Penultimate limiting forms in extreme value theory

Article *in* Annals of the Institute of Statistical Mathematics · December 1984

DOI: 10.1007/BF02481954

Maria Ivette Gomes
University of Lisbon

Ann. Inst. Statist. Math.
36 (1984), Part A, 71-85

PENULTIMATE LIMITING FORMS IN EXTREME VALUE THEORY

M. Ivette Gomes

(Received Sept. 7, 1982; revised Jan. 18, 1983)

Summary

Let $\{X_n\}_{n\geq 1}$ be a sequence of independent, identically distributed random variables. If the distribution function (d.f.) of $M_n = \max(X_1, \cdots, X_n)$, suitably normalized with attraction coefficients $\{a_n\}_{n\geq 1}$ $(a_n>0)$ and $\{b_n\}_{n\geq 1}$, converges to a non-degenerate d.f. $G(x)$, as $n\to\infty$, it is of interest to study the rate of convergence to that limit law and if the convergence is slow, to find other d.f.'s which better approximate the d.f. of $(M_n-b_n)/a_n$ than $G(x)$, for moderate n. We thus consider differences of the form $F^n(a_n x+b_n)-G(x)$, where $G(x)$ is a type I d.f. of largest values, i.e., $G(x)\equiv\Lambda(x)=\exp(-\exp(-x))$, and show that for a broad class of d.f.'s F in the domain of attraction of Λ, there is a penultimate form of approximation which is a type II $[\Phi_\alpha(x)=\exp(-x^{-\alpha}),\ x>0]$ or a type III $[\Psi_\alpha(x)=\exp(-(-x)^\alpha),\ x<0]$ d.f. of largest values, much closer to $F^n(a_n x+b_n)$ than the ultimate itself.

1. Introduction and preliminaries

Let $\{X_n\}_{n\geq 1}$ be a sequence of independent, identically distributed (i.i.d.) random variables (r.v.'s) with common distribution function (d.f.) $F(x)$. If the d.f. of $M_n = \max_{1\leq i\leq n} X_i$, suitably normalized with attraction coefficients $\{a_n\}_{n\geq 1}$ $(a_n>0)$ and $\{b_n\}_{n\geq 1}$, converges to a non-degenerate d.f. $G(x)$, i.e. $\lim_{n\to\infty} F^n(a_n x+b_n)=G(x)$ for all x in the set of continuity points of $G(x)$, we say that $F(\cdot)$ belongs to the domain of attraction of $G(\cdot)$ and denote this fact by $F\in\mathcal{D}(G)$. The only d.f.'s with non-empty domains of attraction are of one of the following three types (Gnedenko [7]):

$$\text{Type I}\ :\ \Lambda(x)=\exp(-\exp(-x))\,, \qquad x\in R$$

AMS 1980 subject classifications: Primary 62E20; Secondary 60F05, 65U05.
Key words and phrases: Extreme value theory, penultimate approximations, rates of convergence, order statistics.

$$\text{Type II}: \quad \Phi_\alpha(x)=\begin{cases} 0 & \text{if } x<0 \\ \exp(-x^{-\alpha}) & \text{if } x\geqq 0 \ (\alpha>0) \end{cases}$$

$$\text{Type III}: \quad \Psi_\alpha(x)=\begin{cases} \exp(-(-x)^{\alpha}) & \text{if } x<0 \\ 1 & \text{if } x\geqq 0 \ (\alpha>0)\,. \end{cases}$$

It is then of practical interest to study the rate of convergence to that limit law $G(\cdot)$ and if convergence is slow, to find other d.f.'s which better approximate the d.f. of $(M_n-b_n)/a_n$ for moderate n, than $G(x)$—the so-called penultimate behaviour of $F^n(\cdot)$ first pointed out by Fisher and Tippett [5]. We shall here deal with these two problems when $G(\cdot)$ is a type I d.f. of largest values.

Related work appears in Uzgören [16], Dronkers [4] and Haldane and Jayakar [11] with no indication of the closeness of the approximations provided. Anderson [1], [2] considers differences of the form $F^n(a_nx+b_n)-G(x)$, $F\in\mathcal{D}(G)$ and finds that under certain conditions on $F(x)$, $F^n(a_nx+b_n)-G(x)=\gamma(x)d(n)g(x)+o(d(n))$ uniformly over finite intervals of x, where both $\gamma(\cdot)$ and $d(\cdot)$ depend on the parent F and $g(x)=G'(x)$. Galambos ([6], pp. 111–116) provides exact estimates, in terms of n and given through inequalities for $|F^n(a_nx+b_n)-G(x)|$. Such estimates depend mainly on the speed of convergence to 0 of $n(1-F(a_nx+b_n))+\log G(x)$, as $n\to\infty$ and hold uniformly on finite intervals of the real line. Hall [12], [13] deals with the rate of convergence of normal extremes towards $\Lambda(x)$ and Reiss [15] deals with asymptotic expansions of the distribution of extreme order statistics, providing error bounds for the approximations derived. None of these authors refers again to the penultimate form of approximation of n r.v.'s with d.f. $F\in\mathcal{D}(\Lambda)$ as being a type II or a type III d.f. of maxima. Gomes [10] considers such a case and Cohen [3] develops a systematic study of a penultimate form of approximation to normal extremes expressed in terms of type III and type II extreme value d.f.'s, together with error bounds for such approximations.

In Section 2 of this paper we shall consider Fisher and Tippett's results concerning the penultimate behaviour of the maximum of normal r.v.'s and generalizing their ideas, we shall consider a broad class $\mathcal{L}$ of d.f.'s attracted to $\Lambda(x)$ and show that there is numerical evidence that for d.f.'s $C(\cdot)$ in that class, $C^n(\cdot)$ is nearer either to an appropriate type III or type II penultimate form than to the ultimate type I limit d.f. according as $d\log(a_n)/d\log(n)<0$ or $d\log(a_n)/d\log(n)>0$, a_n the scale attraction coefficient.

Since the numerical results obtained call for an explanation that, we believe, lies much deeper than the one put forward by Fisher and Tippett, in Section 3 of this paper we shall prove that for a d.f. $F\in\mathcal{D}(\Lambda)$

for which von-Mises' condition together with some mild regularity conditions on $k(x)=-d\log(-\log F(x))/dx$ hold, $F^n(a_nx+b_n)$ is closer to a type III (to a type II) d.f. when $k'(x)>0$ for $x\geqq x_1$ (when $k'(x)<0$ for $x\geqq x_2$) than to the ultimate limiting form $\Lambda(x)$.

In Section 4 of this paper we study the influence of the attraction coefficients upon the rate of convergence and in Section 5 we generalize the results obtained in Section 3, to obtain the penultimate behaviour of the joint d.f. of the i largest order statistics (o.s.), suitably normalized.

Finally, in Section 6, we study the d.f.'s belonging to the class $\mathcal{L}$ previously referred to under the general theory. For this particular class $\mathcal{L}$ the error bounds are valid on the whole real line.

2. Penultimate behaviour of the maximum of i.i.d. r.v.'s attracted to Λ and with d.f. in a class $\mathcal{L}$

Fisher and Tippett, in [5], show that the normal d.f. $\Phi\in\mathcal{D}(\Lambda)$, and consequently the d.f. of the maximum of n independent, normally distributed r.v.'s should be approximated, when n is large, by $\Lambda(x;b_n,a_n)=\Lambda((x-b_n)/a_n)$, where $\sqrt{2\pi}\,b_n\exp(b_n^2/2)=n$ and $a_n=b_n/(b_n^2+1)$ (since $\lim_{x\to\infty}(1-\Phi(x))/\Phi'(x)=1$, this choice of attraction coefficients is asymptotically equivalent to the one suggested by Gnedenko [7], $1-\Phi(b_n^*)=1/\mathrm{n}$, $a_n^*=1/(n\Phi'(b_n^*))$).

They also remark that the convergence of $\Phi^n(a_nx+b_n)$ to $\Lambda(x)$ is extremely slow and they conclude their paper by showing that $\Phi^n(x)$ is "closer" (skewness and kurtosis coefficients are used as an indicator of closeness) to a suitable type III $\Psi_{\alpha_n}(x;\lambda_n,\delta_n)=\Psi_{\alpha_n}((x-\lambda_n)/\delta_n)$, than to $\Lambda(x;b_n,a_n)$, even for $n=10^{12}$.

The justification given by Fisher and Tippett is then essentially the following: it is obvious that any extreme value d.f. belongs to its own domain of attraction, and

(i) if $\{X_j\}_{j\geq 1}$ are i.i.d. r.v.'s with d.f. $\Lambda(x)$, then $M_n\overset{d}{=}X_j+\log(n)$ (there is a change in location only),

(ii) if $\{X_j\}_{j\geq 1}$ are i.i.d. r.v.'s with d.f. $\Phi_\alpha(x)$, then $M_n\overset{d}{=}n^{1/\alpha}X_j$ (scale is $n^{1/\alpha}\to\infty$, as $n\to\infty$),

(iii) if $\{X_j\}_{j\geq 1}$ are i.i.d. r.v.'s with d.f. $\Psi_\alpha(x)$, then $M_n\overset{d}{=}n^{-1/\alpha}X_j$ (scale is $n^{-1/\alpha}\to 0$, as $n\to\infty$).

As in the approach $\Lambda(x;b_n,a_n)$ suggested by Fisher and Tippett (or any ultimate approach $\Lambda(x;b_n^*,a_n^*)$, $\{a_n^*,b_n^*\}_{n\geq 1}$ asymptotically equivalent to the first choice of attraction coefficients), the scale a_n goes to zero, as $n\to\infty$, as happens to the scale $n^{-1/\alpha}$ of the maximum of n i.i.d. type III r.v.'s with index α, they suggest that one should equate the rates at which the scale decreases in both cases, i.e., put $-1/\alpha_n=d\log(a_n)/$

$d \log(n)$, from which follows $\alpha_n=(b_n^2+1)^2/(b_n^2-1)$, and use $\Psi_{\alpha_n}(x;\lambda_n,\delta_n)$, where $\delta_n=a_n\alpha_n$, $\lambda_n=b_n+\delta_n$, are chosen so that $\lim_{n\to\infty}[\Lambda(x;b_n,a_n)-\Psi_{\alpha_n}(x;\lambda_n,\delta_n)]=0$, as a penultimate form of approximation for $\Phi^n(x)$—bearing in mind, of course, that the ultimate limiting form is a type I d.f. of largest values.

The results of Fisher and Tippett suggest, at first glance, the following conjecture: Let $F\in\mathcal{D}(\Lambda)$ be such that $F^n(x)$ is approximated by $\Lambda(x;b_n,a_n)$ ($\{a_n,b_n\}_{n\geq 1}$, possible attraction coefficients of F to the law Λ).

(a) if $a_n\to 0$, as $n\to\infty$, then there is a type III d.f. of largest values, which provides a better approximation to $F^n(x)$ than $\Lambda(x;b_n,a_n)$,

(b) if $a_n\to\infty$, as $n\to\infty$, then there is a type II d.f. of largest values, which provides a better approximation to $F^n(x)$ than $\Lambda(x;b_n,a_n)$,

(c) if $a_n\to c$, $0<c<\infty$, $\Lambda(x;b_n,a_n)$ is closer to $F^n(x)$ than any type II or type III d.f. of largest values.

Later on, we shall turn back to this conjecture. Now and in what concerns the normal d.f., we present the Kolmogorov-Smirnov distances between $\Phi^n(x)$, the d.f. of the maximum of n i.i.d. normal r.v.'s and the penultimate and ultimate forms of approximation. Here, the attraction coefficients are the ones suggested by Gnedenko, $1-\Phi(b_n)=1/n$, $a_n=1/(n\Phi'(b_n))$, the parameters in $\Psi_{\alpha_n}(x;\lambda_n,\delta_n)$ being $\alpha_n=1/(1-a_nb_n)$, $\delta_n=a_n\alpha_n$, $\lambda_n=b_n+\delta_n$. $k_{0,n}$ denotes the Kolmogorov-Smirnov distance between $\Phi^n(x)$ and $\Lambda(x;b_n,a_n)$, $k_{1,n}$ denotes the Kolmogorov-Smirnov distance between $\Phi^n(x)$ and $\Psi_{\alpha_n}(x;\lambda_n,\delta_n)$ and $P_n=\{(k_{0,n}-k_{1,n})/\max(k_{0,n},k_{1,n})\}\cdot 100\%$.

Table 1. Comparison of $\Phi^n(x)$ with suitable ultimate and penultimate forms of approximation

n	10	10^2	10^3	10^4	10^5	10^6
$k_{0,n}$	0.0522	0.0272	0.0183	0.0137	0.0109	0.0091
$k_{1,n}$	0.0264	0.0052	0.0024	0.0013	0.0009	0.0006
P_n	49.30	80.76	87.07	90.27	92.17	93.47

The numerical results obtained in Gomes [8] for other possible choices of attraction coefficients show that in any circumstances we achieve identical results. It could however be argued that we are fitting the distributions with parameters whose validity is asymptotic and that this fact accounts for the unexpected result. However if we fit the distributions with parameters determined, for instance by a quantile method, i.e., if we put $\Phi^n(b_n)=\Psi_{\alpha_n}(b_n;\lambda_n,\delta_n)=\Lambda(b_n;b_n,a_n)=\exp(-1)$, $\Phi^n(a_n+b_n)=\Psi_{\alpha_n}(a_n+b_n;\lambda_n,\delta_n)=\Lambda(a_n+b_n;b_n,a_n)=\exp(-\exp(-1))$, the parameter α_n being determined as before by $(-d\log(a_n)/d\log(n))^{-1}$, P_n is an increasing function of n, and $P_{10}=85.74$, $P_{10^6}=95.26$. If we choose $\{a_n, b_n\}_{n\geq 1}$ in order to minimize $k_{0,n}$ and $\{\lambda_n, \delta_n, \alpha_n\}_{n\geq 1}$ in order to minimize $k_{1,n}$, we reach exactly the same conclusions, and P_n varies monotonically from 87.75 to 97.25 as n varies from 10 to 10^6.

We shall now consider a class $\mathcal{L}$ of d.f.'s attracted to Λ. Before defining such a class, we state without proof the following result— Let F be a d.f. with right endpoint $x_0^F=\sup\{x: F(x)<1\}$ at infinity. Then for some real constants p and q, β and α $(q, \alpha>0)$, $\lim_{x\to\infty}(1-F(x))/((\alpha x+\beta)^p\exp(-(\alpha x+\beta)^q))=A$, $0<A<\infty$, if and only if $\lim_{n\to\infty}F^n(a_nx+b_n)=\Lambda(x)$, with $\{a_n, b_n\}_{n\geq 1}$ given by $a_n=\{q\alpha(\log n)^{(q-1)/q}\}^{-1}$, $b_n=(\log n)^{1/q}/\alpha+(\log A+p\log\log n/q)/(q\alpha(\log n)^{(q-1)/q})-\beta/\alpha$.

It is then natural to put:

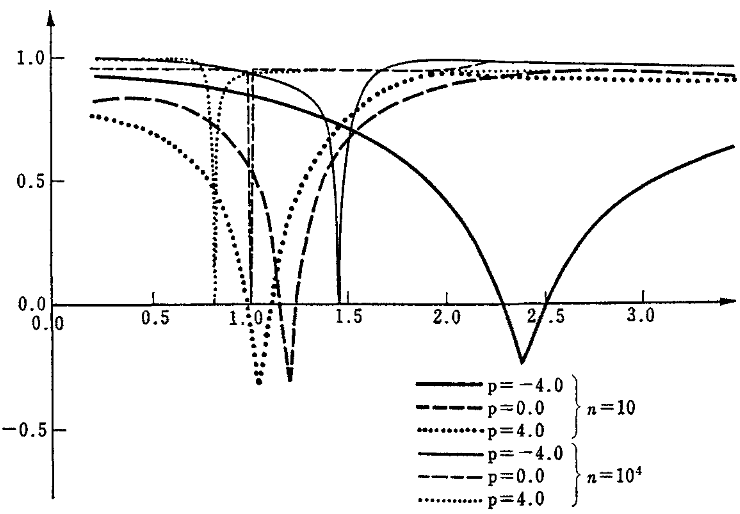
DEFINITION 1. $\mathcal{L}$ is the class of the d.f.'s of the form $\{1-A(\alpha x+\beta)^p\exp(-(\alpha x+\beta)^q)(1+\varepsilon(x))\}I_{[y_0,\infty)}$, where y_0 is the greatest real solution of $A(\alpha x+\beta)^p\exp(-(\alpha x+\beta)^q)(1+\varepsilon(x))=1$, A chosen in such a way that such a solution does exist, β and p real, q and α positive real, $\varepsilon(x)\to 0$, as $x\to\infty$.

We shall often use the notation $C(x; \alpha, \beta; A; p, q; \varepsilon(x))$, which is self explanatory, to denote a member of the class $\mathcal{L}$. The normal d.f. is obviously a member of $\mathcal{L}$: $\Phi(x)\equiv C(x; 1/\sqrt{2}, 0; 1/2\sqrt{\pi}; -1, 2; \varepsilon(x))$, with $\varepsilon(x)=\sum_{n\geq 1}(-1)^n 1\cdots(2n-1)/x^{2n}$.

It is then possible to show that there is numerical evidence that for several d.f.'s in the class $\mathcal{L}$, the d.f. $C^n(x)$ is nearer to an appropriate type III (or type II) penultimate form, than to the ultimate type I d.f.

Note that $a_n\to 0$ iff $q>1$, iff there exists n_0 such that $d\log(a_n)/d\log(n)<0$ for $n\geq n_0$, and $a_n\to\infty$ iff $q<1$, iff there exists n_1 such that $d\log(a_n)/d\log(n)>0$ for $n\geq n_1$. Consequently, according to the conjecture previously put forward, it seems sensible that if $q>1$ there is a penultimate type III approximation, if $q<1$ there is a penultimate type II approximation and if $q=1$ there is no penultimate at all, i.e., Λ itself is approached rapidly. But if we have a closer look at it, we remark that, following Fisher and Tippett, α_n should be taken equal to $|d\log(a_n)/d\log(n)|^{-1}$. Moreover, as $n\to\infty$, $d\log(a_n)/d\log(n)<0$ if $q>1$ and $d\log(a_n)/d\log(n)>0$ if $q<1$; but there are obviously finite values $n_0(q)$ and $n_1(q)$ such that $d\log(a_n)/d\log(n)>0$ for $n\leq n_0(q)$ though $q>1$, and $d\log(a_n)/d\log(n)<0$ for $n\leq n_1(q)$ though $q<1$. So, it seems more sensible to conjecture that if $d\log(a_n)/d\log(n)<0$ there is a type III penultimate approximation and that if $d\log(a_n)/d\log(n)>0$ there is a type II penultimate form of approximation. The numerical results obtained support this conjecture.

For simplicity, we have restricted ourselves to the elements of $\mathcal{L}$ of the form $C(x;1,0;3;p,q;0)$, i.e., to d.f.'s of the type $F_{p,q}(x)=\{1-3x^p\exp(-x^q)\}I_{[y_0,\infty)}$, where y_0, the greatest real solution of $3x^p\exp(-x^q)=1$, does always exist because we have chosen $A=3>\exp(1)$, and we shall present graphically the results obtained for $p=-4, 4(4)$, $q=0.2, 4(0.1)$. In Graph 1 (a continuous version of the discretized graph really obtained for $q=0.2, 4(0.1)$), we plot for values of $p=-4, 0, 4$, q against $R=(k_0-k_1)/\max(k_0, k_1)$, for $n=10$ and $n=10^4$, where k_0 is the Kolmogorov-Smirnov distance between the actual d.f. of the maximum of n i.i.d. r.v.'s with d.f. $F_{p,q}(x)$ and the ultimate type I d.f., i.e., $k_0=\sup|F^n_{p,q}(x)-\Lambda(x;b_n,a_n)|$, k_1 is the Kolmogorov-Smirnov distance between $F^n_{p,q}(x)$ and the penultimate form of approximation, i.e., $k_1=\sup|F^n_{p,q}(x)-A_n(x)|$, where $A_n(x)\equiv\Phi_{\alpha_n}(x;\lambda_n,\delta_n)$ if $d\log(a_n)/d\log(n)>0$ and $A_n(x)\equiv\Psi_{\alpha_n}(x;\lambda_n,\delta_n)$ if $d\log(a_n)/d\log(n)<0$. For this graph, the parameters $\{a_n, b_n\}_{n\geq 1}$ and $\{\lambda_n, \delta_n, \alpha_n\}_{n\geq 1}$ were chosen by a quantile method, i.e., $F^n_{p,q}(b_n)=\Psi_{\alpha_n}(b_n;\lambda_n,\delta_n)=\Lambda(b_n;b_n,a_n)=\exp(-1)\ (=\Phi_{\alpha_n}(b_n;\lambda_n,\delta_n))$, $F^n_{p,q}(a_n+b_n)=\Psi_{\alpha_n}(a_n+b_n;\lambda_n,\delta_n)=\Lambda(a_n+b_n;b_n,a_n)=\exp(-\exp(-1))\ (=\Phi_{\alpha_n}(a_n+b_n;\lambda_n,\delta_n))$, being $\alpha_n=|d\log(a_n)/d\log(n)|^{-1}=|\{n\cdot\exp(1)\cdot a_n\cdot(\exp(1/(n\exp(1)))-1)\cdot(\exp(1/n)-1)\cdot(q(a_n+b_n)^q-p)\cdot(qb_n^q-p)\}/\{b_n\exp(1)\cdot(\exp(1/(n\exp(1)))-1)\cdot(q(a_n+b_n)^q-p)-(a_n+b_n)\cdot(qb_n^q-p)\cdot(\exp(1/n)-1)\}|$, although analogue results have been obtained for other choices of the parameters. There is always a penultimate type II or type III approximation nearer to the actual $F^n_{p,q}(x)$ than the ultimate type I d.f., except when $d\log(a_n/d\log(n)$ is quite close to zero. For instance, for $(p, q, n)=(4, 1, 10)$, $d\log(a_n)/d\log(n)=0.0021$.



Graph 1. $R=(k_0-k_1)/\max(k_0, k_1)$ plotted against q, for particular values of p and n.

3. Rate of convergence and penultimate behaviour

As we have already seen, the reason given by Fisher and Tippett for approximating $\Phi^n(x)$, the d.f. of the maximum of n independent, standard normal r.v.'s, by a type III d.f. of largest values, is the fact that the attraction coefficient a_n, in $\Phi^n(a_n x+b_n)$, decreases to zero as $n \to \infty$, as also happens to the scale of the maximum of n i.i.d. type III r.v.'s. It seems to us that the numerical results of Section 2 call for a deeper explanation of the problem.

Let us suppose for simplicity that the right endpoint x_0^F of the d.f. F is infinity and that von Mises' condition holds for the parent d.f. $F(\cdot)$, that is, $F(x)$ is twice differentiable, at least for large values of x, and

$$(3.1) \qquad \lim_{x\to\infty} \frac{d}{dx}\{(1-F(x)/f(x)\}=0$$

where $f(x)=F'(x)$ (then $F\in \mathcal{D}(\Lambda)$, von Mises [14]).

Putting $k(x)=-d\log(-\log(F(x)))/dx$, it is easily seen that von Mises' condition is equivalent to put

$$(3.2) \qquad \lim_{x\to\infty} d(1/k(x))/dx=0\,.$$

Anderson [1] considers differences of the form $F^n(a_n x+b_n)-\Lambda(x)$, where $F\in\mathcal{D}(\Lambda)$ and finds that under certain conditions on the function $k(x)$, $F^n(a_n x+b_n)-\Lambda(x)=\gamma(x)\ d(n)\lambda(x)+o(d(n))$ uniformly over finite intervals of x. Both $\gamma(\cdot)$ and $d(\cdot)$ depend on the parent d.f. $F(\cdot)$ and $\lambda(x)=\Lambda'(x)$. Using a technique similar to the one used by Anderson, we shall derive the following result:

THEOREM 1. *Suppose that von Mises' condition holds for the d.f. $F(\cdot)$ and let the attraction coefficients $\{a_n\}_{n\geq 1}$ $(a_n>0)$ and $\{b_n\}_{n\geq 1}$ to the limiting law $\Lambda(\cdot)$ be then defined by*

$$(3.3) \qquad F(b_n)=\exp(-1/n)\,;\ a_n=1/k(b_n)$$

with $k(x)=-d\log(-\log(F(x)))/dx$.

If

(3.4) *$k'(x)$ is of constant sign for all x large enough*

and

$$(3.5) \qquad \lim_{x\to\infty} (k(x)k''(x))/(k'(x))^2=c<\infty\,,$$

putting $\alpha_n=|(d\log(a_n)/d\log(n))^{-1}|=k^2(b_n)/|k'(b_n)|$, there exists n_0 such that for $n\geqq n_0$, (3.6) $F^n(a_nx+b_n)=\Psi_{\alpha_n}(x/\alpha_n-1)+O(1/\alpha_n^2)$ *if* $k'(b_n)>0$

(3.7) $F^n(a_nx+b_n)=\Phi_{\alpha_n}(x/\alpha_n+1)+O(1/\alpha_n^2)$ *if* $k'(b_n)<0$

uniformly over finite intervals of x.

PROOF. We have $F^n(a_nx+b_n)=\Lambda(x+d_n(x))$, with $d_n(x)=-x-\log(n)-\log(-\log(F(a_nx+b_n)))$. From (3.3) and from Taylor's expansion for $d_n(x)$, $d_n(x)=-\log(-\log(F(b_n)))+a_nxk(b_n)+(a_nx)^2k'(b_n+a_n\theta_n(x))/2-x-\log(n)$, $0<\theta_n(x)<x$, it follows that $d_n(x)=(a_nx)^2k'(b_n+a_n\theta_n(x))/2$. Also from Taylor's expansion for $1/k(b_n+a_ny)$, $1/k(b_n+a_ny)=1/k(b_n)+a_ny\{-k'(b_n+a_n\psi_n(y))/k^2(b_n+a_n\psi_n(y))\}$, $|\psi_n(y)|<|y|$, we derive that $\lim_{n\to\infty} k(b_n)/k(b_n+a_ny)=1$, uniformly over finite intervals of y. On the other hand, the validity of (3.4) enables us to write $\log|k'(b_n+a_ny)|=\log|k'(b_n)|+a_nyk''(b_n+a_n\beta_n(y))/k'(b_n+a_n\beta_n(y))$, $|\beta_n(y)|<|y|$, and condition (3.5) implies that

$$d_n(x)=x^2k'(b_n)(1+O(k'(b_n)/k^2(b_n)))/(2k^2(b_n)) \tag{3.8}$$

uniformly over finite intervals of x.

From (3.3) it follows that $d\log(a_n)/d\log(n)=-k'(b_n)/k^2(b_n)$.

If $k'(x)>0$ for large values of x, there is n_1, such that for $n\geqq n_1$, $k'(b_n)>0$, and putting $\alpha_n=k^2(b_n)/k'(b_n)$ it follows from (3.8) that $d_n(x)=x^2/(2\alpha_n)+O(1/\alpha_n^2)$ uniformly over finite intervals of x, from which (3.6) follows.

Analagously, if $k'(x)<0$ for large values of x, $\alpha_n=-k^2(b_n)/k'(b_n)$, $n\geqq n_2$, and consequently for $n\geqq n_2$, $F^n(a_nx+b_n)=\Lambda(\alpha_n\log(1+x/\alpha_n))+O(1/\alpha_n^2)$ uniformly over finite intervals of x, from which (3.7) follows at once.

In order to make a comparison we remark that the formula derived by Anderson [1] was, under the same conditions of Theorem 1, $F^n(a_nx+b_n)=\Lambda(x)-x^2\lambda(x)/(2\alpha_n)+o(1/\alpha_n)$, uniformly over finite intervals of x.

So, even for large values of n, $F^n(a_nx+b_n)$ is closer to a type III (to a type II) d.f. when $k'(x)>0$ for $x\geqq x_1$ (when $k'(x)<0$ for $x\geqq x_2$), than to the ultimate limiting form $\Lambda(x)$.

It is also worth remarking that for the intermediate case $\lim_{x\to\infty} k(x)=\alpha$, $0<\alpha<\infty$, although (3.4) may hold we cannot immediately claim its validity, and we could not get any type II or type III penultimate form of approximation of $F^n(a_nx+b_n)$. In such a situation it seems sensible to conjecture that $\Lambda(x)$ is closer to $F^n(a_nx+b_n)$ than any type II or type III d.f. of largest values. For results in this direction see Anderson [1] and Gomes [8]. Moreover, we should need extra conditions, in order to have analogue results valid not only on compact sets but on the whole real line as, for instance, $k^{(j)}(x)/k^{j+1}(x)=O((k'(x)/k^2(x))^j)$, as $x\to\infty$, for all $j\geqq 2$.

An interesting result on speed of convergence towards limiting forms appears in Galambos ([6], Theorem 2.10.1): under mild conditions, $|F^n(a_nx+b_n)-G(x)|\leqq G(x)[r_{1,n}(x)+r_{2,n}(x)+r_{1,n}(x)r_{2,n}(x)]$ where $G(\cdot)$ is one of the possible extreme value distributions (the hypothesis that $F\in\mathcal{D}(G)$ is not used in the body of the proof, and hence the theorem is of interest in the study of the penultimate behaviour), $r_{1,n}(x)=2z_n^2(x)/n+2z_n^4(x)/(n^2(1-q))$, $r_{2,n}(x)=|\rho_n(x)|+\rho_n^2(x)/(2(1-s))$, with $z_n(x)=n(1-F(a_nx+b_n))$, $\rho_n(x)=z_n(x)+\log G(x)$, $q<1$ and $s<1$ such that $2z_n^2(x)/(3n)\leqq q$ and $|\rho_n(x)/3|\leqq s$ respectively. Hall [12] and Cohen [3] papers, referred to in more detail in other points of the present work, make explicit use of Galambos' general results to study rate of convergence problems.

Observe that, though implicitly, Galambos' inequalities are given in terms of the attraction coefficients, while our results make explicit use of the coefficients. In fact, in all possible cases, $r_{1,n}(x)=O(1/n)$; on the other hand the rate of convergence of $r_{2,n}(x)$ towards zero depends on the behaviour of $\rho_n(x)=n(1-F(a_nx+b_n))+\log G(x)$, and in this sense on the choice of the (either ultimate or penultimate) limiting form $G(x)$ and on the choice of the attraction coefficients a_n and b_n. Hence, a comparison of Galambos' results with the results in the present paper is not difficult to achieve and, though tedious, seems worth to be presented elsewhere in detail. In the light of what we have just said, it is obvious that the heart of the matter lies on the faster speed of convergence of $n(1-F(a_nx+b_n))+\log G(x)$ towards zero, when we replace $G(x)$, $F\in\mathcal{D}(G)$, by a suitable penultimate approximation.

4. Influence of the attraction coefficients upon the rate of convergence

Theorem 1 was proved for a particular choice of attraction coefficients $\{a_n, b_n\}_{n\geq 1}$ $(a_n>0)$ defined by (3.3) of the d.f. F to the extreme value d.f. Λ. We are now interested in what happens if we choose new attraction coefficients $\{a_n^*, b_n^*\}_{n\geq 1}$ $(a_n^*>0)$, that is, real constants such that $F^n(a_n^*x+b_n^*)\to\Lambda(x)$, as $n\to\infty$, and for all real x.

THEOREM 2. *Let $F(x)$ be a d.f. satisfying von Mises' condition and let the attraction coefficients $\{a_n, b_n\}_{n\geq 1}$ of F to the law Λ be defined by* (3.3). *Then, under the conditions of Theorem* 1 *and assuming additionally that for the choice of attraction coefficients $\{a_n^*, b_n^*\}_{n\geq 1}$, asymptotically equivalent to the choice* (3.3), *the following condition holds*

(4.1) $\lim_{n\to\infty}(da_n^*/dn)/(da_n/dn)$ *exists, and* $\lim_{n\to\infty}a_n$ *is equal to either* 0 *or* ∞

we have for all sufficiently large x and uniformly over finite intervals of x

(4.2) $F^n(a_n^*x+b_n^*)=\Psi_{\alpha_n^*}(x/\alpha_n^*-1)+x(A_n-1)+B_n+O(1/\alpha_n^{*2})$ *if* $a_n \to 0$

(4.3) $F^n(a_n^*x+b_n^*)=\Phi_{\alpha_n^*}(x/\alpha_n^*+1)+x(A_n-1)+B_n+O(1/\alpha_n^{*2})$ *if* $a_n \to \infty$

with $\alpha_n^*=(|d \log(a_n^*)/d \log(n)|)^{-1}$, $A_n=a_n^*/a_n$, *and* $B_n=(b_n^*-b_n)/a_n$.

PROOF. Convergence of types theorem asserts that $A_n \to 1$ and $B_n \to 0$, as $n \to \infty$. Besides, since $\alpha_n^*/\alpha_n=(a_n/a_n^*)((d\ a_n^*/dn)/(d\ a_n/dn))$, the validity of (4.1) and l'Hôpital's rule imply that $\alpha_n^*/\alpha_n \to 1$, as $n \to \infty$.

But $a_n^*=a_nA_n$, $b_n^*=b_n+a_nB_n$ gives $F^n(a_n^*x+b_n^*)=F^n(a_n(A_nx+B_n)+b_n)$ and so, if $a_n \to 0$, as $n \to \infty$, $k'(b_n)<0$ for $n \geqq n_0$ and then $F^n(a_n^*x+b_n^*)=\Lambda(x+x(A_n-1)-(A_nx+B_n)^2(1+O(1/\alpha_n))/(2\alpha_n)+B_n)$. Since $A_nx+B_n \to x$, as $n \to \infty$, uniformly over finite intervals of x, and $\lim_{n\to\infty} \alpha_n^*/\alpha_n=1$, it follows that $F^n(a_n^*x+b_n^*)=\Lambda(-\alpha_n^* \log(1-x/\alpha_n^*)+O(1/\alpha_n^{*2})+x(A_n-1)+B_n)$, from which (4.2) follows. (4.3) follows analagously.

Consequently, if $\alpha_n^{*2}(A_n-1)$ and $\alpha_n^{*2}B_n$ go to $c<\infty$ as $n \to \infty$, the overall rate of convergence is still $O(1/\alpha_n^{*2})$. If either A_n-1 or B_n converges to zero more slowly than $1/\alpha_n^{*2}$, than the convergence is of smaller order than $1/\alpha_n^{*2}$, and this happens for instance when, for normal extremes, we consider the attraction coefficients $\{a_n^*, b_n^*\}_{n\geq 1}$ given by $2\pi b_n^{*2} \exp(b_n^*)=n^2$, $a_n^*=1/b_n^*$ (the ones considered by Hall [12]).

In Section 1 we have mainly worked with the attraction coefficients $\{a_n^*, b_n^*\}_{n\geq 1}$ given by

(4.4) $$F(b_n^*)=1-1/n\,;\ a_n^*=1/(n\ f(b_n^*))$$

for which $\alpha_n^*=|(d \log(a_n^*)/d \log(n))^{-1}|=|([d((1-F(x))/f(x))/dx]_{x=b_n^*})^{-1}|$. We thus state the following result:

THEOREM 3. *For the attraction coefficients* $\{a_n^*, b_n^*\}_{n\geq 1}$ *given by* (4.4) *and under the conditions of Theorem* 2 *and the additional condition*

(4.5) $$\lim_{x\to\infty} \frac{1}{1-F(x)} \left\{\frac{d}{dx}\frac{1-F(x)}{f(x)}\right\}^2>0$$

we have $\lim_{n\to\infty} \alpha_n^*(A_n-1)=\lim_{n\to\infty} \alpha_n^*B_n=0$, *where* A_n *and* B_n *are defined as in Theorem* 2, *and so the overall rate of convergence in* (4.2) *and in* (4.3) *is still* $O(1/\alpha_n^{*2})$.

PROOF. Since $\exp(-1/n)>1-1/n$, we have $b_n>b_n^*$, and so $(b_n-b_n^*)/a_n=B_n^*>0$, with $\lim_{n\to\infty} B_n^*=0$. From $b_n=b_n^*+a_n^*B_n^*$, we get $F(b_n)=F(b_n^*)+a_n^*B_n^*$ $f(b_n^*+a_n^*\theta_n)$, $0<\theta_n<B_n^*$, and from the fact that $1/f(b_n^*+a_n^*\theta_n)=1/f(b_n^*)+n$ $a_n^*\theta_n[(-f'/(nf^2))]_{b_n^*+a_n^*\varphi_n}$, $0<\varphi_n<\theta_n$, and that $b_n^*+a_n^*\varphi_n \to \infty$, as $n \to \infty$, we get $f(b_n)/f(b_n^*+a_n^*\theta_n)=1+\theta_n[(-f'/(nf^2))]_{b_n^*+a_n^*\varphi_n}=1+\theta_nO(1)=1+o(1)$, from which follows that $n\ B_n^*=n^2(F(b_n)-F(b_n^*))/(1+\theta_nO(1))$. Since $n^2(F(b_n)-$

$F(b_n^*))=n^2(\exp(-1/n)-1+1/n)\to 1/2$, as $n\to\infty$, we get that $B_n^*=O(1/n)$ and consequently the same happens to $B_n=-(a_n^*/a_n)B_n^*$. On the other hand, $A_n^*=a_n/a_n^*=\exp(-1/n)\ f(b_n^*)/f(b_n)$, and since $1/f(b_n)=1/f(b_n^*)+na_n^*B_n^*[(-f'/(nf^2))]_{b_n^*+a_n^*\phi_n}$, $0<\phi_n<B_n^*$, we get $f(b_n^*)/f(b_n)=1+B_n^*O(1)$, and consequently $n(A_n^*-1)=n(\exp(-1/n)-1)+n\exp(-1/n)B_n^*O(1)$, which goes to 1/2, as $n\to\infty$, the same obviously happening to $A_n-1=-(A_n^*-1)a_n^*/a_n$.

Since condition (4.5) implies that $\alpha_n^{*2}/n\to c^*<\infty$, as $n\to\infty$, we get the result.

5. Penultimate behaviour of the joint d.f. of the i largest order statistics

Theorem 1 has an immediate analogue if instead of $\{a_n, b_n\}_{n\geq 1}$ given by (3.3) we choose $\{a_n^*, b_n^*\}_{n\geq 1}$ given by (4.4), and if instead of $F^n(a_nx+b_n)$ we consider $n(1-F(a_n^*x+b_n^*))$. Then, we get uniformly over finite intervals of x,

$$n(1-F(a_n^*x+b_n^*))=(1-x/\alpha_n^*)^{\alpha_n^*}+O(1/\alpha_n^{*2})\qquad \text{if } k'(x)>0 \text{ for large values of } x \tag{5.1}$$

$$n(1-F(a_n^*x+b_n^*))=(1+x/\alpha_n^*)^{-\alpha_n^*}+O(1/\alpha_n^{*2})\qquad \text{if } k'(x)<0 \text{ for large values of } x\,. \tag{5.2}$$

Besides, (5.1) and (5.2) remain valid if $\{a_n^*, b_n^*, \alpha_n^*\}_{n\geq 1}$ is replaced by $\{a_n, b_n, \alpha_n\}_{n\geq 1}$ with $\{a_n, b_n\}_{n\geq 1}$ given by (3.3), and if we further assume the validity of (4.1) and (4.5).

We may thus generalize Theorem 1 for the joint d.f. of the i largest order statistics $(M_n^{(1)}, \cdots, M_n^{(i)})$, suitably normalized, of the sample $(X_1, \cdots, X_n)$, with the XX's i.i.d. r.v.'s with d.f. F, i a fixed integer. For simplicity of notation we state the following result for $i=2$, although there is an immediate analogue for any fixed i.

Related work appears in Reiss [15] who considers asymptotic penultimate expansions for the joint d.f. of the i largest order statistics $(M_n^{(1)}, \cdots, M_n^{(i)})$, i fixed, together with estimates of the remainder terms. The approximations hold uniformly over all Borel sets.

THEOREM 4. *Under the conditions of Theorem* 1 *and assuming additionally that* (4.5) *holds and that for* $\{a_n, b_n\}_{n\geq 1}$ *given by* (4.4), (4.1) *holds, we have for* $\{\tilde{a}_n, \tilde{b}_n, \tilde{\alpha}_n\}_{n\geq 1}$ *either equal to* $\{a_n, b_n, \alpha_n\}_{n\geq 1}$ *or to* $\{a_n^*, b_n^*, \alpha_n^*\}_{n\geq 1}$ *and uniformly over finite intervals of* x,

$$P(M_n^{(1)}<\tilde{a}_nx_1+\tilde{b}_n,\ M_n^{(2)}<\tilde{a}_nx_2+\tilde{b}_n) \tag{5.3}$$

$$=\begin{cases} \Psi_{\tilde{\alpha}_n}(x_1/\tilde{\alpha}_n-1)+O(1/\tilde{\alpha}_n^2) & \text{if } x_1\leqq x_2 \\ \Psi_{\tilde{\alpha}_n}(x_1/\tilde{\alpha}_n-1)[1+\log(\Psi_{\tilde{\alpha}_n}(x_1/\tilde{\alpha}_n-1)/\Psi_{\tilde{\alpha}_n}(x_2/\tilde{\alpha}_n-1))]+O(1/\tilde{\alpha}_n^2) & \\ & \text{if } x_1>x_2 \end{cases}$$

if $k'(x)>0$ *for large values of* x.

(5.4) $$P(M_n^{(1)}<\tilde{a}_n x_1+\tilde{b}_n,\ M_n^{(2)}<\tilde{a}_n x_2+\tilde{b}_n)$$

$$=\begin{cases} \Phi_{\tilde{\alpha}_n}(x_1/\tilde{\alpha}_n+1)+O(1/\tilde{\alpha}_n^2) & \text{if } x_1\leqq x_2 \\ \Phi_{\tilde{\alpha}_n}(x_1/\tilde{\alpha}_n+1)[1+\log(\Phi_{\alpha_n}(x_1/\tilde{\alpha}_n+1)/\Phi_{\tilde{\alpha}_n}(x_2/\tilde{\alpha}_n+1))]+O(1/\tilde{\alpha}_n^2) & \\ & \text{if } x_1>x_2 \end{cases}$$

if $k'(x)<0$ *for large values of* x.

We just remark here that if $F\in\mathcal{D}(G)$ for maxima, in the context defined in Section 1, with attraction coefficients $\{a_n, b_n\}_{n\geqq 1}$ $(a_n>0)$, the joint limiting d.f. of the i largest order statistics suitably narmalized, i.e., $\lim_{n\to\infty} P\left(\bigcap_{j=1}^{i} M_n^{(j)}\leqq a_n x_j+b_n\right)$ is given by

$$G(\min_{1\leqq j\leqq i} x_j)\sum_{\substack{r_j\leqq r_{j+1}\leqq j\\ 1\leqq j\leqq i-1}}\cdots\sum\prod_{j=1}^{i-1}\{\log(G(\min_{1\leqq k\leqq j} x_k)/G(\min_{1\leqq k\leqq j+1} x_k))\}^{r_{j+1}-r_j}/(r_{j+1}-r_j)!$$

with $r_1=0$, where $G(\cdot)$ is obviously the limiting d.f. of $(M_n^{(1)}-b_n)/a_n$. This joint d.f. is to be taken 0 or 1 if one of the x_j's is such that $G(x_j)=0$ or if all the x_j's are such that $G(x_j)=1$, $1\leqq j\leqq i$, respectively. For details see Gomes [8], [9] and Weissman [17].

6. Distribution functions in the class $\mathcal{L}$

THEOREM 5. *If* $F(x)\in\mathcal{L}$, *given in Definition* 1 *and* $\varepsilon(x)$ *is defined and bounded for large values of* x, $\varepsilon'(x)$ *and* $\varepsilon''(x)$ *exist,* $\varepsilon'(x)$ *is a continuous function of* x *and* $\varepsilon''(x)$ *is monotone, than for* $(p, q)\neq(0, 1)$, *assumptions* (3.4), (3.5) *and* (4.5) *hold, and for* $p\neq 0$ *and* $\{a_n^*, b_n^*\}_{n\geqq 1}$ *given by* (4.4), *condition* (4.1) *holds. Consequently, Theorem* 1 *and* 3 *hold for those parent d.f.'s. Moreover, for* $p\neq 0$, *the optimal overall rate of convergence in* (3.6) *and* (3.7) *is* $O(1/(\log n)^2)$ *if* $q\neq 1$ *and* $O(1/(\log n)^4)$ *if* $q=1$, *uniformly on the whole real line.*

PROOF. We first note that the conditions imposed on $\varepsilon(x)$ imply $x\varepsilon'(x)\to 0$ and $x^2\varepsilon''(x)\to 0$, as $x\to\infty$. From

(6.1) $$F(x)=1-A(\alpha x+\beta)^p\exp(-(\alpha x+\beta)^q)(1+\varepsilon(x))=1-y(x)$$

we get

$$(6.2)\qquad \begin{aligned} f(x) &= y(x)\{q\alpha(\alpha x+\beta)^{q-1} - p\alpha/(\alpha x+\beta) - \varepsilon'(x)/(1+\varepsilon(x))\} \\ &= y(x)B(x) \end{aligned}$$

from which follows that $f'(x) = -f(x)B(x) + f(x)B'(x)/B(x)$, and consequently, after some manipulation, we get

$$(6.3)\qquad \begin{aligned} f'(x)/f(x) = &-q\alpha(\alpha x+\beta)^{q-1} + (p+q-1)\alpha/(\alpha x+\beta) + p\alpha/(\alpha x+\beta)^{q-1} \\ &+ \varepsilon'(x)/(1+\varepsilon(x)) + o(1/x^{q+1}) \end{aligned}$$

$$(6.4)\qquad f'(x)/f^2(x) = \{-1 + (q-1)/(q(\alpha x+\beta)^q) + p/(\alpha x+\beta)^{2q} + o(1/x^{2q})\}/y(x)$$

$$(6.5)\qquad \begin{aligned} (f'(x)/f(x))' = &-q(q-1)\alpha^2(\alpha x+\beta)^{q-2} - p\alpha^2/(\alpha x+\beta)^{q+2} \\ &-(p+q-1)\alpha^2/(\alpha x+\beta)^2 + (\varepsilon'(x)/(1+\varepsilon(x)))' + o(1/x^{q+2}) . \end{aligned}$$

On the other hand, since $k(x) = d(x)B(x)$, where $d(x) = -y(x)/((1-y(x))\log(1-y(x))) \to 1$, as $x \to \infty$, $d(x)-1$, $d'(x)$ and $d''(x)$ being of smaller order than $1/x^\nu$ for every finite arbitrarily chosen ν. We thus have

$$(6.6)\qquad k(x) = q\alpha(\alpha x+\beta)^{q-1} - p\alpha/(\alpha x+\beta) - \varepsilon'(x)/(1+\varepsilon(x)) + o(1/x^\nu) .$$

It immediately follows that if $q>1$, $k(x) \to \infty$ and, as $k(x)>0$, $k'(x)$ is positive for large values of x; if $q<1$, $k(x) \to 0$ and so $k'(x)<0$ for large values of x; if $q=1$, $k(x) \to 1$ and then $k'(x)$ has the sign of p for large values of x if $p \neq 0$; only if $p=0$, $q=1$ can we not guarantee that condition (3.4) holds.

Concerning von Mises' condition, we get from (6.3) and (6.6) that $f'/(fk) = -1 + (q-1)/(q(\alpha x+\beta)^q) + O(1/x^{2q})$, and since

$$(6.7)\qquad k'(x)/k^2(x) = f'(x)/(f(x)k(x)) + 1 + \log(F(x))$$

we have $(1/k(x))' = -\log(F(x)) + O(1/x^q)$, if $q \neq 1$ and $(1/k(x))' = -\log(F(x)) + O(1/x^2)$ if $q=1$, and consequently $(1/k(x))'$ goes to zero, as $x \to \infty$, for all values of p and q.

To check the validity of condition (3.5) we use the relation

$$(6.8)\qquad \begin{aligned} k''/(kk') = &f'/(fk) + 2(1+\log(F)) \\ &+ [(f'/f)'/k - k\log(F)]/[f'/f + (1+\log(F))k] \end{aligned}$$

and the fact that $kk''/k'^2 = (k''/(kk'))/(k'/k^2)$.

From (6.3) and from the fact that $(1+\log F)k = q\alpha(\alpha x+\beta)^{q-1} - (q-1)/(\alpha x+\beta) - \varepsilon'(x)/(1+\varepsilon(x)) + o(1/x^\nu)$, we get $f'/f + (1+\log F)k = (q-1)\alpha/(\alpha x+\beta) + p\alpha/(\alpha x+\beta)^{q+1} + o(1/x^{q+1})$, and since $k(x)\log(F(x)) = o(1/x^\nu)$ for every ν finite and arbitrarily chosen, we get that $k\log(F)/(f'/f + (1+\log(F))k) \to 0$, as $x \to \infty$. But $f'/(fk) \to -1$, as $x \to \infty$ and $2(1+\log(F)) \to 2$, as $x \to \infty$, and so it follows from (6.8) that the behaviour of $k''/(kk')$, as $x \to \infty$, depends on the behaviour of $(f'/f)'/(k(f'/f + (1+\log(F))k))$.

But, from the fact that $(f'/f+(1+\log(F))k)=q(q-1)\alpha^2(\alpha x+\beta)^{q-2}+q\alpha^2/(\alpha x+\beta)^2+o(1/x^2)$ and from (6.5) we get the validity of condition (3.5) for all values of p, q, α, β, except for $p=0$, $q=1$. From now on, we consider $(p, q)\neq(0, 1)$. In this particular case, since $[k^{(j)}(b_n)/(k(b_n))^{j+1}]=O((k'(b_n)/k^2(b_n))^j)$, we have $d_n(x)=x^2/(2\alpha_n)+O(1/\alpha_n^2)$ over the whole real line. Besides, since $((1-F)/f)'=-B'(x)/B^2(x)=-y(x)f'(x)/f^2(x)$, it follows from (6.4) that $((1-F)/f)'=-(q-1)/(q(\alpha x+\beta)^q)-p/(\alpha x+\beta)^{2q}+o(1/x^{2q})$. So, $(((1-F)/f')^2/(1-F)$ goes to infinity, as $x\to\infty$ and $((1-F)/f)'/((1/k)')$ goes to 1, as $x\to\infty$, which implies the validity of (4.1) for $\{a_n^*, b_n^*\}_{n\geq 1}$ given by (4.4).

We now need to find an explicit form for the rate of convergence $1/\alpha_n=[|d(1/k(x))/dx|]_{x=b_n}$.

First of all, from $F(b_n)=\exp(-1/n)$ and from (6.1) we get $A(\alpha b_n+\beta)^p\exp(-(\alpha b_n+\beta)^q)(1+\varepsilon(b_n))=1-\exp(-1/n)$ and so $nA(\alpha b_n+\beta)\exp(-(\alpha b_n+\beta)^q=n(1-\exp(-1/n))/(1+\varepsilon(b_n))\to 1$, as $n\to\infty$, from which follows $\log(n)+\log(A)+p\log(\alpha b_n+\beta)-(\alpha b_n+\beta)^q=o(1)$. Working out this last identity, we get

$$(6.9)\qquad (\alpha b_n+\beta)^q=\log(n)+\log(A)+p\log(\log(n))/q+o(1)\,.$$

Besides, from (6.7) and (6.4) we get $[(-1/k)']_{b_n}=1-1/n-\exp(-1/n)/(n(1-\exp(-1/n)))-\exp(-1/n)\{(q-1)/(q(\alpha b_n+\beta)^q)+p/(\alpha b_n+\beta)^{2q}+o(1/b_n^{2q})\}/(n(1-\exp(-1/n)))$, and recalling (6.9) and the fact that $1-1/n-\exp(-1/n)/(n(1-\exp(-1/n)))=O(1/\log^2(n))$, we get $[(-1/k)']_{b_n}=(q-1)(1-p\log(\log(n))/(q\log(n)))+O(1/\log^2(n))$ and hence the result.

Acknowledgements

I was very fortunate indeed working on extreme value theory, first with Professor J. Tiago de Oliveira, later on with Dr. C. W. Anderson. With both of them I had stimulating discussions, that deeply influenced my understanding of the matter. To both of them I express my gratitude.

I am grateful to Calouste Gulbenkian Foundation for finantial support during the preparation of the first draft of the present paper.

I am grateful to the referee for drawing my attention to Cohen's and to Reiss' papers, and to the importance of Galambos' results in the context of penultimate approximations.

Faculdade de Ciências
Da Universidade de Lisboa

References

[1] Anderson, C. W. (1971). Contributions to the asymptotic theory of extreme values, Ph.D. Thesis, Imperial College, London.

[2] Anderson, C. W. (1976). Extreme value theory and its approximations, *Proc. Symp. Reliability Technology*, Bradford, U.K.

[3] Cohen, J. P. (1982). The penultimate form of approximation to normal extremes, *Adv. Appl. Prob.*, **14**, 324-339.

[4] Dronkers, J. J. (1958). Approximate formulae for the statistical distributions of extreme values, *Biometrika*, **45**, 447-470.

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

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

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

[8] Gomes, M. I. (1978). Some probabilistic and statistical problems in extreme value theory, Ph.D. Thesis, Sheffield, England.

[9] Gomes, M. I. (1981a). An i-dimensional limiting distribution function of largest values and its relevance to the statistical theory of extremes, *Statistical Distributions in Scientific Work*, (eds. C. Taillie et al.), Vol. 6, D. Reidel, 389-410.

[10] Gomes, M. I. (1981b). Closeness of penultimate approximations in extreme value theory, *Abst. 14th European Meeting of Statisticians*, Wroclaw.

[11] Haldane, J. B. and S. D. Jayakar (1963). The distribution of extremal and nearly extremal values in samples from a normal population, *Biometrika*, **50**, 89-94.

[12] Hall, P. (1979). On the rate of convergence of normal extremes, *J. Appl. Prob.*, **16**, 433-439.

[13] Hall, P. (1980). Estimating probabilities for normal extremes, *Adv. Appl. Prob.*, **12**, 491-500.

[14] Mises von, R. (1936). La distribution de la plus grande de n valeurs, *Rev. Math. Union Interbalkanique*, **1**, 141-160.

[15] Reiss, R. D. (1981). Uniform approximation to distributions of extremes order statistics, *Adv. Appl. Prob.*, **13**, 533-547.

[16] Uzgören, N. T. (1954). The asymptotic development of the distribution of the extreme values of a sample, *Studies in Mathematics and Mechanics Presented to R. von Mises*, Academic Press, New York.

[17] Weissman, I. (1975). Multivariate extremal processes generated by independent non-identically distributed random variables, *J. Appl. Prob.*, **12**, 477-487.