FITTING EXTREME VALUE COPULAS WITH UNIMODAL CONVEX POLYNOMIAL REGRESSION USING BERNSTEIN POLYNOMIALS

Danielle Gonçalves de Oliveira PRADO [1]
Lucas Monteiro CHAVES[2]
Devanil Jaques de SOUZA[3]
Eleanderson Campos EUGÊNIO FILHO[3]

- ABSTRACT: Bernstein polynomials are suitable for performing shape-constrained regressions, in particular, for unimodal convex regression. The Pickands function is convex and unimodal, being a fundamental element in the theory of extreme value copulas. The purpose of this article is to explain in details the use of Bernstein polynomials in the estimation of Pickands function and to establish a new test of significance for extreme value copulas.
- KEYWORDS: Bernstein polynomials, Pickands function, extreme value copula.

1 Introduction

Bernstein polynomials are suitable for performing shape-constrained regressions, particularly for unimodal convex regression. The Pickands function (PICKANDS, 1981), a fundamental element in the theory of extreme value copulas, is convex and unimodal. The purpose of this article is to explain in details the use of Bernstein polynomials (available: https://www2.math.upenn.edu/~kadison/bernstein.pdf, accessed 10-13-2021) in the estimation of the Pickands function and to establish a new test of significance for extreme value copulas. The theoretical aspects regarding copulas theory are maintained at elementary level. The reader can find other results in the texts by (NELSEN, 2013) and (JOE, 1997). For purpose of completeness some nice properties of Bernstein polynomials are presented.

2 Bernstein polynomials

Each term in the Newton's binomial expansion $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$, taken in $x+y=1$, with $0<x<1$, defines the so-called Bernstein polynomials:

[1]Universidade Tecnológica Federal do Paraná - UTFPR, Campus Apucarana, CEP: 86812-460, Apucarana, PR, Brasil. E-mail: danigoprado@hotmail.com
[2] Universidade Federal de Lavras - UFLA, Instituto de Ciências Exatas e Tecnológicas, Departamento de Matemática e Matemática Aplicada, CEP: 37200-900, Lavras, MG, Brasil. E-mail: lucas@ufla.br
[3] Universidade Federal de Lavras - UFLA, Instituto de Ciências Exatas e Tecnológicas, Departamento de Estatística, CEP: 37200-900, Lavras, MG, Brasil. E-mail: devaniljaques@ufla.br; eleandersoncampos@gmail.com

$$b_{n,k}(x)=\binom{n}{k}x^k(1-x)^{n-k}$$

It follows immediately that, for all $x\in[0,1]$, $\sum_{k=0}^{n} b_{n,k}(x)=1$ and, as $b_{n,k}(x)>0,$ the set $\{b_{n,k}(x),\ k=0,\ldots,n\}$ constitutes a partition of the unit.

These polynomials have a long history and they have a number of properties similar to those of the binomial numbers:

- The elements of $\{b_{n,k}(x),\ k=0,\ldots,n\}$ form a basis for the space of polynomials of degree less than or equal to n.

- Bernstein's polynomials and the Beta probability distribution are related:

$$\begin{aligned} b_{n,k}(x) &= \binom{n}{k}x^k(1-x)^{n-k} \\ &= \frac{1}{n+1}\frac{1}{B\left(k+1,n-k+1\right)}x^{(k+1)-1}\left(1-x\right)^{(n-k+1)-1} \end{aligned}$$

- Bernstein's polynomials can be defined recursively:

$$\begin{aligned} b_{n,k}(x) &= \binom{n}{k}x^k(1-x)^{n-k}=\left[\binom{n-1}{k}+\binom{n-1}{k-1}\right]x^k\left(1-x\right)^{n-k} \\ &= \binom{n-1}{k}x^k\left(1-x\right)^{n-k}+\binom{n-1}{k-1}x^k\left(1-x\right)^{n-k} \\ &= \left(1-x\right)\binom{n-1}{k}x^k\left(1-x\right)^{n-1-k}+x\binom{n-1}{k-1}x^{k-1}\left(1-x\right)^{n-1-(k-1)} \\ &= \left(1-x\right)b_{n-1,k}\left(x\right)+x\,b_{n-1,k-1}\left(x\right) \end{aligned}$$

- The derivatives of the Bernstein polynomials are given by:

$$\begin{aligned} \frac{d}{dx}b_{n,k}\left(x\right) &= \frac{d}{dx}\left[\binom{n}{k}x^k\left(1-x\right)^{n-k}\right] \\ &= \binom{n}{k}\left[k\,x^{k-1}\left(1-x\right)^{n-k}-x^k\left(n-k\right)\left(1-x\right)^{n-k-1}\right] \end{aligned}$$

$$
\begin{aligned}
&= \frac{n!}{(k-1)!(n-k)!} x^{k-1}(1-x)^{n-k} - \frac{n!}{k!(n-k-1)!} x^{k}(1-x)^{n-k-1} \\
&= n\left[\binom{n-1}{k-1} x^{k-1}(1-x)^{n-1-(k-1)} - \binom{n-1}{k} x^{k}(1-x)^{n-1-k}\right] \\
&= \begin{cases} -n(1-x)^{n-1} & \text{para} \quad k=0 \\ nb_{n-1,k-1}(x) - b_{n-1,k}(x) & \text{para} \quad k=1,2,\ldots,n-1 \\ nx^{n-1} & \text{para} \quad k=n \end{cases}
\end{aligned}
$$

For the second derivative we must observe that:

$$
\begin{aligned}
k=1 \qquad & \Rightarrow \frac{d}{dx} b_{n,1}(x) = n(1-x)^{n-1} - n(n-1)x(1-x)^{n-2} \\
& \Rightarrow \frac{d^2}{dx^2} b_{n,1}(x) = -2n(n-1)(1-x)^{n-2} \\
& \qquad\qquad + n(n-1)(n-2)x(1-x)^{n-3} \\
k=n-1 \qquad & \Rightarrow \frac{d}{dx} b_{n,n-1}(x) = n(n-1)x^{n-2}(1-x) - n\,x^{n-1} \\
& \Rightarrow \frac{d^2}{dx^2} b_{n,n-1}(x) = n(n-1)(n-2)x^{n-3}(1-x) \\
& \qquad\qquad - 2n(n-1)x^{n-2} \\
1<k<n-1 \qquad & \Rightarrow \frac{d^2}{dx^2} b_{n,k}(x) = n\left[\frac{d}{dx} b_{n-1,k-1}(x) - \frac{d}{dx} b_{n-1,k}(x)\right] \\
& \qquad = n(n-1)\big[b_{n-2,k-2}(x) - b_{n-2,k-1}(x) \\
& \qquad\qquad - \left(b_{n-2,k-1}(x) - b_{n-2,k}(x)\right)\big] \\
& \qquad = n(n-1)\left[b_{n-2,k-2}(x) - 2b_{n-2,k-1} + b_{n-2,k}(x)\right].
\end{aligned}
$$

All together,

$$\frac{d^2}{dx^2}b_{n,k}(x)=n(n-1)\begin{cases}(1-x)^{n-2} & \text{for} \quad k=0\\ (n-2)x(1-x)^{n-3}-2(1-x)^{n-2} & \text{for} \quad k=1\\ \left[b_{n-2,k-2}(x)-2b_{n-2,k-1}+b_{n-2,k}(x)\right] & \text{for} \quad 2\le k\le n-2\\ (n-2)x^{n-3}(1-x)-2x^{n-2} & \text{for} \quad k=n-1\\ x^{n-2} & \text{for} \quad k=n\end{cases}$$

- One of the most important property of Bernstein's polynomials is to approximate functions uniformly. For a function $f(x)$ with domain at $[0,1]$, defining $B_n(f)(x)=\sum_{k=0}^{n} f(k/n)b_{n,k}(x) \quad (0\le x\le 1)$ as the Bernstein polynomial of $f(x)$, we have:

2.1 Bernstein-Weierstrass approximation theorem

Theorem 2.1.1: If f is a real-valued bounded function with domain in the interval $[0,1]$, then for each point x where f is continuous, $B_n(f)(x)\to f(x)$ as n $\to\infty$. If f is continuous on $[0,1]$, then the Bernstein polynomial $B_n(f)(x)$ tends uniformly to f as $n\to\infty$ (available: https://www2.math.upenn.edu/~kadison/bernstein.pdf, accessed 10-13-2021).

The restriction of the domain to the interval $[0,1]$ can easily be extended to $[0,\tau]$, by redefining the Bernstein's polynomials as:

$$b_{n,k}(x)=\binom{n}{k}\left(\frac{x}{\tau}\right)^k\left(1-\frac{x}{\tau}\right)^{n-k}.$$

Actually, the property of interest here is the fact that Bernstein's polynomials uniformly approximate continuous functions because it is easy to characterize increasing or even unimodal polynomials when they are expressed by linear combinations of Bernstein's polynomials.

Consider the polynomial given by a linear combination of Bernstein's polynomials $P_a(x)=\sum_{k=0}^{n} a_k b_{n,k}(x)$ with $a=(a_o,a_1,...,a_n)$ and $x\in[0,1]$. Observe that:

$$P_a'(x) \;=\sum_{k=0}^{n} a_k\frac{d}{dx}b_{n,k}(x)$$

$$
\begin{aligned}
&= n\left[-a_0 b_{n-1,0}(x) + \sum_{k=1}^{n-1} a_k \left(b_{n-1,k-1}(x) - b_{n-1,k}(x)\right) + a_n b_{n-1,n-1}(x)\right] \\
&= n\Big[-a_0 b_{n-1,0}(x) + a_1 b_{n-1,0}(x) - a_1 b_{n-1,1}(x) + \ldots \\
&\qquad \ldots + a_{n-1} b_{n-1,n-2}(x) - a_{n-1} b_{n-1,n-1}(x) + a_n b_{n-1,n-1}(x)\Big] \\
&= n\Big[\Big[(a_1 - a_0) b_{n-1,0}(x) + \ldots + (a_n - a_{n-1}) b_{n-1,n-1}(x)\Big] \\
&= n \sum_{k=0}^{n-1} (a_{k+1} - a_k) b_{n-1,k}(x) \\
&= n\left[(a_1 - a_0)(1-x)^{n-1} + \sum_{k=1}^{n-2} (a_{k+1} - a_k) b_{n-1,k}(x) + (a_n - a_{n-1}) x^{n-1}\right].
\end{aligned}
$$

In a similar way:

$$
\begin{aligned}
P_a''(x) &= \sum_{k=0}^{n} a_k \frac{d^2}{dx^2} b_{n,k}(x) \\
&= n(n-1)\Big\{a_0 (1-x)^{n-2} + a_1\left[(n-2)x(1-x)^{n-3} - 2(1-x)^{n-2}\right] \\
&\qquad + \sum_{k=2}^{n-2} a_k \left[b_{n-2,k-2}(x) - 2 b_{n-2,k-1} + b_{n-2,k}(x)\right] \\
&\qquad + a_{n-1}\left[(n-2) x^{n-3} (1-x) - 2 x^{n-2}\right] + a_n x^{n-2}\Big\} \\
&= n(n-1)\Big\{a_0 b_{n-2,0}(x) + a_1 b_{n-2,1}(x) - 2a_1 b_{n-2,0}(x) \\
&\qquad + a_2 b_{n-2,0}(x) - 2a_2 b_{n-2,1}(x) + a_2 b_{n-2,2}(x) + \ldots \\
&\qquad \ldots + a_{n-2} b_{n-2,n-4}(x) - 2 a_{n-2} b_{n-2,n-3}(x) + a_{n-2} b_{n-2,n-2}(x) \\
&\qquad + a_{n-1} b_{n-2,n-3}(x) - 2a_{n-1} b_{n-2,n-2}(x) + a_n b_{n-2,n-2}(x)\Big\} \\
&= n(n-1)\Big\{(a_0 - 2a_1 + a_2) b_{n-2,0}(x) + (a_1 - 2a_2 + a_3) b_{n-2,1}(x) + \ldots \\
&\ldots + (a_{n-3} - 2a_{n-2} + a_{n-1}) b_{n-2,n-3}(x) + (a_{n-2} - 2a_{n-1} + a_n) b_{n-2,n-2}(x)\Big\} \\
&= n(n-1) \sum_{k=0}^{n-2} (a_{k+2} - 2a_{k+1} + a_k) b_{n-2,k}(x).
\end{aligned}
$$

It follows from these two results:

Proposition 2.1.1:

i. If $a_0 \le a_1 \le \ldots \le a_n$ then $P_a'(x) \ge 0$, that is, the polynomial $P_a(x)$ is monotonous non-decreasing.

ii. If $a_1 - a_o < 0$, $a_n - a_{n-1} > 0$ and $a_{k+2} + a_k \ge 2a_{k+1}$ $(k = 0,\ldots,n-2)$, then $P_a'(0) < 0$, $P_a'(1) > 0$ and $P_a''(x) \ge 0$, that is, the polynomial $P_a(x)$ is unimodal convex; (derivatives at points 0 and 1 are lateral derivatives).

(CHANG *et al.*, 2007)

As an application of this result, we obtain a polynomial regression with shape constraint as:

Given a data set (x_i, y_i) $i = 1,\ldots,m$, a unimodal convex polynomial of degree n, $P_n(x) = \sum_{k=0}^{n} a_k b_{n,k}(x)$, that best fits this data, in the sense of the minimum squares, is obtained as follows.

The regression matrix is:

$$X = \begin{pmatrix} b_{n,0}(x_1) & b_{n,1}(x_1) & b_{n,2}(x_1) & \ldots & b_{n,n}(x_1) \\ b_{n,0}(x_2) & b_{n,1}(x_2) & b_{n,2}(x_2) & \ldots & b_{n,n}(x_2) \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ b_{n,0}(x_m) & b_{n,1}(x_m) & b_{n,2}(x_m) & \ldots & b_{n,n}(x_m) \end{pmatrix} \qquad a = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

That is, $y = X\,a + \varepsilon$.

The restrictions, according to *(ii)* of Lemma 1, can be described by the vector:

$$A = \begin{pmatrix} a_0 - a_1 \\ a_0 - 2a_1 + a_2 \\ \vdots \\ a_{n-2} - 2a_{n-1} + a_n \\ -a_{n-1} + a_n \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \\ a_n \end{pmatrix} \ge 0$$

Observe, then, that Bernstein Polynomials allowed us to reduce a polynomial regression problem to a linear regression problem with constrains.

$$max\|Xa - y\|^2$$
$$restrito\ A \ge 0$$

This optimization problem is a quadratic programming (BOYD, VANDENBERGHE, 2004, p.152), and is a basic problem in the area of optimization, with a bunch of efficient algorithms to solve it.

3 Extreme value copulas and the Pickands function

Definition 3.1: The copula *C(u, v)* is said to be an extreme value copula if there is a copula C_F such that $\lim_{n\to\infty} C_F\left(u^{1/n}, v^{1/n}\right)^n = C(u,v)$ for all $(u,v)\in[0,1]^2$.

Pickands Theorem: The bivariate copula C is of extreme value if, and only if, for all $(u,v)\in(0,1]^2 - \{(1,1)\}$, it can be expressed in the form:

$$C(u,v) = (uv)^{A(\ln(v)/\ln(uv))} = \exp\left\{ln[uv]\,A\left(\frac{ln[v]}{ln[uv]}\right)\right\}$$

in which the so-called Pickands function $A(.)$ is a convex function with domain in $[0,1]$ and image in $[1/2,1]$, and satisfies $\max\{t,1-t\}\le A(t)\le 1$ and $-1\le A'(t)\le 1$.

The Pickands Theorem can also be stated as:

The bivariate copula C is of extreme value if, the transformation $T_C:[0,1]^2\to[0,1]\times[0,+\infty)$, defined for all $(u,v)\in[0,1]^2$, as $(u,v)\to(t,z)=\left(\frac{ln(v)}{ln(uv)}, \frac{ln(C(u,v))}{ln(uv)}\right)$, results in a convex curve, bounded upper by $z=1$ and lower by $z=\max(t,1-t)$.

3.1 Bernstein's polynomials in the estimation of the Pickands function

The fact that the Pickands function is convex and unimodal is a natural justification for it to be estimated using Bernstein's polynomials (suggestion by Professor Dr. Yan of the University of Connecticut in Storrs).

The estimation of the Pickands function is implemented using the T_C transformation shown below. Given a copula $C(u,v)$ consider the transformation:

$$T_C:[0,1]^2\to[0,1]\times[0,+\infty)\ ,\ (u,v)\to\left(t=\frac{ln(v)}{ln(uv)}, z=\frac{ln(C(u,v))}{ln(uv)}\right).$$

If the copula is of extreme value this transformation degenerates and the image of T_C is no longer $[0,1]\times[0,+\infty)$ and became the graph of the Pickands function. Indeed,

$$C(u,v) = uv^{A\left(\frac{ln(v)}{ln(uv)}\right)} \text{ and,}$$

$$T_C(u,v) = \left(t = \frac{ln(v)}{ln(uv)}, z = \frac{ln\left(uv^{A\left(\frac{ln(v)}{ln(uv)}\right)} \right)}{ln(uv)} \right) = (t, A(t)).$$

3.1.1 Estimation procedure

If $(u_1,v_1),(u_2,v_2),...,(u_N,v_N)$ is a sample of an extreme value copula $C(u,v)=uv^{A\left(\frac{ln(v)}{ln(uv)}\right)}$, then $(t_1,z_1)=T_C(u_1,v_1),...,(t_N,z_N)=T_C(u_N,v_N)$ are points on the graph $(t,A(t))$. If the function $A(t)$ is unknown it can be estimated with the points $(t_1,z_1),(t_2,z_2),...,(t_N,z_N)$ by a linear regression using the Bernstein polynomials with the degree *n* and constraints, obtaining a polynomial $\hat{P}_n(t)$. According to Bernstein-Weierstrass approximation theorem, $\hat{P}_n(t)$ approaches $A(t)$ and the mean of the sum of residues, $T_N = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{P}_n(t_i)-z_i\right)^2$, is expected to be small. Otherwise, if the copula is not an extreme value one this sum is expected to be greater. Appropriate degree *n* of the polynomial can be obtained by model selection methods.

The Pickands function, besides being convex, satisfies $A(0)=A(1)=1$ and $-1\leq A'(t)\leq 1$ (Figure 1).

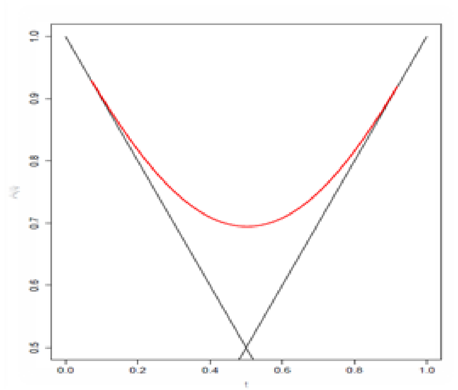

Figura 1 - Graph of a typical Pickands Function.

Thus, the regression should be restricted to the unimodal and convex polynomials in $[0,1]$, according to (ii) of Proposition 1, with $P_n(0)=P_n(1)=1$ and satisfying the inequalities $-1\leq P_n'(0)$ e $P_n'(1)\leq 1$.

$$P_n(x) \quad =\sum_{k=0}^{n} a_k b_{n,k}(x)=a_0(1-x)^n+a_n x^n+\sum_{k=1}^{n-1} a_k \binom{n}{k} x^k (1-x)^{n-k} \text{ ,}$$

$$P_n(0)=1 \qquad \Rightarrow \qquad a_0=1$$

$$P_n(1)=1 \qquad \Rightarrow \qquad a_n=1$$

Therefore, the problem of estimation of the Pickands function using Bernstein's polynomials is a linear regression problem consisting of minimizing the sum of squares $\sum_{i=1}^{m}\left(P_n(t_i)-z_i\right)^2$, with the set of constraints expressed in matrix form as:

$$
\begin{bmatrix}
1 & 0 & 0 & & & & \cdots & 0 \\
1 & -1 & 0 & & & & \cdots & 0 \\
1 & -2 & 1 & 0 & & & \cdots & 0 \\
0 & 1 & -2 & 1 & 0 & & \cdots & 0 \\
\vdots & & & & & \ddots & & \vdots \\
0 & \cdots & & & 0 & 1 & -2 & 1 \\
0 & \cdots & & & & 0 & -1 & 1 \\
0 & \cdots & & & & & 0 & 1
\end{bmatrix}
\begin{bmatrix} a_0 \\ a_1 \\ \\ \\ \vdots \\ \\ a_{n-1} \\ a_n \end{bmatrix}
\geq
\begin{bmatrix} 1 \\ -1/n \\ 0 \\ \\ \vdots \\ 0 \\ -1/n \\ 1 \end{bmatrix}
$$

in which, the inequalities in the first and last lines are equalities (CHANG *et al.*, 2007).

It is also considered the inequality that limits the size of the parameters:

$$\sum_{k=0}^{n} \left| a_k \right| < M_n .$$

This problem can be solved by using the *"solve.QP"* routine of the *quadprog R* package.

3.1.2 Implementation of the Pickands function estimator

For the simulation process, the steps are:

Step 1) Take a sample $(x_1, y_1),(x_2, y_2),...,(x_N, y_N)$ from the random pair (X,Y).

Step 2) For each pair (x_i, y_i), consider (n_i, m_i), in which n_i is the rank of x_i and m_i is the rank of y_i, when sequences $x_1, x_2,..., x_N$ and $y_1, y_2,..., y_N$ are sorted in ascending order.

Step 3) Obtain the sample in the unit square $(0,1)\times(0,1)$ through the transformation $\left(\frac{n_1}{N+1}, \frac{m_1}{N+1}\right),...,\left(\frac{n_N}{N+1}, \frac{m_N}{N+1}\right)$. The division by $N+1$ is used to avoid the point $(1,1)$.

Step 4) Apply the transformation T_C in $\left(\frac{n_1}{N+1}, \frac{m_1}{N+1}\right),...,\left(\frac{n_N}{N+1}, \frac{m_N}{N+1}\right)$.

Step 5) The linear regression, using Bernstein polynomials and its respective restrictions, are obtained via function "*solve.QP*" of the *quadprog R* package.

4 A test of significance for extreme value copulas

The new method of estimation of the Pickands function using Bernstein polynomials allows to create a test of significance that checks if a dataset might be appropriately represented by an extreme value dependence model.

This test is based on the concept of A-plot, and follows the same procedures as in (CORMIER *et al.*, 2014). A briefly description of an A-plot follows:

Let $\{x_1, x_2, \dots x_N\}$ and $\{y_1, y_2, \dots y_N\}$ be realizations of the random pair (X, Y). The marginal distribution functions of X and Y are estimated by their respective empirical versions:

$$F_N(t) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\,(X_i \leq t) \qquad G_N(t) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\,(Y_i \leq t).$$

Let $\hat{U}_i = F_N(X_i)$, $\hat{V}_i = G_N(Y_i)$, be random variables and observe that, according to the probability integral transformation, the variables $\hat{U}_i$ and $\hat{V}_i$ are close to the uniform(0,1) distribution. The idea now is to use the empirical copula $\hat{C}_N$ as the joint distribution of $(\hat{U}_1, \hat{V}_1), \dots, (\hat{U}_N, \hat{V}_N)$. If $(u_1, v_1), \dots, (u_N, v_N)$ are realizations of the random pair $(\hat{U}_i, \hat{V}_i)$, then:

$$\hat{C}_N(u, v) = \frac{1}{N+1} \#\{(u_i, v_i),\ \ u_i \leq u, v_i \leq v\}.$$

The division by $N+1$ avoids that $\hat{C}_N$ takes the value 1.

The pairs $(t_1, z_1), \dots, (t_N, z_N)$ are build:

$$t_i = \frac{\ln(v_i)}{\ln(u_i v_i)}, \quad z_i = \frac{\ln\left(\hat{C}_N(u_i, v_i)\right)}{\ln(u_i v_i)}.$$

As described above, if C is an extreme value copula, the points (t_i, z_i) should be near the graph of the Pickands function of the copula C. This procedure is denominated A-plot (Figure 2).

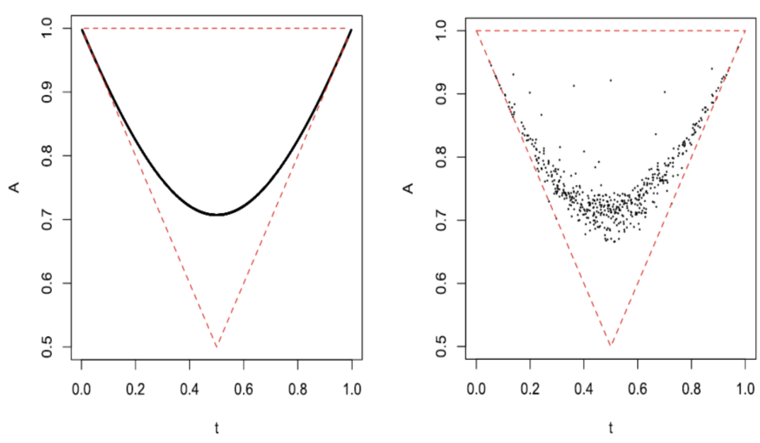

Figura 2 - A-plot for an extreme value copula.

The null hypothesis H_0 states that the observed pairs (u_i, v_i) are realizations of an extreme value copula C. With the points (t_i, z_i), the Bernstein polynomial $\hat{P}_n(t)$ is fitted as an estimative of the Pickands function.

The test statistic is the average of the sum of squares of deviations between the observed points and the fitted points:

$$T_N = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{P}_n(t_i) - z_i\right)^2.$$

The test consists in rejecting H_0 for T_N sufficiently large. The distribution of the statistic T_N under the null hypothesis is not known and is approximated via parametric bootstrap method. The recipe is as follows:

1 – Using the data $(u_1, v_1), \ldots, (u_N, v_N)$, obtain the Kendall's tau which will be used as an estimative of its population version.

2 – Choose a representative of the extreme value copulas family. In this work, the Gumbel copula was selected.

3 – Define the copula's parameter with the Kendall's tau value obtained in step 1 and use it to generate a sample $(\tilde{u}_1, \tilde{v}_1), (\tilde{u}_2, \tilde{v}_2), \ldots, (\tilde{u}_N, \tilde{v}_N)$.

4 – Using the Pickands function transformation, obtain the points $(t_1, z_1), \ldots, (t_N, z_N)$.

5 – Fit a Bernstein polynomial $\hat{P}_n(t)$ to the points of the previous step.

6 – Calculate the test statistic:

$$T_N = \frac{1}{N}\sum_{i=1}^{N}\left\{z_i - \hat{P}_N(t_i)\right\}^2.$$

7 – Steps 3, 4, 5 and 6 are repeated N_b times.

8 – With the N_b values of T_N, build its empirical distribution.

9 – Reject H_0 if the p-value is less than the adopted significance level α. Do not reject H_0 otherwise.

In order to verify type I error rates, 1000 samples of size 200 from the Gumbel copula are generated (extreme value copula) using the values $\tau = 0{,}25$, $\tau = 0{,}50$, $\tau = 0{,}75$ for Kendall's tau. The test is carried out at a nominal level of 5%.

To verify type II error rates, samples from non-extreme value copulas Clayton, Frank, Gaussian and t_4 are generated, using the same configurations of repetitions, sample size and Kendall's tau values from the type I error case.

(Table 1) is used to compare the performance of the proposed test with type I and II error rates of several tests in extreme value copulas organized by (CORMIER *et al.*, 2014, p. 649).

Table 1 - Rejection rates of the null hypothesis (in %) at a nominal level of 5% and sample size of 200

τ	Modelo	*proposed test*	KY	DN	BGN	BDV	KSY	CGN
	Gumbel	5.5	3.8	5.2	5.4	4.5	5.0	4.7
	Clayton	92.5	98.4	96.7	98.0	87.4	94.6	97.7
0.25	Frank	6.5	58.3	57.0	38.4	29.1	66.1	18.7
	Gaussiana	12.5	36.5	40.3	37.3	16.8	38.7	25.5
	t_4	21.0	23.9	19.6	26.2	10.5	26.6	37.7
	Gumbel	6.0	3.9	5.0	5.1	2.9	4.0	5.4
	Clayton	100.0	100.0	100.0	100.0	100.0	100.0	100.0
0.50	Frank	63.0	95.7	84.8	59.4	73.0	96.5	87.8
	Gaussiana	38.0	61.8	61.7	62.6	23.7	51.0	59.4
	t_4	39.5	50.1	45.3	56.0	15.8	52.7	58.6
	Gumbel	5.0	3.2	5.3	4.9	2.5	2.3	6.2
	Clayton	100.0	100.0	100.0	100.0	100.0	100.0	100.0
0.75	Frank	73.5	99.9	92.9	58.5	78.3	99.0	98.3
	Gaussiana	13.5	66.5	71.1	75.2	8.4	46.7	56.5
	t_4	25.5	50.6	55.8	67.8	4.6	69.2	45.8

5 Results and discussion

The results obtained by the new test are comparable to those obtained by other tests. However, the proposed test did not obtain a good control of the type II error rate, except when the data comes from Clayton family.

The runtime with Bernstein's polynomial method adjustment is substantially lower than the others at Cobs bundle in free software R.

The use of Bernstein's polynomials is an efficient way of estimating the Pickands function. With this estimation process it is possible to obtain a new test of significance for extreme value copulas that presents performance compatible with other tests already established in the literature.

Acknowledgements

Thanks to Capes for the financial support provided.

PRADO, D. G. O.; CHAVES, L. M.; SOUZA, D. J.; EUGÊNIO FILHO, E. C. Regressão polinomial convexa unimodal utilizando polinômios de Bernstein no ajuste de cópulas. *Braz. J. Biom*. Lavras, v.40, n.2, p.152-165, 2022.

▪ *RESUMO: Os polinômios de Bernstein são adequados para realizar regressões com restrição de forma, em particular, regressão convexa unimodal. A função de Pickands é convexa e unimodal, sendo um elemento fundamental na teoria das cópulas de valores extremos. O objetivo deste artigo é explicar em detalhes o uso de polinômios de Bernstein na estimação da função de Pickand e estabelecer um novo teste de significância para cópulas de valores extremos.*

▪ *PALAVRAS-CHAVE: Polinômios de Bernstein. Função de Pickands. Cópulas de valor extremo.*

References

BOYD, S.; VANDENBERGHE, L. *Convex optimization*. Cambridge university press, 2004.

CHANG, I. S.; CHIEN, L. C.; HSIUNG, C. A.; WEN, C.C.; WU, Y. J. *Shape restricted regression with random Bernstein polynomials.* In: COMPLEX DATASETS AND INVERSE PROBLEMS. Institute of Mathematical Statistics, 2007. p.187-202.

CORMIER, E.; GENEST, C.; NEŠLEHOVÁ, J. G. Using B-splines for nonparametric inference on bivariate extreme value copulas. *Extremes*, v.17, n.4, p.633-659, 2014.

JOE, H. *Multivariate models and multivariate dependence concepts*. CRC Press, 1997.

KADISON, R.V.; LIU, Z., *Bernstein Polynomials and Approximation*. Available: https://www2.math.upenn.edu/~kadison/bernstein.pdf (accessed 10-13-2021)

NELSEN, R. B. *An introduction to copulas.* Springer Science & Business Media, 2013.

PICKANDS, J. *Multivariate extreme value distributions*. In: PROCEEDINGS 43RD SESSION INTERNATIONAL STATISTICAL INSTITUTE. p.859-878, 1981.

Received in 02.01.2021

Approved after revised in 11.08.2021