Journal of Data Science **10**(2012), 711-722

The Chi-plot and Its Asymptotic Confidence Interval for Analyzing Bivariate Dependence: An Application to the Average Intelligence and Atheism Rates across Nations Data

Vitor A. A. Marchi[1], Francisco A. R. Rojas[1] and Francisco Louzada[2*]
[1] *Universidade Federal de São Carlos and* [2] *Universidade de São Paulo*

Abstract: Bivariate data analysis plays a key role in several areas where the variables of interest are obtained in a paired form, leading to the consideration of possible association measures between them. In most cases, it is common to use known statistics measures such as Pearson correlation, Kendall's and Spearman's coefficients. However, these statistics measures may not represent the real correlation or structure of dependence between the variables. Fisher and Switzer (1985) proposed a rank-based graphical tool, the so called chi-plot, which, in conjunction with its Monte Carlo based confidence interval can help detect the presence of association in a random sample from a continuous bivariate distribution. In this article we construct the asymptotic confidence interval for the chi-plot. Via a Monte Carlo simulation study we discovery the coverage probabilities of the asymptotic and the Monte Carlo based confidence intervals are similar. A immediate advantage of the asymptotic confidence interval over the Monte Carlo based one is that it is computationally less expensive providing choices of any confidence level. Moreover, it can be implemented straightforwardly in the existing statistical softwares. The chi-plot approach is illustrated in on the average intelligence and atheism rates across nations data.

Key words: Analysis of dependence, chi-plot, confidence intervals.

1. Introduction

The analysis of dependence between two random variables is primordial in several areas of knowledge. There are several procedures to identify or classify the kinds of dependence between variables and also several measures to quantify the magnitude of dependence. Although, measures and procedures for dependence analysis usually reveal certain characteristics, but not others. For instance, for analyze the data from two continuous random variables, the Pearson correlation coefficient, Spearman's ρ and Kendall's τ coefficients, are commonly used.

*Corresponding author.

However, these correlation coefficients may not represent the real correlation or the real structure of dependence between the variables. Furthermore, following (Genest and Favre, 2007) many alternative rank-based procedures have been proposed in the statistical literature for testing of independence between two random variables.

Besides the scatter plot of ranks, a graphical tool for detecting dependence, inspired from control charts and based on the chisquare statistic for independence in a two-way table, was proposed by Fisher and Switzer (1985) and fully illustrated in Fisher and Switzer (2001), namely, chi-plots. These graphs depend only on the data through their ranks and produce diagrams that are approximately horizontal under independence. The chi-plot is easy to check and it interprets the sign and the measure in a local dependence base. To help identify departures from independence the authors suggest superimposing horizontal guidelines on the plot, obtained via Monte Carlo simulation, the confidence interval (CI) bounds for the chi-plot.

Moreover, it is common to find different parts of a data set that appear to have different associations Ballester *et al.* (1997). For the human eye judge sometimes is a bit difficult detect the dependence in the scatterplot and in unusual situations of dependence the classical statistic tests can provide at best a single piece of information about a single form of association. However, the chi-plot is a rich graphical procedure for association details in contrast with usual statistical measures of dependence and tests.

In this paper we construct the chi-plot asymptotic confidence intervals (ACI) and compare the CI and ACI coverage probabilities via a Monte Carlo simulation study. An advantage of the ACI over the CI is that it provides choices of any confidence level directly from a normal distribution, whereas is need a Monte Carlo simulation for establish different c_p values for the CI because Fisher and Switzer (2001) just calculates three confidence levels. Moreover, it can be implemented straightforwardly in the existing statistical softwares. Section 2 approaches the construction of the chi-plot. Section 3 presents the construction of the ACI. Section 4 contains the results of Monte Carlo simulation study on the coverage probabilities comparison of the CI and ACI. Section 5 contain a real data example. Section 6 contain some final remarks.

2. Chi-plot

The chi-plot is a graphical representation of the measures of local dependence with an easy interpretation and with more information regarding the usual measures of correlation. Moreover, its use may be more advantageous than other techniques, as it shows the local dependence sign and the measure of dependence can be interpreted locally.

Let $(x_1, y_1), \cdots, (x_n, y_n)$ be a random sample of (X, Y) and $I(A)$ the function indicator of event A. For each data point (x_i, y_i) we have

$$\chi_i = \frac{H_i - F_i G_i}{(F_i(1-F_i)G_i(1-G_i))^{1/2}},$$
$$F_i = \frac{1}{n-1}\sum_{j\neq i} I(x_j \leq x_i),$$
$$H_i = \frac{1}{n-1}\sum_{j\neq i} I(x_j \leq x_i, y_j \leq y_i),$$
$$G_i = \frac{1}{n-1}\sum_{j\neq i} I(y_j \leq y_i),$$
$$\lambda_i = 4 * \mathrm{sign}_i * \max\{(F_i - 0.5)^2, (G_i - 0.5)^2\},$$

and

$$\mathrm{sign}_i = signal\{(F_i - 0.5)(G_i - 0.5)\}.$$

The chi-plot is the scatter-plot of (λ_i, χ_i), for all $|\lambda_i| < 4((1/(n-1)) - 0.5)^2$, with CI given by

$$\left(-\frac{c_p}{\sqrt{n}}, \frac{c_p}{\sqrt{n}}\right), \tag{1}$$

where c_p is 1.54, 1.78 and 2.18 correspond respectively, to $p = 0.90$, 0.95 and 0.99 of confidence (Fisher and Switzer, 2001), which is roughly the proportion of the observations falling within the horizontal bounds, as might be expected under independence.

Following Fisher and Switzer (1985), if y_i is a strictly increasing function of x_i, we have $\chi_i = 1$ and if y_i is a strictly decreasing function of x_i, we have $\chi_i = -1$. On the other hand, if the random variables are independent, when $n \to \infty$ the asymptotic distribution of λ_i is uniformly distributed in $\pm 4((1/(n-1)) - 0.5)$. When there is independence between x_i and y_i, χ_i is randomly distributed around zero. If y_i is increasing (decreasing) compared to x_i, we have $\lambda_i > 0$ (< 0). If Y is positively (negatively) associated with X, i.e. $Cov(Y, X) > 0$ (< 0), there is a tendency that most values of λ are larger (smaller) than zero. The χ_i is the correlation coefficient ϕ for dichotomous variables, which reduces the interpretation of χ_i to the locally Pearson correlation coefficient.

3. Chi-plot Asymptotic Confidence Interval

This section presents the development of the asymptotic distribution of statistic χ_i by similarity to the asymptotic distribution chi-square on a contingency table 2×2. From this distribution we obtain the asymptotic confidence interval for the values of χ_i.

Theorem 1. Consider the chi-plot as defined as in Section 2. If the random variables X and Y are independent, then asymptotically $\chi \sim N(0, 1/(n-1))$, where $N(0, 1/(n-1))$ is the normal distribution with zero mean and variance $1/(n-1)$.

Proof. For each pair of points (x_i, y_i) of the sample $(x_1, y_1), \cdots, (x_n, y_n)$, we can partition the plane (x, y) in four quadrants, represented by the sets A_i, B_i, C_i and D_i:

$$A_i = \{x \leq x_i \cap y \leq y_i\}, \qquad B_i = \{x \leq x_i \cap y > y_i\},$$
$$C_i = \{x > x_i \cap y > y_i\}, \quad \text{and} \quad D_i = \{x > x_i \cap y \leq y_i\}.$$

Let a_i, b_i, c_i and d_i be the numbers of the sample points in the sets A_i, B_i, C_i and D_i, respectively, excluding the point (x_i, y_i).

Table 1 shows the numbers in each quadrant represented by the relationship of the points (x, y) in relation of the point (x_i, y_i).

Table 1: Numbers of points in the sets A_i, B_i, C_i and D_i

	below or equal to x_i	higher than x_i
below or equal to y_i	a_i	d_i
higher than y_i	b_i	c_i

Using the observed frequency, we can rewrite the empirical equations of Section 2 as

$$F_i = \frac{a_i + b_i}{(n-1)}, \quad G_i = \frac{a_i + d_i}{(n-1)}, \quad \text{and} \quad H_i = \frac{a_i}{(n-1)}.$$

By replacing F_i, G_i and H_i in χ_i equation of chi-plot, we obtain:

$$\chi_i = \frac{\dfrac{a_i}{(n-1)} - \dfrac{(a_i+b_i)}{(n-1)}\dfrac{(a_i+d_i)}{(n-1)}}{\sqrt{\dfrac{(a_i+b_i)}{(n-1)}\dfrac{(c_i+d_i)}{(n-1)}\dfrac{(a_i+d_i)}{(n-1)}\dfrac{(b_i+c_i)}{(n-1)}}}$$
$$= \frac{\dfrac{(n-1)a_i - (a_i+b_i)(a_i+d_i)}{(n-1)^2}}{\dfrac{\sqrt{(a_i+b_i)(c_i+d_i)(a_i+d_i)(b_i+c_i)}}{(n-1)^2}},$$

and since $(n-1) = a_i + b_i + c_i + d_i$, then

$$\chi_i = \frac{a_i c_i - b_i d_i}{\sqrt{(a_i + b_i)(c_i + d_i)(a_i + d_i)(b_i + c_i)}}.$$

For Table 1, we can write the chi-square statistic as

$$\chi_1^2 = \frac{(n-1)(a_i c_i - b_i d_i)^2}{(a_i + b_i)(c_i + d_i)(a_i + d_i)(b_i + ci)},$$

in which one has an asymptotic chi-square distribution with one degree of freedom. From literature we know that χ_i is the ϕ coefficient. Therefore, $(n-1)(\chi_i)^2 = \chi_1^2$. From Conover (1971) [p. 182] we conclude that $\chi \sim N(0, 1/(n-1))$. In others words, $(n-1)(\chi_i)^2$ is asymptotically chi-square with one degree of freedom and asymptotically $\chi \sim N(0, 1/(n-1))$.

$\square$

From the Theorem 1, we can define the ACI with $(100 - \alpha/2)\%$ of confidence for χ_i with limits given by

$$\left(-\frac{z_{(1-\alpha/2)}}{\sqrt{(n-1)}}, \frac{z_{(1-\alpha/2)}}{\sqrt{(n-1)}} \right), \tag{2}$$

where $z_{(1-\alpha/2)}$ denotes the $(1-\alpha/2)$ quantile of the standard normal distribution. For instance, for the confidence levels of 90%, 95% and 99% we have the quantile values of 1.64, 1.96 and 2.57, respectively.

4. Comparing the Confidence Intervals

An convenient benefit of the ACI over CI is that we may perform any confidence interval using the normal quantiles. Moreover, the ACI are obtained in a closed form, are computationally extremely less expensive, and may be implemented straightforwardly in the existing statistical softwares. In this section we present the results of a simulation study performed in order to compare the CI and the ACI according to their coverage probabilities.

We consider the proportion of points outside of the CI and ACI, calculated over 300 bivariate independent samples (X, Y).

For the independence case, four independent samples (X, Y) were generated according to $X \sim N(50, 7^2)$ - $Y \sim N(30, 2^2)$, $X \sim U(0, 1)$ - $Y \sim U(0, 1)$, $X \sim \text{Exp}(1)$ - $Y \sim N(0, 1)$ and $X \sim \text{Exp}(1)$ - $Y \sim \text{Beta}(1, 2)$ for sample sizes of $n =$ 50, 100, 200 and 500. Figure 1 shows one simulated random sample for each configuration of independent samples (X, Y). Table 2 shows the mean for the proportion of points outside the CI and ACI of 90%, 95% and 99% of confidence

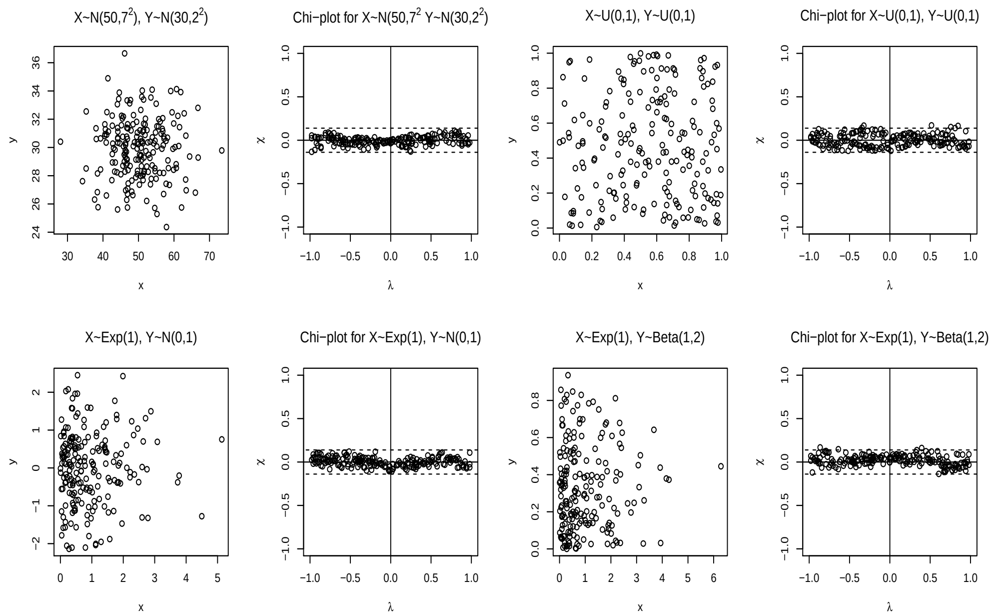

Figure 1: One selected random sample of each configuration of independence samples used in Table 2 with ACI of 95% of confidence for chi-plots

to bivariate independent samples (X, Y). In mean, the proportion of numbers of points outside the CI/ACI ranges are close to the nominal values (0.10, 0.05 and 0.01), but is obvious that we have a significant improvement from CI to ACI between the coverage probabilities.

For the dependence case, four dependent samples (X, Y) were generated according to $(X, Y) \sim$ Gumbel$(2, 3, -0.5)$, $(X, Y) \sim$ Gumbel$(1, 6, 0.9)$, $X \sim U(8, 12)$ and $Y \sim -(X - 10)^2 + N(0.8, 1) + 20$ (Quadratic) and samples with four normal different groups means for sample sizes of $n = 50$, 100, 200 and 500. Here, we consider a chi-square test of independent with significance $\alpha = 0.05$. The Pearson correlation in the Gumbel(α, β, λ) distribution (Gumbel, 1960, (3.1)) is $\lambda/4$. The samples with four normal different groups means each are obtained by making $(X_i, Y_i) \sim N(5, 5, 0.6, 0.6, 0.5)$; $i = 1, \cdots, 75$, $(X_i, Y_i) \sim N(5, 10, 0.6, 0.6, 0.5)$; $i = 76, \cdots, 150$, $(X_i, Y_i) \sim N(10, 5, 0.6, 0.6, 0.5)$; $i = 151, \cdots, 225$ and $(X_i, Y_i) \sim N(10, 10, 0.6, 0.6, 0.5)$; $i = 226, \cdots, 300$, where $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ is the bivariate normal whit marginals $X \sim N(\mu_1, \sigma_1)$, $Y \sim N(\mu_2, \sigma_2)$ and correlation coefficient ρ. Figure 2 shows one random simulated sample of each configuration of dependent samples (X, Y) that compound the Table 3. Table 3 shows the mean for the proportion of points outside the CI and ACI of 90%, 95% and 99% of confidence to bivariate dependent samples (X, Y). In mean, the proportion of numbers of points outside of the CI/ACI ranges are similar, increasing with the sample size as expected.

Table 2: Independence cases. Mean proportion of points outside CI and ACI ranges for independent samples

	$X \sim N(50,7^2)$ - $Y \sim N(30,2^2)$		
Conf. Level	90%	95%	99%
Inter.	CI/ACI	CI/ACI	CI/ACI
(n) 50	0.1131/0.0895	0.0650/0.0418	0.0340/0.0098
100	0.1202/0.0966	0.0722/0.0482	0.0305/0.0090
200	0.1191/0.0960	0.0725/0.0483	0.0272/0.0083
500	0.1181/0.0961	0.0718/0.0481	0.0288/0.0087
	$X \sim U(0,1)$ - $Y \sim U(0,1)$		
Conf. Level	90%	95%	99%
Inter.	CI/ACI	CI/ACI	CI/ACI
(n) 50	0.1215/0.0954	0.0714/0.0450	0.0305/0.0115
100	0.1239/0.0992	0.0758/0.0512	0.0320/0.0121
200	0.1228/0.0991	0.0741/0.0496	0.0344/0.0106
500	0.1222/0.0987	0.0722/0.0475	0.0322/0.0109
	$X \sim \text{Exp}(1)$ - $Y \sim N(0,1)$		
Conf. Level	90%	95%	99%
Inter.	CI/ACI	CI/ACI	CI/ACI
(n) 50	0.1319/0.1087	0.0853/0.0542	0.0307/0.0098
100	0.1170/0.0931	0.0704/0.0460	0.0323/0.0111
200	0.1222/0.0980	0.0733/0.0475	0.0320/0.0102
500	0.1204/0.0973	0.0724/0.0473	0.0312/0.0101
	$X \sim \text{Exp}(1)$ - $Y \sim \text{Beta}(1,2)$		
Conf. Level	90%	95%	99%
Inter.	CI/ACI	CI/ACI	CI/ACI
(n) 50	0.1196/0.0952	0.0706/0.0466	0.0307/0.0101
100	0.1194/0.0952	0.0719/0.0471	0.0301/0.0087
200	0.1205/0.0973	0.0725/0.0488	0.0361/0.0118
500	0.1175/0.0953	0.0705/0.0468	0.0348/0.0105

5. Application

As our first example consider the dataset extracted from Lynn *et al.* (2008), which consists of the association between the intelligence quotient (IQ) and the percentage of the population which do not believe in God, in several world nations. Here we use X to denotes the IQ and Y to denotes the percentage of the population which do not believe in God, respectively. The chi-plot in Figure 3 shows a global positive dependence. Although, in a general ground, the

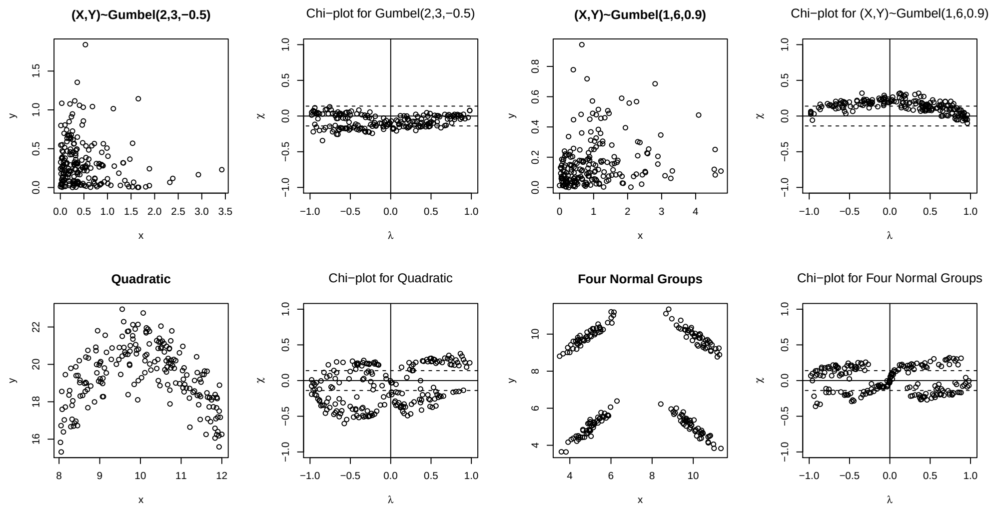

Figure 2: One selected random sample of each configuration of dependence samples used in Table 3 with ACI of 95% of confidence for chi-plots

correlation coefficients are given by r-Pearson $= 0.5973$, ρ-Spearman $= 0.7761$ and τ-Kendall $= 0.6039$.

The chi-plot in the left panel in Figure 3 suggests different association regions with a general behavior of linear dependence. The points in the chi-plot were classified by chi-plot quadrants with the origin $(\lambda, \chi) = (0, 0)$: open balls for $\lambda > 0$ and $\chi > 0$, open squares for $\lambda < 0$ and $\chi > 0$, open triangle for $\lambda > 0$ and $\chi < 0$. In the quadrant formed by $\lambda < 0$ and $\chi < 0$ no points were observed. Moreover, we can see a empty region on the first quadrant in the middle of the point cloud. This empty region subjectively indicates that in the data set more than one type of dependence behavior can be present, that is, for this data set we observe group of countries that have a different behavior in comparison with the others when the variables in study are the percentage of the population which do not believe in God and the IQ. Those were marked as solid balls in the chi-plot. The separation of the regions are made subjectively. For readers interested in the behaviors of Chi-plot to determine particulate relationships between the variables, we suggest reading Fisher and Switzer (1985) and Fisher and Switzer (2001).

When the solid balls in the Chi-plot of Figure 3 are marked and plotted in a scatterplot, as in the right panel of Figure 3, we have a clear separation in three regions. From scatterplot in the left panel of the Figure 3 we have evidence for three different regimes of dependence: $x < 86$: open balls, $86 \leq x < 98$: dark balls and $x \geq 98$: open squares. For the dark balls (region $86 \leq x < 98$) we have $\tau = 0.5706$, $\rho = 0.7219$ and $r = 0.6682$, directing to a positive dependence between the variables. While for the open balls (region $x < 86$) we have $\tau =$

Table 3: Dependence cases. Mean proportion of points outside CI and ACI ranges for dependent samples

	Gumbel$(2, 3, -0.5)$		
Conf. Level Inter.	90% CI/ACI	95% CI/ACI	99% CI/ACI
(n) 50	0.1975/0.1662	0.1341/0.0957	0.0678/0.0261
100	0.2450/0.2105	0.1730/0.1278	0.1024/0.0446
200	0.3735/0.3381	0.2923/0.2359	0.1641/0.0800
500	0.5608/0.5270	0.4815/0.4230	0.3930/0.2641
	Gumbel$(1, 6, 0.9)$		
Conf. Level Inter.	90% CI/ACI	95% CI/ACI	99% CI/ACI
(n) 50	0.2936/0.2563	0.2171/0.1648	0.1471/0.0724
100	0.4855/0.4445	0.3968/0.3297	0.2738/0.1593
200	0.6578/0.6293	0.5884/0.5327	0.4770/0.3462
500	0.8375/0.8226	0.8024/0.7720	0.7344/0.6495
	Quadratic		
Conf. Level Inter.	90% CI/ACI	95% CI/ACI	99% CI/ACI
(n) 50	0.5873/0.5544	0.5100/0.4483	0.4007/0.2741
100	0.7133/0.6894	0.6604/0.6200	0.5857/0.4790
200	0.7960/0.7809	0.7616/0.7359	0.7132/0.6478
500	0.8734/0.8653	0.8544/0.8402	0.8224/0.7859
	Four Normal Groups		
Conf. Level Inter.	90% CI/ACI	95% CI/ACI	99% CI/ACI
(n) 50	0.4979/0.4524	0.3969/0.3273	0.1765/0.0924
100	0.6481/0.6092	0.5633/0.4943	0.4240/0.2614
200	0.7990/0.7797	0.7528/0.7144	0.6771/0.5667
500	0.9028/0.8948	0.8831/0.8673	0.8492/0.8059

0.1394, $\rho = 0.1657$ and $r = 0.1803$, and for the open squares (region $x \geq 98$) we have $\tau = -0.1525$, $\rho = -0.1757$ and $r = -0.0846$. Both leading to non significant dependence between IQ and percentage of the population which do not believe in God. Overall, Lynn *et al.* (2008) concludes for the presence of positive dependence between IQ and no religious belief between nations. Our results support his finds but strongly for the region $86 \leq \text{IQ} < 98$. Hence, the chi-plot approach directs for the present of significant positive dependence between IQ and religious belief but in a specific range of IQ.

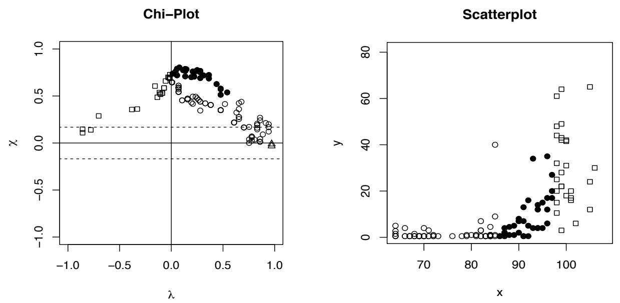

Figure 3: Left panel: chi-plot with points classified by chi-plot quadrants with the origin $(\lambda, \chi) = (0, 0)$: balls for $\lambda > 0$ and $\chi > 0$, squares for $\lambda < 0$ and $\chi > 0$, triangle for $\lambda > 0$ and $\chi < 0$, and diamond for $\lambda < 0$ and $\chi < 0$. Right panel: scatterplot with three regimes of dependence

Figure 4 intends to be a suggestion on how the world is divided by the dependence between IQ and the proportion of the people that do not believe in God considering the chi-plot results, leading to three different regions of dependence. First, the gray region (IQ < 86), where even a large variation in the IQ values correspond to small variation in the percentage of the people that do not believe in God, which is generally composed by nations of Africa, Central America and Caribbean, north west of South America and Central Asia. The dark gray region ($86 \leq$ IQ < 98) with a positive dependence between IQ and the proportion of the people that do not believe in God, which includes most of the nations of South America, USA, Russia, East Europe, East Asia, Portugal, Spain and Ireland. And the light gray region (IQ ≥ 98), where small variation in the IQ values correspond to large variation in the proportion of the people that do not believe in God, though such proportions are high. Such region is mainly composed by Europe, Canada and nations of Australia and Oceania. Blank regions were not sampled.

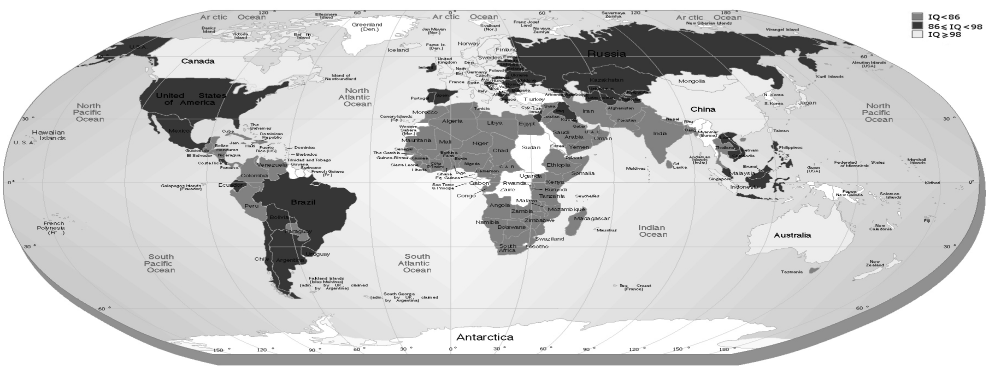

Figure 4: Suggestion on how the world map is divided by considering the relationship between IQ and the proportion of people that not believe in God. Blank regions were not sampled

6. Concluding Remarks

The chi-plot in conjunction with its ACI can help to detect the presence of association in a random sample from a continuous bivariate distribution. It is useful in practical studies where it is common to find different parts of the data set that appear to have different associations. For the human eye judge sometimes is a bit difficult detect the dependence in a scatterplot and in unusual situations of dependence the classical statistic tests can provide at best a single piece of information about a single form of association. However, the chi-plot is a rich graphical procedure for association details in contrast with formal tests like seen in Fisher and Switzer (2001) for mixtures.

Via a Monte Carlo simulation study we discovered that the coverage probabilities have a improvement from CIs to ACIs. Furthermore, the ACIs have a advantage over the Monte Carlo based one since it is computationally less expensive providing choices of any confidence level. Moreover, the ACI can be implemented straightforwardly in the existing statistical softwares. Interested readers can ask the authors for the R codes for constructing the chi-plot and its ACI.

As far as it is concerned, the chi-plot is a visual tool, which is bounded by the limitations of this type of statistical methodology. However, in our data analysis, the chi-plot has proven to be an important statistical tool to reveal interesting features of the data sets. Although we understand other possible techniques, probably those directed to relations of cause and effect, need to considered to endorse the dependency relationship among the considered variables, as well as, the data enrichment with other variables, such as those of economic nature. Particularly, in the IQ and percentage of people that do not believe in God example, which, for sure, may end up in controversial scenarios, which are out of the scope of the paper.

References

Ballester, F., Corella, D., Pérez-Hoyos, S., Sáez, M. and Hervás, A. (1997). Mortality as a function of temperature. A study in Valencia, Spain, 1991-1993. *International Journal of Epidemiology* **26**, 551-561.

Conover, W. R. (1971). *Practical Nonparametric Statistics*, 1st edition. Wiley Series in Probability and Mathematical Statistics, New York.

Fisher, N. I. and Switzer, P. (1985). Chi-plots for assessing dependence. *Biometrika* **72**, 253-265.

Fisher, N. I. and Switzer, P. (2001). Graphical assessment of dependence: Is a picture worth 100 tests? *The American Statistician* **55**, 233-239.

Genest, C. and Favre, Anne-Catherine (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* **12**, 347-368.

Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association* **55**, 698-707.

Lynn, R., Harvey, J. and Nyborg, H. (2008). Average intelligence predicts atheism rates across 137 nations. *Intelligence* **37**, 11-15.

Received March 8, 2012; accepted June 5, 2012.

Vitor A. A. Marchi
Department of Statistics
Universidade Federal de São Carlos
P.O. Box 676, Sao Carlos, 13.565-905, Brazil
vitor_marchi@yahoo.com.br

Francisco A. R. Rojas
Department of Statistics
Universidade Federal de São Carlos
P.O. Box 676, Sao Carlos, 13.565-905, Brazil
rojas@ufscar.br

Francisco Louzada
Department of Applied Mathematics and Statistics
Universidade de São Paulo
P.O. Box 668, São Carlos, 13.566-590, Brazil
louzada@icmc.usp.br