NISS

On High Level Exceedance Modeling and Tail Inference

M.R. Leadbetter

Technical Report Number 8
December, 1993

National Institute of Statistical Sciences
19 T. W. Alexander Drive
PO Box 14006
Research Triangle Park, NC 27709-4006
www.niss.org

On High Level Exceedance Modeling and Tail Inference *

M. R. Leadbetter
Department of Statistics
University of North Carolina
and
National Institute of Statistical Sciences

Abstract

This paper discusses a general framework common to some varied known and new results involving high values of stationary stochastic sequences. In particular these concern

(a) **Probabilistic modeling** of infrequent but potentially damaging physical events such as storms, high stresses, high pollution episodes, describing both repeated occurrences and associated "damage" magnitudes

(b) **Statistical estimation** of "tail parameters" of a stationary stochastic sequence $\{X_j\}$. This includes a variety of estimation problems and in particular, cases such as estimation of expected lengths of clusters of high values (e.g. storm durations), of interest in (a).

"Very high" values (leading to Poisson–based limits) and "high" values (giving normal limits) are considered and exhibited as special cases within the general framework of central limit results for "random additive interval functions". The case of array sums of dependent random variables is revisited within this framework, clarifying the role of dependence conditions and providing minimal conditions for characterization of possible limit types.

*Research supported in part by the Office of Naval Research Grant No. N00014 93 1 0043 and EPA Cooperative agreement CR 819638 01 0.
Key words: Extreme values, Exceedance modeling, tail estimation, central limit theory.
AMS 1991 subject classification: 60F05, 60G70, 62G05, 60G55

1 Introduction

The purpose of this paper is to present a unifying framework and viewpoints for several apparently quite distinct areas, including the modeling of extreme events, statistical tail inference, and central limit theory for (dependent) array sums. It is hoped that thereby useful insights will be provided for the known and new results discussed.

By way of specific example consider r.v.'s $X_1, X_2, \ldots, X_n$ (from an iid or stationary sequence) and write

$$Z_n = \sum_{i=1}^{n} (X_i - u_n)_+$$

where $\{u_n\}$ is a sequence of constants ("levels"), $u_n \to \infty$. From a modeling viewpoint events $\{X_i > u_n\}$ ("exceedances") may represent "damage", (e.g. from storms, structural stresses, pollution in excess of a regulatory threshold) measured as the excess value $X_i - u_n$. Z_n is the total accumulated damage from $X_1, X_2, \ldots, X_n$. On the other hand, Z_n is an array sum and, as a statistic, provides the heart of the so-called "Hill estimator" for the parameter of an exponentially decaying tail distribution, its variants for regularly varying tails, and tail and quantile estimation.

In modeling, Z_n arises from a point process of exceedance locations each "marked" with associated damage $X_i - u_n$. For fast rates of convergence $u_n \to \infty$ this converges in distribution (under suitable normalizations) to a Compound Poisson Process and in particular the total damage Z_n has a Compound Poisson distributional limit. This models damage at very high levels. For lower levels, (with $u_n \to \infty$ more slowly), a normal limit may be obtained for Z_n. These latter limits seem particularly relevant to practical modeling, though little investigated to date.

As a statistic for tail inference, it is the *normal* limit for Z_n which has received considerable study – providing the asymptotic distributions needed for tail estimation.

Most of this activity (as evidenced by other papers in this volume) has focussed on iid assumptions, but with attention ([4], [14]) on dependent cases. Interesting connections between the very high level (Poisson based) and lower level (normal) limits concern the necessity to use lower levels for estimation of parameters relevant to the very high level behavior, for estimation consistency. These are explored in Section 5.

The asymptotic distributional results needed for modeling and statistical estimation fall within the "central limit problem" for array sums (of dependent r.v.'s). The development of this theory is outlined in Section 2 using the framework of "random additive interval functions" which are equivalent to such array sums for discrete time problems, but which also permit application (as very briefly indicated in Section 6) to continuous time cases.

Following the general framework of Section 2, specific application to array sums of dependent r.v.'s is made in Section 3 with the aim of providing a perspective and clarification of dependence assumptions of earlier literature, and covering an apparent gap in that literature regarding characterization of possible limits.

Section 4 concerns the modeling problem – exhibiting three specific models for "damage" from high level exceedances (including that (Z_n) introduced above. Emphasis here is on very high levels and corresponding models from "Poisson-based" distributional convergence.

Section 5 contains a discussion of lower level modeling and tail inference based on slower convergence rates $u_n \to \infty$, and relationships between the fast and slow convergence rates in inference problems. Finally related continuous time problems are very briefly indicated in Section 6.

2 General framework and a central limit theorem.

The underlying theory is briefly described here without proofs – full details may be found in the recent paper [11].

As noted by Bergström [1] it is natural to regard an array sum $S_n = \sum_{i=1}^{n} X_{n,i}$ as a random additive function of intervals $I \subset (0,1]$ viz.

$$Z_n(I) = \sum\{X_{ni} : i/n \in I\} \tag{2.1}$$

Formally a *random additive interval function* (r.i.f.) may be defined as a family of random variables $Z_n(I)$ for intervals $I = (a,b] \subset (0,1]$, satisfying for $0 \leq a < b < c \leq 1$, $n = 1, 2, \ldots$,

$$Z_n(a,c] = Z_n(a,b] + Z_n(b,c].$$

It will be assumed that the r.i.f. is (array) strongly mixing in the specific following sense. Define the σ-fields

$$\begin{aligned} \mathcal{B}_{s,t}^{(n)} &= \sigma\{Z_n(a,b] : s < a < b \leq t\}, \qquad 0 \leq s \leq t \leq 1 \\ \alpha_{n,\ell} &= \sup\{|P(A \cap B) - P(A)P(B)| : \ A \in \mathcal{B}_{0,s}^{(n)}, B \in \mathcal{B}_{s+\ell,1}^{(n)}, \ s + \ell < 1\} \end{aligned} \tag{2.2}$$

That is in this the events A belong to the "past" and B to the "future" from $s + \ell$ to 1, separated by ℓ. Then $\{Z_n\}$ will be called strongly mixing if

$$\alpha_{n,\ell_n} \to 0, \ \text{some} \ \ell_n \to 0 \tag{2.3}$$

Note that the future ends at 1, and the gap ℓ_n between past and future tends to zero rather than infinity, as more customary. These are due to time normalizations, as will be seen in applications. It should also be noted that (a) strong mixing is a much weaker restriction than many other forms of mixing commonly used and (b) in applications where

Z_n is obtained from an underlying stochastic sequence $\{X_n\}$, strong mixing for $\{Z_n\}$ can be implied by weaker conditions for $\{X_n\}$.

The effect of the mixing is substantially through a *standard sequence* $\{k_n\}$ defined to be any integer sequence $k_n = o(n)$ with

$$k_n(\ell_n + \alpha_{n,\ell_n}) \to 0 \tag{2.4}$$

This holds for bounded sequences k_n but also for $k_n \to \infty$ up to the rate limited by (2.4). The less the long range dependence the faster the possible rate for $k_n \to \infty$ up to the limit $k_n = n$ applicable under independence with $\alpha_{n,\ell} = \ell_n = 0$.

The importance of a standard sequence $\{k_n\}$ is that if $I_1, I_2, \ldots, I_{k_n}$ are disjoint intervals (which can change with n), then the k_n quantities $Z_n(I_j)$ are asymptotically independent, i.e. for real t_j (which can also change with n),

$$\mathcal{E}\exp\{i\sum_1^{k_n} t_j Z_n(I_j)\} - \prod_1^{k_n} \mathcal{E}\{\exp it_j Z_n(I_j)\} \to 0.$$

This is shown in [11], using the classical method of "clipping" an amount ℓ_n from each I_j to give separated intervals and hence approximate independence.

The main underlying result is most conveniently stated in terms of associated independent arrays. To that end a *k_n-partition of an interval I* is any partition of I into k_n disjoint subintervals. If $\{\xi_{n,j}\}$ are independent, $1 \le j \le k_n$, $\xi_{n,j} \stackrel{d}{=} Z_n(I_j)$, then $\{\xi_{n,j}\}$ will be termed an *independent array associated with* $Z_n(I)$.

Further the partition $\{I_j\}$ will be called *uniformly asymptotically negligible* (u.a.n.) if

$$\max_{1\le j\le k_n} P\{|Z_n(I_j)| \ge \epsilon\} \to 0, \quad \text{each } \epsilon > 0, \quad \text{i.e. } \{\xi_{n,j}\} \text{ is uan.}$$

The main result follows simply from the asymptotic independence of $Z_n(I_j)$ (see [11] for proof details):

Theorem 2.1 *Let $\{X_{n,j}\}$ be an independent array associated with $Z_n(I)$, Z_n being a (strongly mixing) r.i.f. and I a fixed interval. Then $Z_n(I)$ has the same limit in distribution (if any) as $\sum_{j=1}^{k_n} \xi_{n,j}$. In particular if $\{I_j\}$ is u.a.n., any limit is infinitely divisible.*

This result shows that limiting distributions for $Z_n(I)$ under strong mixing are just those for iid array sums. The classical criteria (cf. [10]) may be applied to distributions of $Z_n(I_j) \stackrel{d}{=} X_{n,j}$ to determine which limit holds. For example the "normal convergence criterion" of [10] has the following form here.

Corollary 2.2 If $Z_n(I) \stackrel{d}{\to} \eta$, a r.v. then η is normal and $\{I_j\}$ is u.a.n. if and only if $\sum P\{|Z_n(I_j)| \geq \epsilon\} \to 0$, each $\epsilon > 0$. Then $\eta = N(\alpha, \sigma^2)$ with

$$\begin{aligned} \alpha &= \lim_{n\to\infty} \sum_j a_{n,j}(\tau), \quad \tau > 0 \\ a_{n,j}(\tau) &= \mathcal{E}\{Z_n(I_j) \ 1(|Z_n(I_j)| < \tau)\} \\ \sigma^2 &= \lim_{n\to\infty} \sum_j [\mathcal{E} Z_n^2(I_j) \ 1\{|Z_n(I_j)| < \tau\} - a_{n,j}^2(\tau)] \end{aligned}$$

3 Strongly mixing array sums.

Central limit theory for arrays of dependent r.v.'s $\{X_{n,j}\}$ saw strong activity in the late 1960's and early 1970's, (see e.g. [12], [1], [2], [6]) and some more recent revisitations ([15], [3], [16]). A good deal of this work was directed towards finding sufficient conditions on the distribution of X_{nj} under which $\sum_j X_{n,j}$ has the same limiting distribution as if the $X_{n,j}$ were independent. As a result stringent conditions – typically involving uniform types of mixing assumptions – were imposed. It was also often assumed that the $X_{n,j}$ had finite second moments.

While sufficient conditions for particular limits will necessarily be complicated, much can be said and the problem illuminated simply from Theorem 2.1, without any moment conditions and without dependence restrictions beyond strong mixing alone. Specifically let the array $\{X_{n,j}\}$ be strongly mixing, and the r.i.f. Z_n defined by (2.1), i.e. $Z_n(I) = \sum(X_{n,i} : i/n \in I)$. Let $\{k_n\}$ be a standard sequence (Eqn. (2.4)), $r_n = [n/k_n]$. Writing $I_j = ((j-1)r_n/n,\ jr_n/n]$ and defining independent r.v.'s $Y_{n,j} \stackrel{d}{=} \sum_{i=(j-1)r_n+1}^{jr_n} X_{n,i}$ it follows simply from Theorem 2.1 that $S_n = \sum_{j=1}^{n} X_{n,j} = Z_n(0,1]$ has the same limit in distribution (if any) as the independent array sum $\sum Y_{n,j}$.

It thus follows that, under strong mixing alone, the possible limiting distributions for array sums are the same as under classical independence assumptions. Further the particular limiting distribution which applies is determined from the distributions of $Y_{n,j}$'s via the classical domain of attraction criteria. All this holds under strong mixing alone, and exhibits its central role in the theory.

Thus sufficient conditions for particular domains of attraction involve the distributions of the $Y_{n,j}$, which are sums of groups of (r_n) consecutive $X_{n,j}$. Hence the provision of sufficient conditions in terms of the original $X_{n,j}$ must obviously require more restrictive dependence assumptions. As noted this problem is considered in the literature (e.g. [12]) under finite moment restrictions. While clearly domain of attraction conditions on the original $X_{n,j}$ are desirable, the above recognition of the component parts of the problem does seem to us to shed light on the roles of the conditions as well as providing the characterization of possible limits, under strong mixing alone.

4 Exceedance modeling.

We now consider special r.i.f.'s based on high values of a stationary sequence $X_1, X_2, \ldots$. This section focuses on their interpretation as models e.g. describing occurrence and magnitude of damage, and Poisson-based limits at very high levels. Lower levels – giving normal approximations and their use in tail inference will be discussed in Section 5.

4.1 Clustering of exceedances.

Typically positive dependence between neighboring terms of a stationary sequence $X_1, X_2, \ldots$ leads to the clustering of exceedances $X_i > u_n$ of a level u_n, since one high value tends to "attract" another. Clusters are naturally regarded as runs of consecutive exceedances and may be so defined, but a (typically asymptotically equivalent) definition is more convenient for our purposes.

Specifically let $\{k_n\}$, $k_n \to \infty$ $k_n = o(n)$, be a sequence (to be chosen) of integers, $r_n = [n/k_n]$ and divide the integers $1, 2, \ldots, k_n r_n (\sim n)$ into k_n consecutive groups (or "blocks") $((i-1)r_n + 1, (i-1)r_n + 2, \ldots i r_n)$ $1 \leq i \leq k_n$. Then the exceedances (if any) in such a block will be called a *cluster*. These "block clusters" may differ from "run clusters" in that they may have gaps with $X_i < u_n$, a run cluster may be split into two block clusters by an endpoint $i r_n$, or a block cluster may consist of two or more run clusters (occurring in the same block). However judicious choice of k_n (as a *standard sequence* in the sense of Section 2) will ensure asymptotic equivalence of the definitions, and its asymptotic independence of choice of k_n beyond a maximal rate of convergence of $k_n \to \infty$.

Corresponding to this (block) definition, the *cluster size* distribution $\pi_n(r)$, $r = 1, 2, \ldots$ is defined to be the probability of r exceedances in a block given at least one. Stationarity implies that this is independent of the particular block chosen.

4.2 Exceedance point process.

For modeling of exceedances it is convenient to normalize their occurrence times $\{i : X_i > u_n, \ 1 \leq i \leq n\}$ by the factor n to obtain a point process N_n on $(0, 1]$ consisting of the points i/n for which $X_i > u_n$. That is for $B \subset (0,1]$ $N_n(B)$ is the number of such normalized exceedance points $i/n \in B$, i.e.

$$N_n(B) = \#\{\frac{i}{n} \in B : X_i > u_n,\ \ 1 \le i \le n\} = \#\{i \in nB : X_i > u_n,\ \ 1 \le i \le n)$$

N_n will be referred to as the *exceedance point process* on $(0,1]$, consisting of the (normalized) exceedances among $X_1, X_2, \ldots, X_n$. Note that its intensity $\mathcal{E}N_n(0,1] = n(1 - F(u_n))$ where F is the d.f. of each X_i.

N_n obviously defines a r.i.f., $\{N_n(I)\}$ which will be assumed strongly mixing. Clusters will then be defined as above using any fixed standard sequence k_n obtained from the mixing condition as in Section 2. It will also be convenient to refer to clusters in the normalized setting as the groups of (normalized) exceedances in intervals $I_i = ((i-1)r_n/n, ir_n/n],\ \ 1 \le i \le k_n$ which, together with a small interval $(k_n r_n/n, 1]$ form a k_n-partition of $(0,1]$. The cluster size distribution π_n then clearly satisfies

(4.1) $$\pi_n\{r\} = P\{N_n(J_1) = r | N_n(J_1) > 0\},\ \ r = 1, 2, \ldots$$

4.3 Compound Poisson limits for the exceedance point process.

The models for exceedances of high values result from limiting theorems as $u_n \to \infty$. For Poisson-type models fast convergence (i.e. very high levels) is required, such that $n(1 - F(u_n))$ converges to a finite limit. As noted, lower levels (with $n(1 - F(u_n)) \to \infty$) will be considered in Section 5 giving normal limits which are relevant also for inference.

For iid $X_1, X_2, \ldots$ and $n(1 - F(u_n)) \to \tau$, $N_n(0,1]$ is binomial, $B(n, p_n = 1 - F(u_n))$ and has a Poisson distributional limit with mean τ. Similar consideration of the joint independent distributions of $N_n(B_i)$ for disjoint B_i, show very simply that in fact the exceedance point process N_n converges in distribution to a Poisson process with intensity

τ on $(0,1]$.

For stationary sequences $\{X_n\}$ (with N_n subject to the r.i.f. strong mixing condition) one has clusters of exceedances each of which converges to a single point after the time normalization ($r_n/n = [n/k_n]/n \sim 1/k_n \rightarrow 0$). Thus it may be expected that the locations of the clusters tend to become Poisson and with size distribution being the limit (if any) of π_n. That is one may expect that N_n converges in distribution to a Compound Poisson Process based on a Poisson Process of limiting cluster positions and independent distributions $\pi = \lim \pi_n$ for cluster sizes. Further if the limiting mean cluster size $\lim_{n\to\infty} \sum_1^\infty j\pi_n(j) = \theta^{-1}$ then $\tau/\theta^{-1} = \theta\tau$ clusters may be expected (the total expected number of exceedances being $n(1-F(u_n)) \rightarrow \tau$). This intuitive reasoning is summarized more formally as follows:

Theorem 4.1 *Let $(X_n, n = 1, 2, \ldots)$ be stationary and u_n levels with $n(1-F(u_n)) \rightarrow \tau$. Suppose the cluster size distribution $\pi_n \xrightarrow{w} \pi$, a probability distribution and let the exceedance point process N_n be strongly mixing and the mean cluster size $\mu_n = \sum_{j=1}^\infty j\pi_n(j) \rightarrow \theta^{-1}$, $0 < \theta \leq 1$. Then the exceedance point process N_n converges in distribution to a Compound Poisson Process $N = CP(\theta\tau, \pi)$ based on a Poisson Process with intensity $\theta\tau$ and event multiplicity distribution π.*

This result is a variant of Theorem 4.2 of [5] and is simply shown from that theorem. The parameter θ – which has important uses in extremal theory – is termed the "extremal index" of the sequence $\{X_n\}$. The conclusion then is that the exceedances of a very high level are potentially well modeled by (limiting) clusters with independent size distributions π, located at the points of a Poisson Process with mean $\theta\tau$.

The quantity $N(B)$ is the (limiting) number of (normalized) exceedances in the set B. If unit damage is associated with each exceedance, then $N(B)$ is also the total damage in B. Put in another way, the magnitudes associated with each Poisson point (which has distribution π) can be regarded as the damage from individual clusters, in the limit.

4.4 Excess height damage measure.

Here we return to the example of Section 1, where the damage associated with an exceedance $X_i > u_n$ is proportional to the excess value $(X_i - u_n)_+$.

The resulting damage may be regarded as a point process Z_n with points at the normalized exceedances $\{i/n : X_i > u_n\}$ having multiplicities $a_n(X_i - u_n)$ for suitably chosen constants a_n. It should perhaps be noted that in a point process with multiple events the multiplicities are usually integers. However events may be permitted to have non integer valued multiplicities provided the *number* of events (and therefore also their total mass) in any bounded interval, is finite. Alternatively Z_n may be regarded as the exceedance "marked" point process with marks given by the damages $a_n(X_i - u_n)$.

Assuming the r.i.f. strong mixing conditions for Z_n, let $\{k_n\}$ be a standard sequence and define k_n intervals $J_i \subset (0,1]$ exactly as above for the exceedance point process. Write now π'_n for the *cluster damage distribution* defined by the d.f.

$$P\{Z_n(J_1) \leq x | Z_n(J_1) > 0\} = P\{Z_n(J_1) \leq x | N_n(J_1) > 0\}$$

The following result then corresponds to Theorem 4.1.

Theorem 4.2 *Let $\{X_n,\ n = 1, 2, \ldots\}$ be stationary and u_n levels with $n(1 - F(u_n)) \to \tau$. Let the r.i.f. Z_n defined above be strongly mixing, the mean exceedance cluster size $\mu_n \to \theta^{-1}$, and the cluster damage distribution $\pi'_n \xrightarrow{w} \pi'$, a probability distribution. Then the damage point process Z_n converges in distribution to a Compound Poisson Process $CP(\theta\tau, \pi')$.*

This result may be proved by similar methods to the previous one. Note that strong mixing for Z_n implies strong mixing for N_n and consequently a standard sequence $\{k_n\}$ for Z_n is also a standard sequence for N_n.

4.5 Peaks over threshold modeling.

The two previous damage mechanisms, viz. unit damage and damage $X_i - u_n$ respectively, impose potentially complicated estimation requirements to determine the limiting cluster size and cluster damage distributions π_n, π'_n. The third example is much simpler in this regard, and in fact the limiting damage distribution involves only the marginal distribution F and has a parametric (Pareto) form, requiring estimation of only two parameters.

In further contrast to the previous cases the damage here does not arise from individual exceedances but from assumed proportionality of cluster damage to the maximum height above the level in the cluster i.e. $W_i = \max\{X_j - u_n;\ \ j \in C_i\}$ for a cluster C_i. The corresponding point process Z_n^* is not defined at all exceedance points but rather simply plots $a_n W_i$ (for some constant a_n) at the location of the cluster. This location may be any cluster point - e.g. the first - since the (normalized) cluster coalesces to a single point as $n \to \infty$. The model has been traditionally used widely without complete justification in hydrology – where the damage may represent e.g. flood depth.

It will be assumed that the marginal d.f. F satisfies the tail condition ($\overline{F} = 1 - F$).

(4.2) $$\overline{F}(u + xg(u))/\overline{F}(u) \to \overline{G}(x) \quad \text{as} \quad u \to \infty$$

for some function $g(u) > 0$, some d.f. G and all positive x in the range where $0 < F(x) < 1$. This includes a very wide class of d.f.'s F. Further it is known [13] that any such limit

G must have "generalized Pareto" form viz.

$$(4.3) \qquad G(x) = G_{\alpha,\beta}(x) = 1 - (1 + \alpha x/\beta)^{-1/\alpha} \quad \beta > 0 \;\; \alpha \neq 0$$
$$= 1 - e^{-x/\beta} \qquad \beta > 0 \;\; \alpha = 0$$

the range of x being $(0, \infty)$ if $\alpha > 0$ and $(0, -\alpha^{-1}\beta)$ if $\alpha < 0$.

Assuming that Z_n^* is strongly mixing, intervals J_i corresponding to a standard sequence are defined as before, leading to the *cluster damage distribution* π_n^* defined as the distribution of $Z_n^*(J_1)$ given $N_n(J_1) > 0$.

The following result then holds (cf. Theorem 3.2 of [7])

Theorem 4.3 *Let X_n be a stationary sequence with marginal d.f. F satisfying (4.2) and u_n levels with $n(1 - F(u_n)) \to \tau$. Let the r.i.f. Z_n^* defined above with $a_n = g(u_n)$ (g from (4.2)) be strongly mixing, and the mean exceedance cluster size $\mu_n \to \theta^{-1}$, $0 < \theta < 1$. Then the cluster damage distribution $\pi_n^* \xrightarrow{w} \pi_G$, the probability distribution with d.f. G from (4.2), and the "peaks over thresholds" point process Z_n^* converges to a Compound Poisson process $CP(\theta\tau, \pi_G)$.*

Thus in this case the limiting cluster damage distribution has Pareto form and is determined by two parameters.

5 Lower levels and test statistics.

The previous sections exhibit Compound Poisson models for damage by exceedances of very high levels, obtained as limits as the level $u_n \to \infty$ at the fast rate dictated by the existence of a finite limit for $n(1 - F(u_n))$. In practice one may be concerned with levels that are high, but not sufficiently high for a good Compound Poisson approximation. For example, environmental regulation may be written in terms of a lower level for particularly harmful pollutants, or it may be convenient for enforcement monitoring to regulate by allowing more exceedances of a lower level.

Further it is often the case that one wishes to model $X_1, X_2, \ldots$ as a (stationary) *normal* sequence. However normal sequences do not exhibit clustering at the very high levels required for Poisson exceedance behavior. In fact it can be shown that the limiting mean cluster size θ^{-1} of Section 4 is 1 for such sequences under commonly used dependence conditions. Hence clustering is incompatible with normality at those high levels even though e.g. serial correlation at unit lag may be very high. Hence to use underlying normal models for the sequence $X_1, X_2, \ldots$ it is necessary to consider lower exceedance levels if clustering is an essential feature.

The Poisson-type approximation relies on the convergence rate of $u_n \to \infty$ to yield a finite limit τ for $n(1 - F(u_n))$. If $u_n \to \infty$ more slowly, so that $n(1 - F(u_n)) \to \infty$, a normal limit may be expected for damage measures $Z_n(B)$ of the types considered. (This is certainly intuitively clear in the simplest iid cases from Poisson and normal convergence for binomial r.v.'s). In such cases one of course does not have point process limits, but rather a normal limit for a standardized version of $N_n(B)$ and independent normal limits for such r.v.'s evaluated for disjoint sets B.

It is of course not automatically true that a normal approximation will be better for modeling at lower levels. Certainly for a given n (sample size or time period of interest) the corresponding u_n will tend to be much higher for the Poisson vis-à-vis normal approximation, but it may be that normal limits require much larger values of n to provide good, finite n approximations. However unless this is known to be the case it would seem sensible to try the normal approximation in practical cases involving less extreme levels.

The use of such normal approximations does not seem to be as extensive as is perhaps desirable in high level exceedance modeling. Some cases have been investigated for statistical purposes and could be used in the modeling context. Specifically in [14] asymptotic statistics of the form

$$Z_n = \sum \psi(X_i - u_n)_+$$

for strongly mixing X_i and levels u_n satisfying $n(1 - F(u_n)) \to \infty$, are investigated for a class of functions ψ. Specialization to $\psi(x) = x$ yields the Hill estimator and equivalently the damage model of Section 4.4, and specialization to $\psi(x) = 1\{x > 0\}$ gives again the exceedance point process (and its limiting distribution for the lower u_n-levels).

This latter choice of ψ is also important for tail inference. The asymptotic mean cluster size θ^{-1} for very high level exceedances ($n(1 - F(u_n)) \to \tau$) is, as noted, an important parameter. The simplest estimate of θ is simply the ratio of the number of observations clusters to the number of exceedances of the level u_n. However this is not a consistent estimator, due to the fact that as n increases the number of "observations" (e.g. exceedances) does not tend to infinity but has a non degenerate (approximately Poisson) distribution with mean τ.

Consistency can of course be attained by replication to consider exceedances of u_n (rather than u_{nm_n}) by say nm_n X_i's where $m_n \to \infty$. But this is equivalent to the use of the original n X_i's and a lower level u_n, satisfying $n(1 - F(u_n)) \to \infty$. That is while θ is a parameter relevant to exceedance properties at very high levels, exceedances at lower levels must be used for its consistent estimation.

Finally we note that the field of tail inference for dependent sequences is a developing one, and refer to [14] and [4] for currently available techniques and results.

6 Continuous time.

It is perhaps worth indicating, without detailed discussion, the relevance of the r.i.f. framework to corresponding continuous time problems. The analog of the exceedance point process for a stationary continuous parameter process $\{X_t, t \geq 0\}$ is the *exceedance random measure*

$$Z_T(B) = \int_{T.B} 1\{X_t > u_T\}dt$$

i.e. the time in the set $T.B$ ($B \subset (0,1]$) which the process spends above u_T. This defines (replacing T by n, at least) an r.i.f. to which the limit theory again applies under strong mixing. Specifically for appropriately high levels, this random measure converges in distribution to a Compound Poisson Process, where the Poisson events correspond to high level upcrossings (normalized by T), marked by the (asymptotic) extent of the excursion above u_T from that upcrossing to the next downcrossing.

This provides a convenient model for e.g. occurrence and damage from storms where damage is measured as storm duration. The replacement of $1\{X_t > u_T\}$ by $(X_t - u_T)_+$ would give a similar model where the damage from an excursion by X_t above u_T is the area formed above u_T in that excursion.

Results of this type for exceedance random measures have been obtained in [8]. For Gaussian processes more explicit results can be found ([8]) and this is also true for a potentially useful class of processes with "deterministic peaks" studied in [9]. Potential uses for statistical inference have not been investigated but these and similar statistics will clearly be relevant for estimation of tail properties for continuous time processes.

References

[1] Bergström, H., A comparisons method for distributions functions of sums of independent and dependent random variables, *Theor. Probab. Appl.*, **15** (1970), 430-457.

[2] Bergström, H., Reduction of the limit problem for sums of random variables under a mixing condition, *Proc. 4th Conf. on Probab. Theory*, Brasov (1973), 107-120.

[3] Denker, M., Uniform integrability and the central limit theorem for strongly mixing processes. In: *Dependence in Probability and Statistics*, Eds. Eberlein, E. and Taqqu, M.S., Birkhäuser, Boston (1986), 269-274.

[4] Hsing, T., On tail index estimation using dependent data, *Ann. Statist.*, **19**, 1991, 1547-1569.

[5] Hsing, T., Hüsler, J., Leadbetter, M.R., On the exceedance point process for a stationary sequence, *Probab. Theory and Related Fields*, **78**, 1988, 97-112.

[6] Ibragimov, I. and Linnik, Yu.V., *Indepedent and Stationary Sequences of Random Variables*, Walter Noordhoff, Groningen (1971).

[7] Leadbetter, M.R., On a basis for 'Peaks over Threshold' modeling, *Statist. and Probab. Letters*, **12**, 1991, 357-362.

[8] Leadbetter, M.R. and Hsing, T., Limit theorems for strongly mixing stationary random measures, *Stoch. Proc. Appl.*, **36**, (1990), 231-243.

[9] Leadbetter, M.R. and Hsing, T., On multiple level excursions by stationary processes with deterministic peaks, to appear.

[10] Loéve, M. *Probability Theory*, 4th ed., Springer, New York (1977).

[11] Leadbetter, M.R. and Rootzén, H., On central limit theory for families of strongly mixing additive random functions, *Festschrift in Honour of G. Kallianpur*, S. Cambanis, et al., eds., Springer-Verlag, NY 1993, 211-223.

[12] Philipp, W., The central limit problem for mixing sequences of random variables, *Z. Wahrsch. verw. Geb.*, **12** (1969), 155-171.

[13] Pickands, J., Statistical inference using extreme order statistics, *Ann. Statist.*, **3**, 1975, 119-131.

[14] Rootzén, H., Leadbetter, M.R., de Haan, L., Tail and quantile estimation for strongly mixing stationary sequences, UNC Center for Stochastic Processes Technical Report No. 292, April 1990.

[15] Samur, J., Convergence of sums of mixing triangular arrays of random vectors with stationary rows, *Ann. Probab.*, **12** (1984), 390-426.

[16] Yori, K. and Yoshihara, K.-I., A note on the central limit theorem for stationary strong-mixing sequences, *Yokohama Math. J.*, **34** (1986), 143-146.