Lambda)$, then it is well-known that F is relative stable (Gnedenko [8]), i.e., F satisfies the condition (b) in Theorem 1.3. Thus we have the conclusion. Q.E.D.

EXAMPLE 4.2. For $a > 0$, $c > 0$, $\alpha > 0$ and $b \in \mathbf{R}$, let f be a density such that $f(x) \sim ax^b e^{-cx^\alpha}$ and the corresponding distribution function is F. Then by Theorem 3.8 in Section 1.3, it holds that $F \in \mathbf{D}(\Lambda)$. Thus we have $F \in$ {ROR}.

By Theorem 1.4 it is easy to see that F belongs to the class I if $\alpha > 1$, and to the class III if $\alpha \leq 1$.

CHAPTER 4

ORDER STATISTICS IN SEQUENTIAL LIFE TESTING AND NON-REGULAR ESTIMATION

4.1. Introduction

In this chapter we deal with the applications of the order statistics in life testing problem and non-regular estimation problem.

Padgett and wei [18] considered the following life testing problem where the test is terminated by time t_0. Let n independent items be put on test at the outset, and when an item fails, it is not replaced. Suppose that the distribution function of the life time of each item is $F_\theta(t)$ and the corresponding density is $f_\theta(t)$. Also, it is assumed that $F_\theta(t)$ is stochastically nondecreasing in θ, i.e., $F_{\theta_1}(t) \geq F_{\theta_2}(t)$, if $\theta_1 < \theta_2$, $t \geq 0$. The hypotheses which were considered are

$$(1.1) \qquad H_0: \theta \geq \theta_0 \quad \text{versus} \quad H_1: \theta \leq \theta_1$$

where $\theta_1 < \theta_0$.

Padgett and Wei [18] proposed and analyzed a one-sided sequential procedure for testing (1.1) based on the random function $X_n(t)$, which represents the number of failures among the n items before or at time $t \leq t_0$. They explained several advantages of using $X_n(t)$ as the test statistic. Their procedure allows a quick rejection of H_0 when H_1 is true.

In Section 4.2 we improve Lemma 1 and Theorem 1 of Padgett and Wei [18], and derive the average sampling time. In Section 4.3 we propose and analyze a sequential procedure which allows a quick acceptance of H_0 when H_0 is true. In Appendix we prove the lemmas in Section 4.3.

Weiss and Wolfowitz [31] developed the following estimation theory. For each positive integer n let $\mathbf{X}(n)$ denote the (finite) vector of random variables of which the estimator is to be a function. ($\mathbf{X}(n)$ need not have n components, nor need its components be independently or identically distributed.) Let $K_n(\mathbf{x}|\theta)$ be the density of $\mathbf{X}(n)$ at the point $\mathbf{x}$ with respect to a σ-finite measure μ_n when θ is the value of the parameter. The latter is a point of the known open set Θ. An estimator (of θ) is a Borel measurable function of $\mathbf{X}(n)$ with values in Θ. We assume, for the sake of simplicity, that Θ is a subset of the real line.

For each n let $k(n)$ be a normalizing factor for the family $K_n(\cdot|\cdot)$. Let R be a bounded, Borel measurable subset of the real line. A maximum probability estimator (m.p.e.) with respect to R is one which maximizes, with respect to d,

$$\int K_n(\mathbf{X}(n)|\theta)d\theta$$

the integral being over the set $\{d - R/k(n)\}$. (For simplicity we assume that there is a unique maximum.)

Let $h > 0$ be any number. We shall say that a sequence $\{\theta_n\}$ is in $H(h)$ (for a special point which below will always be θ_0), if $|k(n)(\theta_n - \theta_0)| \leq h$ for $n = 1, 2, \cdots$.

Weiss and Wolfowitz [31] proved the following theorem.

THEOREM 1.1. *Let M_n be an m.p.e. with respect to R such that: For any $h > 0$ and any sequence $\{\theta_n\}$ in $H(h)$ we have*

$$\lim_{n\to\infty} P_{\theta_n}\{k(n)(M_n - \theta_n) \in R\} = \alpha(\theta_0) \text{ , } \quad \textit{say.}$$

Let ε and δ be arbitrary but positive. For h sufficiently large we have, for any sequence $\{\theta_n\}$ in $H(h)$,

$$\varliminf_{n\to\infty} P_{\theta_n}\{|k(n)(M_n - \theta_n)| < \delta h\} \geq 1 - \varepsilon \text{ .}$$

Let T_n *be any (competing) estimator such that for any* $h > 0$ *and any sequence* $\{\theta_n\}$ *in* $H(h)$ *we have*

$$\lim_{n\to\infty} [P_{\theta_n}\{k(n)(T_n - \theta_n) \in R\} - P_{\theta_0}\{k(n)(T_n - \theta_0) \in R\}] = 0.$$

Then

$$\overline{\lim_{n\to\infty}}\, P_{\theta_0}\{k(n)(T_n - \theta_0) \in R\} \leq \alpha(\theta_0).$$

Weiss and Wolfowitz [31] also showed that a function of extreme order statistic is a m.p.e. in some non-regular case where distributions are concentrated on finite intervals.

In Section 4.4 in order to compare two sequences of m.p.e.'s more precisely, we introduce a notion of the second order efficiency which we apply to some non-regular estimation problems.

4.2. A sequential test in time truncated life testing, I

To test the hypotheses (1.1) Padgett and Wei [18] defined the following stopping rule: Let $T = \inf\{t \geq 0 : X_n(t) \geq at + b\}$ and given t_0, stop testing at time $\min(t_0, T)$ and reject H_0 if and only if $T \leq t_0$. We consider the same stopping rule. The following lemma is an extension of Lemma 1 of Padgett and Wei [18].

LEMMA 2.1. *Let n items be put on test at the outset. Also, let the distribution function of the life times of these n items be $F(t)$ and the corresponding density be $f(t)$. Then the probability that the path $X_n(t)$ lies entirely below the line $y = at + b$ ($a > 0$, $b \geq 0$) for all $t \leq u$, say $P(X_n(t) < at + b$, for all $t \leq u)$, is*

$$\sum_{l=0}^{n} (l!)\binom{n}{l} {}_h\phi_l(u)\cdot(\bar{F}(u))^{n-l}$$

where $\bar{F}(u) = 1 - F(u)$, ${}_h\phi_0(u) = 1$, $h = -b/a$, *for* $u > l\tau + h$

$${}_h\phi_l(u) = \sum_{i=1}^{l+1} {}_hC_i^{l+1} \cdot {}_h\rho_i^{l+1}(u), \text{ and } 0 \text{ otherwise},$$

$${}_h\rho_k^j(\cdot) = (F(\cdot) - F(k\tau + h))^{j-k}/(j-k)!, \quad 1 \leq k \leq j, \quad \tau = 1/a,$$

$${}_hC_r^j = {}_hC_r^{j-1}, \quad 1 \leq r \leq j-1,$$

$${}_hC_j^j = -\sum_{i=1}^{j-2} {}_hC_i^j \cdot {}_h\rho_i^j((j-1)\tau + h), \quad j \leq l+1,$$

$${}_hC_1^1 = 1, \text{ and the empty summation is zero.}$$

REMARK 2.1. If $b = 0$ then we have the same result of Lemma 1 of Padgett and Wei [18]. This lemma is more useful to prove Theorems 2.1 and 2.2 than Lemma 1 of [18] is.

REMARK 2.2. The proof is the same as that of Lemma 1 of Padgett and Wei [18] with $j\tau$ replaced by $j\tau + h$, so we omit it.

Using this result, we have the following theorem for the distribution of T.

THEOREM 2.1.

$$\begin{aligned} H_\theta(s) &= P_\theta(T \le s) = 1 - P_\theta(T > s) \\ &= 1 - P_\theta(X_n(t) < at + b, \textit{ for all } t \le s) \\ &= 1 - \sum_{l=0}^{n} \binom{n}{l} \cdot (l!)\, {}_h\phi_l(s)(\bar{F}_\theta(s))^{n-l} \end{aligned}$$

for $0 \le s$, *where*

$$\bar{F}_\theta(\cdot) = 1 - F_\theta(\cdot),$$

${}_h\phi_l(s)$ *is defined as* ${}_h\phi_l$ *in* Lemma 2.1 *except* $F(\cdot)$ *is replaced by* $F_\theta(\cdot)$.

REMARK 2.3. This expression for H_θ is simpler than that of Theorem 1 of Padgett and Wei [18]. Although both expressions are identical, it is difficult to derive directly the one from the other by simple manipulation.

As in Padgett and Wei [18], the derivative $h_\theta(s)$ of $H_\theta(s)$ at $s(\le t_0)$ exists except for those points t such that $at + b$ is integer. However the probability of having failures exactly at those points is zero. Therefore, we represent the average sampling time of the test procedure as

$$\int_0^{t_0} s h_\theta(s) ds + t_0 P_\theta(T > t_0) = \int_0^{t_0} s H_\theta\{ds\} + t_0[1 - H_\theta(t_0)].$$

Then we have the following theorem.

THEOREM 2.2. *The average sampling time of the test procedure is given by*

$$\int_0^{t_0} sH_\theta\{ds\} + t_0 P_\theta(T > t_0) = \int_0^{t_0} [1 - H_\theta(s)]ds$$

$$= \sum_{l=0}^{n} (l!)\binom{n}{l} \int_0^{t_0} {}_h\phi_l(s)(\bar{F}_\theta(s))^{n-l} ds .$$

PROOF. It is well-known that

$$\int_0^{t_0+} sH_\theta\{ds\} = -t_0[1 - H_\theta(t_0)] + \int_0^{t_0} [1 - H_\theta(s)]ds$$

(see for example Feller [4] (p.150)). Then, by using Theorem 2.1 we can obtain the result. Q.E.D.

4.3. A sequential test in time truncated life testing, II

To test the hypotheses (1.1) we define the following stopping rule: For $t_0 > 0$, let $T^* = \inf\{t: X_n(t) \geq n^* \text{ or } X_n(t) \leq at + b\}$, where $a > 0$, $b < 0$, n^* is a positive integer ($< n$) and $at_0 + b = n^* - 1$. Stop testing at time $\min(t_0, T^*)$, and reject H_0 if and only if $X_n(t) = n^*$. Following lemmas are needed to derive the distribution of T^*. Proofs of those lemmas are shown in Appendix.

LEMMA 3.1. *Let the distribution function of the life time be $F(t)$ and the corresponding density be $f(t)$. We put for $i = 1, 2, \cdots, n$,*

$$\phi_i = \int_0^h \int_{t_1}^{\tau+h} \cdots \int_{t_{i-2}}^{(i-2)\tau+h} \int_{t_{i-1}}^{(i-1)\tau+h} \prod_{k=1}^{i} f(t_k)dt_1 \cdots dt_i ,$$

where $\tau > 0$ and $h > 0$. Then we have

$$\phi_i = \sum_{k=0}^{i} C_k^i \rho_k^i(0) ,$$

where

$$\rho_k^j(\cdot) = \frac{\{F((i-1-k)\tau + h) - F(\cdot)\}^{j-k}}{(j-k)!} , \quad 0 \leq k \leq j \leq i,$$

$$\rho_k^k(\cdot) = 1,$$

$$C_k^j = C_k^{j-1} , \quad 0 \leq k \leq j-1,$$

$$C_j^j = -\sum_{k=0}^{j-2} C_k^j \rho_k^j((i-j)\tau + h),$$

$$C_0^0 = 1,$$

and the empty summation is zero.

LEMMA 3.2. *Let n independent items be put on test at the outset. Also let the distribution function of the life times of these n items be $F_\theta(t)$ and the corresponding density be $f_\theta(t)$. Then it holds that*

$$P_\theta(\ X_n(t) > at + b,\ \ for\ all\ t \le u\ )$$

$$= 1 - \sum_{i=0}^{[au+b]} (1 - F_\theta(i\tau+h))^{n-i}\ {}_nP_i \cdot \phi_i \qquad for\ u\ such\ that\ [au+b] \le n,$$

where $a > 0$, $b < 0$, $h = -b/a$, $\tau = 1/a$, ${}_nP_0 = 1$, $\phi_0 = 1$,

$${}_nP_k = n(n-1)\cdots(n-k+1),\ k = 1, 2, \cdots, n,$$

ϕ_i *is defined as* ϕ_i *in* Lemma 3.1 *except that* F *and* f *are replaced by* F_θ *and* f_θ, *respectively,* $[\cdot]$ *is the greatest integer function, and the empty summation is zero.*

LEMMA 3.3. *Under the same assumptions of* Lemma 3.2, *for* $u \le t_0$, *we have*

$$P_\theta(at + b < X_n(t) < n^*,\ \ for\ all\ t \le u\ )$$

$$= \sum_{j=l}^{n^*-1} \binom{n}{j} (\bar{F}_\theta(u))^{n-j} \{F_\theta(u)^j - \sum_{i=0}^{l-1} (F_\theta(u) - F_\theta(i\tau+h))^{j-i} \cdot {}_jP_i \cdot \phi_i\},$$

where $l = [au+b]+1$ *and* $\bar{F}_\theta(\cdot) = 1 - F_\theta(\cdot)$.

By using these lemmas, we have the following theorems.

THEOREM 3.1.

$$H^*_\theta(s) = P_\theta(T^* \le s) = 1 - P_\theta(T^* > s)$$

$$= 1 - \sum_{j=[as+b]+1}^{n^*-1} \binom{n}{j} (\bar{F}_\theta(s))^{n-j}$$

$$\times\{F_\theta(s)^j - \sum_{i=0}^{[as+b]} (F_\theta(s) - F_\theta(i\tau+h))^{j-i} \cdot {}_jP_i \cdot \phi_i\} \quad for\ s < t_0$$

and $\quad H^*_\theta(s) = 1 \ \ for\ s \ge t_0$.

PROOF. It is trivial from Lemma 3.3.

THEOREM 3.2. *The O.C. function, $L^*(\theta)$, of this sequential procedure is*

$$L^*(\theta) = 1 - P_\theta(\ X_n(t) > at + b,\ \ for\ all\ \ t \leq t_0\ )$$

$$= \sum_{i=0}^{n^*-1} (1 - F_\theta(i\tau+h))^{n-i} \cdot {}_nP_i \cdot \phi_i\ .$$

PROOF. By using Lemma 3.2 it is easy to show.

Note that the derivative $h_\theta^*(s)$ of $H_\theta^*(s)$ at $s(\leq t_0)$ exists except for those points t such that $at + b$ is an integer. However the probability of having failures exactly at those points is zero. Therefore the average sampling time of the test procedure is given by

$$\int_0^{t_0} sh_\theta^*(s)ds + t_0 P_\theta(T^* > t_0) = \int_0^{t_0} \{1 - H_\theta^*(s)\}ds\ .$$

See Section 4.2.

Then we have the following theorem.

THEOREM 3.3. *The average sampling time of the test procedure is*

$$\int_0^{t_0} \{1 - H_\theta^*(s)\}ds$$

$$= \int_0^{t_0} \sum_{j=[as+b]+1}^{n^*-1} \binom{n}{j} (\bar{F}_\theta(s))^{n-j} \{F_\theta(s)^j - \sum_{i=0}^{[as+b]} (F_\theta(s)$$

$$- F_\theta(i\tau+h))^{j-i} \cdot {}_jP_i \cdot \phi_i\ \}ds\ .$$

PROOF. By using Theorem 3.1 we can obtain the result.

4.4. Non-regular estimation problem

If $\{Z_n^{(1)}\}$ and $\{Z_n^{(2)}\}$ are two sequences of m.p.e.'s, then the asymptotic risks of them are equivalent. To evaluate the relative performance of them for finite observations, we define second order efficiency.

DEFINITION 4.1. *Suppose,* $\{Z_n^{(1)}\}$ *and* $\{Z_n^{(2)}\}$ *are sequences of m.p.e.'s (of* $\theta \in \Theta$, *a parameter space) with risks* $\{\beta_n^{(1)}(\theta)\}$ *and* $\{\beta_n^{(2)}(\theta)\}$, *respectively. If there exists*

$$SE(1,2;\theta) = \lim_{n\to\infty} SE_n(1,2;\theta) = \lim_{n\to\infty} \left(\frac{\beta_n^{(2)}(\theta)}{\beta_n^{(1)}(\theta)}\right)^n$$

$$\geq 1 \quad \text{for all } \theta \in \Theta,$$

and there exists at least one $\theta' \in \Theta$, *such that*

$$SE(1,2;\theta') > 1,$$

then we say $\{Z_n^{(1)}\}$ *is more efficient than* $\{Z_n^{(2)}\}$ *in the sense of second order.*

REMARK 4.1. If $SE(1,2;\theta) > 1$, then $\beta_n^{(1)}(\theta) < \beta_n^{(2)}(\theta)$ for sufficiently large n.

Now we consider a non-regular estimation problem as follows.

Suppose $X_1, X_2, \cdots, X_n$ are independent real-valued random variables with a common distribution function $F(x|\theta)$ whose density is $f(x|\theta)$, where θ is a real parameter. Suppose also that for any θ, $\theta < A(\theta)$ (a known function of θ), $A'(\theta) = dA(\theta)/d\theta$ is positive,

$$f(x|\theta) = 0, \qquad x < \theta \quad \text{or} \quad x > A(\theta)$$

$$f(\theta|\theta) = g(\theta) > 0, \qquad f(A(\theta)|\theta) = h(\theta) > 0$$

and there exist $g'_+(\theta) = f'_+(\theta|\theta)$, $h'_-(\theta) = f'_-(A(\theta)|\theta)$ and $A''(\theta)$. For detail see Weiss and Wolfowitz [31] (pp. 46-48). In this case, by Corollary 2.2 in Section 1.2 $n(\min X_i - \theta)$ and $n(\theta - A^{-1}(\max X_i))$ are asymptotically exponentially distributed.

THEOREM 4.1. *Suppose* r *is a positive constant and* $C(\theta) = g(\theta)-h(\theta)A'(\theta) = 0$. *Then* $Z_n^{(1)} = \min X_i - r/n$ *and* $Z_n^{(2)} = A^{-1}(\max X_i) + r/n$ *are m.p.e.'s with respect to* R *defined by*

$$R = (-r,\ r) \tag{4.1}$$

(*see* Weiss and Wolfowitz [31] (p.48)), *and we have*

$$SE(1,2;\theta) = \exp\{\ 2r^2[g_+'(\theta)+h_-'(\theta)A'^2(\theta)+h(\theta)A''(\theta)]\}. \tag{4.2}$$

PROOF. From Taylor expansion, we have

$$\int_\theta^t f(x|\theta)dx = g(\theta)(t-\theta) + \tfrac{1}{2}g_+'(\theta)(t-\theta)^2 + o(1)(t-\theta)^2$$

as $t \to \theta+0$. By

$$\beta_n^{(i)}(\theta) = P_\theta\{n(Z_n^{(i)} - \theta\ ) \notin (-r,\ r)\}\ ,\quad i = 1,\ 2,$$

we have

$$\begin{aligned}
\beta_n^{(1)}(\theta) &= P_\theta\{n(\min X_i - r/n - \theta) \notin (-r,\ r)\} \\
&= P_\theta\{\min X - \theta \geq 2r/n\} \\
&= P_\theta^n\{\ X - \theta \geq 2r/n\} \\
&= [1 - P_\theta\{\ X - \theta < 2r/n\}]^n \\
&= [1 - 2rg(\theta)/n - \{2r^2g_+'(\theta)+o(1)\}/\ n^2]^{\ n} \\
&= \{1 - 2rg(\theta)/n\}^n \left(1 - \frac{2r^2g_+'(\theta)+o(1)}{n^2(1-2rg(\theta)/n)}\right)^n.
\end{aligned}$$

Similarly, we have

$$\beta_n^{(2)}(\theta) = \left\{1 - \frac{2rh(\theta)A'(\theta)}{n}\right\}^n \left\{1 + \frac{2r^2[h'_-(\theta)A'^2(\theta)+h(\theta)A''(\theta)]}{n^2(1-2rh(\theta)A'(\theta)/n)}\right\}^n .$$

Taking the ratio of the last two expressions and noting that $C(\theta) = 0$, we obtain (4.2). Q.E.D.

COROLLARY 4.2. *Consider a simple case,*

$$f(x|\theta) = f(x-\theta),$$

$$f(x) \begin{cases} > 0 & \text{if } a \le x \le b \\ = 0 & \text{if } x < a \text{ or } x > b, \end{cases}$$

where a and b are known constants and we assume that there exist $f'_+(a)$ *and* $f'_-(b)$. *Suppose* $f(a) = f(b) = c > 0$, *then* $Z_n^{(1)} = \min X_i - a - r/n$ *and* $Z_n^{(2)} = \max X_i - b + r/n$ *are m.p.e.'s with respect to R defined in* (4.1), *and we have*

$$SE(1,2;\theta) = \exp\{2r^2(f'_+(a) + f'_-(b))\}$$

independently of θ.

Appendix

Proof of Lemma 3.1. For fixed i, define

$$\phi_i^j(t_{i-j}) = \int_{t_{i-j}}^{(i-j)\tau+h} \cdots \int_{t_{i-1}}^{(i-1)\tau+h} \prod_{k=i-j+1}^{i} f(t_k)dt_{i-j+1}\cdots dt_i \ ,$$

$$j = 1, \cdots , i,$$

where $t_0 = 0$. Then we have

$$\phi_i^j(t_{i-j}) = \sum_{k=0}^{j} C_k^j \rho_k^j(t_{i-j}). \tag{A.1}$$

The result of Lemma 3.1 is the special case of (A.1), i.e.,

$$\phi_i = \phi_i^i(0) = \sum_{k=0}^{i} C_k^i \rho_k^i(0).$$

The proof of (A.1) is same as that of Lemma 1 of Padgett and Wei [18] and is omitted. Q.E.D.

Proof of Lemma 3.2. Let $T_1 < T_2 < \cdots < T_n$ be the failure times. Define

$$Q(u) = P_\theta(\ X_n(t) > at + b, \text{ for all } t \leq u\ )$$

$$= P_\theta(\ T_i < \tau(i-1)+h,\ i = 1, 2, \cdots , l) = R_l,$$

where $l = [au+b]+1$. Then we have

$$R_l = \int_0^h \int_{t_1}^{\tau+h} \cdots \int_{t_{l-1}}^{(l-1)\tau+h} {}_nP_l(1-F_\theta(t_l))^{n-l} \prod_{i=1}^{l} f_\theta(t_i)dt_1\cdots dt_l$$

$$= R_{l-1} - (1 - F_\theta((l-1)\tau+h)^{n-(l-1)} \cdot {}_nP_{l-1}\cdot\phi_{l-1}.$$

Therefore

$$R_l = Q(u) = 1 - \sum_{i=0}^{l-1} (1-F_\theta(i\tau+h))^{n-i} \cdot {}_nP_i\cdot\phi_i.$$

Q.E.D.

Proof of Lemma 3.3. Suppose that there are exactly j ($\geq l=[au+b]+1$) failures before or at time u. Let $T_1 < T_2 < \cdots < T_j$ be the failure times, where $l \leq j \leq n^*-1$. Then, the conditional joint density of $(T_1, T_2, \cdots, T_j)$ is $j!\Pi_{i=1}^{j} f_\theta(t_i)/(F_\theta(u))^j$. For fixed j,

$$at + b < X_n(t) < n^*, \quad \text{for all } t \leq u$$

if and only if

$$T_i < \tau(i-1)+h, \ i = 1, \cdots, l, \quad T_{l+1} < \cdots < T_j < u,$$

and the probability of this event is

$$\tilde{\phi}_j(u) = \frac{j!}{F_\theta(u)^j} \int_0^h \int_{t_1}^{\tau+h} \cdots \int_{t_{l-1}}^{(l-1)\tau+h} \int_{t_l}^{u} \cdots \int_{t_{j-1}}^{u} \prod_{i=1}^{j} f_\theta(t_i)dt_1 \cdots dt_j$$

$$= \frac{{}_jP_l}{F_\theta(u)^j} \int_0^h \int_{t_1}^{\tau+h} \cdots \int_{t_{l-1}}^{(l-1)\tau+h} (F_\theta(u)-F_\theta(t_l))^{j-l} \cdot \prod_{i=1}^{l} f_\theta(t_i)dt_1 \cdots dt_l .$$

Define $S_l = F_\theta(u)^j \tilde{\phi}_j(u)$. Then we have

$$S_l = S_{l-1} - (F_\theta(u) - F_\theta((l-1)\tau+h))^{j-(l-1)} \cdot {}_jP_{l-1}\phi_{l-1},$$

so

$$S_l = F_\theta(u)^j - \sum_{i=0}^{l-1} (F_\theta(u) - F_\theta(i\tau+h))^{j-i} {}_jP_i \cdot \phi_i .$$

Therefore, the probability of $at+b < X_n(t) < n^*$, for all $t \leq u$ is give by

$$\sum_{j=l}^{n^*-1} \binom{n}{j} P_\theta(T_i < \tau(i-1)+h, \ i = 1,\cdots, l, \ T_{l+1} < \cdots < T_j < u \mid \text{exactly } j \text{ failures before or at time } u)$$

$$\times P_\theta(\text{exactly } j \text{ failures before or at time } u)$$

$$= \sum_{j=l}^{n^*-1} \binom{n}{j} (F_\theta(u))^{n-j} \{F_\theta(u)^j - \sum_{i=0}^{l-1} (F_\theta(u)-F_\theta(i\tau+h))^{j-i} \cdot {}_jP_i \cdot \phi_i\}.$$

Q.E.D.

ACKNOWLEDGEMENT

I wish to express my gratitude to Professor M. Okamoto, Osaka University, for his continuous guidance, valuable comments and encouragement. I am greatful to Professor N. Inagaki, Osaka University, for his useful suggestions and encouragement. I wish to thank members of Statistical Seminar Group in Osaka University, Dr. S. Shirahata and Dr. T. Isogai, Osaka University, and Dr. M. Yamaguchi, Fukui University, for their advice and encouragement. I also wish to thank members of Kansai Applied Statistical Research Group, Professor T. Nagai and Dr. M. Goto for their encouragement. Finally, I would like to thank Professor J. Galambos, Temple University, for suggesting the problem dealt with in Section 2.3.

AUTHOR'S ADDRESS

Department of Applied Mathematics,

Kobe University of Mercantile Marine,

Higashinada, Kobe,

JAPAN.

REFERENCES

[1] Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[2] David, H. A. (1970). *Order Statistics*. Wiley, New York.

[3] Epstein, B. (1954). Truncated life tests in the exponential case. *Ann. Math. Statist.* **25**, 555-564.

[4] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*. Vol.2, 2nd Ed. Wiley, New York.

[5] Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Camb. Phil. Soc.* **24**, 180-190.

[6] Galambos, J. (1975). Order statistics of samples from multivariate distributions. *J. Amer. Statist. Assoc.* **70**, 674-680.

[7] Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*. Wiley, New York.

[8] Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44**, 423-453.

[9] Green, R. F. (1974). A note on outlier-prone families of distributions. *Ann. Statist.* **2**, 1293-1295.

[10] Green, R. F. (1976). Outlier-prone and outlier-resistant distributions. *J. Amer. Statist. Assoc.* **71**, 502-505.

[11] Gumbel, E. J. (1958). *Statistics of Extremes*. Columbia University Press, New York.

[12] Haan, L. de (1970). *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*. Math. Centre Tracts **32**, Amsterdam.

[13] Marcus, M. B. and Pinsky, M. (1969). On the domain of attraction of $\exp(-e^{-x})$. *J. Math. Anal. Appl.* **28** , 440-449.

[14] Marshall, A. W. and Olkin, I. (1983). Domains of attraction of multivariate extreme value distributions. *Ann. Probab.* **11**, 168–177.

[15] Mises, R. von (1936). La distribution de la plus grande de *n* valeurs. Reprinted in *Selected Papers* II. Amer. Math. Soc., Providence, R. I., 1954, pp. 271–294.

[16] Nair, K. A. (1976). Bivariate extreme value distributions. *Comm. Statist.* **5**, 575–581.

[17] Neyman, J. and Scott, E. L. (1971). Outlier proneness of phenomena and of related distributions. In: Rustagi, J. (Ed.), *Optimizing Methods in Statistics*, 413–423. Academic Press, New York.

[18] Padgett, W. J. and Wei, L. J. (1982). A sequential test and interval estimation in time truncated life testing. *Sankhyā Ser. A* **44**, 242–250.

[19] Resnick, S. I. (1971). Tail equivalence and its applications. *J. Appl. Probab.* **8**, 136–156.

[20] Sibuya, M. (1960). Bivariate extreme statistics, I. *Ann. Inst. Statist. Math.* **11**, 195–210.

[21] Singpurwalla, N. D. (1972). Extreme values from a lognormal law with applications to air pollution problems. *Technometrics* **14**, 703–711.

[22] Smirnov, N. V. (1952). Limit distributions for the terms of a variational series. *Amer. Math. Soc. Transl. Ser.* **1**, No. 67. 1–64. Original published in 1949.

[23] Takahashi, R. (1978). Stabilizing constants and asymptotic behaviors of the extreme statistic. *J. Japan Statist. Soc.* **8**, 77–82.

[24] Takahashi, R. (1979). Remark on stabilizing constants of the extreme statistic. *J. Japan Statist. Soc.* **9**, 79–86.

[25] Takahashi, R. (1984). Note on the average sampling time for the sequential test of Padgett and Wei. *J. Japan Statist. Soc.* **14**, 89–91.

[26] Takahashi, R. (1986). A sequential life test using the number of failures. *Review of Kobe Univ. of Mercantile Marine* II, **34**, 93-95.

[27] Takahashi, R. (1986). Some properties of multivariate extreme value distributions and multivariate tail equivalence. To appear in *Ann. Inst. Statist. Math.*

[28] Takahashi, R. (1986). Normalizing constants of a distribution which belongs to the domain of attraction of double exponential distribution. Submitted.

[29] Takahashi, R. (1986). Note on outlier-prone and outlier-resistant distributions. Submitted.

[30] Takahashi, R. (1986). Some properties of multivariate extreme value distributions. Submitted.

[31] Weiss, L. and Wolfowitz, J. (1974). *Maximum Probability Estimators and Related Topics*. Lecture Notes in Math., Vol. 424, Splinger-Verlag.