Annales Univ. Sci. Budapest., Sect. Comp. 34 (2011) 67–86

SOME JOINT WORK WITH JANOS GALAMBOS

Enrique Castillo (Santander, Spain)

Dedicated to Janos Galambos on the occasion of his 70th birthday

Abstract. In this paper the joint work of the author with Janos Galambos is summarized. This includes some extreme value problems, some conditional specified distributions and some characterization of distributions results. In addition, the solution of a functional equation arising in fatigue and crack growth problems is solved and discussed. Finally, some personal experiences shared with Prof. Janos Galambos are reported.

1. Introduction and motivation

Janos Galambos is well known mainly for his seemingly work on extreme value distributions. The field of extremes, maxima and minima of random variables, has attracted the attention of engineers, scientists, probabilists, and statisticians for many years. The fact that engineering works need to be designed for extreme conditions forces us to pay special attention to singular values more than to regular (or mean) values. Since the statistical theory for dealing with mean values is very different from the statistical theory required for extremes, one cannot solve the above indicated problems without a specialized knowledge of statistical theory for extremes. Applications of extremes include all fields of research, for example, ocean, structural and hydraulics

Key words and phrases: Characterization of distributions, conditional specification of distributions, domains of attraction, extreme values, fatigue models.

https://doi.org/10.71352/ac.34.067

engineering, meteorology, and the study of material strength, traffic, corrosion, pollution, and so on.

After finishing my Civil Engineering studies at the University of Cantabria and later my Ph. D. degrees, one awarded by Northwesetern University and another by the Polytechnical University of Madrid, I soon realized that an engineer cannot survive without knowledge of extremes, because engineering design must be based on extraordinary values, such as large loads, winds, waves, earthquakes, etc. or smaller values, such as strengths, water supplies, etc. Consequently, I was motivated to study and analyze the problem of extremes.

One of my first works ([24]) after joining the University of Cantabria consisted of developing a model to reproduce the random wave behavior with the aim of designing maritime works. The model reproduced the occurrence of Poissonian storms having each a random number of sea states with a Poissonian random number of waves, which were assumed to follow a Rayleigh distribution. The aim of this model was to derive the statistical distribution of the largest waves.

Later I was concerned with approximating distribution functions in their tails, motivated by the fact that in the engineering practice you are forced to extrapolate the existing data. More precisely, based on a reduced set of data you are asked to predict the probability of occurrence of very large (larger than the observed) waves. This led to the publication of two papers in which the tail approximation was dealt with from different points of view ([25] and [26]).

During that period I encountered for the first time to Janos, not physically, but throughout his first extremes book ([30]), in which I discovered many interesting results on the problem of extremes. I must recognize here with gratitude that having the opportunity to read his book changed my life.

In 1983 the NATO ISI Advanced Study Institute on extremes was held an Vimeiro (Portugal), where, among other important researchers, participated Janos Galambos from Temple University, Barry C. Arnold, from the University of California (Riverside) and Richard Smith, from the Imperial College (London). There I met them for the first time and the possibility of spending a sabbatical year at Temple University arose. Later Janos and Eva Galambos facilitated me to expend my sabbatical year at Temple, where I moved with my wife and four children. Janos was kindly involved in finding a house for my family.

During my sabbatical leave at Temple, we were supposed to write a joint book on extremes, but soon Wiley, the Editorial company of Janos' book, claimed about the possibility of the new book competing with his old book, and Janos, who already had initiated writing the book with me, had to abandon the project. This is indicated in the prolog of my book ([6]), in which it is indicated that some pages are due to Janos. Two years after my stay at Temple, Janos' book became out of print, and he published the new edition ([31]).

During that year I worked very hard in the book, that was finished at the end of my sabbatical stay and published a little later by Academic Press.

In the book I tried to bring to the engineers and applied people the already existing theory on extremes and complementing it with new methods and ideas in order to make the important advances useful to people not expert in Statistics. I also was concerned with including practical examples with real data to illustrate some possible applications and motivate researchers.

One of the main contributions included in the book, and discussed with Janos, is an important theorem, that permits deciding the domain of attraction of a given distribution and provides sets of sequences that allow the convergence to the limit distributions. Before, three theorems were required, one for the Weibull, one for the Gumbel and one for the Frechet types domains of attraction. If instead of the theoretical distribution you have data, which is the usual case in practice, the previous method is useless and you must use the curvature method, that we developed together with J. M. Sarabia and presented at Oberwolfach in 1987 ([20, 19]). We summarize these methods below.

2. Characterization of domains of attraction

Theorem 1. (Characterizing domain of attractions for maxima.) *A necessary and sufficient condition for the continuous cumulative distribution function (cdf) $F(x)$ to belong to the maximal domain of attraction, $H_\kappa(x)$, that is, for*

$$\lim_{n\to\infty} [F(a_n + b_n x)]^n = H_\kappa(x), \quad \forall x \tag{1}$$

is that

$$\lim_{\varepsilon\to 0} \frac{F^{-1}(1-\varepsilon) - F^{-1}(1-2\varepsilon)}{F^{-1}(1-2\varepsilon) - F^{-1}(1-4\varepsilon)} = 2^{-\kappa}, \tag{2}$$

where κ is the shape parameter of the associated limit distribution.

This implies that

1. *If $\kappa > 0,\ F(x)$ belongs to the maximal Weibull domain of attraction.*

2. *If $\kappa = 0,\ F(x)$ belongs to the maximal Gumbel domain of attraction, and*

3. *If $\kappa < 0,\ F(x)$ belongs to the maximal Frechet domain of attraction.*

The sequences a_n and b_n can be chosen as:

1. Weibull:

$$a_n = w(F) \quad \textit{and} \quad b_n = w(F) - F^{-1}\left(1 - \frac{1}{n}\right), \tag{3}$$

2. Gumbel:

$$a_n = F^{-1}\left(1 - \frac{1}{n}\right) \quad \textit{and} \quad b_n = F^{-1}\left(1 - \frac{1}{ne}\right) - a_n, \tag{4}$$

3. Fréchet:

$$a_n = 0 \quad \textit{and} \quad b_n = F^{-1}\left(1 - \frac{1}{n}\right), \tag{5}$$

where $w(F) = \sup\{x|F(x) < 1\}$ is the upper end of $F(x)$.

In addition to the book, I still had some time to work with Janos in some papers, some of them were finished during this year, and some later as [13, 14, 15, 16, 17, 18, 19, 21].

3. Normal conditionals distributions

My first contact with conditional specification took place at Temple University (Philadelphia) in 1985 when I was working with Janos Galambos. One Thursday he invited me to write a paper to be presented at a Conference to start 10 days later. He said:

> "I have been invited to deliver one of the main talks at a Conference on Weighted Distributions at Penn State University in Pittsburgh. Since I am going to Pittsburgh by car, you can come with me, and you can finance your hotel expenses with some small financial support you will receive because of presenting the paper."

Then, I asked Janos "what are weighted distributions?", and joking he answered me: "I don't know". So, I had ten days to write a paper on a topic I did not know.

During that weekend I decided what paper to present and I was working very hard to be able to show Janos some work the next Monday. Since I knew that all conditionals of a bivariate normal were normal, I asked myself about the possibility of existing other bivariate distributions with the same property. Fortunately, I had some experience on functional equations, and I was able to solve the problem producing the first draft. From this draft, we generated the joint paper, which in summary was as follows.

Assume a bivariate absolutely continuous distribution whose marginal and conditional probability density functions, respectively, are $f_{(X,Y)}(x,y)$, $g(x)$, $h(y)$, $f_{X|Y}(x|y)$ and $f_{Y|X}(y|x)$. It is evident that

$$f_{(X,Y)}(x,y) = f_{X|Y}(x|y)h(y) = f_{Y|X}(y|x)g(x). \tag{6}$$

If we assume that all conditional distributions are normal, we must have

$$\begin{aligned} f_{Y|X}(y|x)g(x) &= \frac{\exp\left\{-\frac{1}{2}\left[\frac{y-a(x)}{b(x)}\right]^2\right\}}{b(x)}g(x) = \\ = f_{X/Y}(x|y)h(y) &= \frac{\exp\left\{-\frac{1}{2}\left[\frac{x-d(y)}{c(y)}\right]^2\right\}h(y)}{c(y)}, \end{aligned} \tag{7}$$

where $b(x) > 0, c(y) > 0$ and $a(x)$ and $d(y)$ define the regression lines and $b(x)$ and $c(y)$ are the corresponding standard deviations. Solving this equation, we have:

$$a(x) = \frac{-(A+Bx+Cx^2)}{(D+2Ex+Fx^2)}; \quad d(y) = \frac{-(H+By+Ey^2)}{(J+2Cy+Fy^2)}; \tag{8}$$

(9)

$$b^2(x) = \frac{1}{(D+2Ex+Fx^2)}; \quad c^2(y) = \frac{1}{(J+2Cy+Fy^2)}; \tag{10}$$

(11)

$$u(x) = -\tfrac{1}{2}\left[G+2Hx+Jx^2 - \frac{(A+Bx+Cx^2)^2}{(D+2Ex+Fx^2)}\right]; \tag{12}$$

(13)

$$v(y) = -\tfrac{1}{2}\left[G+2Ay+Dy^2 - \frac{(H+By+Ey^2)^2}{(J+2Cy+Fy^2)}\right]; \tag{14}$$

(15)

$$g(x) = (D+2Ex+Fx^2)^{-1/2}\exp\left\{-\frac{G+2Hx+Jx^2-\frac{(A+Bx+Cx^2)^2}{(D+2Ex+Fx^2)}}{2}\right\}; \tag{16}$$

(17)

$$h(y) = (J+2Cy+Fy^2)^{-1/2}\exp\left\{-\frac{G+2Ay+Dy^2-\frac{(H+By+Ey^2)^2}{(J+2Cy+Fy^2)}}{2}\right\}; \tag{18}$$

(19)

- $f(x,y) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{G}{2}\right\}$

- $\exp\left\{-\frac{1}{2}[2Hx+2Ay+Jx^2+Dy^2+2Bxy+2Cx^2y+2Exy^2+Fx^2y^2]\right\}.$

For the function $f(x,y)$ to be a density function, constants $\{A,B,C,D,E,F,G,H,J\}$ must satisfy the following conditions:

$$\begin{array}{l} \text{(i)}\ \ F=E=C=0,\quad D>0,\quad J>0,\quad B^2<DJ, \\ \text{(ii)}\ \ F>0,\quad FD>E^2,\quad JF>C^2. \end{array}$$

Model (i) is the bivariate normal model, and Model (ii) has the following properties:

- Regression lines are not straight lines
- Marginal distributions are not normal.
- The mode or the modes are at the intersection of the regression lines.

The surprising result is that there exists another family with normal conditionals. Figure 1 shows three examples of distributions with normal conditionals, where the upper figure corresponds to the normal case and the other two are not normal, one being unimodal and the other an interesting bimodal example.

These results were published at that conference ([13]). Professor Barry Arnold from the University of California (Riverside) was at that conference and two months later wrote Janos to ask for permision to publish a paper inspired in ours but dealing with exponential conditionals instead of normals (see [5]). Our results were also published in a congress in Cairo ([15]) and one extension to characterize normal distributions in [17]. A little later we solved the problem of Weibull conditionals in [18], gamma conditionals in [21] and beta kind two in [27].

I must recognize here the importance of being familiar with functional equations and the important role played by Janos Aczél's book ([1]) on making this possible. I cannot understand how functional equations, which are as important and powerful as differential equations (see, [10], [9]) are not studied in regular courses at the graduate level.

When returning to Spain, we invited Janos to deliver a course at the University of Cantabria in Spain, and the Estadística Española Journal invited Janos, Sarabia and I to write a paper ([21]) in which we dealt with the characterization of bivariate models with gamma conditionals. Later, I met Janos in Budapest at the Loránd Eötvös University, where he started before his stays in two African universities just before arriving in Philadelphia (Temple University). There I contacted Imre Kátai, the organizer of this special issue dedicated to Janos, and worked with some colleagues in ([2]).

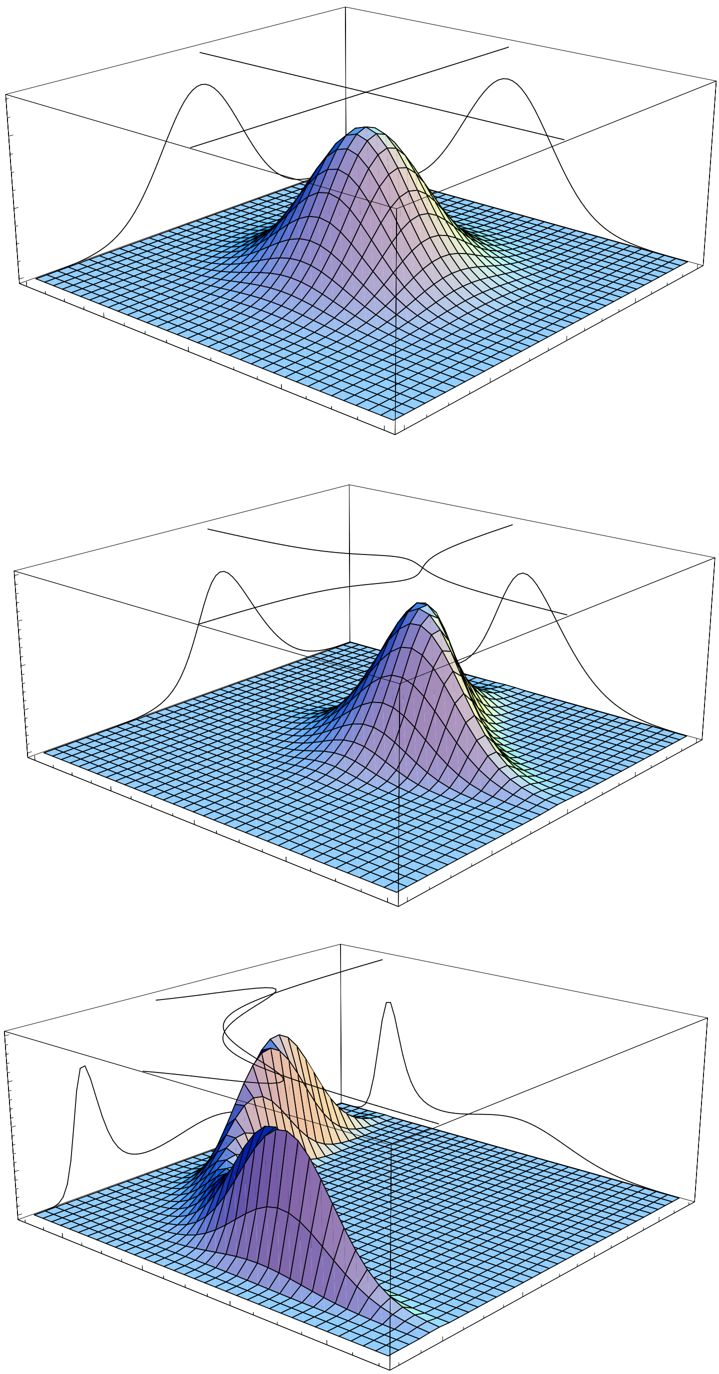

Figure 1. Examples of distributions with normal conditionals showing the joint densities, the marginals as horizontal projections, and the regression lines on a top projection. The upper figure corresponds to the normal case. The other two are not normal, one being unimodal and the other an interesting bimodal example.

In 1993 Janos Galambos, who was a consultant at NIST (National Institute of Standards and Technology in Gaithersburg, Maryland (US)) organized a Conference on Extremes, and he invited me to teach a course on extremes, to deliver one of the main invited talks ([7]), and to participate in the panel discussion on the future of extreme value theory and its applications, together with Laurens de Haan, Lucien La Cam and Richard L. Smith. At the conference I also presented one expert system for extreme value analysis ([8]), in which all the steps to be followed to select the adequate distribution for engineering design based on extremes are explained and the user is guided to avoid common errors.

Soon, I was invited to write a paper on extreme wave heights by the Journal of Research of the National Institute of Standards and Technology ([29]) and later another in [28].

Next, I worked on extremes with Barry Arnold and José María Sarabia in [3] and several years later I published the second edition of my book ([23]) with Ali S. Hadi (Cornell University), N. Balakrishnan (Mc Master University) and J. M. Sarabia (University of Cantabria).

4. The Curvature Method

The problems of checking whether a sample comes from a maximal Gumbel family of distributions, and the problem of determining whether the domain of attraction of a given sample is maximal Gumbel, are different. In the former case, the whole sample is expected to follow a linear trend when plotted on a maximal Gumbel probability paper, while only the upper tail is expected to exhibit that property for the latter.

Consequently, it is incorrect to reject the assumption that a maximal Gumbel domain of attraction does not fit the data only because the data do not show a linear trend when plotted on a maximal Gumbel probability paper. This is because two distributions with the same tail, even though the rest be completely different, lead to exactly the same domain of attraction, and should be approximated by the same model.

Since only the right (left) tail governs the domain of attraction for maxima (minima), then one should focus only on the behavior of the high- (low-) order statistics. The remaining data values are not needed. In fact, they can be more of a hinderance than of a help in solving the problem. The problem is then to determine how many data points should be considered in the analysis or to decide which weights must be assigned to each data point.

When selecting the weights for different data points, the following considerations should be taken into account:

1. The tail quantiles estimates have larger variance than those not in the tail.

2. Data on the tail of interest have more information on the limit distribution than those not in the tail.

These are contradicting considerations and should be balanced. However, on the other tail the weights must be very small if not zero.

With the aim of illustrating graphically the role of the tails in the extreme behavior, Figures 2 and 3 show well-known distributions on maximal and minimal Gumbel probability papers, respectively. Note that those who are not Gumbel exhibit non-linear trends. Figure 2 shows an almost linear trend for the right tail of the normal distribution, confirming that the normal belongs to the maximal Gumbel domain of attraction. Note also the curvatures in the right tail of the uniform (positive curvature and vertical asymptote) and Cauchy distributions (negative curvature and horizontal trend) showing Weibull and Fréchet domains of attractions for maxima, respectively.

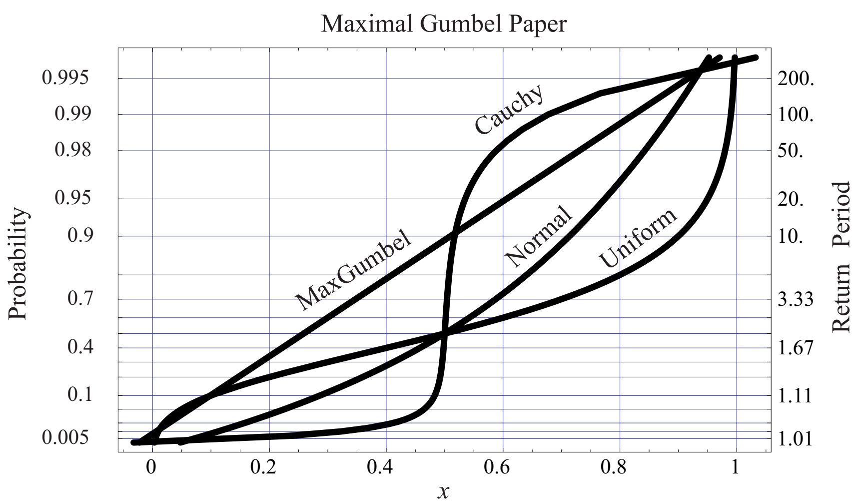

Figure 2. The uniform, normal, Cauchy, and maximal Gumbel distributions plotted on a maximal Gumbel probability paper.

Figure 3 shows an almost linear trend for the left tail of the normal distribution, confirming that it belongs to the minimal Gumbel domain of attraction. Note also the curvatures in the left tail of the uniform (negative curvature and vertical asymptote) and the reversed Cauchy distributions (positive curvature and horizontal trend) showing minimal Weibull and Fréchet domains of attractions, respectively.

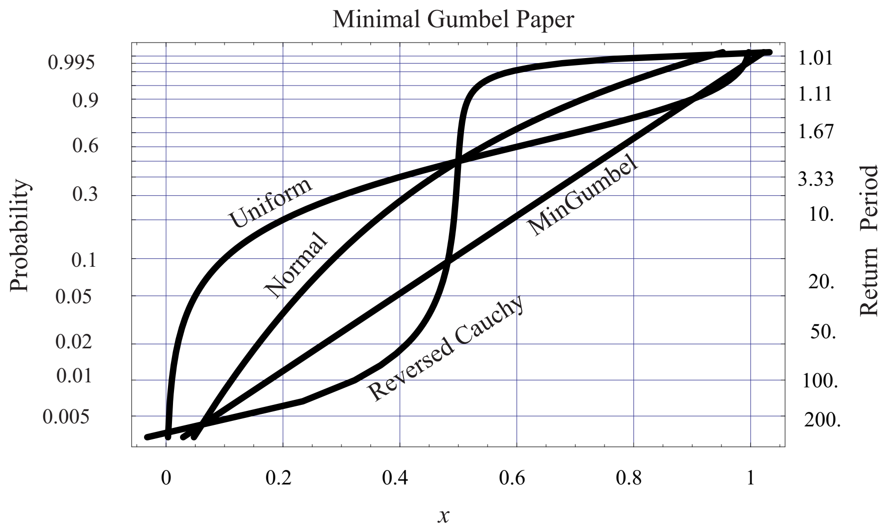

Figure 3. The uniform, normal reversed Cauchy, and minimal Gumbel distributions plotted on a minimal Gumbel probability paper.

The method to be described below was done with Janos Galambos and has the same appealing geometrical property of the basic idea that is used for the probability paper method, i.e., the statistics upon which a decision will be made is based on the tail curvature (see [20]). This curvature can be measured in different ways. For example, by the difference or the quotient of slopes at two points. In addition, any of these two slopes can be measured by utilizing two or more data points. The latter option seems to be better in order to reduce variances. Here we propose to fit two straight lines, by least-squares, to two tail intervals and to use the quotient of their slopes to measure the curvature. More precisely, we use the statistic

$$S = \frac{S_{n_1,n_2}}{S_{n_3,n_4}}, \tag{20}$$

where $S_{i,j}$ is the slope of the least-squares straight line fitted on Gumbel probability paper, to the rth order statistics with $i \leq r \leq j$. Thus, we can write

$$S_{ij} = \frac{m\Sigma_{11} - \Sigma_{10}\Sigma_{01}}{m\Sigma_{20} - \Sigma_{10}\Sigma_{01}}, \tag{21}$$

where $m = n_j - n_i + 1$ and

$$\Sigma_{10} = \sum_{k=n_i}^{n_j} \left\{ -\log\left[-\log\left(\frac{k-0.5}{n} \right) \right] \right\}, \tag{22}$$

$$\Sigma_{01} = \sum_{k=n_i}^{n_j} x_k, \tag{23}$$

$$\Sigma_{11} = \sum_{k=n_i}^{n_j} \left\{ -x_k \log\left[-\log\left(\frac{k-0.5}{n} \right) \right] \right\}, \tag{24}$$

$$\Sigma_{20} = \sum_{k=n_i}^{n_j} \left\{ -\log\left[-\log\left(\frac{k-0.5}{n} \right) \right] \right\}^2, \tag{25}$$

and n is the sample size.

An important property of the least squares slope S_{ij} is that it is a linear combination of order statistics with coefficients which add up to zero. This property makes the statistic S location and scale invariant.

The selection of n_1, n_2, n_3 and n_4 must be based on the sample size and the speed of convergence to the asymptotic distribution, which sometimes can be inferred from the sample. Apart from speed of convergence considerations, we have selected the following values when the right tail is of interest

$$n_1 = n + 1 - [2\sqrt{n}], \qquad n_4 = n, \tag{26}$$

$$n_2 = n_3 = n + 1 - \frac{[2\sqrt{n}]}{2}, \tag{27}$$

where $[x]$ means the integer part of x. The $\sqrt{n}$ is selected to ensure using only high order statistics.

According to the above theory and with the values in (27), if the statistic S is well above 1 we can decide that the domain of attraction is Weibull type. And, if it is well below 1, the decision is in favor of a Fréchet type. However, in order to be able to give significance levels of the test, we need to know the cdf of S. The asymptotic properties of this method have been studied by [20].

If we are interested in the left tail, we can make the change of variables $X = -X$ and use the previous method for the resulting right tail.

5. Gamma Conditionals family

The family of gamma distributions with scale and shape parameters forms a two-parameter exponential family of the form

$$f(x; \theta_1, \theta_2) = x^{-1} e^{\theta_1 \log x - \theta_2 x} \theta_2^{\theta_1} \left[\Gamma(\theta_1)\right]^{-1} I(x > 0). \tag{28}$$

If X has its density of the form (28) then we write $X \sim \Gamma(\theta_1, \theta_2)$. If we require all conditionals to be in the family (28) then the general gamma conditionals class of densities is given by

$$f(x,y) = (xy)^{-1} \exp\left\{ \begin{pmatrix} 1 & -x & \log x \end{pmatrix} M \begin{pmatrix} 1 \\ -y \\ \log y \end{pmatrix} \right\}, \quad x > 0, \quad y > 0, \tag{29}$$

where M is a matrix of parameters.

It remains only to determine appropriate values of the parameters M to ensure integrability of this joint density. Such conditions were provided by Castillo, Galambos, and Sarabia (1990) using a different parametrization. For fixed y, the density $f(x,y)$ is of the form $c(y)x^{\alpha(y)-1}e^{-\beta(y)x}$ for suitable $\alpha(y)$ and $\beta(y)$. For this to be integrable $\alpha(y)$ and $\beta(y)$ must both be positive. The conditional distributions corresponding to (29) are of the form

$$X|Y = y \sim \Gamma(m_{20} + m_{22}\log y - m_{21}y, m_{10} - m_{11}y + m_{12}\log y) \tag{30}$$

and

$$Y|X = x \sim \Gamma(m_{02} + m_{22}\log x - m_{12}x, m_{01} - m_{11}x + m_{21}\log x). \tag{31}$$

Thus, our parameters must be such that all the expressions on the right-hand sides of (30) and (31) are positive. In addition, we must verify that the marginal densities thus obtained are themselves integrable. We have

$$f_X(x) = x^{-1} \frac{\Gamma(m_{02} + m_{22}\log x - m_{12}x) e^{m_{00} - m_{10}x + m_{20}\log x}}{(m_{01} - m_{11}x + m_{21}\log x)^{m_{02} + m_{22}\log x - m_{12}x}}, \quad x > 0, \tag{32}$$

and an analogous expression for $f_Y(y)$. It turns out that under the parametric conditions sufficient to ensure positivity of the gamma parameters in (30) and (31), the function $f_X(x)$ is bounded in a neighborhood of the origin and, for large x, is bounded by $x^{1/2}e^{-\delta x}$ for some $\delta > 0$. Consequently it is integrable. The requisite conditions for a proper density $f(x,y)$ in (29) may be summarized as follows, where five different families are defined:

MODEL I (in this case, X and Y are independent):

$$\begin{array}{llll} m_{11} = 0, & m_{12} = 0, & m_{21} = 0, & m_{22} = 0, \\ m_{10} > 0, & m_{20} > 0, & m_{01} > 0, & m_{02} > 0. \end{array} \tag{33}$$

MODEL II:

$$\begin{array}{llll} m_{11} < 0, & m_{12} = 0, & m_{21} = 0, & m_{22} = 0, \\ m_{10} > 0, & m_{20} > 0, & m_{01} > 0, & m_{02} > 0. \end{array} \tag{34}$$

MODEL IIIA:

$$(35)\qquad \begin{array}{llll} m_{11} < 0, & m_{12} = 0, & m_{21} < 0, & m_{22} = 0, \\ m_{10} > 0, & m_{20} > 0, & m_{02} > 0, & m_{01} > m_{21}\left(1 - \log \dfrac{m_{21}}{m_{11}}\right). \end{array}$$

MODEL IIIB:

$$(36)\qquad \begin{array}{llll} m_{11} < 0, & m_{12} < 0, & m_{21} = 0, & m_{22} = 0, \\ m_{20} > 0, & m_{01} > 0, & m_{02} > 0, & m_{10} > m_{12}\left(1 - \log \dfrac{m_{12}}{m_{11}}\right). \end{array}$$

MODEL IV:

$$(37)\qquad \begin{array}{llll} m_{01} > m_{21}\left(1 - \log \dfrac{m_{21}}{m_{11}}\right), & m_{11} < 0, & m_{12} < 0, & m_{21} < 0, \\ m_{10} > m_{12}\left(1 - \log \dfrac{m_{12}}{m_{11}}\right), & m_{20} > 0, & m_{02} > 0, & m_{22} = 0. \end{array}$$

and finally MODEL V:

$$(38)\qquad \begin{array}{ll} m_{11} < 0, & m_{10} > m_{12}\left(1 - \log \dfrac{m_{12}}{m_{11}}\right), \\ m_{12} < 0, & m_{20} > m_{22}\left(1 - \log \dfrac{m_{22}}{m_{21}}\right), \\ m_{21} < 0, & m_{01} > m_{21}\left(1 - \log \dfrac{m_{21}}{m_{11}}\right), \\ m_{22} < 0, & m_{02} > m_{22}\left(1 - \log \dfrac{m_{22}}{m_{12}}\right). \end{array}$$

The regression functions for the gamma conditionals distribution are of course generally nonlinear. We have

$$(39)\qquad E(X|Y = y) = \frac{m_{20} + m_{22}\log y - m_{21}y}{m_{10} + m_{12}\log y - m_{11}y}$$

and

$$(40)\qquad E(Y|X = x) = \frac{m_{02} + m_{22}\log x - m_{12}x}{m_{01} + m_{21}\log x - m_{11}x},$$

which are obtained using (30) and (31). Expressions for the conditional variances can also be written by referring to (30) and (31). As a curiosity, we may note that certain fortuitous parametric choices can lead to $E(X|Y = y) = c_1$ and $E(Y|X = x) = c_2$, i.e., $m_{20}/m_{10} = m_{22}/m_{12} = m_{21}/m_{11}$, etc. These will generally not correspond to independent marginals since the corresponding conditional variances will not be constant.

The modes of this distribution are at the intersection of the conditional mode curves, that is, they are the solution of the system

$$\begin{aligned} x &= \frac{m_{20} + m_{22}\log y - m_{21}y - 1}{m_{10} + m_{12}\log y - m_{11}y}, \\ y &= \frac{m_{02} + m_{22}\log x - m_{12}x - 1}{m_{01} + m_{21}\log x - m_{11}x}, \end{aligned}$$

where we assume that $m_{20} > 1$ and $m_{02} > 1$.

A more detailed analysis of these models are in [4].

6. Weibull Conditionals family

Definition 1. *Weibull distribution We say that X has a Weibull distribution if*

$$P(X > x) = \exp[-(x/\sigma)^{\gamma}], \quad x > 0, \tag{41}$$

where $\sigma > 0$ and $\gamma > 0$.

If (41) holds we write $X \sim \text{Weibull}(\sigma, \gamma)$. The case $\gamma = 1$, corresponds to the exponential distribution. Our goal is to characterize all bivariate distributions with Weibull conditionals, a problem solved in [18]. That is, random variables (X, Y) such that

$$X|Y = y \sim \text{Weibull}(\sigma_1(y), \gamma_1(y)), \quad y > 0, \tag{42}$$

and

$$Y|X = x \sim \text{Weibull}(\sigma_2(x), \gamma_2(x)), \quad x > 0. \tag{43}$$

Interest centers on nontrivial examples in which $\gamma_1(y)$ and $\gamma_2(x)$ are not constant. Writing the joint density corresponding to (42) and (43) as products of marginals and conditionals yields the following functional equation, valid for $x, y > 0$:

$$f_Y(y)\frac{\gamma_1(y)}{\sigma_1(y)}\left[\frac{x}{\sigma_1(y)}\right]^{\gamma_1(y)-1} \exp\left[-\left(\frac{x}{\sigma_1(y)}\right)^{\gamma_1(y)}\right] =$$

$$= f_X(x)\frac{\gamma_2(x)}{\sigma_2(x)}\left[\frac{y}{\sigma_2(x)}\right]^{\gamma_2(x)-1} \exp\left[-\left(\frac{y}{\sigma_2(x)}\right)^{\gamma_2(x)}\right]. \tag{44}$$

For suitably defined functions $\phi_1(y), \phi_2(y), \psi_1(x)$ and $\psi_2(x)$, (44) can be written in the form

$$\phi_1(y)x^{\gamma_1(y)} \exp\left[-\phi_2(y)x^{\gamma_1(y)}\right] = \psi_1(x)y^{\gamma_2(x)} \exp\left[-\psi_2(x)y^{\gamma_2(x)}\right]. \tag{45}$$

This equation is not easy to solve. We may indicate the nature of solutions to (45) by assuming differentiability (a differencing argument would lead to the same conclusions albeit a bit more painfully). Take logarithms on both sides of (45) then differentiate once with respect to y to get a new functional equation, and then again with respect to y to get another functional equation. Now set $y = 1$ in these two equations and in (45). This yields the following three equations for $\psi_1(x), \psi_2(x)$ and $\gamma_2(x)$:

$$
\begin{aligned}
&(46) & c_1 x^{\gamma_1} e^{-c_2 x^{\gamma_1}} &= \psi_1(x) e^{-\psi_2(x)},\\
&(47) & c_3 + c_4 \log x &= -c_5 x^{\gamma_1} - c_6(\log x)x^{\gamma_1} + \gamma_2(x)\left[1 - \psi_2(x)\right],\\
& & c_7 + c_8 \log x &= -c_9 x^{\gamma_1} - c_{10}(\log x)x^{\gamma_1} - c_{11}(\log x)^2 x^{\gamma_1}\\
&(48) & & \quad - \gamma_2(x) - \psi_2(x)\gamma_2(x)\left[\gamma_2(x) - 1\right].
\end{aligned}
$$

These three equations may be solved to yield $\gamma_2(x), \psi_1(x)$, and $\psi_2(x)$. See [18] for details. A similar approach will yield expressions for $\phi_1(y), \phi_2(y)$, and $\gamma_1(y)$. Unfortunately, it appears to be difficult to determine appropriate values for all of the constants appearing in the solutions to guarantee compatibility in a nontrivial situation.

7. Solving a functional equation arising in fatigue models

In 1984, working with Alfonso Fernández Canteli at the ETH in Zürich, on a fatigue problem we arrived at the following functional equation (see [12], [22], [11]):

$$1 - \exp\left\{-\left[a(x)y + b(x)\right]^{c(x)}\right\} = 1 - \exp\left\{-\left[d(y)x + e(y)\right]^{f(y)}\right\}; \quad \forall x, y \tag{49}$$

which states a compatibility condition on two Weibull distributions (see Figure 4).

At that time, using a theorem of [1], Alfonso and I were able to solve it for the particular case of $c(x) = constant$ and $f(y) = constant$ and obtained very interesting results for modeling the fatigue Wöhler fields.

Later with Janos, and using asymptotic theory, we were able to solve this functional equation in general ([16]) for the case in which

$$1 - \exp\left\{-\left[a(x)y + b(x)\right]^{c(x)}\right\} = 1 - \exp\left\{-\left[d(y)x + e(y)\right]^{f(y)}\right\}; \quad \forall x, y \tag{50}$$

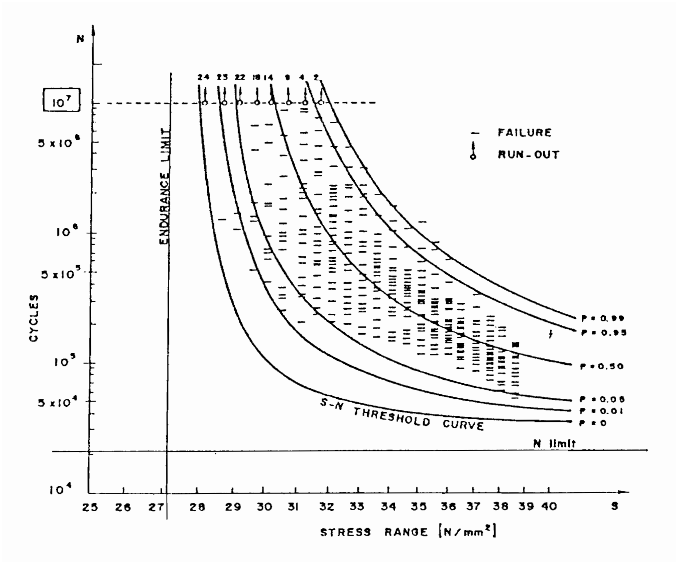

Figure 4. Illustration of the S-N field with the percentile curves associated with Weibull distributions.

is a cdf as a function of y for all x and, at the same time, a cdf as a function of x for all y. The main difficulty arises when $c(x)$ and $f(y)$ are not constant or even identical. The final result is that the cdf in (50) must belong to one of the two families:

$$1 - \exp\left[-D(Ex+F)^{C\log(Ay+B)}\right] \tag{51}$$

or

$$1 - \exp\left\{-\left[C(x-A)(y-B)+D\right]^{E}\right\}, \tag{52}$$

where A, B, C, D, E and F are constants.

One application of the first model is given in [11]. The second model is being studied at the present time.

Recently, we have discovered that Equation (50) has interest and arises in crack and damage growth models. This allows us to conclude that for large size of long pieces of material it is justified the use of straight lines or hyperboles as crack or damage growth curves.

8. Some characterizations of the normal distribution

In 1989 we published a joint work ([17]) in which we characterized the normal distribution based on conditionals and other properties. In particular we stated the following theorem.

Theorem 2. *Assume that $f(x, y)$ is a bivariate density whose conditionals $f_x(x)$ and $g_y(y)$ are univariate normal densities. Denote the expectation and variance parameters as $m_1(y)$ and $\sigma_1^2(y)$ for $g_y(x)$ and by $m_2(x)$ and $\sigma_2^2(x)$ for $f_x(y)$. Then, $f(x, y)$ is normal if, and only if, any of the following properties holds:*

1. *$\sigma_1(y)$ or $\sigma_2(x)$ is constant.*
2. *as $y \to +\infty$, $y^2\sigma_1(y) \to +\infty$ (or similar property for $\sigma_2(x)$ holds),*
3. *as $y \to +\infty$, $\liminf \sigma_1(y) \neq 0$ (or a similar property for $\sigma_2(x)$).*

Under any one of these conditions, $m_1(y)$ and $m_2(x)$ are linear.

9. Epilogue

It has been a pleasure to participate in this special issue of the Annales Universitatis Scientiarum Budapestinensis de Rolando Eotvos Nominatae Sectio Computatorica to honor Prof. Janos Galambos on the occasion of his 70th birthday. The scientific community must recognize the role played by Janos on extremes and in other areas of Probability, Statistics and Mathematics (see [32, 33, 34]). I have considered only some of his extreme value contributions, but many other must be acknowledged. I must congratulate Imre Kátai for having the nice idea of this special issue, and thank all participants to make this possible.

The last time a saw Janos was in 2007 on the occasion of a trip to the University of Maryland. From there I traveled to Philadelphia to see Janos. I look forward to see Janos again as soon as possible and transmit him personally my gratitude. Some of my successes have been possible because of you, and thus, they are yours too.

Janos, we wish you a happy birthday in the company of Eva and all of us.

References

[1] **Aczél, J.,** *Lectures on functional equations and their applications*. Mathematics in Science and Engineering, Vol. 19. New York: Academic Press, (1996). Translated by Scripta Technica, Inc. Supplemented by the author. Edited by Hansjorg Oser.

[2] **Alvarez, E., A. Benczúr, E. Castillo and J. Sarabia,** The problem of learning concepts. A probabilistic view, *Annales Univ. Sci. Budapest., Sect. Comp.*, **13** (1992), 179–194.

[3] **Arnold, B.C., E. Castillo and J.M. Sarabia,** Some alternative bivariate Gumbel models, *Environmetrics*, **9**(6) (1998), 599–616.

[4] **Arnold, B.C., E. Castillo and J.M. Sarabia,** Conditionally specified distributions: an introduction, *Statist. Sci.*, **16**(3) (2001), 249–274. With comments and a rejoinder by the authors.

[5] **Arnold, B.C. and D. Strauss,** Bivariate distributions with exponential conditionals, *Journal of the American Statistical Association*, **83** (1998), 522–527.

[6] **Castillo, E.,** *Extreme value theory in Engineering*, California: Academic Press., San Diego (1988).

[7] **Castillo, E.,** Extremes in engineering applications, extreme value theory and applications In: *Proceedings of the Conference on Extreme Value Theory and Applications*, J. Galambos, J. Lechner, E. Simiu, and N. Macri (Eds.), Gaithersburg Maryland 1993, Kluwer Academic Publishers, pp. 15–42.

[8] **Castillo, E., E. Alvarez, A. Cobo and T. Herrero,** An expert system prototype for the analysis of extreme value problems, *National Institute of Standards and Technology, NIST Special Publication* 866, (1994) 85–93.

[9] **Castillo, E., A. Cobo, J.M. Gutirrez and E. Pruneda,** Functional networks: A new network-based methodology, *Computer-Aided Civil and Infrastructure Engineering*, **15**(2) (2000), 90–106.

[10] **Castillo, E., A. Cobo, J. Manuel Gutirrez and E. Pruneda,** Working with differential, functional and difference equations using functional networks, *Applied Mathematical Modelling*, **23**(2) (1999), 89–107.

[11] **Castillo, E., A. Fernández-Canteli and A.S. Hadi,** On fitting a fatigue model to data, *International Journal of Fatigue*, **21** (1999), 97–106.

[12] **Castillo, E. and A. Fernndez-Canteli,** A general regression model for lifetime evaluation and prediction, *International Journal of Fracture*, **107**(2) (2001), 117–137.

[13] **Castillo, E. and J. Galambos,** Modeling and estimation of bivariate distributions with normal conditionals based on their marginals. In: *Conference on weighted distributions*, Penn. State University, 1985.

[14] **Castillo, E. and J. Galambos,** The characterization of a regression model associated with fatigue problems, In: *Conference on Weighted Distributions*, Penn. State University, 1986.

[15] **Castillo, E. and J. Galambos,** Bivariate distributions with normal conditionals, In: *Proceedings of the IASTED International Symposium: Simulation, Modelling and Development SMD '87*, Cairo, pp. 59–62. ACTA Press, 1987.

[16] **Castillo, E. and J. Galambos,** Lifetime regression models based on a functional equation of physical nature, *Journal of Applied Probability*, **24** (1987), 160–169.

[17] **Castillo, E. and J. Galambos,** Conditional distributions and the bivariate normal distribution, *Metrika, International Journal for Theoretical and Applied Statistics*, **36**(3) (1989), 209–214.

[18] **Castillo, E. and J. Galambos,** Bivariate distributions with Weibull conditionals, *Analysis Mathematica*, **16**(1) (1990), 3–9.

[19] **Castillo, E., J. Galambos and J. Sarabia,** The selection of the domain of attraction of an extreme value distribution from a set of data, proceedings, oberwolfach, Extreme Value Theory, *Lecture Notes in Statistics 51*.

[20] **Castillo, E., J. Galambos and J.M. Sarabia,** The selection of the domain of attraction of an extreme value distribution from a set of data, In: *Extreme value theory (Oberwolfach, 1987)*, Volume 51 (1989), pp. 181–190, New York: Springer.

[21] **Castillo, E., J. Galambos and J.M. Sarabia,** Caracterización de modelos bivariantes condistribuciones condicionadas tipo gamma, *Estadística Española*, **32** (1990), 439–450.

[22] **Castillo, E. and A. Hadi,** Modelling lifetime data with applications to fatigue models, *Journal of the Americal Statistical Association*, **90** (1995), 1041–1054.

[23] **Castillo, E., A.S. Hadi, N. Balakrishnan and J.M. Sarabia,** *Extreme Value and Related Models with Applications in Engineering and Science*, Wiley, New York 2004.

[24] **Castillo, E., M. Losada and J. Puig-Pey,** Probabilistic analysis of the number of waves and their influence on the design wave height of marine structures under dynamic loading, In: *Safety of structures under dynamic loading*, Trondheim (Norway).

[25] **Castillo, E., E. Moreno and J. Puig-Pey,** Criterios minimo-cuadráticos de ajuste de distribuciones de probabilidad a datos experimentales, *Revista de Obras Públicas* 129 (1982), 433–439.

[26] **Castillo, E., E. Moreno and J. Puig-Pey,** Nuevos modelos de distribuciones de extremos basados en aproximaciones en las ramas, *Trabajos de Estadística e Investigación Operativa*, **34** (1983), 6–24.

[27] **Castillo, E. and J.M. Sarabia,** Bivariate distributions with second kind beta conditionals, *Communications in Statistics, Theory and Methods*, **19** (1990), 3433–3445.

[28] **Castillo, E. and J.M. Sarabia,** Engineering analysis of extreme value data, *Journal of Waterway, Port, Coastal, and Ocean Engineering, ASCE*, **118** (1992), 129–146.

[29] **Castillo, E. and J.M. Sarabia,** *Extreme value analysis of wave heights*, *Journal of Research of the National Institute of Standards and Technology*, **99** (1994), 445–454.

[30] **Galambos, J.,** *The Asymptotic Theory of Extreme Order Statistics*, John Wiley and Sons, New York 1978.

[31] **Galambos, J.,** *The Asymptotic Theory of Extreme Order Statistics*, (Second ed.). Malabar, Florida: Robert E. Krieger, 1987.

[32] **Galambos, J.,** *Advanced probability theory* (Second ed.), Volume 10 of Probability: Pure and Applied. New York: Marcel Dekker Inc.

[33] **Galambos, J. and I. Simonelli,** *Bonferroni-type inequalities with applications*, Probability and its Applications (New York). New York: Springer-Verlag, 1996.

[34] **Galambos, J. and I. Simonelli,** *Products of random variables*, Volume 268 of Monographs and Textbooks in Pure and Applied Mathematics, New York: Marcel Dekker Inc. Applications to problems of physics and to arithmetical functions, 2004.

E. Castillo
Department of Applied Mathematics and Computational Sciences
University of Cantabria
39005 Santander Spain
enrique.castillo@unican.es