Revista Colombiana de Estadística
January 2018, Volume 41, Issue 1, pp. 53 to 73
DOI: http://dx.doi.org/10.15446/rce.v41n1.64900

Estimating the Gumbel-Barnett Copula Parameter of Dependence

Estimación del parámetro de dependencia de la copula Gumbel-Barnett

Jennyfer Portilla Yela[a], José Rafael Tovar Cuevas[b]

Escuela de Estadística, Facultad de Ingeniería, Universidad del Valle, Cali, Colombia

Abstract

In this paper, we developed an empirical evaluation of four estimation procedures for the dependence parameter of the Gumbel-Barnett copula obtained from a Gumbel type I distribution. We used the maximum likelihood, moments and Bayesian methods and studied the performance of the estimates, assuming three dependence levels and 20 different sample sizes. For each method and scenario, a simulation study was conducted with 1000 runs and the quality of the estimator was evaluated using four different criteria. A Bayesian estimator assuming a $\text{Beta}(a, b)$ as prior distribution, showed the best performance regardless the sample size and the dependence structure.

***Key words*:** Bayesian; Copula; Correlation; Dependence; Estimation; GB copula; Simulation.

Resumen

En este artículo, desarrollamos una evaluación empírica de cuatro procedimientos de estimación para el parámetro de dependencia, de la función copula Gumbel-Barnett obtenida a partir de la distribución Gumbel tipo I. Se usó el método de estimación por momentos, el método de la máxima verosimilitud y dos aproximaciones Bayesianas. Se estudió el comportamiento de las estimaciones asumiendo tres niveles de dependencia y 20 tamaños de muestra distintos. Para cada método y escenario formado entre el nivel de dependencia y el tamaño de muestra, se desarrolló un estudio de simulación con 1000 repeticiones y el comportamiento de las estimaciones fue evaluado usando cuatro criterios. El estimador obtenido asumiendo una distribución $\text{Beta}(a, b)$ para modelar la información previa, presentó el mejor desempeño sin importar el tamaño de muestra y la estructura de dependencia.

***Palabras clave*:** bayesiana; copula Gumbel Barnett; correlación; dependencia copula; estimación; simulación.

[a]Msc Student. E-mail: jennyfer.portilla@correounivalle.edu.co
[b]PhD. E-mail: jose.r.tovar@correounivalle.edu.co

1. Introduction

The origin of the Gumbel-Barnett copula function dates to 1960, when Gumbel developed a cumulative function of probability assuming two dependent exponential standard distributions as marginals, which was called the Gumbel Type I distribution. The distribution was later studied by Barnett (1983), who discussed the estimation of the parameter θ using the maximum likelihood method, as well as a method based on the product-moment correlation. When the marginals were transformed using the integral transformation of probability, the distribution was indexed by the parameter θ, which ranges in the interval $(0, 1)$ and models the dependence structure existing between the two random variables. Many authors refer to the copula function obtained after applying Sklar's theorem to the cumulative distribution function of the Gumbel Type I distribution as another Gumbel family copula, but Balakrishnan & Lai (2009) called it the Gumbel-Barnett (GB) copula because the characteristics of this copula function assuming different marginal distributions (including the uniform distribution) were first discussed by Barnett (1980).

In this work, we assume the definition given by Balakrishnan & Lai (2009) and refer to the function as the Gumbel-Barnett copula function obtained from the Gumbel Type I distribution. For dependencies that can be modeled using this copula function when $\theta = 0$, Pearson's coefficient (ρ_s) takes the value of zero and the variables under study can be assumed independent. When θ increases, ρ_s decreases, reaching a value of -0.404 when $\theta = 1$. Genest & Mackay (1986) remarked that although the survival copula of the Gumbel-Barnett (GB) copula belongs to the Archimedean copulas, the Gumbel-Barnett copula should not be included. In accordance with our literature review, there are several studies where the authors used the survival form of the copula, all of which referred to it as the Gumbel-Barnett copula rather than the cumulative form copula.

An interesting characteristic of the Gumbel-Barnett copula function is that it models weak and not necessarily linear dependencies, i.e., when the usual plot used to identify dependence between the random variables shows a structure similar to that observed when the variables are independent and the hypothesis tests associated with commonly used indexes, such as Pearson's or Spearman's coefficients, do not reject the independence hypothesis. This characteristic was explored within a clinical diagnosis framework by Tovar & Achcar (2011, 2012, 2013).

Many other authors have studied additional theoretical characteristics of the GB copula function; for instance, Klein & Christa (2011) studied the weighted geometric and harmonic means from different copula families, including the survival Gumbel-Barnett. Omidi & Mohammadzadeh (2015) obtained the stationary spatial covariance functions for five archimedean copula families, including the GB family. Louie (2014) evaluated ten copula functions, including the GB, to identify the copula families that were best suited to model dependence structures for wind power. Martinez & Achcar (2014) used the GB as well as the Farlie Gumbel Morgerstern and Clayton copulas to model the dependence structure between survival times, assuming a Lindley distribution for the marginal distributions.

In this paper, we study the empirical performance of different approaches to obtain estimates of the dependence parameter of the Gumbel-Barnett copula function using the maximum likelihood, moments and Bayesian methods. A simulation study was performed assuming three different levels of dependence and different sample sizes (50 to 1000 in steps of 50).

2. Estimating the Gumbel-Barnett Copula Dependence Parameter

2.1. Maximum Likelihood Approach

Let two random variables X_1 and X_2 have cumulative distribution function $F_{X_1,X_2}(x_1,x_2)$, which can be fitted by a Gumbel Type I bivariate exponential distribution. We can define a parameter θ to model the dependence structure between X_1 and X_2, and we can use the integral transformation theorem to obtain $u = F(x_1)$ and $v = F(x_2)$, the normalized marginals such that u and v range in the interval $(0,1)$. With the transformed variables, by applying Sklar's theorem, it is possible to express the joint distribution of X_1 and X_2 as a copula function, as follows:

$$C(u,v) = u + v - 1 + (1-u)(1-v)e^{-\theta\ln(1-u)\ln(1-v)} \quad 0 \le \theta \le 1 \tag{1}$$

whose density function is:

$$c(u,v) = e^{-\theta\ln(1-u)\ln(1-v)}\left[(\theta\ln(1-v)-1)(\theta\ln(1-u)-1)-\theta\right] \tag{2}$$

Consider a data set of observations (x_{1i}, x_{2i}); $i = 1, 2, \ldots, n$ with normalized marginal distributions $u = F_{X_1}(x_1)$, $v = F_{X_2}(x_2)$ and a dependence structure that can be modeled using a GB copula function through a parameter θ. The likelihood and log likelihood functions are given by:

$$L(\theta|\mathbf{u},\mathbf{v}) = \prod_{i=1}^{n} e^{\beta_i}\left((\theta\ln(1-u_i)-1)(\theta\ln(1-v_i)-1)-\theta\right) \tag{3}$$

where $\beta_i = -\theta\ln(1-u_i)\ln(1-v_i)$ and,

$$\begin{aligned}\ell(\theta|\mathbf{u},\mathbf{v}) = -&\sum_{i=1}^{n}\theta\ln(1-u_i)\ln(1-v_i) + \\ &\sum_{i=1}^{n}\ln\left((\theta\ln(1-u_i)-1)(\theta\ln(1-v_i)-1)-\theta\right)\end{aligned} \tag{4}$$

It is possible to obtain the estimate of θ from the corresponding survival copula, which has the following analytical form (see, Barnett 1980):

$$C'(u,v) = uve^{-\theta\ln(u)\ln(v)} \tag{5}$$

with density function

$$\mathrm{c}'(\mathrm{u},\mathrm{v}) = \lambda e^{-\theta\ln u\ \ln v} \tag{6}$$

where, $\lambda = \left(1 - \theta \ln u \; - \theta \ln v - \theta + \theta^2 \ln u \ln v\right)$.

Then, from (5) and (6), the likelihood and log likelihood functions are:

$$L(\theta|\mathbf{u}, \mathbf{v}) = \prod_{i=1}^{n} e^{-\theta \ln u_i \; \ln v_i} (\lambda_i) \tag{7}$$

and

$$l(\theta|\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} -\theta \ln u_i \ln v_i + \sum_{i=1}^{n} ln\left(\lambda_i\right) \tag{8}$$

Given that there is no closed way to obtain a maximum for the equations (4) and (8), it is necessary to use numerical algorithms to approximate a solution. We used the *optim* function implemented in R software, which uses the algorithms of Brent, BFGS and Nelder-Mead to obtain the root that maximizes the log likelihood function in (8).

2.2. Moments Approach

To obtain the moments estimator, we used the mathematical relationship between Spearman's rho (ρ_s) and the Kendall's tau (τ) with the GB dependence parameter. The associated expressions for both indexes are (Nelsen 2006):

$$\rho_s = 12 \int_0^1 \int_0^1 C_{\mathbf{Y}}(u_1, u_2)\, du_1 du_2 - 3 \tag{9}$$

$$\tau = 4 \int_0^1 \int_0^1 C_{\mathbf{Y}}(u_1, u_2) c_{\mathbf{Y}}(u_1, u_2)\, du_1 du_2 - 1 \tag{10}$$

In both cases, the equation includes the exponential integral function $Ei(\cdot)$

$$\rho_s = 12 \left[-\frac{e^{\frac{4}{\theta}}}{\theta} Ei\left(-\tfrac{4}{\theta}\right) \right] - 3$$

$$\tau = \left[7 - \tfrac{2}{\theta}\right] e^{\frac{2}{\theta}} Ei\left(-\tfrac{2}{\theta}\right) - 4e^{\frac{1}{\theta}} Ei\left(-\tfrac{1}{\theta}\right) + 4e^{\frac{4}{\theta}} Ei\left(-\tfrac{4}{\theta}\right) - 1$$

which makes it very difficult to compute the chosen index. As an alternative, we employed the relationship between the Pearson's correlation coefficient (ρ_p) and the parameter of the Type I bivariate exponential function to obtain the Gumbel-Barnett copula (θ). This relationship was studied by Gumbel (1960):

$$\rho_p = -1 + \int_0^{\infty} \frac{e^{-z}}{1 + \theta z}\, dz$$

In this way, the estimation of the dependence parameter θ is the solution to the following equation that approaches zero:

$$\rho_p + 1 - \int_0^\infty \frac{e^{-z}}{1+\theta z}\, dz \approx 0$$

2.3. Bayesian Approach

For the Bayesian approach, three different prior distributions were assumed. Using the likelihood function obtained in (3) and prior distribution $\pi(\theta)$, it is possible to obtain the posterior distribution as follows:

$$\pi(\theta|\mathbf{u},\mathbf{v}) \propto \pi(\theta)L(\theta|\mathbf{u},\mathbf{v}) \tag{11}$$

As an informative prior distribution ($\pi_1(\theta)$), a Beta distribution with parameters α and β was assumed. To obtain the values of the hyperparameters, we did not have prior information obtained from a specialist or alternative source, so we used the procedure developed by Tovar (2012) considering fixed intervals and the Chebychev inequality to obtain the parameters of the prior distribution. Then, a Beta$(3.4, 23.6)$, Beta$(7.5, 7.5)$ and Beta$(23.6, 3.4)$ were used as informative prior distributions for weak, moderate and strong dependence levels, respectively. A Beta$(1, 1)$ and a Beta$(0.5, 0.5)$ distributions were used as noninformative prior distributions.

The Bayes estimates were obtained using MCMC methods implemented in the interface between `R` and OpenBUGS software.

2.4. Confidence Intervals and Credibility Regions

We obtained 0.95 confidence intervals for the estimates obtained using the maximum likelihood and moments methods and 0.95 credibility regions for the Bayes estimates.

The confidence intervals for the maximum likelihood (θ_{ML}) and moments estimates (θ_{Mom}) were obtained using bootstrap methods, as in Efron (1992), as follows:

Let $(u_1, v_1), \dots, (u_n, v_n)$ be a random bivariate vector of data with dependence structure that can be modelled with a GB copula function. It is possible to generate B samples with replacement from the vector of dependent bivariate data assuming that n is the population size. For each vector, the estimates $\hat{\theta}_{bML}$ and $\hat{\theta}_{bMom}$, $b = 1, \dots, B$ are obtained. To obtain the confidence intervals, we used the formula:

$$IC[\theta]_{1-\alpha} = [\hat{\theta}_{\frac{\alpha}{2}}, \hat{\theta}_{1-\frac{\alpha}{2}}] \tag{12}$$

The average length and the probability of coverage were used to compare the performance of the confidence intervals and credibility regions. The average length is defined as:

$$L_{aver} = \frac{\sum_{i=1}^{M}(L_{s_i} - L_{I_i})}{M} \tag{13}$$

where L_{s_i} is the upper limit and L_{I_i} is the lower limit of the ith interval. The probability of coverage is defined as:

$$CP = \frac{\sum_{i=1}^{M} I_i(\theta \, \epsilon \, IC_{ik})}{M} \tag{14}$$

where $I_i(\theta \epsilon IC_{ik})$ is the indicator function, which takes the value of one if the ith interval contains the true value of the parameter and is otherwise zero.

3. Simulation Study

To simulate vectors of pairs of data with a GB structure of dependence, we wrote a program in R language to implement an algorithm based on the inverse transformation of the cumulative distribution function of the variable X, given a value of the variable Y, as follows:

- Given a value of $\theta \in (0, 1)$
- Simulate $w_1 \sim uniform(0, 1)$ and simulate $w_2 \sim uniform(0, 1)$
- Then $y = -\ln(1 - w_1)$
- Substitute y and w_2 in

  $$A = 1 - w_2 - (1 + \theta x)e^{-(1+\theta y)x}$$

  and obtain x doing $A = 0$

The x and y values are realizations of the random variables X and Y whose natural behavior can be modeled using a Gumbel Type I distribution under the assumption that $X \sim exponential(1)$ and $Y \sim exponential(1)$. Carrying out $u = 1 - e^{-x}$ and $v = 1 - e^{-y}$, we obtain normalized values of two variables whose dependence structure can be modeled using a GB copula function.

The simulation procedure was based on scenarios defined by three dependence levels and 20 different sample sizes taking values between 50 and 1000 and increasing each 50. We assumed that, the dependence level is weak when $\theta = 0.2$, moderate for $\theta = 0.5$ and strong if $\theta = 0.9$. Each scenario was simulated 1000 times, and for each repetition of the procedure, we obtained the maximum likelihood, moments and Bayes estimators (including credible regions). For each simulated sample, we applied nonparametric boostrap to obtain the estimation intervals and for each interval, we determined whether it contained the true value of the parameter and computed its length. Finally, we obtained the proportion of times (among 1000 samples) that the interval contained the true value of the parameter, and we computed the average of the lengths.

4. Results

We computed the bias, mean square error (MSE), average length and the probability of coverage for each estimate, based on the sample size used. The performance of the bias obtained is showed in Figures 1-3. As can be observed when the dependence level is weak, the different estimators showed little bias and tended to overestimate the parameter when the sample sizes were lower than 200. The Bayesian estimators that assumed Beta(a, b) distributions had the highest bias levels when the sample size was the lowest of those studied; while the estimators of maximum likelihood and moments had the lowest bias. For moderate dependency levels, the estimator obtained using the moments method overestimated the value of the parameter with the greatest bias observed, regardless of sample size; whereas the Bayesian estimator obtained assuming an a priori Beta(a, b) distribution had the lowest bias observed for the smallest sample size. The estimates obtained with the other methods were all very similar in their behavior, but the Bayesian estimates assuming Beta$(1, 1)$ and Beta$(0.5, 0.5)$ distributions overlap. In the case of strong dependencies, the tendency is opposite to that observed for weak dependencies (all estimators underestimate the parameter); nevertheless, it is important to take into account that the Moments estimator shows this behavior until the sample size reaches 300, after which it begins to overestimate the value real (Figures 1-3).

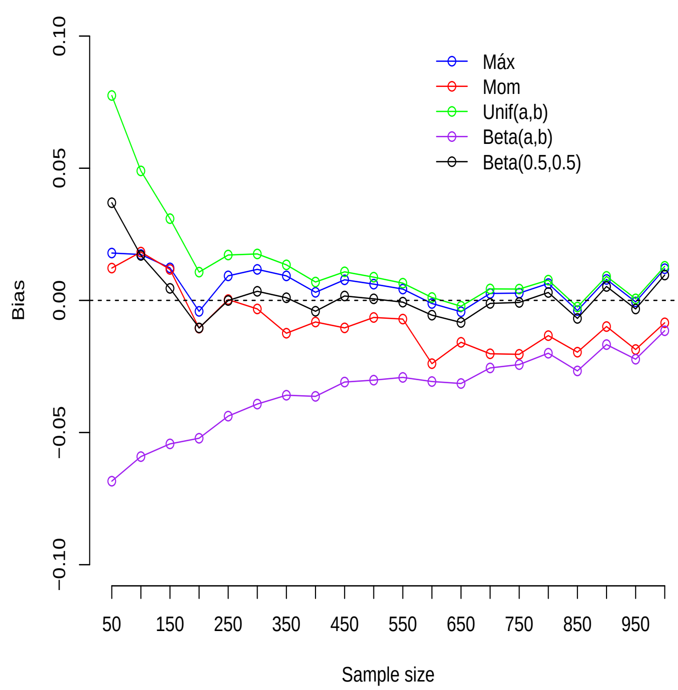

Figure 1: Bias of the estimates when the dependence is weak.

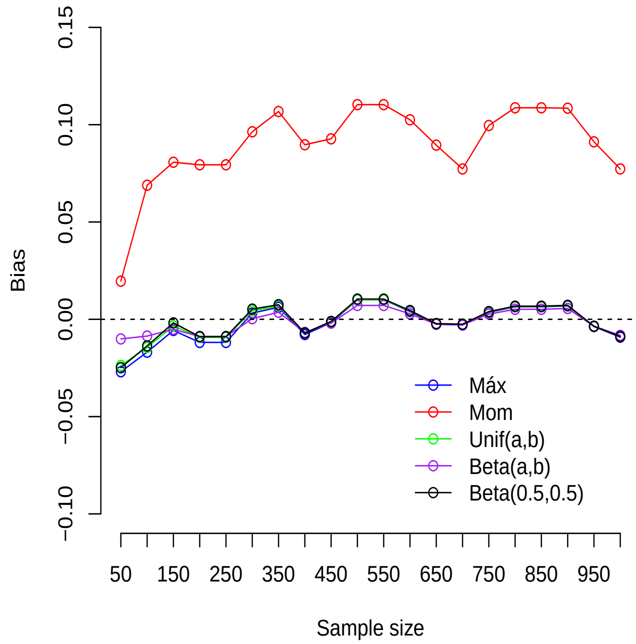

FIGURE 2: Bias of the estimates when the dependence is moderate.

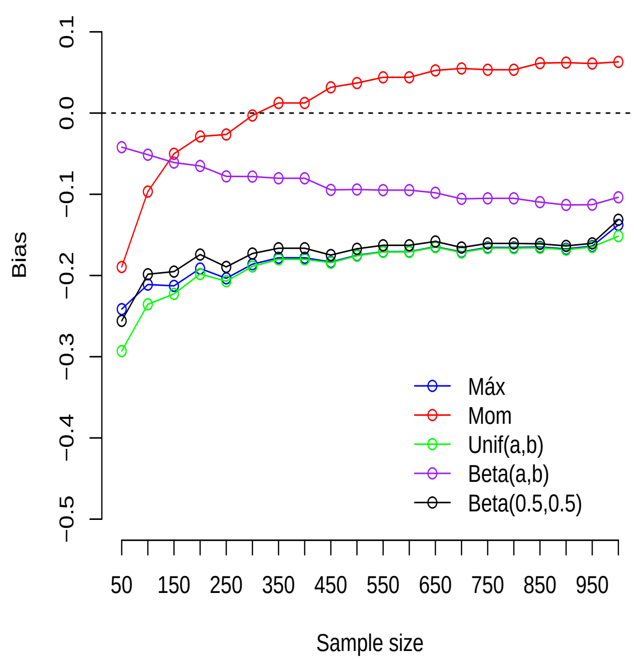

FIGURE 3: Bias of the estimates when the dependence is strong.

The performance of the MSE obtained for the estimates is showed in Figures 4-6. If the structure of a Gumbel Barnett dependency can be considered weak or moderate y regardless of the sample size, the MSE behaved similarly to the estimator for maximum likelihood and the Bayesian estimators assuming $\text{Beta}(1,1)$ and $\text{Beta}(0.5,0.5)$ distributions. The estimator obtained with the moments method had the highest MSE observed; while the Bayesian estimator obtained assuming an a priori $\text{Beta}(a,b)$ distribution had the lowest value. With strong dependency structures, the non-informative ML and Bayesian estimators —assuming a priori $\text{Beta}(1,1)$ and $\text{Beta}(0.5,0.5)$ distributions— had a higher MSE for all sample sizes. The estimator of Moments had higher MSE values when the sample sizes were below 200; for larger sample sizes, the MSE was lower than that observed for the three previous estimators. The Bayesian estimator with a priori $\text{Beta}(a,b)$ distribution had the lowest MSE for all the sample sizes evaluated when the sample sizes were under 400. The MSE tended towards zero; but for the larger sample sizes, the tendency was to increase, but not significantly.

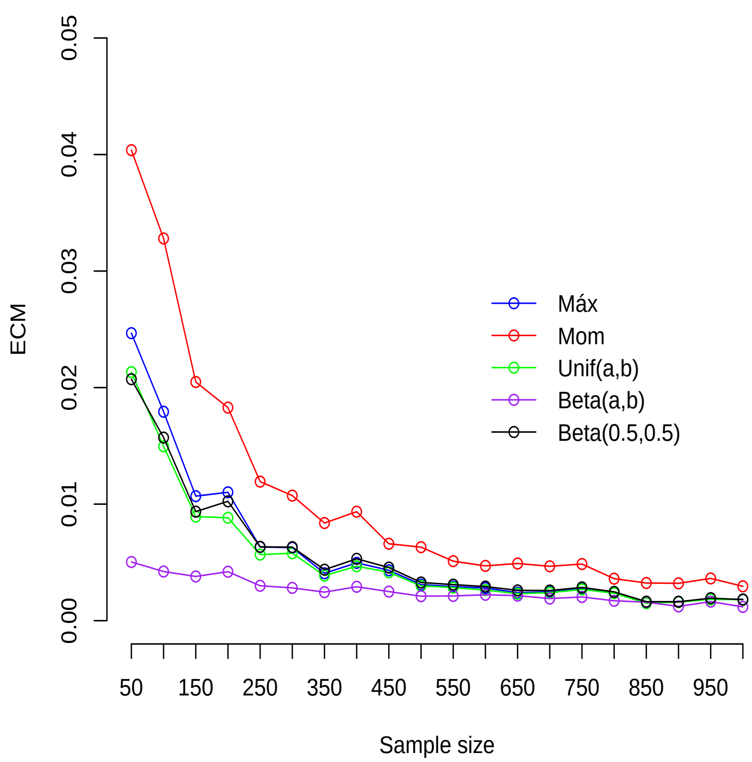

Figure 4: MSE of the estimates when the dependence is weak.

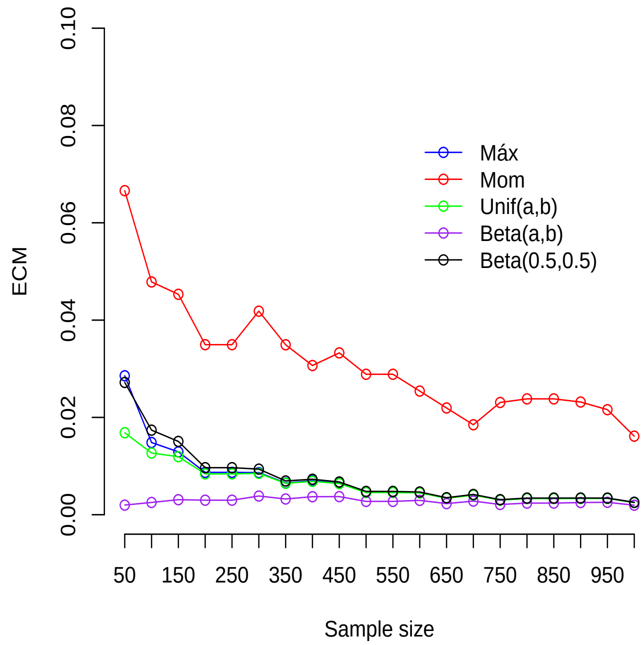

Figure 5: MSE of the estimates when the dependence is moderate.

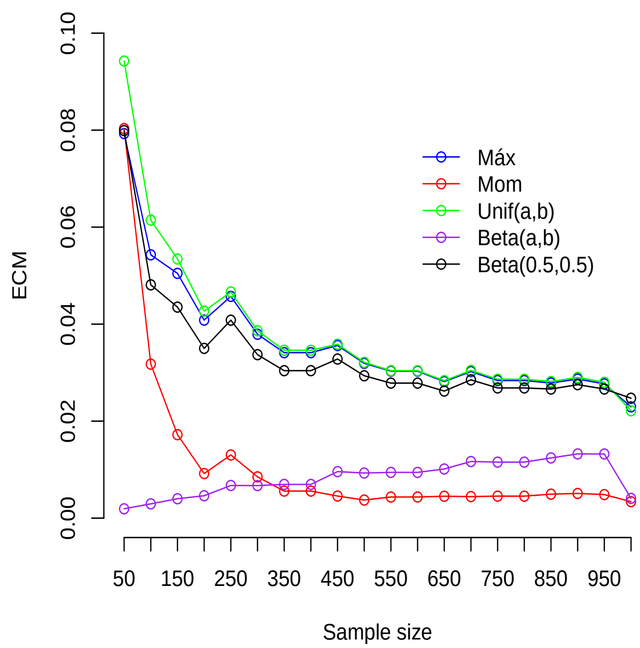

Figure 6: MSE of the estimates when the dependence is strong.

In accordance with our evaluation criteria, we expected that the real value of the parameter would be at least 95% of the confidence intervals or credibility regions. The probabilities of coverage observed are shown in Figures 7-9. When the dependency structure was weak, the estimators obtained using classical methods (ML and Moments) had the lowest probabilities of coverage; but with the three estimators obtained using the Bayesian methods, their behavior was more in line with what was expected (Figure 7) If the dependency is moderate and the sample size is more than 250, the probabilities of coverage of the Moments estimator tend to be lower (Figure 8). In the case of strong dependencies, the probabilities of coverage for the different estimators have variable behaviors: when the sample size is under 100, the probability of coverage tends to be near zero for all the estimators. The Moments estimator performed best with respect to this quality indicator. The probabilities of coverage of the maximum likelihood and non-informative Bayesian estimators performed similarly, tending to decrease as sample size increases. The Bayesian estimator assuming an a priori Beta(a, b) distribution had probabilities of coverage within what was expected when the sample sizes ranged from $100 - 400$; from 450 upwards, the probabilities dropped towards zero (Figure 9),

The other criterion used to evaluate the quality of the estimators was the average length of the confidence interval or region of credibility. Given that the shortest length possible is observed when the dependence is weak or moderate (Figures 10-11), the Bayesian estimator with a priori Beta(a, b) performs best for this indicator in all the sample sizes. The indicator for the Moments estimator has the worst behavior for all sample sizes. If the level of dependence is strong, all estimators had average lengths greater than 0.2, the moments estimator has a highly erratic behavior; while the Bayesian estimator with a priori Beta(a, b) distribution has the stablest behavior (Figure 12). The dependence is weak or moderate (Figures 10-11). The Bayesian estimator with a priori Beta(a, b) distribution behaved best for the indicator in all sample sizes. The indicator for the Moments estimator behaved the worst for all sample sizes. If the level of dependence was strong, all the estimators had average lengths greater than 0.2. The Moments estimator was characterized by a very erratic behavior; whereas the Bayesian estimator with an a priori Beta(a, b) distribution had the stablest behavior (Figure 12).

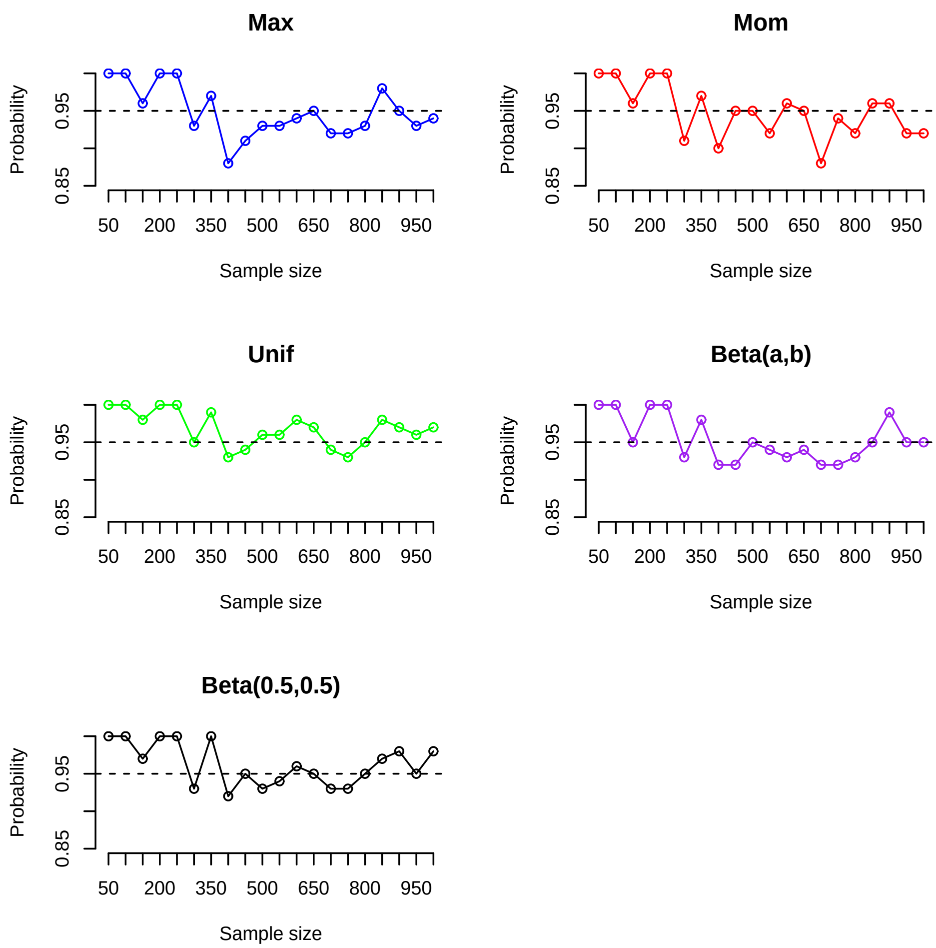

Figure 7: Probability of coverage for the confidence intervals and credibility regions when dependence is weak.

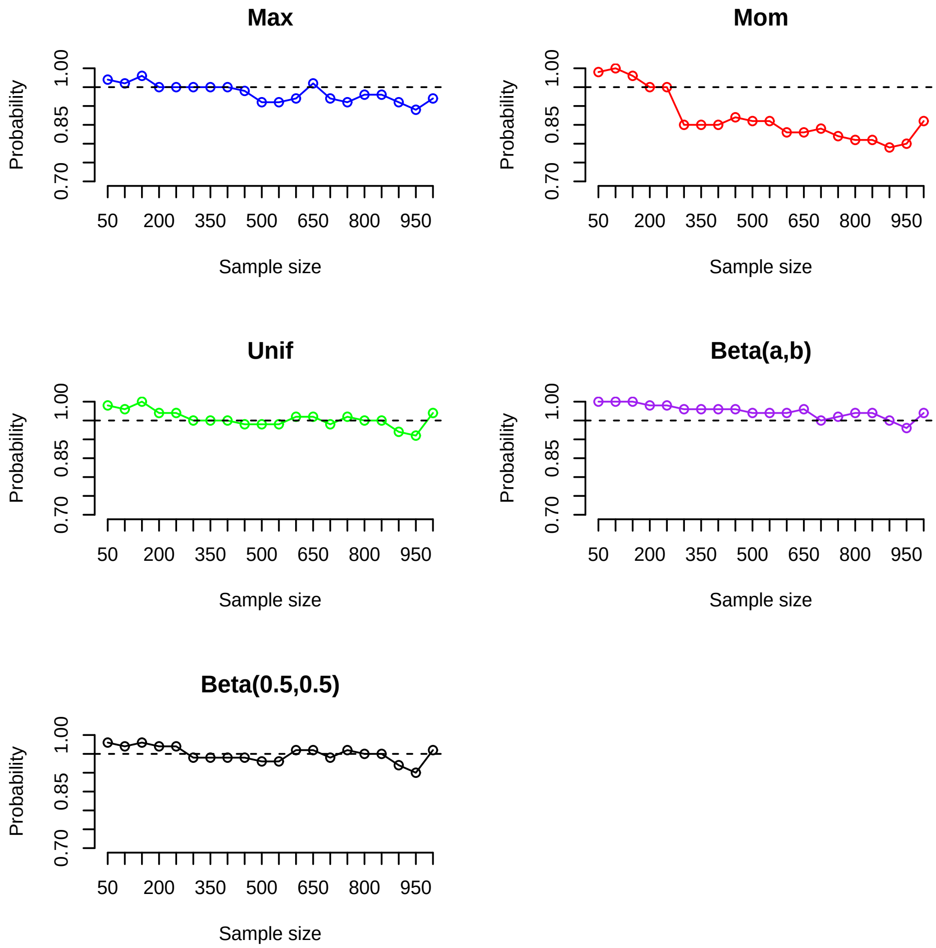

FIGURE 8: Probability of coverage for the confidence intervals and credibility regions when dependence is moderate.

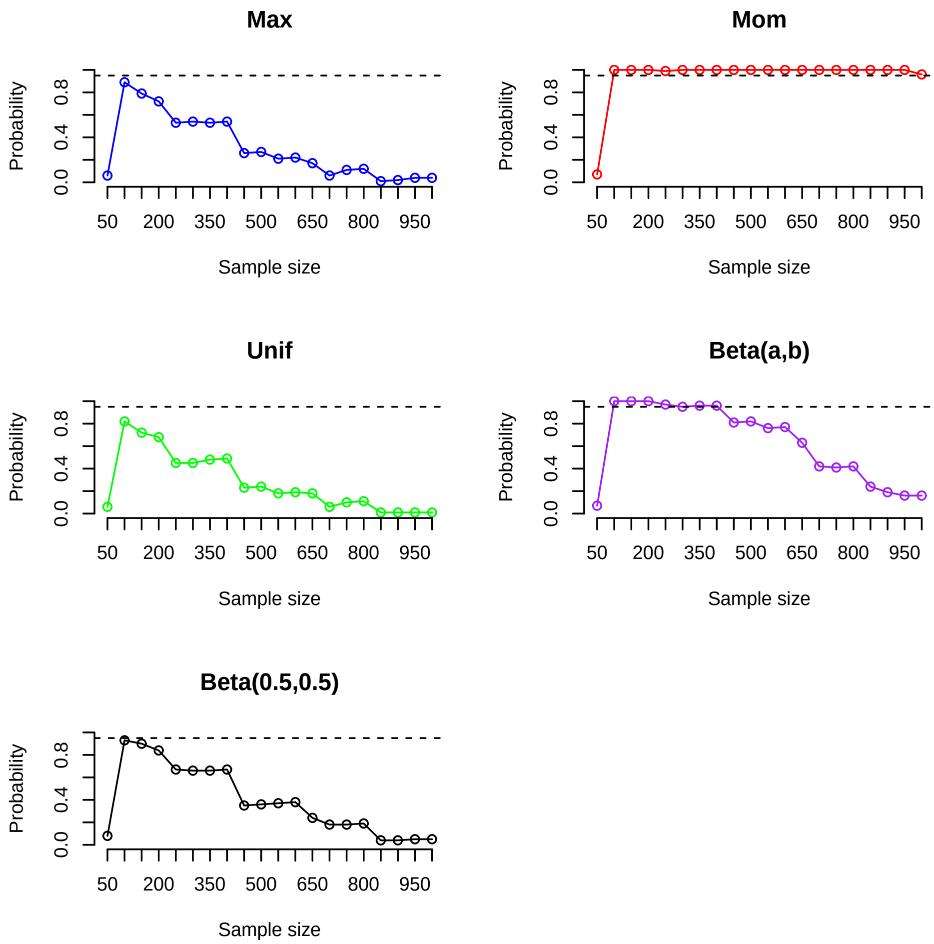

Figure 9: Probability of coverage for the confidence intervals and credibility regions when dependence is strong.

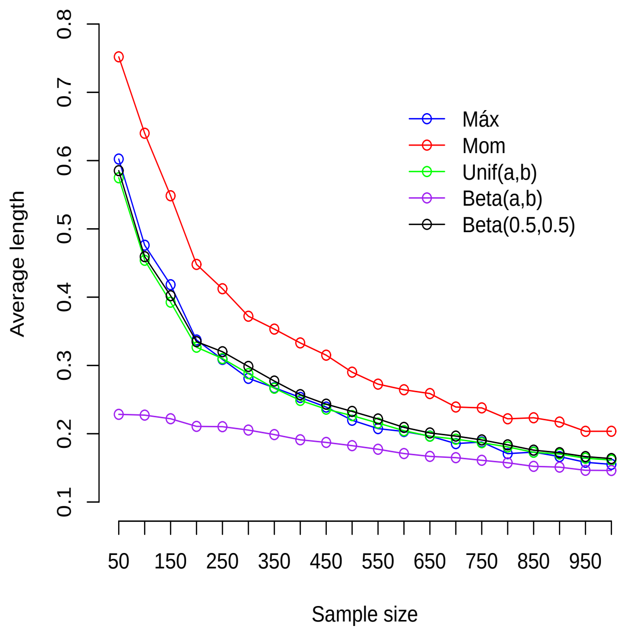

FIGURE 10: Average length of the confidence and credibility regions: weak dependence.

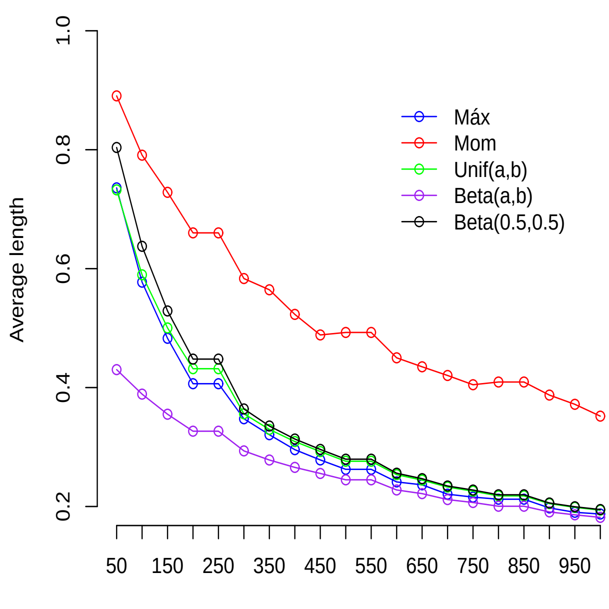

FIGURE 11: Average length of the confidence and credibility regions: moderate dependence.

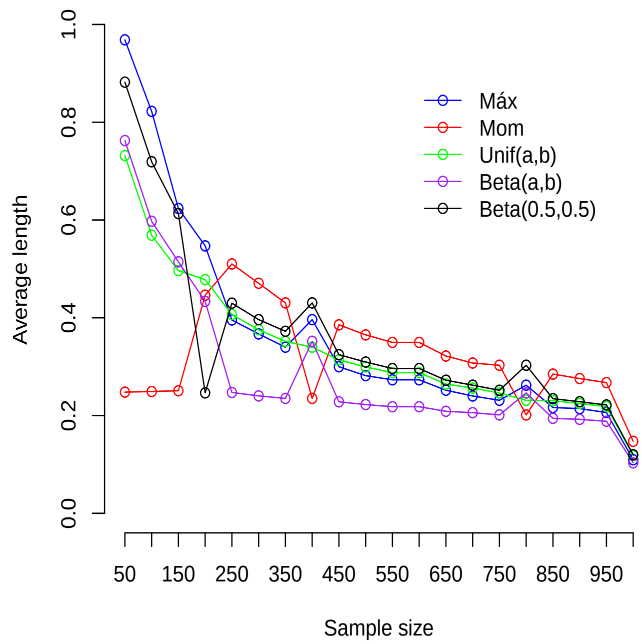

Figure 12: Average length of the confidence and credibility regions: strong dependence.

5. Discussion and conclusions

Many authors, including Genest, Ghoudi & Rivest (1995), Hoff (2007), Menger sen, Pudlo & Robert (2013), Min & Czado (2010) and Oh & Patton (2013), have focused their studies on developing procedures to estimate the dependence parameters of copula families; however, none of them has studied the empirical performance of their proposed estimators. Some authors have studied the estimation problem for the Gumbel Hougaard form (e.g., Weiß 2011), but none used the non-parametric method of moments as an estimation procedure. Estimating the dependence parameter of a Gumbel-Barnett copula is a complicated task that involves the use of elaborate computational methods of approximation, given that the analytical structure of the copula function, the likelihood and log likelihood functions have complex and non-closed forms. In this paper, we developed an empirical study of the performance of the estimates obtained from five different procedures to estimate the dependence parameter of the Gumbel-Barnett copula function, considering three dependence levels and 20 sample sizes. We used the analytical copula form obtained from the Gumbel Type 1 distribution with exponential marginals. In our review of the literature, we did not find classic or Bayesian studies on the empirical performance of different estimators for the Gumbel-Barnett copula dependence parameter.

Based on our results, we can conclude that the performance of the different estimators depends on the sample size and the dependence level. In situations where the sample size is relatively small (approx. 50), conventional estimation methods, such as moments and maximum likelihood, does not perform well when estimating weak or moderate GB dependencies. As the sample size increases, all the estimators show good performance. When the dependence is strong and the sample size is larger than 300 the moments estimator is the best option. In general, a estimator obtained assuming a $\mathrm{Beta}(a, b)$ as prior informative distribution, presents a nice performance to estimate the GB dependence parameter independent of the sample size and the dependence level.

Given that the dependence parameter of the Gumbel Barnett copula function ranges in the interval $(0; 1)$, it would be natural to think that the distribution $\mathrm{Beta}(a, b)$ is a good option for a prior distribution. Considering the analytical form of the likelihood function, the posterior distribution has a structure of an additive mixture, whose components are products of Beta distributions with an exponential function and an additional normalization constant to be computed. The approximation of that posterior distribution is very complex given the difficulty of establishing a candidate distribution to use the Metropolis-Hastings algorithm. Wherever, our results allow us conclude that it is possible to obtain estimates of very good performance for the dependence parameter of the Gumbel-Barnett copula, using the $\mathrm{Beta}(a, b)$ as prior distribution.

In this paper, we used the classic confidence intervals (CI) and Bayesian credibility regions (CR) in a similar way to evaluate the quality of our estimates, although the CI and the CR are concepts with entirely different interpretations. The CI establishes two values for the estimator within which it is expected to find the unknown quantity (i.e., the parameter): a specified percentage of times; whereas the CR should provide a proportion of values of the unknown quantity equal to an established probability. Although conceptually the confidence intervals and the credibility regions are different, for practical purposes, their average lengths and percentages of coverage obtained after many repetitions of the estimation procedure can be used as indicators of the performance of the estimates obtained. It is important to consider that our proposed estimation procedures use normalized uniform data. If unknown marginal distributions are assumed, the marginal distribution parameters should be estimated in addition to the dependence parameter, which involves a little more work.

It is important to consider that our proposed estimation procedures use normalized uniform data. If unknown marginal distributions are assumed, the marginal distribution parameters should be estimated in addition to the dependence parameter, which involves a little more work.

Acknowledgement

This work was supported by a grant for the project (code 3179) of the Foundation for the Promotion of Research and Technology of the Bank of the Republic in Colombia. The participation of the first author was partially financed by a scholarship from the Virginia Gutierrez de Pineda Program for Young Researches and Innovators of the Administrative Department of Science, Technology and Innovation (Colciencias) in Colombia.

[Received: May 2017 — Accepted: November 2017]

References

Balakrishnan, N. & Lai, C.-D. (2009), *Continuous bivariate distributions*, Springer Science & Business Media.

Barnett, V. (1980), 'Some bivariate uniform distributions', *Communications in statistics-theory and methods* **9**(4), 453–461.

Barnett, V. (1983), 'Reduced distance measures and transformations in processing multivariate outliers', *Australian & New Zealand Journal of Statistics* **25**(1), 64–75.

Efron, B. (1992), Bootstrap methods: another look at the jackknife, *in* 'Breakthroughs in Statistics', Springer, pp. 569–593.

Genest, C., Ghoudi, K. & Rivest, L.-P. (1995), 'A semiparametric estimation procedure of dependence parameters in multivariate families of distributions', *Biometrika* pp. 543–552.

Genest, C. & Mackay, R. J. (1986), 'Copules archimédiennes et families de lois bidimensionnelles dont les marges sont données', *Canadian Journal of Statistics* **14**(2), 145–159.

Gumbel, E. J. (1960), 'Bivariate exponential distributions', *Journal of the American Statistical Association* **55**(292), 698–707.
*http://www.jstor.org/stable/2281591

Hoff, P. D. (2007), 'Extending the rank likelihood for semiparametric copula estimation', *The Annals of Applied Statistics* pp. 265–283.

Klein, I. & Christa, F. (2011), Families of copulas closed under the construction of generalized linear means, Technical report, IWQW discussion paper series.

Louie, H. (2014), 'Evaluation of bivariate archimedean and elliptical copulas to model wind power dependency structures', *Wind Energy* **17**(2), 225–240.

Martinez, E. Z. & Achcar, J. A. (2014), 'Bayesian bivariate generalized lindley model for survival data with a cure fraction', *Computer methods and programs in biomedicine* **117**(2), 145–157.

Menger sen, K. L., Pudlo, P. & Robert, C. P. (2013), 'Bayesian computation via empirical likelihood', *Proceedings of the National Academy of Sciences* **110**(4), 1321–1326.

Min, A. & Czado, C. (2010), 'Bayesian inference for multivariate copulas using pair-copula constructions', *Journal of Financial Econometrics* **8**(4), 511–546.

Nelsen (2006), *An introduction to copulas*, Vol. 139, Springer Science & Business Media.

Oh, D. H. & Patton, A. J. (2013), 'Simulated method of moments estimation for copula-based multivariate models', *Journal of the American Statistical Association* **108**(502), 689–700.

Omidi, M. & Mohammadzadeh, M. (2015), 'A new method to build spatiotemporal covariance functions: analysis of ozone data', *Statistical Papers* (3), 1–15.

Tovar, J. (2012), 'Eliciting beta prior distributions for binomial sampling', *Revista Brasileira de Zoologia* **30**(1), 159–172.

Tovar, J. & Achcar, J. (2011), 'Indexes to measure dependence between clinical diagnostic tests: A comparative study', *Revista Colombiana de Estadıstica* **34**, 433–450.

Tovar, J. & Achcar, J. (2012), 'Two dependent diagnostic tests: Use of copula functions in the estimation of the prevalence and performance test parameters', *Revista Colombiana de Estadística* **35**(3), 331–347.

Tovar, J. & Achcar, J. (2013), 'Dependence between two diagnostic tests with copula function approach: a simulation study', *Communications in Statistics-Simulation and Computation* **42**(2), 454–475.

Weiß, G. (2011), 'Copula parameter estimation by maximum-likelihood and minimum-distance estimators: a simulation study', *Computational Statistics* **26**(1), 31–54.

Appendix

```
#GUMBEL MODEL ON OPENBUGS WITH A PRIORI Uniform(0,1)
sink("modelo1.txt")
cat("  model{
    for(i in 1 : n) {
    zeros[i] <- 0
    phi[i] <- -log(L[i])
    zeros[i] ~ dpois(phi[i])
    L[i]<- exp(-t1[i]- t2[i]- rho*t1[i]*t2[i] +
    log(1-rho+rho*t1[i] +rho*t2[i]+pow(rho,2)*t1[i]*t2[i]))}

    rho ~ dunif(0, 1)}}",fill=TRUE)
sink()
```

```
#GUMBEL MODEL ON OPENBUGS WITH A PRIORI Beta$(a,b)$
sink("modelo2.txt")
cat(" model{
    for(i in 1 : n) {
    zeros[i] <- 0
    phi[i] <- -log(L[i])
    zeros[i] ~ dpois(phi[i])
    L[i]<- exp(-t1[i]- t2[i]- rho*t1[i]*t2[i] +
    log(1-rho+rho*t1[i] +rho*t2[i]+pow(rho,2)*t1[i]*t2[i]))
    }

    rho ~ dBeta$(a,b)$} }",fill=TRUE)
sink()

#GUMBEL MODEL ON OPENBUGS WITH A PRIORI Beta(0.5,0.5)
sink("modelo3.txt")
cat(" model {
    for(i in 1 : n) {
    zeros[i] <- 0
    phi[i] <- -log(L[i])
    zeros[i] ~ dpois(phi[i])
    L[i]<- exp(-t1[i]- t2[i]- rho*t1[i]*t2[i] +
    log(1-rho+rho*t1[i] +rho*t2[i]+pow(rho,2)*t1[i]*t2[i]))
    }
     rho ~ dBeta(0.5,0.5)}}",fill=TRUE)
sink()

#ALGORITHM TO SIMULATE GB DATA
GB=function(theta,n)
{x=0; u=0
  for(i in 1:n){
    u1=runif(1)
    w=runif(1)
    fun=function(x,u1,w,theta){
      y=-log(1-u1)
      1-(1+theta*x)*exp(-(1+theta*y)*x)-w
    }
    x[i]=uniroot(fun,c(0, 20),u=u1,theta=theta,w=w)$root[1]
    u[i]=u1 }
  return(cbind(u,v=1-exp(-x)))}

fi=0.2; n=850; B=1000;N=1000

Estimador=matrix(0,N,5); Intervalos=matrix(0,N,12)
c1=c2=c3=c4=c5=c6=0

for(j in 1:N){
Muestra=GB(fi,n)
  likeGumbel1<-function(theta){
    LikeG<- log(prod(exp(-theta*log(1-Muestra[,1])*log(1-Muestra[,2]))*
 ((((theta*log(1-Muestra[,1]))-1)*
 ((theta*log(1-Muestra[,2]))-1 ))-theta)))}

  Vero<-(optim(runif(1), likeGumbel1, method = "Brent",lower=0,upper=1,
  hessian = TRUE, control = list(fnscale = -1)))

  Est.M=Vero$par
  Var.M=-solve(Vero$hessian)[1]
```

```
  Est.Mom=Sim(Muestra)

 Boot=0;  BootM=0
  for(i in 1:B){
    elegir=sample(1:length(Muestra[,1]),replace=TRUE,size=length(Muestra[,1]))
    BMuestra=Muestra[elegir,]

    likeGumbel2<-function(theta){
      LikeG<- log(prod(exp(-theta*log(1-BMuestra[,1])*log(1-BMuestra[,2]))*
((((theta*log(1-BMuestra[,1]))-1)*((theta*log(1-BMuestra[,2]))-1 ))-theta)))}

    Boot[i]<-(optim(runif(1), likeGumbel2, method = "Brent",lower=0,upper=1,
    control = list(fnscale = -1)))$par
    BootM[i]<-Sim(BMuestra)
  }

 t1=-log(1-Muestra[,1]);t2=-log(1-Muestra[,2])
  Rho=c("rho")

Datos=list(t1=t1,t2=t2,n=n)

  Mod1=bugs(data=Datos,inits=NULL,parameters.to.save=Rho, model.file ="modelo1.txt",
  n.iter=10000,n.burnin=5000,n.chains=1)
  Mod2=bugs(data=Datos,inits=NULL,parameters.to.save=Rho, model.file = "modelo2.txt",
  n.iter=10000,n.burnin=5000,n.chains=1)
  Mod3=bugs(data=Datos,inits=NULL,parameters.to.save=Rho, model.file = "modelo3.txt",
  n.iter=10000,n.burnin=5000,n.chains=1)

  Est.B1=Mod1$summary[1]
  Est.B2=Mod2$summary[1]
  Est.B3=Mod3$summary[1]

  Estimador[j,1]=Est.M
  Estimador[j,2]=Est.B1
  Estimador[j,3]=Est.B2
  Estimador[j,4]=Est.Mom
  Estimador[j,5]=Est.B3
  Intervalos[j,(1:2)]=c(Est.M-1.96*sqrt(Var.M),Est.M+1.96*sqrt(Var.M))
  Intervalos[j,(3:4)]=c(quantile(Boot,0.025)[[1]],quantile(Boot,0.975)[[1]])
  Intervalos[j,(5:6)]=c(Mod1$summary[5],Mod1$summary[13])
  Intervalos[j,(7:8)]=c(Mod2$summary[5],Mod2$summary[13])
  Intervalos[j,(9:10)]=c(quantile(BootM,0.025)[[1]],quantile(BootM,0.975)[[1]])
  Intervalos[j,(11:12)]=c(Mod3$summary[5],Mod3$summary[13])

  if(Intervalos[j,1]<fi && Intervalos[j,2]>fi){c1=1+c1}
  if(Intervalos[j,3]<fi && Intervalos[j,4]>fi){c2=1+c2}
  if(Intervalos[j,5]<fi && Intervalos[j,6]>fi){c3=1+c3}
  if(Intervalos[j,7]<fi && Intervalos[j,8]>fi){c4=1+c4}
  if(Intervalos[j,9]<fi && Intervalos[j,10]>fi){c5=1+c5}
  if(Intervalos[j,11]<fi && Intervalos[j,12]>fi){c6=1+c6}
write(t(Estimador),ncolum=5,file=paste("Estimadores-",n,"-",fi,".txt"))
write(t(Intervalos),ncolum=12,file=paste("Intervalos-",n,"-",fi,".txt"))
}
```