

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
DEPARTAMENTO DE ESTATÍSTICA

ESTUDO DE SIMULAÇÃO PARA COMPARAR O
DESEMPENHO DOS COEFICIENTES DE CORRELAÇÃO TIPO-C
EM RELAÇÃO AO COEFICIENTE DE CORRELAÇÃO DE
PEARSON PARA DADOS CATEGORIZADOS

Sérgio Kakuta Kato

Orientadora: Jandyra M. G. Fachel

Monografia apresentada para a obtenção
do título de Bacharel em Estatística.

Porto Alegre, Dezembro de 1999.

AGRADECIMENTOS

A conclusão de um curso de graduação é apenas um passo no desenvolvimento pessoal e criativo. Aumentar a nossa criatividade significa abrir novas oportunidades o que é fundamental para enfrentar os desafios de nosso tempo. Um tempo em que a única constante é a “mudança”.

Agradeço aos professores do curso de Estatística que foram responsáveis pela minha formação.

A meus familiares que sempre me apoiaram.

A todos os amigos, colegas do curso de Estatística e da Polícia Rodoviária Federal.

À amiga, professora e orientadora Jandyra Fachel pela sua grande dedicação e conhecimento não só na realização desta monografia mas também em projetos de pesquisa, assessoria, e pela minha formação profissional.

À Luciane pela sua dedicação e amor durante quase todo este curso.

Este trabalho de conclusão não teria sido desenvolvido sem a participação dos bolsistas de Iniciação Científica do Departamento de Estatística trabalhando sob orientação da professora Jandyra Fachel, a seguir citados, aos quais agradeço muitíssimo e mais do que isto, este é o resultado de um trabalho de equipe.

Eduardo D’Ávila, pela programação do software Crosspsi o qual facilitou e otimizou a nossa tarefa de simulação.

Hélio Bittencourt e Marcos Martins que iniciaram a técnica de simulação neste projeto.

Mathias A.B. Bressel, pela ajuda no *software* Maple V.

Agradeço também ao formando e secretário do NAE, Manoel Mendonça, responsável pela organização administrativa do projeto de pesquisa e Gilberto Mesquita, estatístico do NAE.

O projeto teve financiamento da PROPESQ/UFRGS e CNPq/UFRGS através das bolsas PIBIC e resultados parciais foram apresentados nos IX, X e XI Salões de Iniciação Científica de 1997 a 1999, tendo sido agraciado com prêmios destaque nos anos de 1997 e 1999.

SUMÁRIO

1- INTRODUÇÃO.....	4
2- RAZÃO DE PRODUTOS CRUZADOS GENERALIZADOS	11
2.1- Distribuição Tipo-C.....	11
2.2- Estimção de Máxima Verossimilhança para ψ	15
2.3 - Coeficientes de Correlação Tipo-C.....	18
3- <i>SOFTWARE</i> PARA O COEFICIENTE DE CORRELAÇÃO TIPO-C	28
3.1 - Manual do <i>software</i> Crosspsi.....	29
4- ESTUDO DE SIMULAÇÃO	37
5- RESULTADOS	46
6- CONCLUSÃO.....	58
7- BIBLIOGRAFIA.....	60
ANEXOS.....	63

1- INTRODUÇÃO

Em diversas áreas do conhecimento, é usualmente necessário medirmos a relação existente entre duas variáveis. Na área médica, por exemplo, uma das medidas de associação utilizada em tabelas de contingência 2x2, é a razão de chances (*odds ratio*), Ψ , uma medida de associação que varia de $[0;\infty)$, é definida pela fórmula

$$\Psi = \frac{a.d}{b.c}$$

onde a, b, c, d estão definidos no exemplo de tabela 2x2 a seguir que ilustra a relação entre o fator tabagismo e câncer no pulmão

Tabagismo (X)	Câncer no pulmão (Y)	
	Tem (+)	Não tem (-)
Fuma (+)	20 (a)	10 (b)
Não fuma (-)	10 (c)	20 (d)

O coeficiente de associação Ψ (*odds ratio*) expressa o número de vezes em que a presença do fator (no exemplo tabagismo) aumenta a chance de ocorrência da doença (câncer no pulmão), comparado com a ausência do fator.

No exemplo, $\Psi = 4$ indica que quem fuma pode ter 4 vezes mais chances de ter câncer no pulmão, comparado com quem não fuma.

É possível expressar a medida de associação em termos de uma medida de correlação, isto é, em termos de um coeficiente de correlação como uma medida de relação. Já no início do século, Yule (1900) propôs a seguinte fórmula, para transformar o coeficiente de associação Ψ , em um coeficiente de correlação variando no intervalo $[-1;1]$.

$$Q = \frac{\Psi - 1}{\Psi + 1}$$

No exemplo acima, $Q = 0,60$ ou seja pode existir uma correlação positiva média (0,60) entre tabagismo e câncer no pulmão.

A medida de associação Ψ havia sido desenvolvida para aplicação apenas em tabelas de contingência 2×2 (variáveis binárias). Porém, na prática, deparamo-nos muitas vezes com variáveis categóricas ordinais politômicas, ou seja, com mais de duas categorias.

Este tipo de escala é amplamente utilizado em áreas como Administração, Psicologia e Sociologia, constituindo-se muitas vezes na única forma possível de mensuração de constructos subjetivos como satisfação, atitude e inteligência, por exemplo, que não são diretamente mensuráveis.

Exemplo de variáveis categóricas ordinais:

1- Avaliação

(1) Ótimo (2) Bom (3) Regular (4) Ruim (5) Péssimo

2- Opinião

(1) Concorda (2) Indiferente (3) Discorda

3- Estágio de uma doença

(1) Progressiva (2) Estacionária (3) Regressiva

Supomos agora, que exista uma distribuição bivariada contínua associada aos dados da tabela de contingência 2×2 e que as dicotomias tenham sido formadas por pontos de corte arbitrários nas distribuições marginais. Seria conveniente que tivéssemos uma distribuição bivariada com a propriedade de que a associação entre as variáveis fosse constante, independentemente dos pontos de corte escolhidos para formar as dicotomias.

Pearson e Heron (1913) mostraram que é sempre possível construir uma superfície para a qual o parâmetro de associação é constante, independentemente dos pontos de corte escolhidos nas marginais X e Y contínuas. Esta distribuição bivariada contínua foi denominada "superfície de associação constante". Plackett

(1965) reintroduziu a mesma distribuição, no contexto de uma família de distribuições bivariadas contínuas com um único parâmetro de associação (Ψ) o “*odds ratio*”, cujos membros da família variam conforme a escolha das distribuições marginais que podem ser normal, uniforme, logística e outras. Mardia (1970) mostrou as propriedades desta família de distribuições e a denominou Distribuição Tipo-Contingência, ou abreviadamente, Distribuição Tipo-C. Esta denominação tem a vantagem de podermos denominar cada membro da família de distribuições com um nome apropriado, como por exemplo, Distribuição Tipo-C Normal, para variáveis com distribuições marginais Normal, Distribuição Tipo-C Logística, quando as distribuições marginais são Logística etc.

O parâmetro de associação desta distribuição é estimado pela razão de produtos cruzados para dados apresentados em tabelas de contingência 2×2 , supondo-se que tenham sido gerados por um dos membros da família Tipo-C.

Para o caso geral, quando os dados são apresentados em uma tabela de contingência $R \times C$, a qual tem variáveis marginais observadas com R e C categorias ordenadas que podem ser oriundas de questionários ou instrumentos de medida com escalas de Likert, não podíamos utilizar a medida de associação Ψ pois ela estava definida apenas para tabelas 2×2 . Fachel (1986) propôs a estimativa por máxima verossimilhança do parâmetro de associação ψ para dados observados em tabelas de contingência $R \times C$, supondo uma distribuição subjacente bivariada contínua dada pela Distribuição Tipo Contingência, ou Superfície de Associação Constante (ver Bartholomew, 1993).

Todos os coeficientes baseados na razão de produtos cruzados para tabelas 2×2 são coeficientes tipo-contingência, pela definição acima. E, a partir da generalização da razão de produtos cruzados para tabelas de contingência $R \times C$ (ver Fachel, 1986), podemos usar o mesmo tipo de coeficientes para dados politômicos a exemplo do que é feito para dados bivariados dicotômicos. Além disto, a própria medida de associação estimada a partir dos dados de uma tabela de contingência $R \times C$ pode ser interpretada como uma generalização da razão de produtos cruzados, podendo desta forma auxiliar na interpretação das relações

entre as categorias na análise de tabelas de contingência e ser de útil aplicação na área de Bioestatística.

Na prática atualmente, para medir correlação entre variáveis categóricas politômicas tem sido aplicado o Coeficiente de Correlação de Pearson diretamente às variáveis politômicas, como se elas fossem contínuas. Para tal é feita a suposição de uma escala com igual distanciamento entre os escores 1, 2, 3 etc. Isto é, para variáveis ou itens medidas como escalas de Likert de 5 pontos, supõe-se que a distância entre 1 e 2 é a mesma do que entre 2 e 3 ou entre 4 e 5 e assim por diante. Esta suposição não é sempre correta. Além disto, supondo que a escala de Likert é uma escala quantitativa e não categórica, ficaria implícito que o escore 4 é o dobro do escore 2 na escala, o que novamente nem sempre é verdadeiro ou, às vezes, não faz sentido, caracterizando claramente que as escalas que estamos trabalhando não têm a propriedade numérica de uma escala quantitativa ou escala de intervalo.

Outros autores têm proposto coeficientes de correlação "policóricos" para o caso de variáveis categóricas ordinais politômicas, generalizando o tradicional coeficiente de correlação tetracórico, o qual havia sido definido na primeira metade do século, para dados observados em tabelas 2x2, mas com a suposição de terem sido gerados por uma distribuição normal bivariada. Os mais importantes trabalhos sobre os coeficientes policóricos são os de Olsson (1979) e Quiroga (1992). Olsson propõe a estimativa de máxima verossimilhança do parâmetro ρ da distribuição Normal supondo que os dados são observados em uma tabela de contingência RxC mas com distribuição subjacente contínua normal bivariada. Os coeficientes de correlação tipo-contingência podem então ser utilizados como medidas de correlação para dados politômicos e podem também ser utilizados em métodos multivariados para dados categorizados, simplificando o processo computacional dos métodos existentes, como foi mostrado em Fachel (1986).

Nesta monografia o objetivo é comparar o desempenho do Coeficiente de Correlação Tipo-C proposto por Fachel (1986), para variáveis politômicas, em relação ao coeficiente de correlação de Pearson, calculado para variáveis categóricas, através de um estudo de simulação Monte Carlo. Utilizaremos no

delineamento da simulação diferentes formas de distribuições marginais, diferentes tamanhos de amostras, e diferentes números de categorias das variáveis marginais.

Esta monografia tem seis capítulos, neste primeiro capítulo fizemos uma introdução histórica sobre coeficientes de correlação para dados categóricos. Introduzimos a origem da idéia de uma Superfície de Associação Constante independente dos pontos de corte das marginais e conduzimos nosso histórico para a generalização da razão de produto cruzado (*odds ratio*) para dados observados não em tabelas 2x2, mas tabelas de contingência com qualquer número de linhas (R) e colunas (C).

No capítulo II, definiremos a distribuição de probabilidade bivariada contínua denominada distribuição Tipo-C que na verdade é uma família de distribuições bivariadas que obtêm formas diferenciadas dependendo da forma de suas marginais. Também no capítulo II apresentaremos resumidamente a derivação do método de máxima verossimilhança e o processo computacional para obtenção desta solução, o qual resulta na estimativa da razão de produtos cruzados generalizados. Através da obtenção deste “odds ratio” global ou “odds ratio” generalizado, transformamos esta medida de associação em medida de correlação da mesma forma como era feito para o caso de tabelas 2x2. Só que agora temos tabelas de contingência RxC. As fórmulas e diversos tipos de coeficientes de correlação Tipo-C são apresentados também neste capítulo com exemplos.

No capítulo III faremos uma retrospectiva dos procedimentos computacionais desenvolvidos especialmente para a obtenção da razão de produtos cruzados generalizados e descreveremos o *software* Crosspsi desenvolvido por D’Avila e Fachel (1998).

No capítulo IV apresentaremos o estudo de simulação, que é o objetivo desta monografia, o qual foi realizado ao longo dos dois últimos anos dentro do projeto de pesquisa financiado pela PROPESQ e CNPq/UFRGS com a participação de dois bolsistas de Iniciação Científica. Além do autor a participação da acadêmica Luciane H. Weber foi muito importante neste estudo de

simulação. Para a elaboração desta monografia foi intensificado o estudo de simulação com o objetivo de completar o delineamento total do estudo. Aproximadamente um terço do trabalho de simulação foi completado nestes últimos meses.

O estudo de simulação Monte Carlo foi delineado com os seguintes parâmetros: foram geradas distribuições de probabilidade marginais normais de tal forma que a distribuição bivariada conjunta tivesse parâmetro de correlação igual a 0,50 ($\rho=0,50$). Fizemos o estudo para dois diferentes tamanhos de amostras ($n=100$ e $n=500$); para cinco diferentes tipos de formas de categorização das distribuições marginais, para diferentes tamanhos de números de categorias (3 a 11 categorias) e para três diferentes números de repetições (200, 400 e 1000 repetições). A justificativa das escolhas destes parâmetros e formas distribucionais é descrita no capítulo IV.

No capítulo V são apresentados os resultados da simulação dos coeficientes de correlação Tipo-C e comparados com o coeficiente de correlação de Pearson para os dados categorizados e para os dados contínuos gerados no estudo de simulação. Os coeficientes de correlação calculados também são comparados com o parâmetro de correlação do estudo, definido como $\rho=0,50$. Os resultados são sintetizados para alguns casos, e as tabelas completas do estudo de simulação são apresentadas em anexo devido ao grande número de tabelas deste estudo.

Finalmente no capítulo VI apresentaremos a conclusão final e passos futuros no delineamento desta pesquisa.

Durante este trabalho devido ao grande porte do estudo de simulação, às dimensões e ao grande número de arquivos elaborados, foi necessário uma definição de rotinas e fórmulas no *software* Excel® para manusear os arquivos após o cálculos dos coeficientes de correlação no *software* Crosspsi. Estas rotinas, têm como resultado a geração de gráficos, cálculos de médias, desvio-padrões, erros-quadráticos e erros quadráticos médios das k repetições dos coeficientes de correlação estudados. No final do processo de simulação o tempo computacional gasto foi muito menor do que no início devido à introdução destas rotinas.

Para dar uma idéia da dimensão deste estudo de simulação, chegamos ao final com uma área de armazenamento de aproximadamente 180Mb no disco rígido ou seja o equivalente a 125 disquetes (capacidade 1,44Mb). Formas de organização destes arquivos para armazenamento também tiveram que ser estipuladas para futuras utilizações destes resultados.

Finalmente gostaríamos de acrescentar que, além do aprendizado na área teórica e empírica, principalmente no que diz respeito às distribuições amostrais dos estimadores, as quais através do estudo de formulação é possível visualizar, tivemos um amadurecimento na experiência de participar em um projeto de pesquisa, experiência esta que será de grande valia em nossa futura vida profissional.

2- RAZÃO DE PRODUTOS CRUZADOS GENERALIZADOS

Neste capítulo definiremos a distribuição Tipo-C com suas funções distribuição e densidade de probabilidade, a derivação do método de máxima verossimilhança para estimar a razão de produtos cruzados generalizados e o processo iterativo utilizado para obtenção dos estimadores.

2.1- DISTRIBUIÇÃO TIPO-C

Supomos duas variáveis aleatórias X e Y contínuas com função distribuição F(x) e G(y) respectivamente e com função distribuição conjunta H(x,y). As distribuições marginais podem ser dicotomizadas em pontos arbitrários x e y formando uma tabela de contingência 2x2, como segue:

$$\begin{array}{c}
 y \\
 \downarrow \\
 \begin{array}{cc|c}
 & H(x,y) & F(x)-H(x,y) & F(x) \\
 x \rightarrow & \hline
 & G(y)-H(x,y) & 1-F(x)-G(y)+H(x,y) & 1-F(x) \\
 & \hline
 & G(y) & 1-G(y) & 1
 \end{array}
 \end{array}$$

A razão de produtos cruzados (*odds ratio*) para esta tabela é dada por

$$\psi = \frac{H(x,y)[1-F(x)-G(y)+H(x,y)]}{[F(x)-H(x,y)][G(y)-H(x,y)]}$$

ou

$$(\psi-1) H^2(x,y) - \{1+[F(x)+G(y)] (\psi-1)\} H(x,y) + \psi F(x) G(y) = 0 \quad \psi > 0$$

Mardia (1967) mostrou que a única raiz possível da equação quadrática acima é dada por:

$$H(x, y) = \begin{cases} \frac{\left[S(x, y) - \left\{ S^2(x, y) - 4\psi(\psi-1)F(x)G(y) \right\}^{1/2} \right]}{\{2(\psi-1)\}}, & (\psi \neq 1) \\ F(x)G(y), & \psi = 1 \end{cases}$$

$$\text{onde, } S(x, y) = 1 + (F(x) + G(y))(\psi - 1) \quad \begin{array}{l} F(x) = P(X \leq x) \\ G(y) = P(Y \leq y) \end{array}$$

H é a função distribuição bivariada Tipo-C, para obtermos sua função densidade de probabilidade, derivamos e obtemos $H'(x, y)$ dada por:

$$H'(x, y) = \frac{1}{2.(\psi-1)^2} \left[\frac{S(x, y) - 2(\psi-1)F(x)G(y)}{\sqrt{A(x, y)}} - 1 \right] \quad \begin{array}{l} x = 1, \dots, r-1 \\ y = 1, \dots, c-1 \end{array}$$

onde:

$$S(x, y) = 1 + (\psi-1)(F(x) + G(y))$$

$$A(x, y) = S^2(x, y) - 4\psi(\psi-1)F(x)G(y)$$

Esta expressão define a função distribuição Tipo-C. Se X e Y são variáveis Normais, a distribuição dada por esta expressão é denominada distribuição Tipo-C Normal. Se as distribuições marginais são Uniformes, temos distribuição Tipo-C Uniforme. Se X e Y são Logísticas, temos distribuição Tipo-C Logística. O parâmetro de associação para qualquer membro desta família de distribuição bivariada contínua é ψ que é a razão de produtos cruzados generalizados. Para dados apresentados em tabelas 2x2, este parâmetro é estimado pela razão de produtos cruzados amostral dado por

$$\hat{\psi} = \frac{m_{11} m_{22}}{m_{12} m_{21}}$$

onde m_{ij} são as frequências observadas, $i, j = 1, 2$.

Como Pearson e Heron (1913) mostraram, sempre é possível construir esta superfície de associação constante para a qual o parâmetro de associação ψ independe dos pontos de dicotomia nas marginais.

Utilizando o *software* Maple V, obtivemos os gráficos das funções densidade de probabilidade bivariada contínua para três tipos de distribuição. A Figura 2.1 mostra a distribuição normal bivariada com parâmetros $\mu_x=0$; $\sigma_x=1$; $\mu_y=0$; $\sigma_y=1$ e $\rho_{xy}=0,50$, utilizando a amplitude $[-4, 4]$ para os eixos x e y , para uma melhor visualização da curva. A Figura 2.2 mostra a função densidade de probabilidade bivariada contínua Tipo-C Normal com parâmetro $\psi=4,41$, o qual corresponde a $\rho_{xy}=0,50$ para possível comparação visual. Esta correspondência foi obtida a partir da transformação proposta por Chambers (1982). A amplitude utilizada em todos os gráficos é a mesma. Finalmente na Figura 2.3 mostraremos a função densidade de probabilidade bivariada contínua Tipo-C Uniforme com marginais Uniforme no intervalo $[-4, 4]$ e parâmetro de associação $\psi=4,41$, também escolhido pelo mesmo motivo.

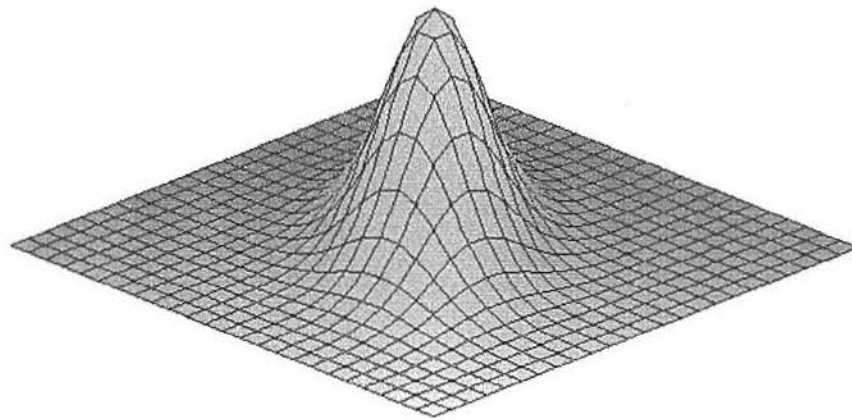


Figura 2.1 – Gráfico da função densidade de probabilidade normal bivariada com os parâmetros $N(\mu_x=0; \sigma_x=1; \mu_y=0; \sigma_y=1; \rho_{xy}=0,50)$.

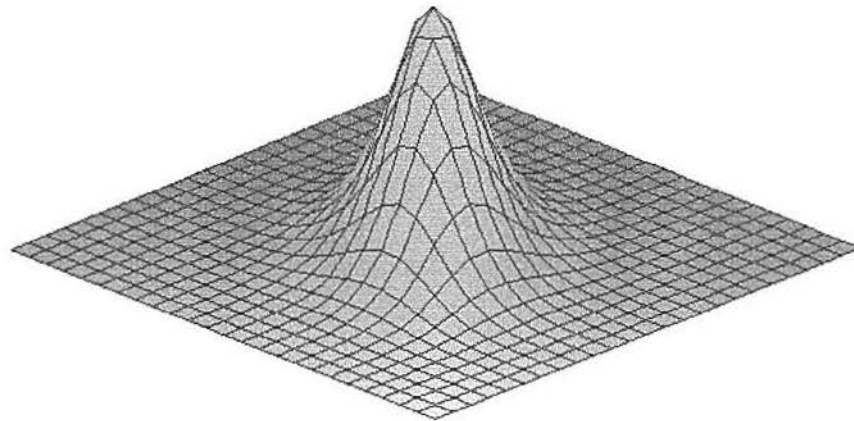


Figura 2.2 – Gráfico da função densidade de probabilidade bivariada Tipo-C Normal com os parâmetros: $N(\mu_x=0; \sigma_x=1)$, $N(\mu_y=0; \sigma_y=1)$ e $\psi=4,41$.

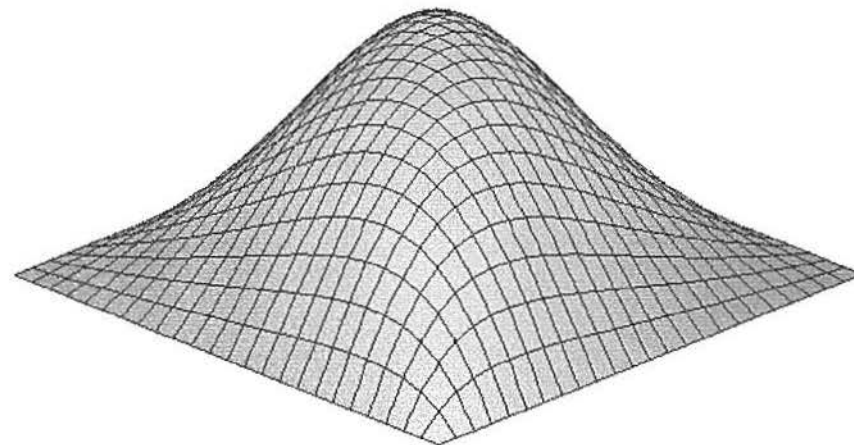


Figura 2.3 – Gráfico da função densidade de probabilidade bivariada Tipo-C Uniforme com os parâmetros: $U_x[-4,4]$, $U_y[-4,4]$ e $\psi=4,41$.

2.2- ESTIMAÇÃO DE MÁXIMA VEROSSIMILHANÇA PARA ψ

Considere o caso geral quando os dados são apresentados em uma tabela de contingência $R \times C$, com variáveis marginais com R e C categorias ordenadas respectivamente, isto é, quando as variáveis observadas são politômicas em vez de dicotômicas. Para estimar o parâmetro de associação ψ para tabelas $R \times C$, a idéia é obter uma “razão de produtos cruzados global” para dados politômicos.

O problema pode ser resumido da seguinte forma: nós observamos duas variáveis categóricas ordinais U e V , estas variáveis são classificadas em R e C categorias respectivamente. Nós supomos que subjacentes às variáveis U e V existam as variáveis latentes contínuas X e Y as quais têm distribuição bivariada contínua Tipo-C. A relação entre U e X , e entre V e Y , pode ser assim especificada:

$$\begin{array}{ll} U=1 \text{ se } X < x_1 & V=1 \text{ se } Y < y_1 \\ U=2 \text{ se } x_1 \leq X < x_2 & V=2 \text{ se } y_1 \leq Y < y_2 \\ U=3 \text{ se } x_2 \leq X < x_3 & V=3 \text{ se } y_2 \leq Y < y_3 \\ \dots\dots\dots & \dots\dots\dots \\ U=R \text{ se } x_{R-1} \leq X & V=R \text{ se } y_{R-1} \leq Y \end{array}$$

O problema de estimar ψ para este caso é resolvido utilizando-se o método de estimação de máxima verossimilhança.

Os dados são apresentados em uma tabela de contingência $R \times C$ com frequências observadas n_{ij} , $i=1, \dots, R$; $j=1, \dots, C$. As probabilidades p_{ij} são obtidas em função de H , a função distribuição Tipo-C. Então a função de verossimilhança é dada por

$$L = C \prod_{i=1}^R \prod_{j=1}^C p_{ij}^{n_{ij}}$$

onde C é uma constante que não depende do parâmetro ψ e $\sum_i \sum_j p_{ij} = 1$.

Usando logaritmos temos: $l = \log L = \log C + \sum_i \sum_j n_{ij} \cdot \log p_{ij}$

e então
$$\frac{\partial l}{\partial \psi} = \sum_{i=1}^R \sum_{j=1}^C \frac{n_{ij}}{p_{ij}} \frac{\partial p_{ij}}{\partial \psi}$$

Na tabela 2.1 apresentamos as fórmulas necessárias para estimar ψ a partir de dados apresentados em tabelas $R \times C$, sendo f_a e g_b as proporções marginais estimadas através das proporções marginais observadas correspondentes.

Tabela 2.1 – Fórmulas necessárias para estimar ψ para tabelas $R \times C$

n_{ij}	p_{ij}	$\frac{\partial p_{ij}}{\partial \psi}$
n_{11}	$p_{11} = H_{11}$	H'_{11}
n_{12}	$p_{12} = H_{12} - H_{11}$	$H'_{12} - H'_{11}$
n_{21}	$p_{21} = H_{21} - H_{11}$	$H'_{21} - H'_{11}$
H_{ab} $a = 2, \dots, r-1$ $b = 2, \dots, c-1$	p_{ab}	$\frac{\partial p_{ab}}{\partial \psi}$
n_{a1}	$p_{a1} = H_{a1} - H_{a-1,1}$	$H'_{a1} - H'_{a-1,1}$
n_{1b}	$p_{1b} = H_{1b} - H_{1,b-1}$	$H'_{1b} - H'_{1,b-1}$
\vdots	\vdots	\vdots
\vdots	\vdots	\vdots
n_{ab}	$p_{ab} = H_{ab} - H_{a-1,b} - H_{a,b-1} + H_{a-1,b-1}$	$H'_{ab} - H'_{a-1,b} - H'_{a,b-1} + H'_{a-1,b-1}$
\vdots	\vdots	\vdots
\vdots	\vdots	\vdots
n_{ac}	$p_{ac} = f_a - H_{a,c-1} + H_{a-1,c-1}$	$-H'_{a,c-1} + H'_{a-1,c-1}$
\vdots	\vdots	\vdots
\vdots	\vdots	\vdots
n_{rb}	$p_{rb} = g_b - H_{r-1,b} + H_{r-1,c-1}$	$-H'_{r-1,b} + H'_{r-1,c-1}$
n_{1c}	$p_{1c} = f_1 - H_{1,c-1}$	$-H'_{1,c-1}$
n_{r1}	$p_{r1} = g_1 - H_{r-1,1}$	$-H'_{r-1,1}$
n_{rc}	$p_{rc} = 1 - \sum_{i=1}^a f_i - \sum_{j=1}^b g_j + H_{r-1,c-1}$	$H'_{r-1,c-1}$

A solução da equação

$$\frac{\partial l}{\partial \psi} = \sum_{i=1}^R \sum_{j=1}^C \frac{n_{ij}}{p_{ij}} \frac{\partial p_{ij}}{\partial \psi}$$

é obtida pelo método iterativo de escores de Fischer (ver Kendall and Stuart, 1979). O processo iterativo é dado por:

$$\hat{\psi} = t - \left(\frac{\partial \ln L}{\partial \psi} \right)_t / \left[E \left(\frac{\partial^2 \ln L}{\partial \psi^2} \right) \right]_t = t + \left(\frac{\partial \ln L}{\partial \psi} \right)_t (\hat{\text{var}} \psi)_t$$

Para a obtenção da solução numérica da equação de verossimilhança (visto que ela não tem solução analítica), o processo inicia com algum valor inicial t e é repetido até a convergência da estimativa para um desejado grau de precisão. O valor inicial não é crítico pois o processo converge rapidamente para diferentes valores iniciais.

Devido a dificuldades numéricas que apareceram no processo iterativo para valores de ψ na vizinhança de 1, foi necessário buscar uma expressão alternativa para a função distribuição tipo-C, $H(x,y)$. A expressão alternativa foi obtida por Fachel (1986) pela expansão binomial da função distribuição Tipo-C. O resultado obtido foi:

$$H = FG + \frac{\alpha FG \overline{FG}}{1 + \alpha (F\overline{G} + \overline{F}G)} + \frac{\alpha^3 (FG \overline{FG})^2}{[1 + \alpha (F\overline{G} + \overline{F}G)]^3} + \frac{2\alpha^5 (FG + \overline{FG})^3}{[1 + \alpha (F\overline{G} + \overline{F}G)]^5} + \frac{5\alpha^7 (FG + \overline{FG})^4}{[1 + \alpha (F\overline{G} + \overline{F}G)]^7} +$$

$$+ \frac{14\alpha^9 (FG + \overline{FG})^5}{[1 + \alpha (F\overline{G} + \overline{F}G)]^9} + \frac{42\alpha^{11} (FG + \overline{FG})^6}{[1 + \alpha (F\overline{G} + \overline{F}G)]^{11}} + \frac{132\alpha^{13} (FG + \overline{FG})^7}{[1 + \alpha (F\overline{G} + \overline{F}G)]^{13}} + \dots \quad |\alpha| < \frac{1}{(F\overline{G} + \overline{F}G)}$$

Por causa das vantagens numéricas da expressão aproximada acima ela é usada no processo iterativo para valores do parâmetro de associação, no intervalo $[0.98, 1.02]$. Na Figura 2.4, mostramos um exemplo dado por Fachel (1986) onde a aproximação foi utilizada com sensível melhoria nos resultados do processo iterativo.

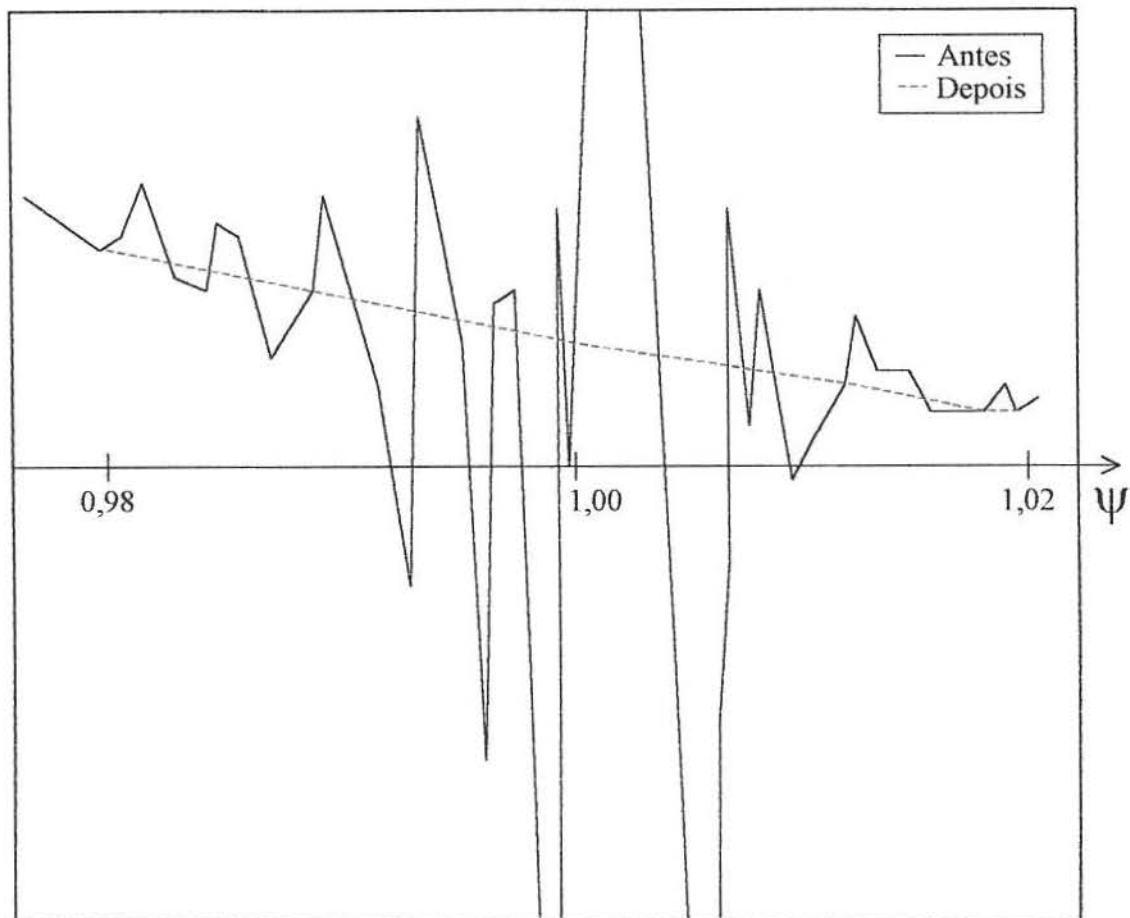


Figura 2.4 - Derivada da função de verossimilhança na vizinhança de $\psi=1$, antes e depois de usar a aproximação para $H(x,y)$.

2.3 - COEFICIENTES DE CORRELAÇÃO TIPO-CONTINGÊNCIA

Medidas de correlação baseadas no parâmetro de associação da distribuição bivariada contínua Tipo-C, a qual supomos gerar os dados observados em uma tabela de contingência $R \times C$ e que na prática é estimado pelo método de verossimilhança proposto por Fachel (1986), são buscadas a partir de transformações propostas na literatura tradicional para tabelas 2×2 .

O coeficiente de correlação Q de Yule e seu coeficiente r_y são exemplos deste tipo de transformações dado por:

$$Q = \frac{\Psi - 1}{\Psi + 1}$$

$$r_y = \frac{(\psi^{0,5} - 1)}{(\psi^{0,5} + 1)}$$

Mardia (1967) estudou o coeficiente de correlação Tipo-C para distribuição Uniforme e obteve:

$$r_u = \frac{\hat{\psi} + 1}{\hat{\psi} - 1} - \frac{2\hat{\psi} \ln \hat{\psi}}{(\hat{\psi} - 1)^2}$$

Chambers (1982) generalizou o coeficiente de correlação de Yule e mostrou que se elevarmos Ψ na potência 0,74 obteremos um bom estimador do coeficiente de correlação para variáveis com distribuição Normal Bivariada.

$$r_{0,74} = \frac{\Psi^{0,74} - 1}{\Psi^{0,74} + 1}$$

Plackett (1965) propôs um estimador do coeficiente de correlação para variáveis com distribuição subjacente Normal Bivariada através da seguinte fórmula:

$$\rho_p(\psi) = -\cos \left[\frac{\pi \sqrt{\hat{\psi}}}{(1 + \sqrt{\hat{\psi}})} \right]$$

Os coeficientes de correlação baseados na medida de associação da Distribuição Tipo-C são denominados, por analogia, Coeficientes de Correlação Tipo-Contingência, ou simplesmente Coeficiente de Correlação Tipo-C.

VARIÂNCIA ASSINTÓTICA DOS ESTIMADORES DOS COEFICIENTES DE CORRELAÇÃO

Uma das vantagens da estimação pelo método de máxima verossimilhança apresentado neste capítulo é que no processo obtemos a variância assintótica para o estimador de ψ , que é dado pela seguinte expressão:

$$Var(\hat{\psi}) = - \frac{1}{\left[E \left(\frac{\partial^2 \ln L}{\partial \hat{\psi}^2} \right) \right]_{\hat{\psi}}}$$

onde,

$$E \left(\frac{\partial^2 \ln L}{\partial \hat{\psi}^2} \right) = n \cdot \sum_{a=1}^{r-1} \sum_{b=1}^{c-1} \left[H_{ab} \left(- \frac{\partial p_{ab}}{\partial \psi} + \frac{1}{p_{a,b+1}} \cdot \frac{\partial p_{a,b+1}}{\partial \psi} + \frac{1}{p_{a+1,b}} \cdot \frac{\partial p_{a+1,b}}{\partial \psi} - \frac{1}{p_{a+1,b+1}} \cdot \frac{\partial p_{a+1,b+1}}{\partial \psi} \right) \right]$$

Para uma função geral $f(\psi)$, a variância assintótica para $f(\hat{\psi})$ é dado por:

$$Var[f(\hat{\psi})] \cong [f'(\hat{\psi})]^2 \cdot Var(\hat{\psi})$$

onde, $f'(\hat{\psi})$ é a derivada de f com relação a ψ . (Kendall and Stuart, 1979, vol.II, pg. 53).

Além da estimação por ponto de ψ é usual obtermos um intervalo de confiança para ψ . Para evitar a obtenção de intervalos de confiança para ψ com limites inferiores negativos, preferimos obter o intervalo de confiança para $\ln \psi$ e após converter este intervalo através de exponenciação dos limites para obtermos o intervalo desejado para ψ . Consequentemente um intervalo para $\ln \psi$ é facilmente obtido usando-se a variância assintótica dos estimadores dada acima, logo:

$$Var(\ln \hat{\psi}) = \frac{1}{\hat{\psi}^2} Var \hat{\psi}$$

então por exemplo um intervalo de 95% de confiança para $\ln \psi$ é dada por

$$\ln \hat{\psi} \pm 1,96 \frac{\sqrt{\widehat{Var} \hat{\psi}}}{\hat{\psi}}$$

Também podemos facilmente derivar a expressão para a variância assintótica dos estimadores dos coeficientes de correlação Tipo-C apresentados anteriormente neste capítulo. Assim, com o uso da notação r para os estimadores, nós temos:

Variância do estimador do coeficiente de correlação Tipo-C Mardia

$$Var[r_u(\hat{\psi})] = \left[\frac{2 \cdot (\hat{\psi}^2 - 1) \cdot \ln \hat{\psi} - 4(\hat{\psi} - 1)^2}{(\hat{\psi} - 1)^4} \right]^2 \cdot Var(\hat{\psi})$$

Variância do estimador do coeficiente de correlação Tipo-C Chamber 0,74

$$Var[r_{0,74}(\hat{\psi})] = \left[\frac{2 \cdot (0,74) \cdot \hat{\psi}^{0,74} - 1}{(\hat{\psi}^{0,74} + 1)^2} \right]^2 \cdot Var(\hat{\psi})$$

Variância do estimador do coeficiente de correlação Tipo-C Plackett

$$Var[r_p(\hat{\psi})] = \left[\frac{\Pi \cdot \text{sen} \left(\frac{\Pi \sqrt{\hat{\psi}}}{1 + \sqrt{\hat{\psi}}} \right)}{2 \cdot \sqrt{\hat{\psi}} \cdot (1 + \sqrt{\hat{\psi}})^2} \right]^2 \cdot Var(\hat{\psi})$$

Obtidas as variâncias assintóticas para os estimadores dos coeficientes de correlação Tipo-C, podemos facilmente obter os intervalos de confiança para os coeficientes de correlação Tipo-C utilizando a distribuição normal como distribuição assintótica.

Por exemplo se quiséssemos um intervalo de 95% de confiança para $r_{0,74}(\psi)$ teríamos:

$$r_{0,74}(\hat{\psi}) \pm 1,96\sqrt{Var[r_{0,74}(\hat{\psi})]}$$

EXEMPLO NUMÉRICO

Nesta seção nós apresentaremos alguns exemplos numéricos do Coeficiente de Correlação Tipo-C em tabelas de contingência, onde compararemos este coeficiente de correlação com o parâmetro de correlação e com o Coeficiente de Correlação de Pearson.

Pearson e Heron (1913) em seus estudos sobre a teoria de associação, usaram como exemplo os dados de altura dos pais e dos filhos, onde observaram 1000 casos em uma tabela de contingência na forma 7x7 (ver Pearson e Heron, 1913, tabela XV, p.220). Nós apresentaremos estes dados na tabela 2.2, parte (a).

Com dados artificiais, Pearson e Heron também apresentaram duas tabelas obtidas através de pontos de cortes nas distribuições marginais Normal com parâmetro de correlação $\rho=0,50$ e $\rho=0,30$. Nós apresentaremos estes dados artificiais na parte (a) das tabelas 2.3 ($\rho=0,50$) e 2.4 ($\rho=0,30$).

Desta forma temos três tabelas, a primeira com dados reais onde assumimos a existência de uma distribuição bivariada associada aos dados da tabela de contingência. Nas tabelas seguintes, foram utilizados dados artificiais para uma superfície Normal. Os resultados dos Coeficientes de Correlação Tipo-C podem ser comparados com o Coeficiente de Correlação de Pearson para os dados reais e também com o parâmetro de correlação nos exemplos com dados artificiais. As frequências esperadas, sob a hipótese de que os dados foram gerados por uma distribuição Tipo-C, estão apresentadas na parte (b) das tabelas. Nós também apresentaremos a estimativa de máxima verossimilhança para ψ com erro-padrão e a estatística do teste de ajustamento Qui-quadrado na parte (c). A

estimativa dos Coeficientes de Correlação Tipo-C com seus erros-padrão estão apresentados na parte (d).

O teste Qui-quadrado de ajustamento (*goodness of fit*) é dado através da seguinte fórmula:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

onde os graus de liberdade são dados por:

$$GL = (r-1)(c-1) - 1$$

Sendo que um grau de liberdade é perdido pois estamos estimando o parâmetro ψ .

Tabela 2.2 – (a) Dados reais: Classe de estatura do pai e do filho.

Classe de estatura do filho	Classe de estatura do pai							Total
	1	2	3	4	5	6	7	
1	4	22	7	-	1	-	-	34
2	23	154	84	26	8	6	-	301
3	8	87	75	66	22	24	2	284
4	1	29	36	37	14	14	6	137
5	-	18	27	26	11	18	5	105
6	-	9	26	19	7	29	8	98
7	-	3	9	6	6	10	7	41
Total	36	322	264	180	69	101	28	1000

Tabela 2.2 -(b) Frequência esperada para tabela (a) supondo a distribuição Tipo-C (dados obtidos através do *software* Crosspsi)

Classe de estatura do filho	Casse de estatura do pai							Total
	1	2	3	4	5	6	7	
1	4.7	19.7	5.7	2.2	0.7	0.8	0.2	34
2	20.3	156.0	73.7	29.6	8.5	10.4	2.5	301
3	6.7	89.4	94.5	53.3	16.1	19.4	4.6	284
4	1.9	26.3	40.2	35.0	12.9	16.7	4.0	137
5	1.1	15.2	24.9	27.7	12.6	18.6	4.9	105
6	0.9	11.3	18.5	23.5	12.9	23.5	7.4	98
7	0.4	4.1	6.5	8.7	5.3	11.6	4.4	41
Total	36	322	264	180	69	101	28	1000

Tabela 2.2 -(c) Estimativa de máxima verossimilhança (EMV) para a razão de produtos cruzados generalizada ψ , erro-padrão e teste Qui-quadrado, para os dados da tabela (a) (dados obtidos através do *software* Crosspsi)

EMV para ψ	Qui-quadrado	Significância
$4,833 \pm 0,463$	49,80 (GL= 35)	P= 0,05

Tabela 2.2 -(d) Estimativa dos Coeficientes de Correlação Tipo-C e de Pearson, para os dados da tabela (a) (dados obtidos através do *software* Crosspsi)

Coeficiente de correlação	Correlação \pm Erro-padrão
Tipo-C Mardia	0.485 ± 0.025
Tipo-C Placket	0.555 ± 0.027
Tipo-C Chambers	0.524 ± 0.026
Pearson	0.478 ± 0.026

Tabela 2.3– (a) Dados artificiais I: Distribuição normal bivariada para $\rho= 0.50$ ajustado para unidades nas caselas (de Pearson e Heron, 1913, p.220)

	1	2	3	4	5	6	7	Total
1	7	20	5	2	-	-	-	34
2	21	145	79	36	10	9	1	301
3	6	94	85	54	19	22	4	284
4	2	32	39	31	12	17	4	137
5	-	18	28	25	11	18	5	105
6	-	11	22	24	12	22	7	98
7	-	2	6	8	5	13	7	41
Total	36	322	264	180	69	101	28	1000

Tabela 2.3 -(b) Frequência esperada para tabela (a) supondo a distribuição Tipo-C (dados obtidos através do *software* Crosspsi)

	1	2	3	4	5	6	7	Total
1	4.5	19.4	5.9	2.4	0.7	0.9	0.2	34
2	20.0	153.3	74.4	30.7	8.9	11.0	2.7	301
3	7.0	89.9	92.7	53.2	16.4	20.0	4.8	284
4	2.0	27.1	39.9	34.3	12.8	16.8	4.1	137
5	1.2	15.9	25.2	27.2	12.3	18.3	4.9	105
6	0.9	12.0	19.1	23.5	12.6	22.8	7.1	98
7	0.4	4.4	6.8	8.7	5.3	11.2	4.2	41
Total	36	322	264	180	69	101	28	1000

Tabela 2.3 -(c) Estimativa de máxima verossimilhança (EMV) para a razão de produtos cruzados generalizada ψ para os dados da tabela (a) (dados obtidos através do *software* Crosspsi)

EMV para ψ
4,486

Tabela 2.3 -(d) Estimativa dos Coeficientes de Correlação Tipo-C e de Pearson, para os dados da tabela (a) (dados obtidos através do *software* Crosspsi)

Coeficiente de correlação	Correlação
Tipo-C Mardia	0.465
Tipo-C Placket	0.534
Tipo-C Chambers	0.504
Pearson	0.459
Parâmetro de correlação	0.50

Tabela 2.4 – (a) Dados artificiais I: Distribuição normal bivariada para $\rho = 0.30$ ajustado para unidades nas caselas (de Pearson e Heron, 1913, p.220)

	1	2	3	4	5	6	7	Total
1	4	17	8	3	1	1	-	34
2	17	123	80	45	15	18	3	301
3	9	93	78	53	19	26	7	284
4	3	38	37	27	11	16	4	137
5	2	25	28	22	9	15	4	105
6	1	20	24	21	10	16	6	98
7	-	6	9	9	4	9	4	41
Total	36	322	264	180	69	101	28	1000

Tabela 2.4 -(b) Frequência esperada para tabela (a) supondo a distribuição Tipo-C (dados obtidos através do *software* Crosspsi)

	1	2	3	4	5	6	7	Total
1	2.6	16.4	7.9	3.8	1.2	1.6	0.4	34
2	16.6	128.9	78.3	41.1	13.6	17.9	4.6	301
3	9.2	92.0	80.8	51.8	18.4	25.2	6.6	284
4	3.1	34.9	37.5	29.0	11.4	16.6	4.5	137
5	2.1	23.2	27.0	23.3	9.9	15.3	4.4	105
6	1.7	19.2	23.3	21.9	10.1	16.7	5.0	98
7	0.7	7.4	9.2	9.1	4.4	7.8	2.5	41
Total	36	322	264	180	69	101	28	1000

Tabela 2.4 -(c) Estimativa de máxima verossimilhança (EMV) para a razão de produtos cruzados generalizada ψ para os dados da tabela (a) (dados obtidos através do *software* Crosspsi)

EMV para ψ
2.353

Tabela 2.4 -(d) Estimativa dos Coeficientes de Correlação Tipo-C e de Pearson, para os dados da tabela (a) (dados obtidos através do *software* Crosspsi)

Coeficiente de correlação	Correlação
Tipo-C Mardia	0.278
Tipo-C Placket	0.325
Tipo-C Chambers	0.306
Pearson	0.269
Parâmetro de correlação	0.30

Os três exemplos numéricos desta seção mostram que:

1. O método de máxima verossimilhança para estimar o parâmetro ψ da distribuição Tipo-C nos fornece um novo método para calcular coeficientes de correlação para dados politômicos.
2. O Coeficiente de Correlação Tipo-C, obtido através da transformação proposta por Chambers e utilizando a razão de produtos cruzados generalizada é um bom estimador do coeficiente de correlação de uma distribuição normal bivariada para dados apresentados em tabelas de contingência $R \times C$.
3. A distribuição Tipo-C normal é similar à distribuição normal bivariada nos exemplos apresentados.
4. O Coeficiente de Correlação de Pearson, utilizado quando estamos trabalhando com variáveis categóricas, subestimou o parâmetro de correlação.

3- *SOFTWARES* PARA OBTER O COEFICIENTE DE CORRELAÇÃO TIPO-C

Inicialmente um programa de computador, em FORTRAN 77, para o cálculo do coeficiente de associação Tipo-C, foi escrito por Fachel(1986) e o programa foi compilado para o equipamento VAX da Universidade de Londres. Posteriormente para a utilização do programa em microcomputadores foi escrito o programa TAB (Fachel e Machado,1990) em linguagem Pascal. Este programa fornece o valor do Coeficiente de Associação ψ (razão de produtos cruzados generalizados) para tabelas de contingências RxC. A entrada dos dados é em forma de tabela de contingência e o cálculo é feito para cada tabela separadamente. Isto é, se tivéssemos por exemplo 100 tabelas de contingência, as 100 tabelas eram digitadas uma de cada vez e o programa deveria ser executado 100 vezes.

Devido a esta limitação, um novo programa então foi escrito por D'Ávila e Fachel (1998) para ambiente Windows e com possibilidade de importar bancos de dados de outros *softwares* como o SPSS. O programa foi denominado Crosspsi e foi escrito na linguagem Delphi e pode tanto ler os dados através de arquivos na forma de banco de dados como também fornecer a solução para dados em tabelas de contingência, como era feito no Tab. Mais detalhes sobre as funções e forma de utilização do *software* Crosspsi estão descritas a seguir no manual deste *software*, o qual foi escrito a partir da nossa experiência no uso do Programa Crosspsi neste estudo de simulação.

3.1- MANUAL DO SOFTWARE CROSSPSI

Ao executar o *software* CROSSPSI, entrará a seguinte janela.

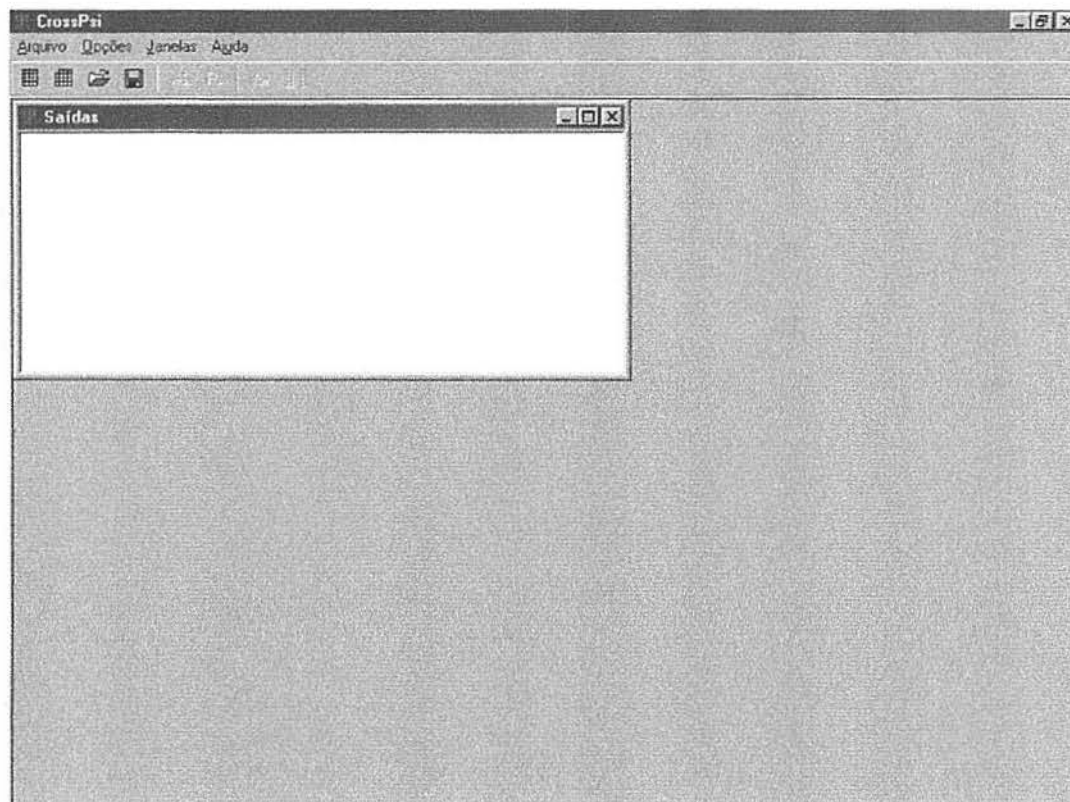


Figura 3.1 – tela de entrada do software CROSSPSI



FORMAS DE ENTRADA DE DADOS

A entrada de dados pode ser feita das seguintes formas:

1. tabela de contingência
2. banco de dados
3. importação de banco de dados externos

1- *Entrada de dados em forma de tabela de contingência*: este procedimento de entrada de dados é utilizado quando queremos calcular os Coeficientes de Correlação Tipo-C a partir de uma tabela de contingência, ou seja, as variáveis já

estão cruzadas; neste procedimento, devemos digitar cada tabela de contingência separadamente.

O ícone  indica “Nova tabela de contingência”, ao clicar nele, entrará uma tabela com as caselas vazias na forma 3x3. Para alterar o tamanho da tabela, devemos seleccionar o ícone  como mostra a figura 3.2.

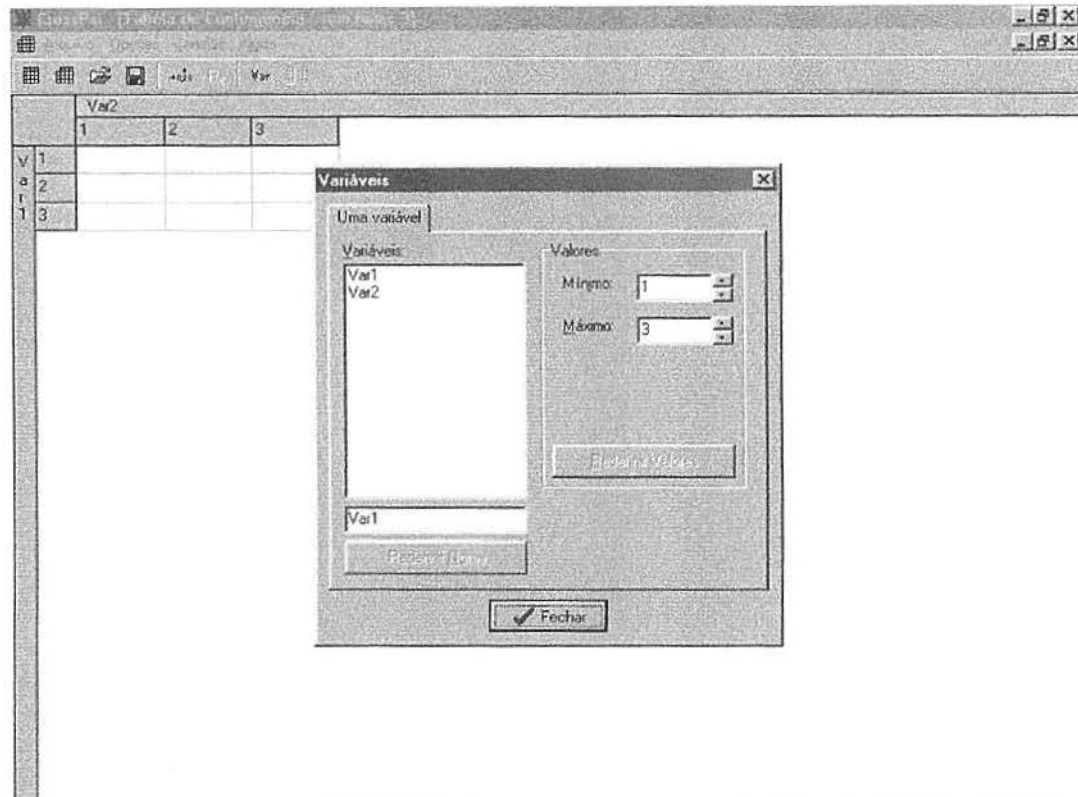




Figura 3.2 – Janela de alteração do tamanho de tabela de contingência

Esta nova janela possibilita não só alterarmos o número de categorias de cada uma das variáveis, mas também redefinirmos o nome das variáveis que estamos cruzando. Após alterarmos a amplitude de cada variável devemos clicar no botão “Redefinir Valores” e se for inserido o nome de cada variável também devemos clicar no botão “Redefinir Nomes”. Após estes procedimentos terem sido feitos para as duas variáveis, devemos clicar no botão “Fechar”. As frequências observadas devem ser digitadas nas caselas, e para passar de uma casela para a outra, deve-se usar as setas do teclado.


2- *Entrada de banco de dados*: um banco de dados, pode ser digitado diretamente no CROSSPSI. Esta opção está disponível no ícone  com o título “Nova tabela de dados”, porém a digitação nesta forma não é recomendada pois quando estamos neste módulo encontramos espaço definido por *default* para 3 variáveis com 4 casos e para aumentar o número de variáveis ou o número de casos, deve-se fazê-lo um a um, ou seja, para acrescentar uma variável devemos posicionar o *mouse* em cima do nome de alguma variável existente e depois clicar com o botão direito do *mouse*, assim aparecerá a opção “inserir variável”; para inserir um caso o procedimento é o mesmo porém o *mouse* deve ser posicionado em cima do “caso 1” a “caso 4”.

3- *Importação de banco de dados externos*: é a forma mais operacional para entrada de dados em forma de banco de dados (planilha), montado nos *softwares* como o SPSS e o EXCEL. Esses arquivos devem ser gravados no formato de “valores separados por tabulação” (.dat). Trazer um banco de dados do *software* SPSS é mais prático, pois os nomes das variáveis que estamos trabalhando são mantidos, no caso do EXCEL, para manter os nomes das variáveis devemos escrever o nome de cada variáveis na primeira linha da planilha. A princípio não existe limitação quanto ao tamanho do banco de dados, em nosso processo de simulação chegamos a trabalhar com 2000 variáveis e 500 casos.

Importando um banco de dados externo, existe a possibilidade de não só calcularmos os Coeficientes de Correlação Tipo-C mas também o Coeficiente de Correlação de Pearson para as variáveis categóricas e contínuas.

Para importar um banco de dados externos, ao rodar o *software* CROSSPSI, deve-se clicar no ícone  onde abrirá uma janela. Nessa janela, temos que alterar o tipo de arquivo que queremos importar, no nosso caso devemos escolher o tipo “valores separados por tabulações” e depois localizar o arquivo desejado e por fim clicar em “Abrir”.


DEFINIÇÃO DA AMPLITUDE DAS VARIÁVEIS

Quando estamos trabalhando na forma de banco de dados, seja o banco de dados montado diretamente no CROSSPSI ou importado de outro *software*, devemos definir a “propriedade das variáveis” no ícone , onde abrirá uma janela na qual podemos escolher se queremos definir cada variável ou todas de uma só vez na opção “intervalo de variáveis”. Para informar os valores possíveis que as variáveis assumem, na opção “uma variável”, devemos selecionar cada variável e informar o valor mínimo e máximo, depois clicar em “redefinir valores”. Na forma “intervalo de variáveis”, devemos indicar a variável inicial, a variável final e o valor mínimo e máximo que estas variáveis assumem. Com este procedimento estaremos definindo o intervalo de variação de todas as variáveis consecutivas neste intervalo. Para ambas situações pode-se usar a opção “coincidir todos os valores existentes como válidos”, porém, para esse procedimento não podem existir valores não informados (*missing data*).

CÁLCULO DOS COEFICIENTES DE CORRELAÇÃO

Este *software*, foi desenvolvido especificamente para calcular os coeficientes de correlação Tipo-C e para otimizar a tarefa de simulação. Foi também implementado no *software*, o cálculo do coeficiente de correlação de Pearson para variáveis categorizadas ou contínuas.

a) Cálculo dos coeficientes de correlação Tipo-C

Após terem sido feitas a entrada de dados e a definição da amplitude de variação das variáveis estudadas, devemos clicar no ícone  onde abrirá uma janela como mostra a figura 3.3, onde deveremos escolher a opção de cruzamento:

- 1- todos possíveis
- 2- somente um
- 3- intervalo

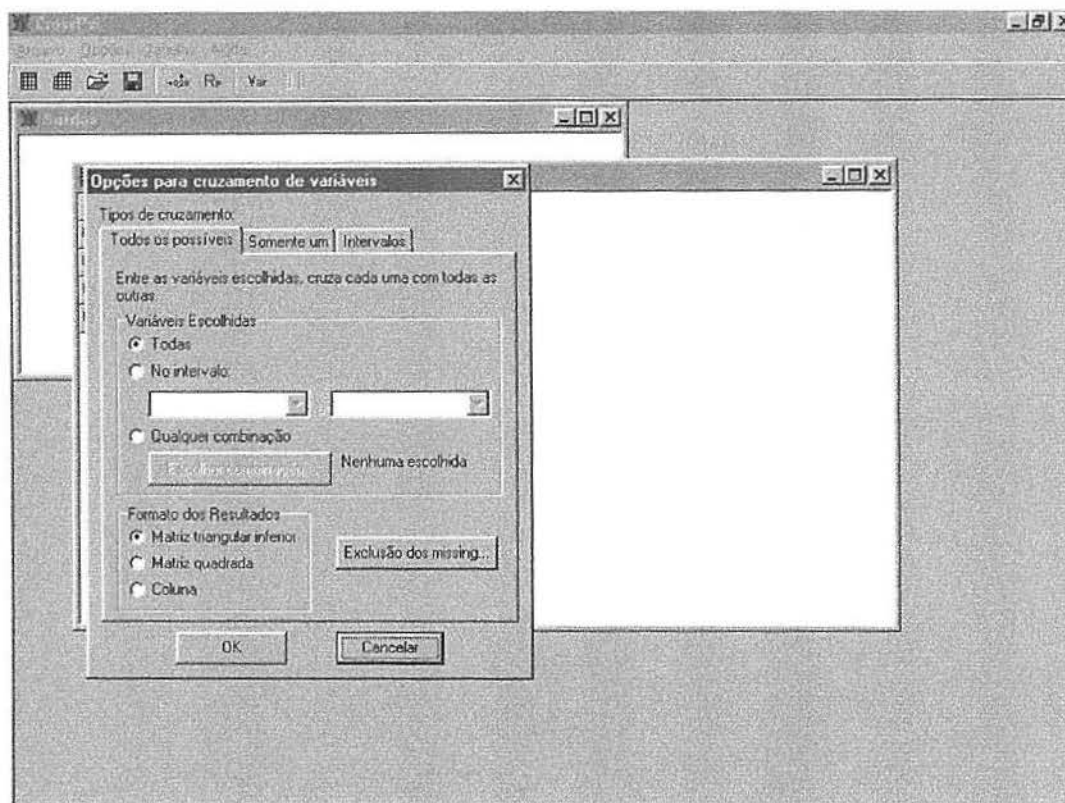



Figura 3.3 – Janela de opções de cruzamento

1- *todos possíveis* – nesta opção, obteremos os coeficientes de correlação Tipo-C e de Pearson para dados categorizados, onde devemos escolher se queremos calcular as correlações de todas as variáveis existentes ou todas dentro de um intervalo de variáveis. Para esta segunda opção devemos indicar o valor inicial e final. Também nesta opção deveremos selecionar a forma de saída destes resultados, seja ela na forma de matrizes triangulares inferiores, quadradas ou na forma de coluna. Na forma de coluna, obteremos uma única coluna para cada tipo de coeficiente de correlação com o número de linhas igual ao número de combinações possíveis das variáveis utilizadas.

2- *Somente um* – escolheremos esta opção quando o interesse é obter os coeficientes de correlação entre apenas um cruzamento (2 variáveis). Nesta opção deveremos indicar apenas as duas variáveis de interesse.

3- *Intervalo* – esta opção foi criada quase que exclusivamente para otimizar a tarefa de simulação. Nela devemos escolher a variável inicial dos dois grupos nos quais queremos obter os coeficientes de correlação e indicar o número de variáveis nestes intervalos. Para ilustrar esta idéia, supondo uma simulação com 10 repetições, ou seja, se tivermos 10 variáveis aleatórias X_i e Y_i e se o objetivo fosse obter os coeficientes de correlação entre as variáveis X_1 e Y_1 , X_2 e Y_2 e assim sucessivamente até X_{10} e Y_{10} , devemos utilizar esta opção onde indicaremos as variáveis iniciais de cada grupo no caso X_1 e Y_1 e o número de variáveis nos intervalos (comparações) que no caso seria 10. Neste módulo, somente teremos os resultados dos coeficientes de correlação na forma de colunas.

b) Cálculo do coeficiente de correlação de Pearson para dados contínuos

Para calcularmos o coeficiente de correlação de Pearson para dados contínuos, devemos importar o arquivo externo, da mesma forma que quando estamos trabalhando com variáveis categorizadas, depois deveremos clicar no ícone , onde aparecerá uma janela com as opções de cruzamento similar ao procedimento utilizado para cruzarmos variáveis categóricas, como ilustrado na figura 3.3.

SAÍDA DO SOFTWARE CROSSPSI

Os resultados dos coeficientes de correlação apresentados pelo *software* CROSSPSI, são os citados a seguir:

- Tipo-C Mardia
- Tipo-C Placket
- Tipo-C Chambers

Também temos o resultado do coeficiente de correlação de Pearson calculado para os dados categorizados, quando a entrada de dados não é feita na forma de tabela de contingência.

A figura 3.4 ilustra os resultados quando a entrada de dados é feita em forma de tabela de contingência, a partir do exemplo de Pearson e Heron (1913) sobre as alturas do pai e do filho, anteriormente citado.

Results de Tabela de Contingência - C:\Pesquisa\PERSON_HERON.ctg
 Psi: 4,832934
 Desvio padrão: 0,463368
 IC 95%: (4,004964;5,832074)

Coefficientes de Correlação Tipo-C:
 Tipo-C_M: 0,465259 IC95%: (0,436111; 0,534406)
 Tipo-C_P: 0,555160 IC95%: (0,502398; 0,607922)
 Tipo-C: 0,524783 IC95%: (0,474401; 0,575165)

Controles do Processo Iterativo:
 Derivada do logaritmo: -0,000010
 Número de iterações: 5

Valores observados:

4,000000	22,000000	7,000000	0,000000	1,000000	0,000000	0,000000
23,000000	154,000000	84,000000	26,000000	8,000000	6,000000	0,000000
8,000000	87,000000	75,000000	66,000000	22,000000	24,000000	2,000000
1,000000	29,000000	36,000000	37,000000	14,000000	14,000000	6,000000
0,000000	18,000000	27,000000	26,000000	11,000000	18,000000	5,000000
0,000000	9,000000	26,000000	19,000000	7,000000	29,000000	8,000000
0,000000	3,000000	9,000000	6,000000	6,000000	10,000000	7,000000

Valores esperados sob hipótese de distribuição Tipo-C bivariada:

4,731771	19,688620	5,647937	2,227469	0,662935	0,834139	0,207133
20,326560	156,048900	73,741990	29,542310	8,476290	10,354140	2,510137
6,722266	89,437040	94,504510	53,308100	16,105550	19,360290	4,562245
1,888493	26,215650	40,170810	35,047340	12,341930	16,703210	4,032574
1,129681	15,177910	24,915960	27,663890	12,584260	18,633940	4,894171
0,875750	11,337120	18,510920	23,530680	12,914070	23,485460	7,346009
0,326178	4,095190	6,507874	8,680222	5,314969	11,628830	4,447732

Figura 3.4 – Saída de dados do software CROSSPSI para tabelas de contingência

Observando a tabela 3.4 , podemos verificar os resultados do *software* CROSSPSI que apresentam a estimativa do coeficiente de associação ψ generalizado com intervalo de 95% de confiança, vários coeficientes de correlação Tipo-C, também com intervalos de confiança, controle do processo iterativo e valores observados e esperados segundo a Distribuição Tipo-C.

Quando estamos trabalhando com os dados na forma de variáveis, temos as opções de resultados na forma de matriz triangular inferior, matriz quadrada e coluna. Quando queremos resultados em forma de matrizes, abre-se uma janela para cada tipo de coeficientes de correlação Tipo-C e para o coeficiente de correlação de Pearson também calculado com os dados categóricos. Na forma de coluna nós obteremos uma única coluna para cada tipo de coeficiente de correlação com o número de linhas igual ao número de combinações possíveis das variáveis utilizadas. Porém, estas formas de resultados não são tão completas em relação ao resultado obtido a partir de uma tabela de contingência como mostrado na figura 3.4, pois não apresenta os intervalos de confiança para os parâmetros.

4- ESTUDO DE SIMULAÇÃO

Supondo a existência de uma variável subjacente contínua associada a variáveis politômicas ordinais, realizamos o processo de simulação no *Software* SPSS (V 8.0), onde foram geradas variáveis aleatórias independentes X_i e ε_i com distribuição $X_i \sim N(0;1)$ e $\varepsilon_i \sim N(0;1)$. Utilizando estas variáveis aleatórias geradas e as relações de regressão linear a seguir definidas, geramos variáveis Y_i de tal forma que Y_i e X_i não fossem variáveis independentes, isto é, fossem variáveis correlacionadas. Estipulamos que a correlação entre Y_i e X_i fosse $\rho=0,50$. Desta forma

$$Y_i = \beta X_i + \varepsilon_i$$

$$E(Y_i) = \beta E(X_i) + E(\varepsilon_i)$$

$$\text{isto é, } E(Y_i) = \beta \cdot 0 + 0 = 0$$

e

$$Var(Y_i) = \sigma_y^2 = \beta^2 Var(X_i) + Var(\varepsilon_i) + 2Cov(X_i, \varepsilon_i)$$

$$\text{como } X_i \text{ e } \varepsilon_i \text{ são independentes} \Rightarrow Cov(X_i, \varepsilon_i) = 0$$

logo

$$Var(Y_i) = \sigma_y^2 = \beta^2 \cdot 1 + 1 + 0 = \beta^2 + 1$$

Sabemos através dos resultados da teoria de regressão linear simples que:

$$\beta = \frac{\sigma_y}{\sigma_x} \rho_{xy}$$

$$\text{fixando, } \rho_{xy} = 0,5 \quad \beta = \frac{\sqrt{\beta^2 + 1}}{1} \cdot 0,5$$

$$\text{ou seja } \beta = 0,57735$$

obtivemos então variáveis aleatórias com distribuição Normal de média zero e desvio-padrão igual a 1,1547 ($Y_i \sim N(0; 1,1547)$). As variáveis aleatórias X_i e Y_i têm parâmetro de correlação igual a 0,5 ($\rho_{xy} = 0,5$). No estudo de simulação, as

variáveis contínuas geradas foram categorizadas em números de categorias variando de 3 a 11, isto é, as variáveis categóricas teriam 3, 4, 5, 6, 7, 8, 9, 10, 11 categorias obtidas através de pontos de corte nas distribuições marginais. Para tanto, foi utilizado o comando *Recode* do SPSS, dando forma às distribuições marginais. Foram geradas 200, 400 e depois 1000 tabelas de contingência cruzando X_i com Y_i a partir das variáveis simuladas. Isto é foram feitos tres estudos de simulação, um com 200 repetições, outro com 400 e finalmente para que o poder do estudo de simulação fosse maior para testar a robustez do estimador de máxima verossimilhança de Ψ , realizamos o estudo com 1000 repetições.

Inicialmente os pontos de corte eram definidos da seguinte forma: suponha que estejamos trabalhando com as variáveis contínuas simuladas com 200 repetições de X_i e 200 repetições de Y_i , todas com tamanho de amostra igual a 500 e queremos gerar variáveis categóricas de 4 categorias com distribuição marginal uniforme, ou seja, pegariamos cada variável e categorizaríamos com 25% de frequência para cada categoria de 1 a 4. O problema é que os pontos de corte de cada variável não são exatamente os mesmos, para isso deveríamos calcular os quartis de cada uma das 200 variáveis X_i e Y_i , mas isso seria inviável. No início, resolvemos então sortearmos 30 variáveis X_i e 30 variáveis Y_i , e destas foram obtidos os quartis, variável por variável e depois foi calculado a média do 1º quartil dos 30 X_i e o mesmo para Y_i , e assim sucessivamente para os 2º e 3º quartis, e estas médias foram utilizadas como pontos de corte para todas variáveis, como ilustra a figura 4.1 e a tabela 4.1.

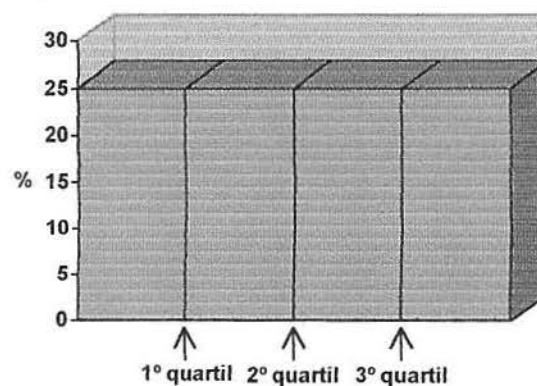


Figura 4.1 – Gráfico ilustrativo dos pontos de corte por quartis.

Tabela 4.1 – Tabela da média das 30 repetições dos quartis das distribuições marginais X e Y

	Média das 30 repetições		
	1º quartil	2º quartil	3º Quartil
X	-0,67441	0,00191	0,67439
Y	-0,77801	0,00298	0,77835

No entanto entendemos que este processo além de ser muito trabalhoso, incluía no estudo um erro amostral devido ao pequeno número de tamanho de amostra utilizado para a obtenção das separatrizes. Foi decidido então utilizar as separatrizes teóricas.

Por estarmos trabalhando com variáveis cujas distribuições subjacentes marginais são conhecidas ($X_i \sim N(0;1)$ e $Y_i \sim N(0;1,1547)$), utilizamos como pontos de corte as separatrizes teóricas (tabeladas), como mostra a tabela 4.2..

Tabela 4.2 – Quartis teóricos das distribuições marginais de X e Y.

	Pontos de cortes teóricos		
	1º quartil	2º quartil	3º Quartil
X	-0,67449	0	0,67449
Y	-0,77883	0	0,77883

Comparando as tabelas 4.1 e 4.2, concluímos que o primeiro procedimento estava bem próximo dos resultados teóricos, mas tínhamos um erro amostral pois para calcular estas médias só foram utilizadas 30 das 200 repetições, outra razão para utilizarmos os pontos de corte teóricos foi que, segundo o nosso delineamento de simulação, tínhamos que calcular uma infinidade de médias de pontos de corte.

A categorização das variáveis normais através de separatrizes nas distribuições subjacentes contínuas gerou variáveis com número de categorias variando de 3 a 11 categorias. Também variamos as formas das distribuições

marginais das variáveis categorizadas definindo os seguintes cinco tipos: formas simétricas suave e concentrada em torno da média, uniforme e assimétricas suave e concentrada à esquerda da distribuição.

Para ilustrar a maneira como isto foi feito, a figura 4.2 apresenta o histograma de uma distribuição marginal contínua que posteriormente foi categorizada de 3 a 11 categorias com formas simétricas suave e concentrada em torno da média, uniforme e assimétricas suave e concentrada à esquerda.

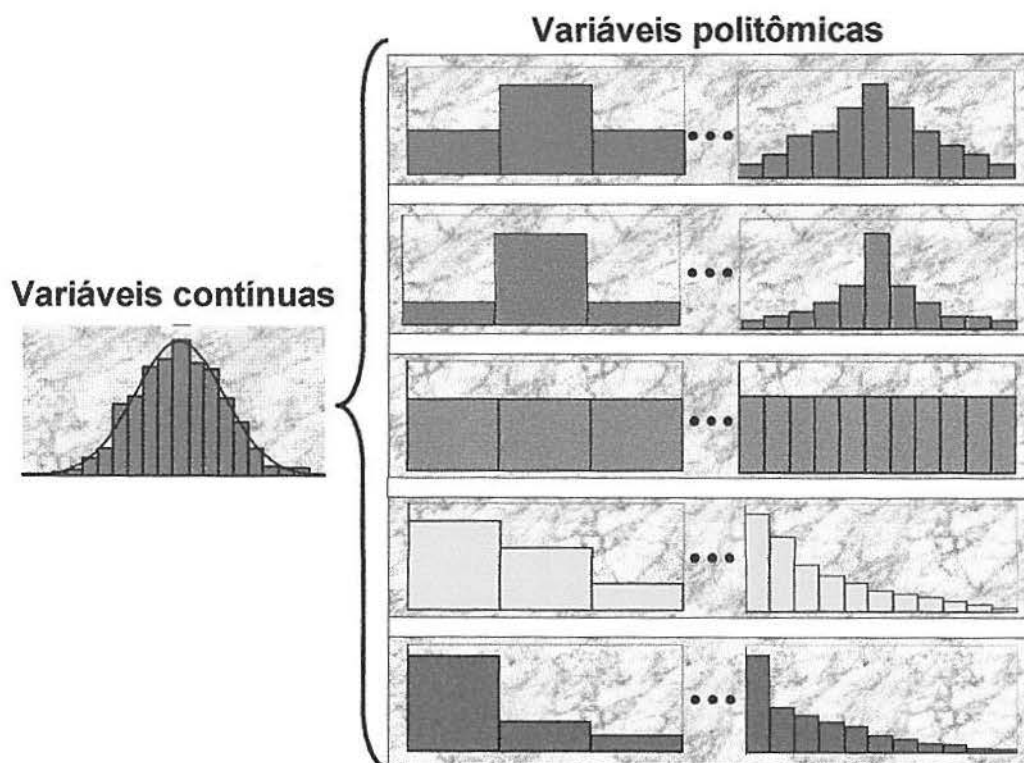


Figura 4.2 – Histograma da distribuição marginal contínua que foi categorizada em 3 e 11 categorias simétricas suave e concentrada em torno da média, uniforme e assimétricas suave e concentrada à esquerda

Nas tabelas 4.3 a 4.11 mostraremos as probabilidades marginais utilizadas no processo de simulação para transformar as variáveis contínuas com distribuições normais em variáveis categóricas com número de categorias variando de 3 a 11.

Tabela 4.3 – Probabilidade em cada uma das 3 categorias das distribuições marginais

Forma das dist. marginais	Probabilidade marginal em cada categoria		
	1	2	3
Simétrica suave	0,25	0,50	0,25
Simétrica concentrada	0,15	0,70	0,15
Uniforme	0,33	0,33	0,33
Assimétrica suave	0,50	0,35	0,15
Assimétrica concentrada	0,70	0,20	0,10

Tabela 4.4 – Probabilidade em cada uma das 4 categorias das distribuições marginais

Forma das dist. marginais	Probabilidade marginal em cada categoria			
	1	2	3	4
Simétrica suave	0,20	0,30	0,30	0,20
Simétrica concentrada	0,15	0,35	0,35	0,15
Uniforme	0,25	0,25	0,25	0,25
Assimétrica suave	0,45	0,30	0,20	0,05
Assimétrica concentrada	0,70	0,15	0,10	0,05

Tabela 4.5 – Probabilidade em cada uma das 5 categorias das distribuições marginais

Forma das dist. marginais	Probabilidade marginal em cada categoria				
	1	2	3	4	5
Simétrica suave	0,10	0,25	0,30	0,25	0,10
Simétrica concentrada	0,05	0,10	0,70	0,10	0,05
Uniforme	0,20	0,20	0,20	0,20	0,20
Assimétrica suave	0,50	0,30	0,10	0,06	0,04
Assimétrica concentrada	0,70	0,15	0,08	0,04	0,03

Tabela 4.6 – Probabilidade em cada uma das 6 categorias das distribuições marginais

Forma das dist. marginais	Probabilidade marginal em cada categoria					
	1	2	3	4	5	6
Simétrica suave	0,10	0,15	0,25	0,25	0,15	0,10
Simétrica concentrada	0,05	0,10	0,35	0,35	0,10	0,05
Uniforme	0,17	0,17	0,17	0,17	0,17	0,17
Assimétrica suave	0,45	0,25	0,15	0,07	0,05	0,03
Assimétrica concentrada	0,70	0,12	0,09	0,05	0,03	0,01

Tabela 4.7 – Probabilidade em cada uma das 7 categorias das distribuições marginais

Forma das dist. marginais	Probabilidade marginal em cada categoria						
	1	2	3	4	5	6	7
Simétrica suave	0,05	0,10	0,20	0,30	0,20	0,10	0,05
Simétrica concentrada	0,03	0,05	0,07	0,70	0,07	0,05	0,03
Uniforme	0,14	0,14	0,14	0,14	0,14	0,14	0,14
Assimétrica suave	0,45	0,25	0,10	0,08	0,07	0,04	0,01
Assimétrica concentrada	0,70	0,10	0,07	0,05	0,04	0,03	0,01

Tabela 4.8 – Probabilidade em cada uma das 8 categorias das distribuições marginais

Forma das dist. marginais	Probabilidade marginal em cada categoria							
	1	2	3	4	5	6	7	8
Simétrica suave	0,05	0,10	0,15	0,20	0,20	0,15	0,10	0,05
Simétrica concentrada	0,03	0,05	0,07	0,35	0,35	0,07	0,05	0,03
Uniforme	0,13	0,13	0,13	0,13	0,13	0,13	0,13	0,13
Assimétrica suave	0,45	0,23	0,10	0,08	0,07	0,04	0,02	0,01
Assimétrica concentrada	0,70	0,08	0,07	0,05	0,04	0,03	0,02	0,01

Tabela 4.9 – Probabilidade em cada uma das 9 categorias das distribuições marginais

Forma das dist. marginais	Probabilidade marginal em cada categoria								
	1	2	3	4	5	6	7	8	9
Simétrica suave	0,03	0,06	0,12	0,17	0,24	0,17	0,12	0,06	0,03
Simétrica concentrada	0,01	0,02	0,05	0,07	0,70	0,07	0,05	0,02	0,01
Uniforme	0,11	0,11	0,11	0,11	0,11	0,11	0,11	0,11	0,11
Assimétrica suave	0,45	0,20	0,10	0,08	0,07	0,04	0,03	0,02	0,01
Assimétrica concentrada	0,64	0,08	0,07	0,06	0,05	0,04	0,03	0,02	0,01

Tabela 4.10 – Probabilidade em cada uma das 10 categorias das distribuições marginais

Forma das dist. marginais	Probabilidade marginal em cada categoria									
	1	2	3	4	5	6	7	8	9	10
Simétrica suave	0,02	0,03	0,05	0,15	0,25	0,25	0,15	0,05	0,03	0,02
Simétrica concentrada	0,01	0,02	0,05	0,07	0,35	0,35	0,07	0,05	0,02	0,01
Uniforme	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10
Assimétrica suave	0,40	0,15	0,11	0,09	0,08	0,07	0,04	0,03	0,02	0,01
Assimétrica concentrada	0,55	0,09	0,08	0,07	0,06	0,05	0,04	0,03	0,02	0,01

Tabela 4.11 – Probabilidade em cada uma das 11 categorias das distribuições marginais

Forma das dist. marginais	Probabilidade marginal em cada categoria										
	1	2	3	4	5	6	7	8	9	10	11
Simétrica suave	0,03	0,05	0,07	0,10	0,15	0,20	0,15	0,10	0,07	0,05	0,03
Simétrica concentrada	0,01	0,02	0,03	0,04	0,05	0,70	0,05	0,04	0,03	0,02	0,01
Uniforme	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0,09
Assimétrica suave	0,40	0,12	0,10	0,09	0,08	0,06	0,05	0,04	0,03	0,02	0,01
Assimétrica concentrada	0,45	0,10	0,09	0,08	0,07	0,06	0,05	0,04	0,03	0,02	0,01

A Figura 4.3 ilustra o delineamento da simulação Monte Carlo que foi utilizado para parâmetro de correlação $\rho=0,50$, para cinco diferentes formas das distribuições marginais, para 9 diferentes tamanhos de números de categorias (3 a 11 categorias), para três diferentes números de repetições (200, 400 e 1000 repetições) e dois diferentes tamanhos de amostras ($n=100$ e $n=500$). Este delineamento foi utilizado para calcularmos os coeficientes de correlação Tipo-C obtidos através das transformações propostas por Chambers ($r_{0,74}$) e por Mardia (r_u), e para o coeficiente de correlação de Pearson (utilizando as variáveis categorizadas). Com este delineamento totalizamos 144.000 coeficientes de correlação para os coeficientes de correlação Tipo-C Mardia, Tipo-C Chambers e Pearson, ou seja 18.000 coeficientes de correlação para o caso de 200 repetições, 36.000 para o caso de 400 repetições e finalmente 90.000 coeficientes para o caso de 1.000 repetições.

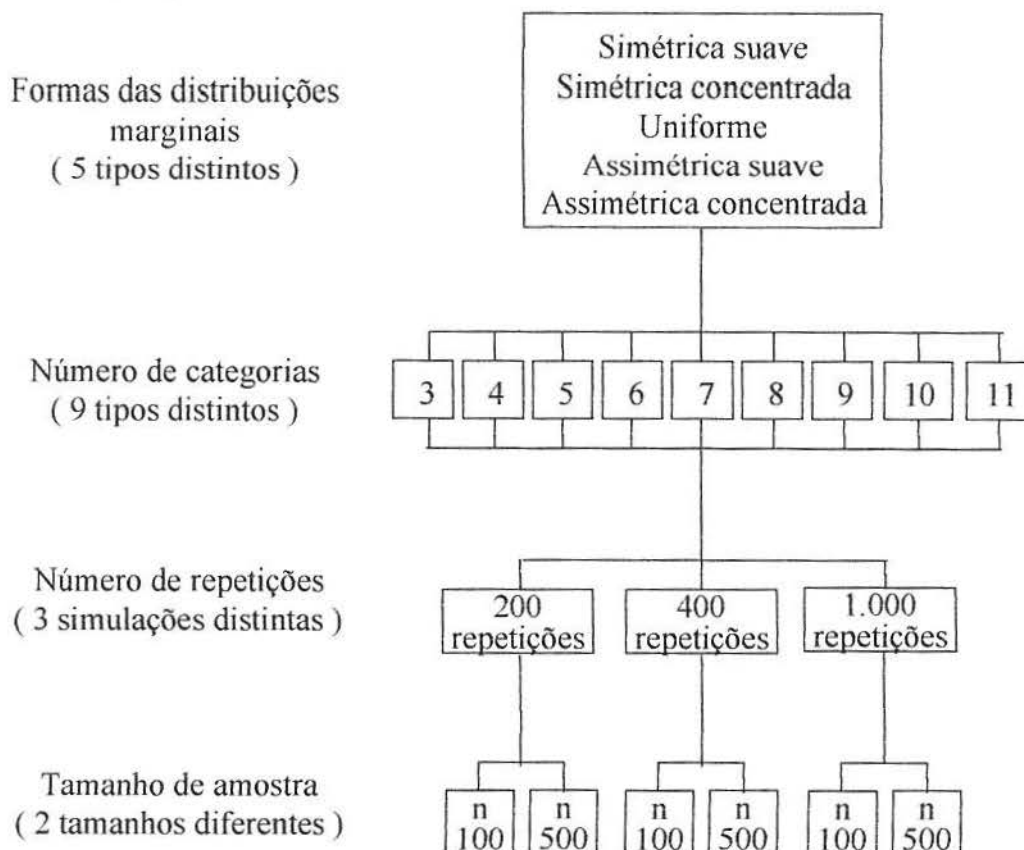


Figura 4.3 – Delineamento da simulação

Nesta monografia, apresentamos somente resultados de correlações de variáveis categóricas ordinais politômicas na forma de tabela quadrada, ou seja, 3x3, 4x4 11x11, pois na prática estes tipos de variáveis são oriundas de escalas de medida onde o número de categorias são os mesmos de um item para outro.

Na psicometria e na sociometria é muito comum o uso de variáveis com escalas tipo Likert, com o objetivo de, por exemplo, medir a quantidade de estresse, a qualidade de vida, a desesperança, a depressão e outros constructos não diretamente mensuráveis onde muitas vezes o interesse do pesquisador é relacionar itens desta mesma escala.

5- RESULTADOS

Nas figuras 5.1 a 5.9 ilustram graficamente as 400 repetições da simulação com tamanho de amostra igual a 100, ordenados ascendentemente em relação ao coeficiente de correlação de Pearson calculado com os dados contínuos (variáveis antes de serem categorizadas). Nós resolvemos apresentar os gráficos para tamanhos de amostra iguais a 100, pois na prática trabalhamos geralmente com amostras não muito grandes, mas em anexo encontram-se também estudos com $n=500$; o número de repetições para os gráficos foram todos de 400 repetições, pois se utilizássemos 200 repetições ficaríamos com poucos pontos e se utilizássemos 1000 repetições ficaríamos com muitos pontos, ou seja, teríamos gráficos mais “poluídos”, mas estes resultados também constam em anexo.

Para todos os gráficos, a linha preta representa o coeficiente de correlação de Pearson calculado a partir das distribuições marginais contínuas, onde a média das 400 simulações para amostras de tamanho 100 foi 0,49551 o qual é muito próximo do parâmetro de correlação 0,50, que é o valor fixado anteriormente na simulação.

A linha vermelha em todos os gráficos, representa o coeficiente de correlação Tipo-C calculado através do *software* Crosspsi para as 400 repetições de amostras de tamanho 100 e diferentes números de categorias das distribuições marginais.

A linha azul representa o coeficiente de correlação de Pearson calculado sobre as variáveis categorizadas (gráfico separado de 3 a 11 categorias) para 400 repetições e amostra de tamanho 100.

O Coeficiente de Correlação Tipo-C apresentado nos seguintes gráficos é obtido através do Coeficiente de Associação Ψ generalizado (Fachel, 1986) que foi transformado em um coeficiente de correlação através da fórmula proposta por Chambers (1982). Em anexo, temos também o Coeficiente de Correlação Tipo-C, onde o Ψ foi transformado através da fórmula proposta por Mardia (1967).

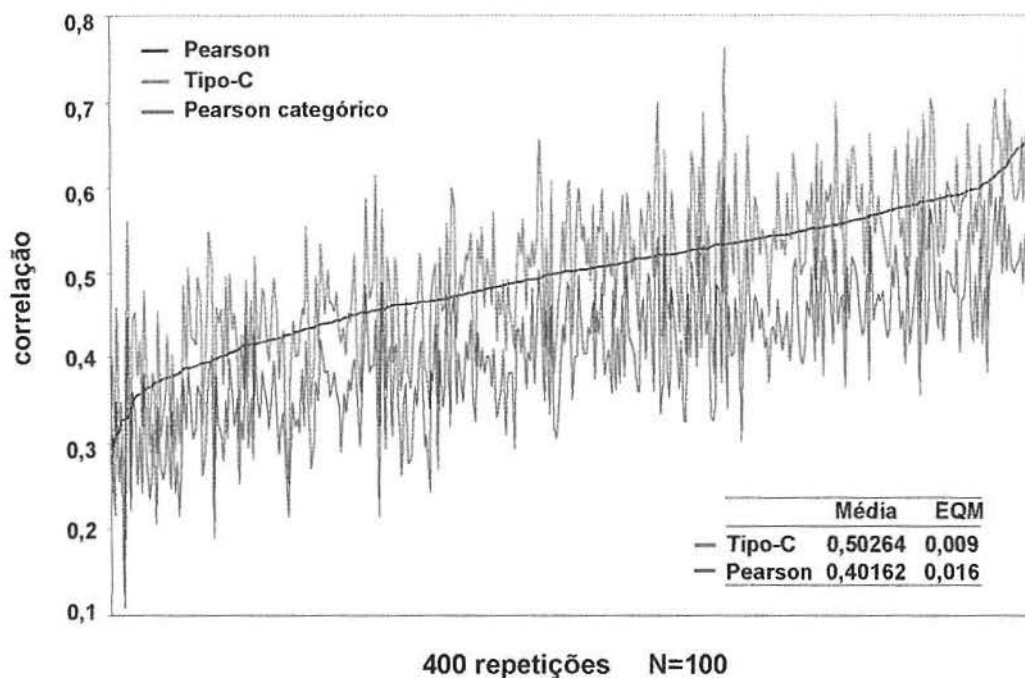


Figura 5.1 - Simulação para variáveis com 3 categorias com distribuições marginais simétrica suave

Neste gráfico, observamos que para variáveis com 3 categorias com distribuições marginais simétricas “suave” em torno da média, o Coeficiente de Correlação Tipo-C (linha vermelha) oscila sempre sobre o coeficiente de Correlação de Pearson calculado com os dados contínuos (linha preta, antes de categorizar as variáveis). A média das 400 repetições foi 0,50264, ou seja, bem próxima do parâmetro de correlação que foi fixado em 0,50; já o coeficiente de correlação de Pearson calculado com os dados categorizados (linha azul) subestimou o parâmetro, pois a média das 400 repetições foi 0,40162, bem abaixo de 0,50 e com um erro quadrático médio (EQM) quase duas vezes maior do que o do Coeficiente de Correlação Tipo-C.

Outra observação que se pode fazer neste gráfico é que os Coeficientes de Correlação Tipo-C e de Pearson, ambos calculados com os dados reduzidos a 3 categorias com distribuições marginais simétricas em torno da média, tiveram sempre o mesmo comportamento (correlação positiva forte entre as estimativas),

ou seja, no gráfico, quando um sobe outro sobe, quando um baixa o outro acompanha, porém em escalas diferentes.

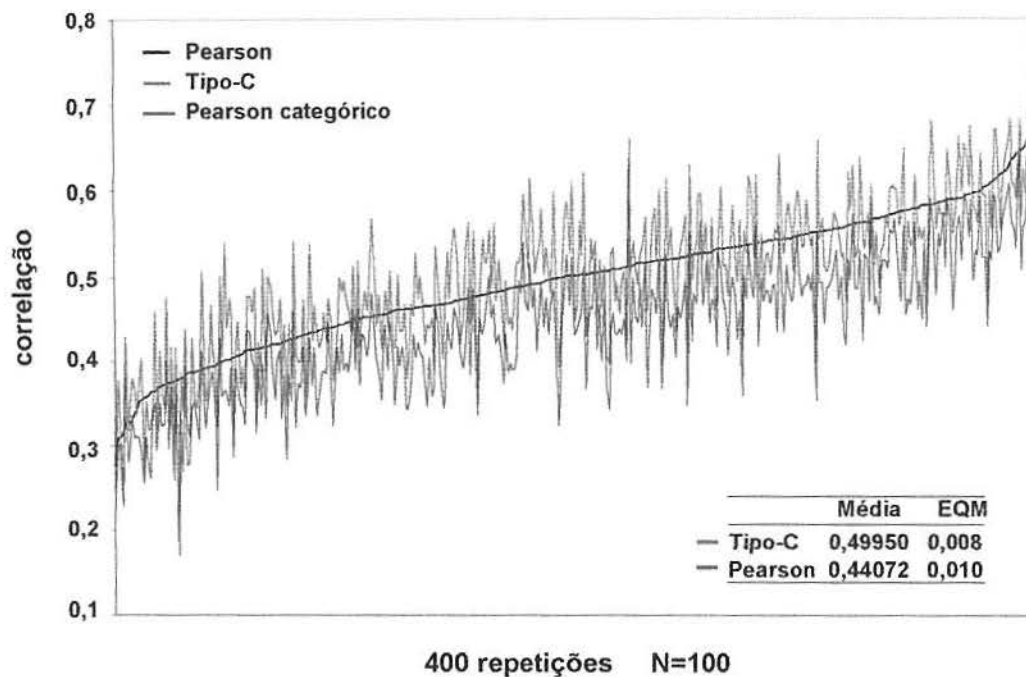


Figura 5.2 - Simulação para variáveis com 4 categorias com distribuições marginais simétrica suave

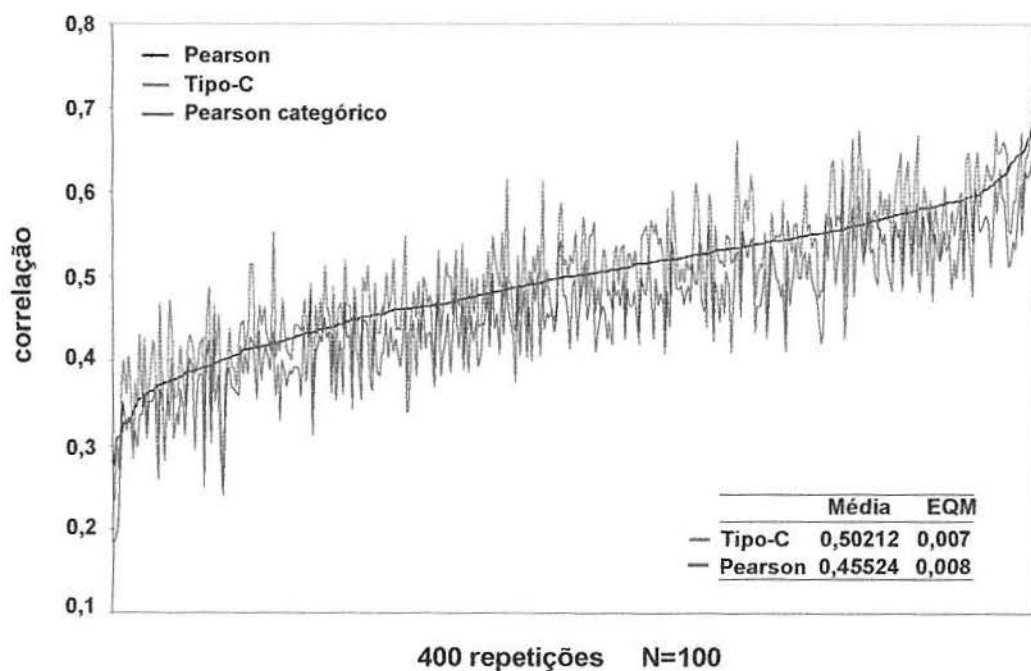


Figura 5.3 - Simulação para variáveis com 5 categorias com distribuições marginais simétrica suave

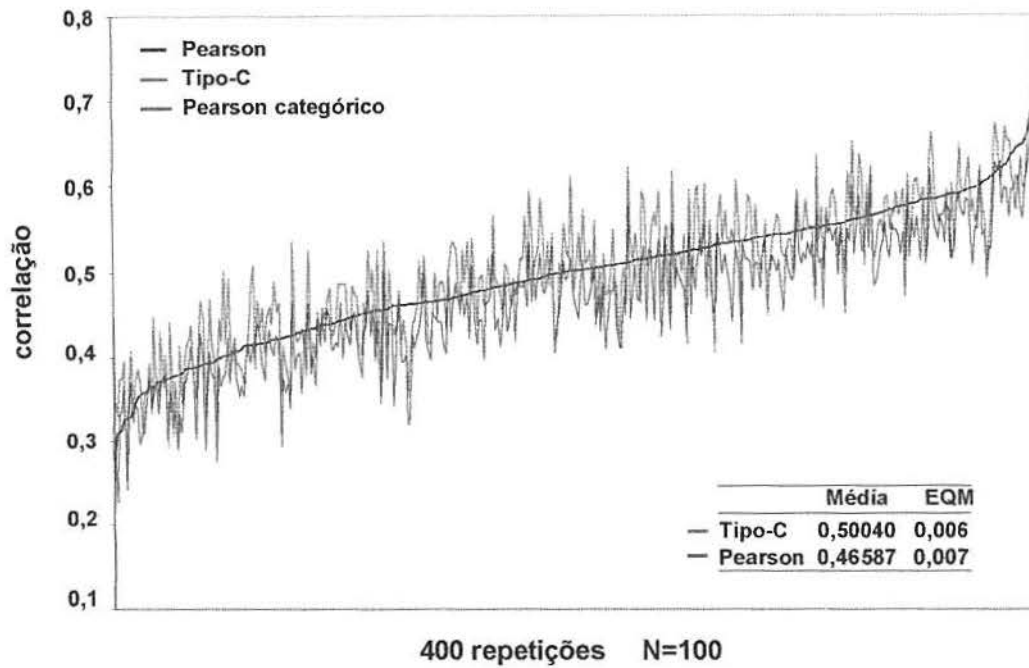


Figura 5.4 - Simulação para variáveis com 6 categorias com distribuições marginais simétrica suave

Para variáveis de 4, 5 e 6 categorias com distribuição marginal simétrica suave em torno da média, observamos através dos gráficos que o Coeficiente de Correlação de Pearson, teve uma melhora em relação aos de 3 categorias, porém, a média do Coeficiente de Correlação Tipo-C está bem mais próxima do parâmetro de correlação que é 0,50. O EQM do Coeficiente de Correlação Tipo-C é menor que o do de Pearson, ambos calculados com os dados categorizados.

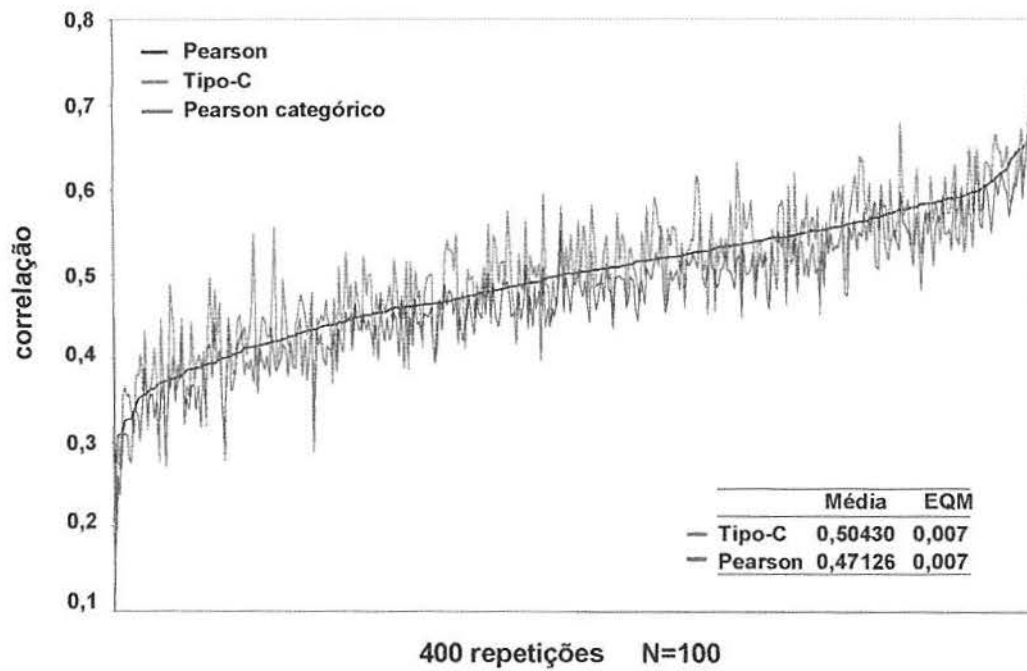


Figura 5.5 - Simulação para variáveis com 7 categorias com distribuições marginais simétrica suave

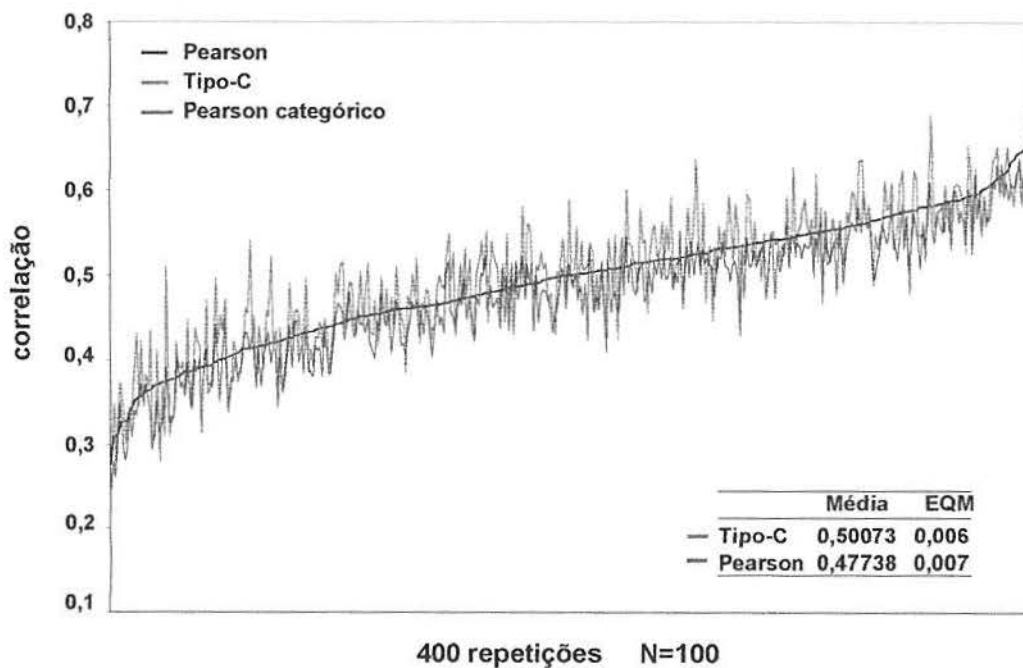


Figura 5.6 - Simulação para variáveis com 8 categorias com distribuições marginais simétrica suave

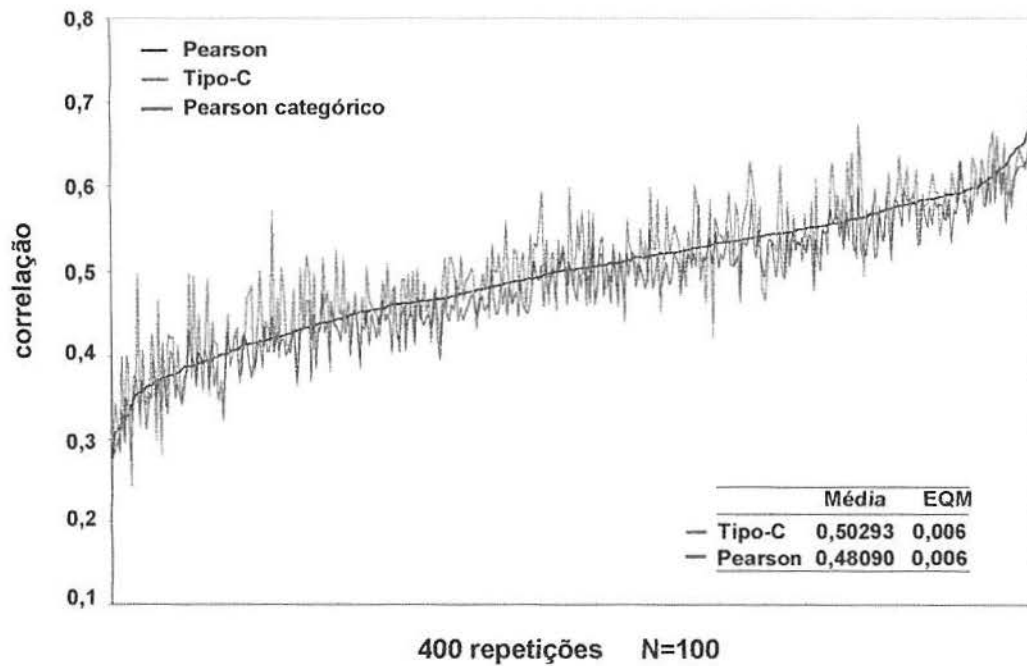


Figura 5.7 - Simulação para variáveis com 9 categorias com distribuições marginais simétrica suave

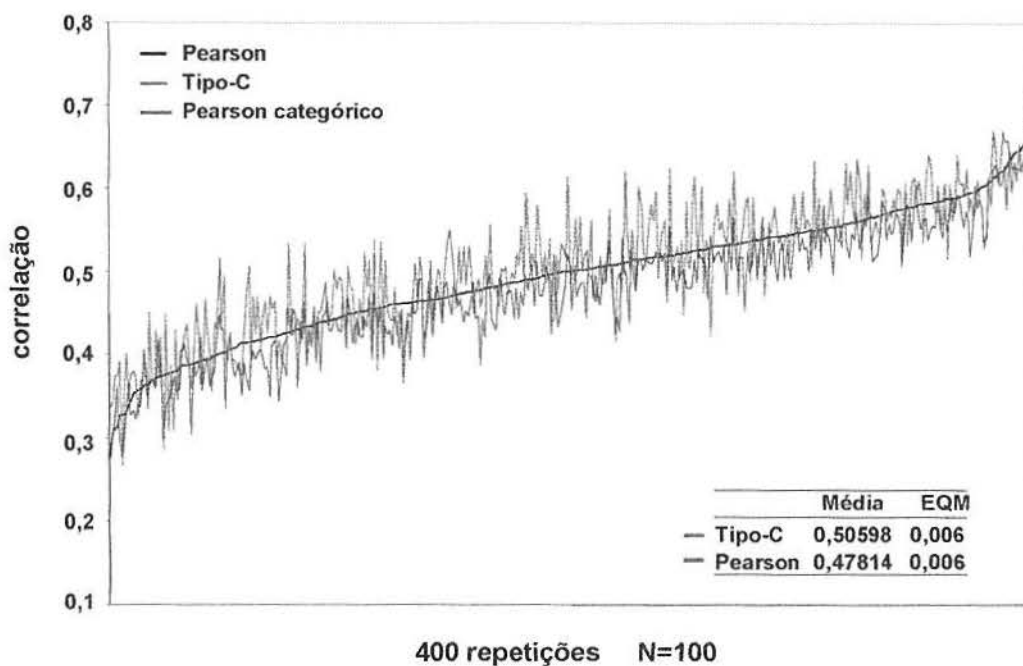


Figura 5.8 - Simulação para variáveis com 10 categorias com distribuições marginais simétrica suave

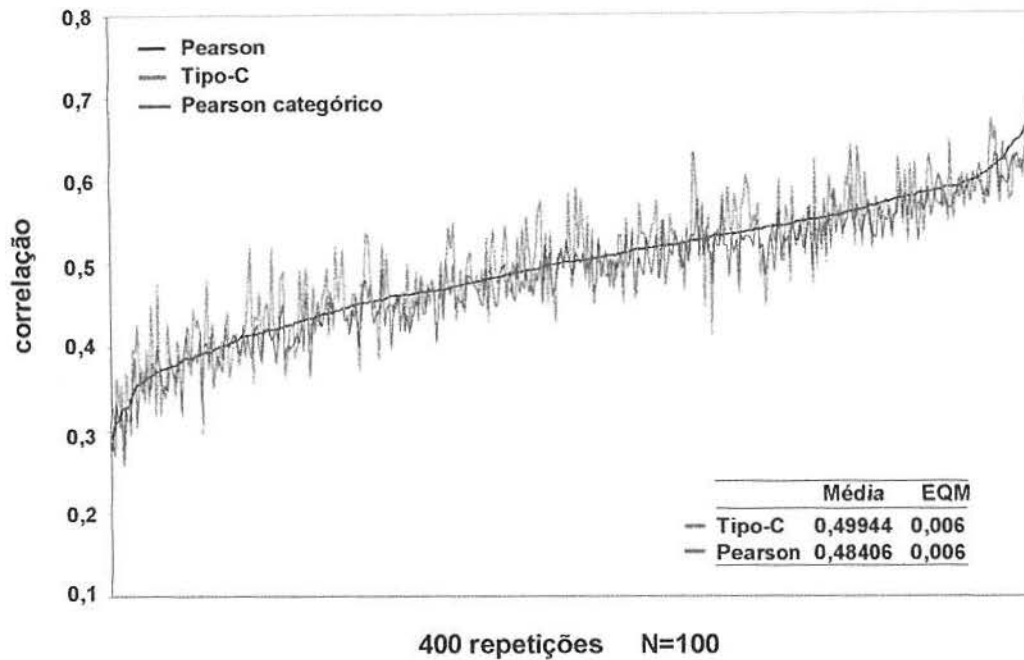


Figura 5.9 - Simulação para variáveis com 11 categorias com distribuições marginais simétrica suave

Para os gráficos onde as variáveis marginais foram reduzidas de 7 a 11 categorias com distribuição marginais simétricas suave em torno da média, observamos que o EQM do Coeficiente de Correlação de Pearson é aproximadamente o mesmo que o do Coeficiente de Correlação Tipo-C, porém para todos estes diferentes números de categorias, a média das 400 repetições do Coeficiente de Correlação Tipo-C está bem mais próximo de 0,50 do que a média do Coeficiente de Correlação de Pearson que está sempre subestimando o parâmetro.

Nestes nove gráficos acima apresentados, percebe-se que a medida que o número de categorias aumenta, o Coeficiente de Correlação de Pearson calculado com os dados categorizados (linha azul) está oscilando cada vez mais próximo do Coeficiente de Correlação de Pearson calculado com os dados contínuos (linha preta), mas este fato é lógico pois quanto maior o número de categorias, mais

próximo da distribuição contínua estamos, isto é, a perda de informação ao categorizarmos as variáveis é também menor.

É importante ressaltar que o Coeficiente de Correlação Tipo-C tem um desempenho muito melhor do que o de Pearson, quando estamos trabalhando com números pequenos de categorias, o que é comum na prática com o uso de instrumentos com escalas tipo Likert com 3 a 5 categorias.

Tabela 5.1 - Tabela comparativa dos coeficientes de correlação para a simulação com distribuições marginais simétrica "suave" em torno da média, tamanho de amostra igual a 100 e número de repetições igual a 400.

Nº de categorias	Coeficientes de correlação			
	Tipo-C		Pearson	
	Média	EQM	Média	EQM
3	0,50264	0,00914	0,40162	0,01649
4	0,49950	0,00757	0,44072	0,00992
5	0,50212	0,00698	0,45524	0,00811
6	0,50040	0,00639	0,46587	0,00717
7	0,50430	0,00671	0,47126	0,00682
8	0,50073	0,00648	0,47738	0,00653
9	0,50293	0,00633	0,48090	0,00601
10	0,50598	0,00636	0,47814	0,00627
11	0,49944	0,00611	0,48406	0,00587

Tabela 5.2 - Tabela comparativa dos coeficientes de correlação para a simulação com distribuições marginais simétrica "concentrada" em torno da média, 400 repetições e tamanho de amostra igual a 100.

Nº de categorias	Coeficiente de correlação			
	Tipo-C		Pearson	
	Média	EQM	Média	EQM
3	0,54668	0,01420	0,35993	0,02701
4	0,50988	0,00783	0,43984	0,00993
5	0,56231	0,01593	0,38529	0,02246
6	0,51952	0,00801	0,45810	0,00817
7	0,56269	0,01566	0,38622	0,02278
8	0,52033	0,00789	0,45533	0,00867
9	0,56396	0,01589	0,38911	0,02310
10	0,52108	0,00796	0,45657	0,00888
11	0,56373	0,01581	0,38684	0,02404

Tabela 5.3 – Tabela comparativa dos coeficientes de correlação para a simulação com distribuições marginais uniforme em torno da média, 400 repetições e tamanho de amostra igual a 100.

Nº de categorias	Coeficiente de correlação			
	Tipo-C		Pearson	
	Média	EQM	Média	EQM
3	0,48124	0,01442	0,34433	0,03237
4	0,48711	0,00771	0,43324	0,01123
5	0,49076	0,00711	0,44858	0,00905
6	0,48813	0,00679	0,45436	0,00856
7	0,49043	0,00647	0,46138	0,00762
8	0,49098	0,00630	0,46457	0,00757
9	0,48926	0,00630	0,46503	0,00743
10	0,49153	0,00614	0,46909	0,00712
11	0,49125	0,00597	0,47051	0,00699

Tabela 5.4 - Tabela comparativa dos coeficientes de correlação para a simulação com distribuições marginais assimétrica "suave" em torno da média, 400 repetições e tamanho de amostra igual a 100.

Nº de categorias	Coeficiente de correlação			
	Tipo-C		Pearson	
	Média	EQM	Média	EQM
3	0,49628	0,01066	0,39556	0,01906
4	0,49572	0,00932	0,41996	0,01401
5	0,50097	0,00997	0,41359	0,01750
6	0,49480	0,00853	0,42727	0,01406
7	0,49473	0,00871	0,42596	0,01466
8	0,49293	0,00870	0,42721	0,01488
9	0,49143	0,00851	0,42759	0,01486
10	0,48714	0,00796	0,43981	0,01216
11	0,48833	0,00780	0,44180	0,01173

Tabela 5.5 - Tabela comparativa dos coeficientes de correlação para a simulação com distribuições marginais assimétrica “concentrada” em torno da média, 400 repetições e tamanho de amostra igual a 100.

Nº de categorias	Coeficiente de correlação			
	Tipo-C		Pearson	
	Média	EQM	Média	EQM
3	0,50664	0,01515	0,35655	0,03176
4	0,50870	0,01468	0,36699	0,02982
5	0,51117	0,01451	0,37115	0,03006
6	0,51283	0,01392	0,37553	0,02863
7	0,51096	0,01434	0,37063	0,03068
8	0,51115	0,01403	0,37181	0,03030
9	0,50069	0,01258	0,38965	0,02440
10	0,49489	0,01003	0,41253	0,01811
11	0,48935	0,00845	0,43272	0,01340

Tabela 5.6 - Tabela comparativa dos coeficientes de correlação para a simulação com distribuições marginais simétrica “suave” em torno da média, 1.000 repetições e tamanho de amostra igual a 500.

Nº de categorias	Coeficiente de correlação			
	Tipo-C		Pearson	
	Média	EQM	Média	EQM
3	0,50854	0,00200	0,40894	0,00981
4	0,50052	0,00152	0,44503	0,00436
5	0,51316	0,00159	0,45797	0,00301
6	0,50341	0,00134	0,47096	0,00210
7	0,50957	0,00135	0,47715	0,00172
8	0,50401	0,00124	0,48258	0,00148
9	0,50555	0,00125	0,48500	0,00141
10	0,50918	0,00139	0,48068	0,00159
11	0,50301	0,00122	0,48851	0,00132

Tabela 5.7 - Tabela comparativa dos coeficientes de correlação para a simulação com distribuições marginais simétrica “concentrada” em torno da média, 1.000 repetições e tamanho de amostra igual a 500.

Nº de categorias	Coeficiente de correlação			
	Tipo-C		Pearson	
	Média	EQM	Média	EQM
3	0,55214	0,00507	0,36580	0,01963
4	0,51114	0,00163	0,44444	0,00438
5	0,56689	0,00670	0,38706	0,01458
6	0,52034	0,00185	0,46067	0,00280
7	0,56824	0,00687	0,38752	0,01462
8	0,52150	0,00190	0,45791	0,00310
9	0,56968	0,00708	0,38959	0,01433
10	0,52233	0,00195	0,45833	0,00313
11	0,56973	0,00707	0,38667	0,01502

Tabela 5.8 - Tabela comparativa dos coeficientes de correlação para a simulação com distribuições marginais uniforme, 1.000 repetições e tamanho de amostra igual a 500.

Nº de categorias	Coeficiente de correlação			
	Tipo-C		Pearson	
	Média	EQM	Média	EQM
3	0,50000	0,00198	0,40757	0,01007
4	0,49058	0,00164	0,43802	0,00528
5	0,49253	0,00143	0,45275	0,00354
6	0,49384	0,00136	0,46143	0,00280
7	0,49404	0,00130	0,46641	0,00240
8	0,49506	0,00126	0,47005	0,00218
9	0,49553	0,00126	0,47252	0,00202
10	0,49564	0,00123	0,47423	0,00191
11	0,49588	0,00120	0,47554	0,00185

Tabela 5.9 - Tabela comparativa dos coeficientes de correlação para a simulação com distribuições marginais assimétrica “suave”, 1.000 repetições e tamanho de amostra igual a 500.

Nº de categorias	Coeficiente de correlação			
	Tipo-C		Pearson	
	Média	EQM	Média	EQM
3	0,49961	0,00203	0,40058	0,01145
4	0,49953	0,00178	0,42466	0,00717
5	0,50350	0,00190	0,41932	0,00838
6	0,49816	0,00162	0,43357	0,00608
7	0,49763	0,00161	0,43036	0,00652
8	0,49759	0,00163	0,43427	0,00613
9	0,49626	0,00158	0,43455	0,00610
10	0,49351	0,00152	0,44665	0,00444
11	0,49307	0,00152	0,44693	0,00439

Tabela 5.10 - Tabela comparativa dos coeficientes de correlação para a simulação com distribuições marginais assimétrica “concentrada” à esquerda, 1.000 repetições e tamanho de amostra igual a 500.

Nº de categorias	Dados categorizados assimetricamente “concentrada”			
	Tipo-C		Pearson	
	Média	EQM	Média	EQM
3	0,51152	0,00293	0,36273	0,02106
4	0,51018	0,00272	0,36457	0,02045
5	0,51672	0,00286	0,37808	0,01748
6	0,51644	0,00279	0,38156	0,01658
7	0,51632	0,00281	0,37888	0,01732
8	0,51638	0,00282	0,38004	0,01710
9	0,50701	0,00230	0,39762	0,01280
10	0,49849	0,00185	0,41915	0,00849
11	0,49370	0,00161	0,43872	0,00542

6- CONCLUSÃO

Nesta monografia podemos verificar através das Figuras 2.1 e 2.2 que a distribuição bivariada Tipo-C Normal e a distribuição Normal bivariada são muito semelhantes e que, portanto, a suposição de dados com distribuição subjacente dada por esta distribuição é possível nas mesmas condições onde supomos normalidade.

As tabelas e as figuras apresentadas no capítulo V demonstram que para 400 repetições e tamanho de amostra igual a 100, independente da forma e do número de categorias das distribuições marginais o Coeficiente de Correlação Tipo-C apresentou resultados sempre melhores quando comparado com o Coeficiente de Correlação de Pearson pois está sempre mais próximo do parâmetro de correlação ($\rho=0,50$) e tem estimativas mais precisas.

Para as variáveis categóricas, observamos que os coeficientes de correlação Tipo-C e de Pearson estão altamente correlacionados, porém em escalas diferentes.

O Coeficiente de Correlação de Pearson para dados categóricos, sempre subestimou o parâmetro de correlação pois a média das 400 repetições sempre ficou abaixo de 0,50. Nos casos onde o número de categorias é pequeno, este coeficiente mostrou um péssimo desempenho pois além de subestimar bastante o parâmetro de correlação, seu erro quadrático médio foi bem alto quando comparado com o Coeficiente de Correlação Tipo-C.

Quando a forma das distribuições marginais é simétrica concentrada, observamos que o coeficiente de correlação Tipo-C superestima o parâmetro de correlação. E neste caso o resultado do coeficiente de correlação de Pearson é também pior do que nos outros casos, subestimando o parâmetro. Supõe-se que este resultado é devido ao fato de muitas células na tabela de contingência, formadas pelo cruzamento de duas distribuições simétricas concentradas terem

frequências muito baixas nas caselas não centrais, o que torna o estudo menos eficiente.

Estas conclusões também são válidas para os demais casos do delineamento da simulação ou seja para 200 e 1000 repetições, com tamanhos de amostra 100 e 500.

Podemos observar através deste estudo que o coeficiente de correlação Tipo-C também apresentou um desempenho muito bom em tabelas com baixa frequências.

Nos resultados anteriormente apresentados, podemos verificar que estamos correlacionando sempre variáveis com as mesmas distribuições marginais (X e Y), mas as tabelas A31 e A32 mostram que para variáveis com o mesmo número de categorias e independente da forma das distribuições marginais, seja X simétrica suave e Y uniforme por exemplo, ou outras combinações, o coeficiente de correlação Tipo-C mantém sempre o mesmo comportamento, ou seja, a sua média sempre próxima do parâmetro de correlação e um erro quadrático sempre menor do que o do coeficiente de correlação de Pearson que continua subestimando o parâmetro. O estudo de simulação para formas diferenciadas de distribuições marginais não é objeto desta monografia e foi realizado apenas em alguns casos, sendo estes resultados apresentados em apêndice com objetivo de reunir o material completo deste tópico. Futuras pesquisas serão realizadas para completar estes estudos de simulação.

7- BIBLIOGRAFIA

- Bartholomew, D.J. (1993) Latent variable models . London: Griffiin
- Bishop, Y.M; Fienberg, S.E. and Holland, P.W. (1975) Discrete multivariate analysis: theory and practice. Cambridge, Mass.: The MIT Press.
- Bittencourt, H.R.; Martins, M.H.M. e Fachel, J.M.G. (1997) Comparação de medidas de associação versus o coeficiente de correlação de Pearson. IX Salão de Iniciação Científica, UFRGS, p.47 (Resumo).
- Chambers, R.G. (1982) Correlation coefficients from 2X2 tables and from biserial data. British Journal of Mathematical and Statistical Psychology , 35, 216-227.
- D'Ávila, E.R. e Fachel, J.M.G. (1998) Programa "Crosspsi" para calcular o coeficiente de correlação Tipo-C. X Salão de Iniciação Científica, UFRGS, p.25 (Resumo).
- Everitt, B.S. (1992) The analysis of contingency tables. London: Chapman & Hall.
- Fachel, J.M.G. (1986) The C-Type distribution as an underlying model for categorical data and its use in Factor Analysis. PhD Thesis. London School of Economics and Political Science. University of London.
- Fachel, J.M.G. and Machado, A.A. (1982) Correlation coefficients as functions of the global cross-product ratio for RxC contingency tables. Biometric Bulletin, Vol.9, nº 2, pg.8.

- Fachel, J.M.G. and Machado, A.A. (1991) Modelos de associação para tabelas de contingência, 36º Reunião Anual da RBRAS da Sociedade Internacional de Biometria, UFG, Goiânia, p.54.
- Fachel, J.M.G. and Machado, A.A. (1990) modelagem da superfície de associação constante para tabelas de contingência RxC. XIII Congresso Nacional de Matemática Aplicada e Computacional, p.45.
- Kato, S.K.; Weber, L.H.; Fachel, J.M.G. (1998) Estudos de simulação para testar uma nova medida de correlação para dados politômicos. X Salão de Iniciação Científica, UFRGS, p.24 (Resumo).
- Kato, S.K. e Fachel, J.M.G. (1999) O efeito do número de categorias das distribuições marginais para o cálculo do coeficiente de correlação politômico Tipo-C: Um Estudo de Simulação. XI Salão de Iniciação Científica, UFRGS, p.57 (Resumo).
- Kendall, M. and Stuart, A. (1979) The advanced theory of statistics ,Vol. 2, 4th ed., London: Charles Griffin & Co.
- Martins, M.H.M.; Bittencourt, H.R.; Fachel, J.M.G. (1997) Simulação de dados para estudar o desempenho de coeficientes de associação. IX Salão de Iniciação Científica, UFRGS, p.47 (Resumo).
- Mardia, K.V. (1967) Some contributions to contingency-type bivariate distributions. Biometrika, 54, 235-249.
- Mardia, K.V. (1970) Families of bivariate distributions. London: Griffin.
- Olsson, U. (1979) Maximum likelihood estimation of the polychoric correlation coefficient. Psychometrika, 44, 443-460.

- Pearson, K. and Heron, D. (1913) On theories of association. Biometrika, 9, 159-315.
- Plackett, R.L. (1965) A class of bivariate distributions, Journal of the American Statistical Association, 60, 516-522.
- Quiroga, A.M. (1982). Studies of the polychoric correlation and other correlation measures for ordinal variables. Doctoral Dissertation, Uppsala University, Acta Universitatis Upsaliensis, nº 29.
- Weber, L.H. e Fachel, J.M.G. (1999). Resultados do desempenho do coeficiente de correlação Tipo-C para escalas de Likert. XI Salão de Iniciação Científica, UFRGS, p.57 (Resumo).

ANEXOS

A1- Simétrico “suave” n=100 200 comparações, 200 x e 200 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49079	0,46365	0,50117	0,39805
	4	0,49079	0,45788	0,49554	0,43627
	5	0,49079	0,45819	0,49595	0,44938
	6	0,49079	0,45980	0,49771	0,46181
	7	0,49079	0,46047	0,49837	0,46658
	8	0,49079	0,46041	0,49836	0,47436
	9	0,49079	0,45980	0,49769	0,47642
	10	0,49079	0,46489	0,50298	0,47340
	11	0,49079	0,45752	0,49538	0,47953
Desvio Padrão	3	0,07016	0,09209	0,09482	0,08168
	4	0,07016	0,08112	0,08397	0,07747
	5	0,07016	0,07797	0,08081	0,07747
	6	0,07016	0,07353	0,07601	0,07556
	7	0,07016	0,07533	0,07813	0,07382
	8	0,07016	0,07295	0,07547	0,07312
	9	0,07016	0,07461	0,07727	0,07224
	10	0,07016	0,07235	0,07464	0,07074
	11	0,07016	0,07225	0,07482	0,07170
Erro Quadrático	3	0,99656	1,95199	1,78932	3,40656
	4	0,99656	1,66441	1,40703	2,00671
	5	0,99656	1,55941	1,30265	1,70692
	6	0,99656	1,39926	1,15071	1,42788
	7	0,99656	1,44184	1,21531	1,30770
	8	0,99656	1,37235	1,13386	1,19556
	9	0,99656	1,43080	1,18913	1,14973
	10	0,99656	1,28815	1,11044	1,13724
	11	0,99656	1,39981	1,11834	1,10671
Erro Quadrático Médio	3	0,00498	0,00976	0,00895	0,01703
	4	0,00498	0,00832	0,00704	0,01003
	5	0,00498	0,00780	0,00651	0,00853
	6	0,00498	0,00700	0,00575	0,00714
	7	0,00498	0,00721	0,00608	0,00654
	8	0,00498	0,00686	0,00567	0,00598
	9	0,00498	0,00715	0,00595	0,00575
	10	0,00498	0,00644	0,00555	0,00569
	11	0,00498	0,00700	0,00559	0,00553

A2 - Simétrico "concentrado" n=100 200 comparações, 200 x e 200 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49079	0,50299	0,54095	0,35513
	4	0,49079	0,46783	0,50583	0,43575
	5	0,49079	0,51833	0,55659	0,38100
	6	0,49079	0,47711	0,51541	0,45416
	7	0,49079	0,51857	0,55682	0,38138
	8	0,49079	0,47774	0,51607	0,45069
	9	0,49079	0,51995	0,55819	0,38386
	10	0,49079	0,47853	0,51687	0,45145
	11	0,49079	0,51924	0,55750	0,38074
Desvio Padrão	3	0,07016	0,10712	0,10969	0,08456
	4	0,07016	0,08001	0,08257	0,07634
	5	0,07016	0,10462	0,10666	0,08948
	6	0,07016	0,07804	0,08033	0,07362
	7	0,07016	0,10471	0,10671	0,09380
	8	0,07016	0,07755	0,07976	0,07571
	9	0,07016	0,10540	0,10736	0,09886
	10	0,07016	0,07807	0,08027	0,07785
	11	0,07016	0,10470	0,10662	0,10015
Erro Quadrático	3	0,99656	2,28515	2,72975	5,62062
	4	0,99656	1,48084	1,36341	1,98525
	5	0,99656	2,24515	2,90432	4,42565
	6	0,99656	1,31677	1,33161	1,49891
	7	0,99656	2,25064	2,91169	4,56489
	8	0,99656	1,29571	1,31769	1,62702
	9	0,99656	2,29024	2,97068	4,64258
	10	0,99656	1,30505	1,33914	1,67743
	11	0,99656	2,25549	2,92327	4,84071
Erro Quadrático Médio	3	0,00498	0,01143	0,01365	0,02810
	4	0,00498	0,00740	0,00682	0,00993
	5	0,00498	0,01123	0,01452	0,02213
	6	0,00498	0,00658	0,00666	0,00749
	7	0,00498	0,01125	0,01456	0,02282
	8	0,00498	0,00648	0,00659	0,00814
	9	0,00498	0,01145	0,01485	0,02321
	10	0,00498	0,00653	0,00670	0,00839
	11	0,00498	0,01128	0,01462	0,02420

A3 - Uniforme n=100 200 comparações, 200 x e 200 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49079	0,43677	0,47337	0,39407
	4	0,49079	0,44822	0,48553	0,42951
	5	0,49079	0,44856	0,48595	0,44259
	6	0,49079	0,44721	0,48464	0,45014
	7	0,49079	0,45004	0,48760	0,45775
	8	0,49079	0,44943	0,48702	0,46024
	9	0,49079	0,44710	0,48461	0,46040
	10	0,49079	0,45065	0,48829	0,46462
	11	0,49079	0,44889	0,48647	0,46509
Desvio Padrão	3	0,07016	0,09373	0,09783	0,08659
	4	0,07016	0,08212	0,08521	0,08186
	5	0,07016	0,07954	0,08273	0,08029
	6	0,07016	0,07605	0,07900	0,07871
	7	0,07016	0,07470	0,07750	0,07576
	8	0,07016	0,07264	0,07539	0,07810
	9	0,07016	0,07250	0,07539	0,07757
	10	0,07016	0,07167	0,07435	0,07719
	11	0,07016	0,07217	0,07497	0,07745
Erro Quadrático	3	0,99656	2,54784	2,04651	3,73649
	4	0,99656	1,87831	1,48663	2,32735
	5	0,99656	1,78840	1,40135	1,94216
	6	0,99656	1,70836	1,28927	1,73026
	7	0,99656	1,60949	1,22594	1,49919
	8	0,99656	1,56139	1,16490	1,52990
	9	0,99656	1,60574	1,17849	1,51096
	10	0,99656	1,50936	1,12750	1,43591
	11	0,99656	1,55887	1,15519	1,43740
Erro Quadrático Médio	3	0,00498	0,01274	0,01023	0,01868
	4	0,00498	0,00939	0,00743	0,01164
	5	0,00498	0,00894	0,00701	0,00971
	6	0,00498	0,00854	0,00645	0,00865
	7	0,00498	0,00805	0,00613	0,00750
	8	0,00498	0,00781	0,00582	0,00765
	9	0,00498	0,00803	0,00589	0,00755
	10	0,00498	0,00755	0,00564	0,00718
	11	0,00498	0,00779	0,00578	0,00719

A4 - assimétrico “suave” n=100 200 comparações, 200 x e 200 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49079	0,45629	0,49342	0,39302
	4	0,49079	0,45643	0,49373	0,41797
	5	0,49079	0,45932	0,49664	0,40953
	6	0,49079	0,45477	0,49217	0,42522
	7	0,49079	0,45403	0,49140	0,42206
	8	0,49079	0,45288	0,49016	0,42420
	9	0,49079	0,44990	0,48713	0,42295
	10	0,49079	0,44447	0,48159	0,43482
	11	0,49079	0,44652	0,48374	0,43738
Desvio Padrão	3	0,07016	0,09779	0,10111	0,08778
	4	0,07016	0,09246	0,09572	0,08684
	5	0,07016	0,09468	0,09772	0,09792
	6	0,07016	0,08729	0,09044	0,08988
	7	0,07016	0,08752	0,09071	0,09378
	8	0,07016	0,08941	0,09265	0,09581
	9	0,07016	0,08755	0,09093	0,09560
	10	0,07016	0,08433	0,08773	0,08996
	11	0,07016	0,08338	0,08663	0,08952
Erro Quadrático	3	0,99656	2,28526	2,04313	3,82248
	4	0,99656	2,08102	1,83127	2,84658
	5	0,99656	2,11484	1,90268	3,54529
	6	0,99656	1,92528	1,64010	2,72613
	7	0,99656	1,94711	1,65238	2,96504
	8	0,99656	2,03468	1,72768	2,97592
	9	0,99656	2,02744	1,67851	3,00603
	10	0,99656	2,03210	1,59932	2,46010
	11	0,99656	1,95536	1,54616	2,37921
Erro Quadrático Médio	3	0,00498	0,01143	0,01022	0,01911
	4	0,00498	0,01041	0,00916	0,01423
	5	0,00498	0,01057	0,00951	0,01773
	6	0,00498	0,00963	0,00820	0,01363
	7	0,00498	0,00974	0,00826	0,01483
	8	0,00498	0,01017	0,00864	0,01488
	9	0,00498	0,01014	0,00839	0,01503
	10	0,00498	0,01016	0,00800	0,01230
	11	0,00498	0,00978	0,00773	0,01190

A5 - assimétrico "concentrado" n=100 200 comparações, 200 x e 200 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49079	0,46866	0,50555	0,35721
	4	0,49079	0,47083	0,50780	0,36727
	5	0,49079	0,47250	0,50954	0,37095
	6	0,49079	0,47259	0,50962	0,37372
	7	0,49079	0,47034	0,50726	0,36810
	8	0,49079	0,47012	0,50708	0,36870
	9	0,49079	0,45984	0,49673	0,38678
	10	0,49079	0,45256	0,48971	0,40942
	11	0,49079	0,44784	0,48503	0,42897
Desvio Padrão	3	0,07016	0,11637	0,12004	0,10127
	4	0,07016	0,11587	0,11951	0,10474
	5	0,07016	0,11496	0,11852	0,10744
	6	0,07016	0,11519	0,11876	0,11001
	7	0,07016	0,11671	0,12041	0,11369
	8	0,07016	0,11532	0,11889	0,11220
	9	0,07016	0,10943	0,11340	0,10700
	10	0,07016	0,09371	0,09731	0,09962
	11	0,07016	0,08652	0,08993	0,09250
Erro Quadrático	3	0,99656	2,89142	2,87363	6,11875
	4	0,99656	2,84189	2,85458	5,70659
	5	0,99656	2,78122	2,81361	5,62784
	6	0,99656	2,79093	2,82501	5,59778
	7	0,99656	2,88678	2,89554	6,05144
	8	0,99656	2,82496	2,82279	5,95341
	9	0,99656	2,70553	2,56119	4,84213
	10	0,99656	2,19772	1,90549	3,61581
	11	0,99656	2,03384	1,65408	2,71170
Erro Quadrático Médio	3	0,00498	0,01446	0,01437	0,03059
	4	0,00498	0,01421	0,01427	0,02853
	5	0,00498	0,01391	0,01407	0,02814
	6	0,00498	0,01395	0,01413	0,02799
	7	0,00498	0,01443	0,01448	0,03026
	8	0,00498	0,01412	0,01411	0,02977
	9	0,00498	0,01353	0,01281	0,02421
	10	0,00498	0,01099	0,00953	0,01808
	11	0,00498	0,01017	0,00827	0,01356

A6 - Simétrico "Suave" n=500 200 comparações, 200 x e 200 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.	
Média	3	0,50010	0,47114	0,50999	0,40983	
	4	0,50010	0,46130	0,49990	0,44435	
	5	0,50010	0,46839	0,50724	0,46145	
	6	0,50010	0,46503	0,50379	0,47099	
	7	0,50010	0,47020	0,50912	0,47713	
	8	0,50010	0,46547	0,50426	0,48250	
	9	0,50010	0,46681	0,50563	0,48507	
	10	0,50010	0,47075	0,50969	0,48082	
	11	0,50010	0,46499	0,50377	0,48922	
	Desvio Padrão	3	0,03261	0,04155	0,04282	0,03805
		4	0,03261	0,03678	0,03797	0,03420
5		0,03261	0,03530	0,03643	0,03434	
6		0,03261	0,03351	0,03458	0,03366	
7		0,03261	0,03424	0,03531	0,03424	
8		0,03261	0,03181	0,03282	0,03152	
9		0,03261	0,03383	0,03493	0,03391	
10		0,03261	0,03316	0,03416	0,03306	
11		0,03261	0,03274	0,03382	0,03322	
Erro Quadrático		3	0,21162	0,51018	0,38486	1,91421
		4	0,21162	0,56883	0,28697	0,85221
	5	0,21162	0,44781	0,27464	0,53190	
	6	0,21162	0,46811	0,24080	0,39377	
	7	0,21162	0,41089	0,26478	0,33789	
	8	0,21162	0,43986	0,21793	0,25900	
	9	0,21162	0,44804	0,24918	0,27341	
	10	0,21162	0,39000	0,25096	0,29102	
	11	0,21162	0,45845	0,23048	0,24286	
	Erro Quadrático Médio	3	0,00106	0,00255	0,00192	0,00957
		4	0,00106	0,00284	0,00143	0,00426
5		0,00106	0,00224	0,00137	0,00266	
6		0,00106	0,00234	0,00120	0,00197	
7		0,00106	0,00205	0,00132	0,00169	
8		0,00106	0,00220	0,00109	0,00130	
9		0,00106	0,00224	0,00125	0,00137	
10		0,00106	0,00195	0,00125	0,00146	
11		0,00106	0,00229	0,00115	0,00121	

A7 - Simétrico "concentrado" n=500 200 comparações, 200 x e 200 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,50010	0,51247	0,55214	0,36647
	4	0,50010	0,47110	0,51004	0,44359
	5	0,50010	0,52738	0,56725	0,38836
	6	0,50010	0,48048	0,51969	0,46038
	7	0,50010	0,52849	0,56837	0,38851
	8	0,50010	0,48141	0,52065	0,45791
	9	0,50010	0,53012	0,57001	0,39123
	10	0,50010	0,48234	0,52160	0,45876
	11	0,50010	0,52983	0,56973	0,38714
Desvio Padrão	3	0,03261	0,04838	0,04924	0,03962
	4	0,03261	0,03441	0,03542	0,03250
	5	0,03261	0,04750	0,04814	0,04286
	6	0,03261	0,03420	0,03512	0,03290
	7	0,03261	0,04659	0,04717	0,04348
	8	0,03261	0,03343	0,03432	0,03357
	9	0,03261	0,04712	0,04767	0,04684
	10	0,03261	0,03358	0,03446	0,03526
	11	0,03261	0,04698	0,04756	0,04704
Erro Quadrático	3	0,21162	0,49680	1,02607	3,87855
	4	0,21162	0,40265	0,26989	0,84660
	5	0,21162	0,59887	1,36551	2,85809
	6	0,21162	0,30894	0,32302	0,52943
	7	0,21162	0,59425	1,37778	2,86228
	8	0,21162	0,29145	0,31961	0,57858
	9	0,21162	0,62321	1,43251	2,80299
	10	0,21162	0,28673	0,32962	0,58754
	11	0,21162	0,61727	1,42251	2,98764
Erro Quadrático Médio	3	0,00106	0,00248	0,00513	0,01939
	4	0,00106	0,00201	0,00135	0,00423
	5	0,00106	0,00299	0,00683	0,01429
	6	0,00106	0,00154	0,00162	0,00265
	7	0,00106	0,00297	0,00689	0,01431
	8	0,00106	0,00146	0,00160	0,00289
	9	0,00106	0,00312	0,00716	0,01401
	10	0,00106	0,00143	0,00165	0,00294
	11	0,00106	0,00309	0,00711	0,01494

A8 - Uniforme n=500 200 comparações, 200 x e 200 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,50010	0,44955	0,48769	0,40796
	4	0,50010	0,45270	0,49102	0,43775
	5	0,50010	0,45454	0,49292	0,45290
	6	0,50010	0,45439	0,49279	0,46035
	7	0,50010	0,45536	0,49380	0,46605
	8	0,50010	0,45621	0,49468	0,46956
	9	0,50010	0,45614	0,49462	0,47140
	10	0,50010	0,45737	0,49589	0,47432
	11	0,50010	0,45743	0,49596	0,47562
Desvio Padrão	3	0,03261	0,04148	0,04301	0,03845
	4	0,03261	0,03600	0,03729	0,03592
	5	0,03261	0,03677	0,03812	0,03622
	6	0,03261	0,03365	0,03485	0,03400
	7	0,03261	0,03333	0,03451	0,03371
	8	0,03261	0,03334	0,03451	0,03414
	9	0,03261	0,03358	0,03477	0,03435
	10	0,03261	0,03320	0,03437	0,03391
	11	0,03261	0,03217	0,03330	0,03400
Erro Quadrático	3	0,21162	0,85150	0,39844	1,98848
	4	0,21162	0,70538	0,29281	1,03185
	5	0,21162	0,68223	0,29912	0,70479
	6	0,21162	0,64141	0,25204	0,54446
	7	0,21162	0,61971	0,24468	0,45670
	8	0,21162	0,60471	0,24262	0,41726
	9	0,21162	0,60903	0,24632	0,39841
	10	0,21162	0,58287	0,23842	0,36075
	11	0,21162	0,56835	0,22395	0,34889
Erro Quadrático Médio	3	0,00106	0,00426	0,00199	0,00994
	4	0,00106	0,00353	0,00146	0,00516
	5	0,00106	0,00341	0,00150	0,00352
	6	0,00106	0,00321	0,00126	0,00272
	7	0,00106	0,00310	0,00122	0,00228
	8	0,00106	0,00302	0,00121	0,00209
	9	0,00106	0,00305	0,00123	0,00199
	10	0,00106	0,00291	0,00119	0,00180
	11	0,00106	0,00284	0,00112	0,00174

A9 - Assimétrico "suave" n=500 200 comparações, 200 x e 200 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,50010	0,45920	0,49772	0,39853
	4	0,50010	0,46051	0,49907	0,42382
	5	0,50010	0,46250	0,50113	0,41654
	6	0,50010	0,45953	0,49809	0,43266
	7	0,50010	0,45929	0,49784	0,43006
	8	0,50010	0,45894	0,49748	0,43313
	9	0,50010	0,45709	0,49557	0,43233
	10	0,50010	0,45422	0,49260	0,44451
	11	0,50010	0,45412	0,49250	0,44534
	Desvio Padrão	3	0,03261	0,03859	0,03985
4		0,03261	0,03794	0,03917	0,03678
5		0,03261	0,03862	0,03983	0,04070
6		0,03261	0,03540	0,03656	0,03701
7		0,03261	0,03600	0,03719	0,03869
8		0,03261	0,03582	0,03700	0,03978
9		0,03261	0,03564	0,03685	0,04043
10		0,03261	0,03548	0,03675	0,03787
11		0,03261	0,03535	0,03661	0,03746
Erro Quadrático		3	0,21162	0,62931	0,31707
	4	0,21162	0,59838	0,30553	1,42991
	5	0,21162	0,57798	0,31599	1,72291
	6	0,21162	0,57704	0,26677	1,17945
	7	0,21162	0,58929	0,27612	1,27635
	8	0,21162	0,59246	0,27374	1,20918
	9	0,21162	0,62105	0,27416	1,24112
	10	0,21162	0,66970	0,27966	0,90134
	11	0,21162	0,66972	0,27797	0,87677
	Erro Quadrático Médio	3	0,00106	0,00315	0,00159
4		0,00106	0,00299	0,00153	0,00715
5		0,00106	0,00289	0,00158	0,00861
6		0,00106	0,00289	0,00133	0,00590
7		0,00106	0,00295	0,00138	0,00638
8		0,00106	0,00296	0,00137	0,00605
9		0,00106	0,00311	0,00137	0,00621
10		0,00106	0,00335	0,00140	0,00451
11		0,00106	0,00335	0,00139	0,00438

A10 - Assimétrico "concentrado" n=500 200 comparações, 200 x e 200 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.	
Média	3	0,50010	0,47259	0,51138	0,36184	
	4	0,50010	0,47632	0,51526	0,37377	
	5	0,50010	0,47773	0,51670	0,37602	
	6	0,50010	0,47734	0,51630	0,37942	
	7	0,50010	0,47696	0,51591	0,37618	
	8	0,50010	0,47734	0,51629	0,37763	
	9	0,50010	0,46835	0,50708	0,39606	
	10	0,50010	0,45882	0,49733	0,41755	
	11	0,50010	0,45485	0,49326	0,43723	
	Desvio Padrão	3	0,03261	0,04858	0,04999	0,04483
		4	0,03261	0,04559	0,04681	0,04432
5		0,03261	0,04579	0,04701	0,04753	
6		0,03261	0,04569	0,04693	0,04916	
7		0,03261	0,04566	0,04691	0,04949	
8		0,03261	0,04639	0,04765	0,05042	
9		0,03261	0,04415	0,04552	0,04667	
10		0,03261	0,03859	0,03989	0,04200	
11		0,03261	0,03527	0,03649	0,03802	
Erro Quadrático		3	0,21162	0,61986	0,52324	4,21744
		4	0,21162	0,52568	0,48269	3,57771
	5	0,21162	0,51641	0,49548	3,52390	
	6	0,21162	0,51824	0,49138	3,38875	
	7	0,21162	0,52101	0,48853	3,55380	
	8	0,21162	0,53096	0,50489	3,50092	
	9	0,21162	0,58823	0,42239	2,59402	
	10	0,21162	0,63542	0,31808	1,71057	
	11	0,21162	0,65520	0,27407	1,07560	
	Erro Quadrático Médio	3	0,00106	0,00310	0,00262	0,02109
		4	0,00106	0,00263	0,00241	0,01789
5		0,00106	0,00258	0,00248	0,01762	
6		0,00106	0,00259	0,00246	0,01694	
7		0,00106	0,00261	0,00244	0,01777	
8		0,00106	0,00265	0,00252	0,01750	
9		0,00106	0,00294	0,00211	0,01297	
10		0,00106	0,00318	0,00159	0,00855	
11		0,00106	0,00328	0,00137	0,00538	

A11 - Simétrico "suave" n=100 400 comparações, 400 x e 400 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.	
Média	3	0,49551	0,46509	0,50264	0,40162	
	4	0,49551	0,46180	0,49950	0,44072	
	5	0,49551	0,46425	0,50212	0,45524	
	6	0,49551	0,46251	0,50040	0,46587	
	7	0,49551	0,46633	0,50430	0,47126	
	8	0,49551	0,46283	0,50073	0,47738	
	9	0,49551	0,46495	0,50293	0,48090	
	10	0,49551	0,46792	0,50598	0,47814	
	11	0,49551	0,46153	0,49944	0,48406	
	Desvio Padrão	3	0,07450	0,09285	0,09571	0,08266
		4	0,07450	0,08426	0,08714	0,08011
5		0,07450	0,08095	0,08361	0,07824	
6		0,07450	0,07744	0,08003	0,07757	
7		0,07450	0,07930	0,08193	0,07752	
8		0,07450	0,07797	0,08060	0,07769	
9		0,07450	0,07703	0,07959	0,07522	
10		0,07450	0,07717	0,07959	0,07621	
11		0,07450	0,07568	0,07824	0,07502	
Erro Quadrático		3	2,22277	3,92763	3,65760	6,59770
		4	2,22277	3,41663	3,02974	3,96630
	5	2,22277	3,12568	2,79090	3,24410	
	6	2,22277	2,95513	2,55577	2,86646	
	7	2,22277	2,96249	2,68581	2,72792	
	8	2,22277	2,97830	2,59230	2,61294	
	9	2,22277	2,85894	2,53107	2,40366	
	10	2,22277	2,78767	2,54204	2,50858	
	11	2,22277	2,87735	2,44249	2,34710	
	Erro Quadrático Médio	3	0,00556	0,00982	0,00914	0,01649
		4	0,00556	0,00854	0,00757	0,00992
5		0,00556	0,00781	0,00698	0,00811	
6		0,00556	0,00739	0,00639	0,00717	
7		0,00556	0,00741	0,00671	0,00682	
8		0,00556	0,00745	0,00648	0,00653	
9		0,00556	0,00715	0,00633	0,00601	
10		0,00556	0,00697	0,00636	0,00627	
11		0,00556	0,00719	0,00611	0,00587	

A12 - Simétrico "concentrado" n=100 400 comparações, 400 x e 400 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49551	0,50866	0,54668	0,35993
	4	0,49551	0,47192	0,50988	0,43984
	5	0,49551	0,52415	0,56231	0,38529
	6	0,49551	0,48131	0,51952	0,45810
	7	0,49551	0,52447	0,56269	0,38622
	8	0,49551	0,48207	0,52033	0,45533
	9	0,49551	0,52575	0,56396	0,38911
	10	0,49551	0,48281	0,52108	0,45657
	11	0,49551	0,52551	0,56373	0,38684
Desvio Padrão	3	0,07450	0,10749	0,10977	0,08608
	4	0,07450	0,08543	0,08807	0,07955
	5	0,07450	0,10814	0,10990	0,09656
	6	0,07450	0,08510	0,08744	0,08017
	7	0,07450	0,10666	0,10843	0,09927
	8	0,07450	0,08424	0,08657	0,08182
	9	0,07450	0,10706	0,10875	0,10409
	10	0,07450	0,08449	0,08680	0,08372
	11	0,07450	0,10684	0,10851	0,10612
Erro Quadrático	3	2,22277	4,64037	5,67942	10,80479
	4	2,22277	3,22726	3,13395	3,97267
	5	2,22277	4,89888	6,37224	8,98412
	6	2,22277	3,02907	3,20341	3,26657
	7	2,22277	4,77901	6,26332	9,11011
	8	2,22277	2,95995	3,15549	3,46943
	9	2,22277	4,83862	6,35532	9,24194
	10	2,22277	2,96671	3,18366	3,55112
	11	2,22277	4,81476	6,32254	9,61502
Erro Quadrático Médio	3	0,00556	0,01160	0,01420	0,02701
	4	0,00556	0,00807	0,00783	0,00993
	5	0,00556	0,01225	0,01593	0,02246
	6	0,00556	0,00757	0,00801	0,00817
	7	0,00556	0,01195	0,01566	0,02278
	8	0,00556	0,00740	0,00789	0,00867
	9	0,00556	0,01210	0,01589	0,02310
	10	0,00556	0,00742	0,00796	0,00888
	11	0,00556	0,01204	0,01581	0,02404

A13 - Uniforme n=100 400 comparações, 400 x e 400 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49551	0,44503	0,48124	0,34433
	4	0,49551	0,44979	0,48711	0,43324
	5	0,49551	0,45324	0,49076	0,44858
	6	0,49551	0,45065	0,48813	0,45436
	7	0,49551	0,45283	0,49043	0,46138
	8	0,49551	0,45334	0,49098	0,46457
	9	0,49551	0,45167	0,48926	0,46503
	10	0,49551	0,45386	0,49153	0,46909
	11	0,49551	0,45356	0,49125	0,47051
Desvio Padrão	3	0,07450	0,11460	0,11875	0,09032
	4	0,07450	0,08383	0,08698	0,08240
	5	0,07450	0,08087	0,08390	0,08012
	6	0,07450	0,07871	0,08166	0,08059
	7	0,07450	0,07709	0,07993	0,07837
	8	0,07450	0,07617	0,07898	0,07955
	9	0,07450	0,07589	0,07872	0,07887
	10	0,07450	0,07530	0,07802	0,07863
	11	0,07450	0,07414	0,07689	0,07833
Erro Quadrático	3	2,22277	6,44852	5,76716	12,94773
	4	2,22277	3,81242	3,08517	4,49181
	5	2,22277	3,48385	2,84248	3,61850
	6	2,22277	3,44582	2,71681	3,42487
	7	2,22277	3,26079	2,58606	3,04723
	8	2,22277	3,18555	2,52156	3,02716
	9	2,22277	3,23181	2,51896	2,97100
	10	2,22277	3,11413	2,45762	2,84870
	11	2,22277	3,05629	2,38943	2,79614
Erro Quadrático Médio	3	0,00556	0,01612	0,01442	0,03237
	4	0,00556	0,00953	0,00771	0,01123
	5	0,00556	0,00871	0,00711	0,00905
	6	0,00556	0,00861	0,00679	0,00856
	7	0,00556	0,00815	0,00647	0,00762
	8	0,00556	0,00796	0,00630	0,00757
	9	0,00556	0,00808	0,00630	0,00743
	10	0,00556	0,00779	0,00614	0,00712
	11	0,00556	0,00764	0,00597	0,00699

A14 - assimétrico "suave" n=100 400 comparações, 400 x e 400 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49551	0,45911	0,49628	0,39556
	4	0,49551	0,45839	0,49572	0,41996
	5	0,49551	0,46358	0,50097	0,41359
	6	0,49551	0,45738	0,49480	0,42727
	7	0,49551	0,45733	0,49473	0,42596
	8	0,49551	0,45558	0,49293	0,42721
	9	0,49551	0,45410	0,49143	0,42759
	10	0,49551	0,44985	0,48714	0,43981
	11	0,49551	0,45098	0,48833	0,44180
Desvio Padrão	3	0,07450	0,09982	0,10333	0,09040
	4	0,07450	0,09341	0,09657	0,08730
	5	0,07450	0,09675	0,09998	0,10027
	6	0,07450	0,08922	0,09235	0,09375
	7	0,07450	0,09013	0,09332	0,09592
	8	0,07450	0,08991	0,09314	0,09801
	9	0,07450	0,08873	0,09196	0,09817
	10	0,07450	0,08517	0,08841	0,09253
	11	0,07450	0,08446	0,08767	0,09143
Erro Quadrático	3	2,22277	4,64427	4,26591	7,62392
	4	2,22277	4,17373	3,72843	5,60310
	5	2,22277	4,26570	3,98879	6,99856
	6	2,22277	3,90266	3,41369	5,62304
	7	2,22277	3,96934	3,48577	5,86424
	8	2,22277	4,01437	3,48113	5,95176
	9	2,22277	3,98413	3,40329	5,94234
	10	2,22277	3,90004	3,18474	4,86506
	11	2,22277	3,80763	3,12149	4,69035
Erro Quadrático Médio	3	0,00556	0,01161	0,01066	0,01906
	4	0,00556	0,01043	0,00932	0,01401
	5	0,00556	0,01066	0,00997	0,01750
	6	0,00556	0,00976	0,00853	0,01406
	7	0,00556	0,00992	0,00871	0,01466
	8	0,00556	0,01004	0,00870	0,01488
	9	0,00556	0,00996	0,00851	0,01486
	10	0,00556	0,00975	0,00796	0,01216
	11	0,00556	0,00952	0,00780	0,01173

A15 - assimétrico "concentrado" n=100 400 comparações, 400 x e 400 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49551	0,46984	0,50664	0,35655
	4	0,49551	0,47177	0,50870	0,36699
	5	0,49551	0,47415	0,51117	0,37115
	6	0,49551	0,47568	0,51283	0,37553
	7	0,49551	0,47393	0,51096	0,37063
	8	0,49551	0,47408	0,51115	0,37181
	9	0,49551	0,46368	0,50069	0,38965
	10	0,49551	0,45766	0,49489	0,41253
	11	0,49551	0,45206	0,48935	0,43272
Desvio Padrão	3	0,07450	0,11939	0,12308	0,10587
	4	0,07450	0,11732	0,12099	0,11028
	5	0,07450	0,11650	0,12007	0,11617
	6	0,07450	0,11405	0,11742	0,11475
	7	0,07450	0,11587	0,11940	0,11825
	8	0,07450	0,11480	0,11809	0,11791
	9	0,07450	0,10866	0,11229	0,11072
	10	0,07450	0,09665	0,10014	0,10240
	11	0,07450	0,08811	0,09144	0,09433
Erro Quadrático	3	2,22277	6,05070	6,06176	12,70363
	4	2,22277	5,81023	5,87139	11,92925
	5	2,22277	5,68257	5,80234	12,02589
	6	2,22277	5,42611	5,56666	11,45102
	7	2,22277	5,62924	5,73617	12,27388
	8	2,22277	5,52675	5,61388	12,11989
	9	2,22277	5,23834	5,03083	9,76196
	10	2,22277	4,44419	4,01152	7,24453
	11	2,22277	4,01686	3,38165	5,36115
Erro Quadrático Médio	3	0,00556	0,01513	0,01515	0,03176
	4	0,00556	0,01453	0,01468	0,02982
	5	0,00556	0,01421	0,01451	0,03006
	6	0,00556	0,01357	0,01392	0,02863
	7	0,00556	0,01407	0,01434	0,03068
	8	0,00556	0,01382	0,01403	0,03030
	9	0,00556	0,01310	0,01258	0,02440
	10	0,00556	0,01111	0,01003	0,01811
	11	0,00556	0,01004	0,00845	0,01340

A16 - Simétrico "suave" n=500 400 comparações, 400 x e 400 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49995	0,47052	0,50936	0,40924
	4	0,49995	0,46205	0,50067	0,44450
	5	0,49995	0,46868	0,50752	0,46159
	6	0,49995	0,46490	0,50365	0,47073
	7	0,49995	0,47039	0,50930	0,47671
	8	0,49995	0,46513	0,50391	0,48210
	9	0,49995	0,46704	0,50585	0,48466
	10	0,49995	0,47057	0,50949	0,48026
	11	0,49995	0,46471	0,50347	0,48827
Desvio Padrão	3	0,03390	0,04146	0,04273	0,03809
	4	0,03390	0,03776	0,03899	0,03621
	5	0,03390	0,03640	0,03757	0,03604
	6	0,03390	0,03478	0,03590	0,03544
	7	0,03390	0,03491	0,03600	0,03529
	8	0,03390	0,03306	0,03412	0,03364
	9	0,03390	0,03421	0,03533	0,03440
	10	0,03390	0,03452	0,03557	0,03471
	11	0,03390	0,03351	0,03461	0,03437
Erro Quadrático	3	0,45851	1,03334	0,76345	3,87415
	4	0,45851	1,14500	0,60662	1,75512
	5	0,45851	0,92109	0,58575	1,10854
	6	0,45851	0,97532	0,51947	0,84387
	7	0,45851	0,83707	0,55180	0,71376
	8	0,45851	0,92239	0,47062	0,57958
	9	0,45851	0,90169	0,51164	0,56630
	10	0,45851	0,82192	0,54090	0,63668
	11	0,45851	0,94613	0,48281	0,52636
Erro Quadrático Médio	3	0,00115	0,00258	0,00191	0,00969
	4	0,00115	0,00286	0,00152	0,00439
	5	0,00115	0,00230	0,00146	0,00277
	6	0,00115	0,00244	0,00130	0,00211
	7	0,00115	0,00209	0,00138	0,00178
	8	0,00115	0,00231	0,00118	0,00145
	9	0,00115	0,00225	0,00128	0,00142
	10	0,00115	0,00205	0,00135	0,00159
	11	0,00115	0,00237	0,00121	0,00132

A17 - Simétrica concentrada n=500 400 comparações, 400 x e 400 y

	Categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49995	0,51106	0,55069	0,36473
	4	0,49995	0,47156	0,51048	0,44350
	5	0,49995	0,52577	0,56560	0,38604
	6	0,49995	0,48052	0,51970	0,45976
	7	0,49995	0,52722	0,56707	0,38660
	8	0,49995	0,48166	0,52087	0,45708
	9	0,49995	0,52881	0,56867	0,38928
	10	0,49995	0,48256	0,52180	0,45791
	11	0,49995	0,52869	0,56855	0,38584
Desvio Padrão	3	0,03390	0,04889	0,04974	0,04058
	4	0,03390	0,03676	0,03785	0,03505
	5	0,03390	0,04810	0,04871	0,04345
	6	0,03390	0,03641	0,03739	0,03527
	7	0,03390	0,04776	0,04830	0,04450
	8	0,03390	0,03618	0,03713	0,03620
	9	0,03390	0,04817	0,04869	0,04709
	10	0,03390	0,03637	0,03731	0,03741
	11	0,03390	0,04784	0,04837	0,04708
Erro Quadrático	3	0,45851	1,00248	2,01470	7,97577
	4	0,45851	0,86286	0,61559	1,76714
	5	0,45851	1,18879	2,66798	5,94829
	6	0,45851	0,68081	0,71302	1,14378
	7	0,45851	1,20648	2,73020	5,93401
	8	0,45851	0,65688	0,72422	1,25997
	9	0,45851	1,25786	2,83210	5,78817
	10	0,45851	0,64950	0,74548	1,26716
	11	0,45851	1,24223	2,81340	6,09740
Erro Quadrático Médio	3	0,00115	0,00251	0,00504	0,01994
	4	0,00115	0,00216	0,00154	0,00442
	5	0,00115	0,00297	0,00667	0,01487
	6	0,00115	0,00170	0,00178	0,00286
	7	0,00115	0,00302	0,00683	0,01484
	8	0,00115	0,00164	0,00181	0,00315
	9	0,00115	0,00314	0,00708	0,01447
	10	0,00115	0,00162	0,00186	0,00317
	11	0,00115	0,00311	0,00703	0,01524

A18 - Uniforme n=500 400 comparações, 400 x e 400 y

	Categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49995	0,45638	0,49461	0,35407
	4	0,49995	0,45288	0,49120	0,43805
	5	0,49995	0,45480	0,49319	0,45272
	6	0,49995	0,45500	0,49342	0,46085
	7	0,49995	0,45559	0,49404	0,46623
	8	0,49995	0,45662	0,49510	0,47002
	9	0,49995	0,45710	0,49559	0,47214
	10	0,49995	0,45743	0,49594	0,47431
	11	0,49995	0,45763	0,49616	0,47564
Desvio Padrão	3	0,03390	0,05166	0,05353	0,04100
	4	0,03390	0,03707	0,03840	0,03732
	5	0,03390	0,03648	0,03780	0,03678
	6	0,03390	0,03465	0,03589	0,03584
	7	0,03390	0,03378	0,03497	0,03532
	8	0,03390	0,03376	0,03495	0,03570
	9	0,03390	0,03431	0,03551	0,03601
	10	0,03390	0,03360	0,03478	0,03544
	11	0,03390	0,03273	0,03387	0,03542
Erro Quadrático	3	0,45851	1,82614	1,15492	9,18874
	4	0,45851	1,43623	0,61940	2,09120
	5	0,45851	1,34826	0,58878	1,43412
	6	0,45851	1,28912	0,53136	1,12569
	7	0,45851	1,24415	0,50209	0,95379
	8	0,45851	1,20772	0,49709	0,86796
	9	0,45851	1,20592	0,51093	0,82795
	10	0,45851	1,17535	0,48929	0,76508
	11	0,45851	1,14548	0,46361	0,73788
Erro Quadrático Médio	3	0,00115	0,00457	0,00289	0,02297
	4	0,00115	0,00359	0,00155	0,00523
	5	0,00115	0,00337	0,00147	0,00359
	6	0,00115	0,00322	0,00133	0,00281
	7	0,00115	0,00311	0,00126	0,00238
	8	0,00115	0,00302	0,00124	0,00217
	9	0,00115	0,00301	0,00128	0,00207
	10	0,00115	0,00294	0,00122	0,00191
	11	0,00115	0,00286	0,00116	0,00184

A19 - Assimétrico "suave" n=500 400 comparações, 400 x e 400 y

	categorias	Pearson	Mardia	Chambers	Pear. cat.
Média	3	0,49995	0,46018	0,49868	0,39902
	4	0,49995	0,46191	0,50049	0,42447
	5	0,49995	0,46427	0,50290	0,41754
	6	0,49995	0,45999	0,49853	0,43253
	7	0,49995	0,46035	0,49890	0,43120
	8	0,49995	0,45928	0,49780	0,43308
	9	0,49995	0,45802	0,49650	0,43342
	10	0,49995	0,45550	0,49390	0,44565
	11	0,49995	0,45505	0,49344	0,44636
Desvio Padrão	3	0,03390	0,04242	0,04382	0,03873
	4	0,03390	0,04017	0,04147	0,03892
	5	0,03390	0,04211	0,04346	0,04394
	6	0,03390	0,03813	0,03938	0,04091
	7	0,03390	0,03867	0,03995	0,04213
	8	0,03390	0,03863	0,03990	0,04340
	9	0,03390	0,03812	0,03942	0,04385
	10	0,03390	0,03752	0,03886	0,04120
	11	0,03390	0,03718	0,03852	0,04080
Erro Quadrático	3	0,45851	1,35206	0,76695	4,67745
	4	0,45851	1,22418	0,68628	2,88604
	5	0,45851	1,21836	0,75682	3,49043
	6	0,45851	1,22039	0,61967	2,48840
	7	0,45851	1,22554	0,63723	2,60177
	8	0,45851	1,25848	0,63727	2,54292
	9	0,45851	1,28454	0,62501	2,54029
	10	0,45851	1,35400	0,61743	1,85889
	11	0,45851	1,35972	0,60934	1,81512
Erro Quadrático Médio	3	0,00115	0,00338	0,00192	0,01169
	4	0,00115	0,00306	0,00172	0,00722
	5	0,00115	0,00305	0,00189	0,00873
	6	0,00115	0,00305	0,00155	0,00622
	7	0,00115	0,00306	0,00159	0,00650
	8	0,00115	0,00315	0,00159	0,00636
	9	0,00115	0,00321	0,00156	0,00635
	10	0,00115	0,00339	0,00154	0,00465
	11	0,00115	0,00340	0,00152	0,00454

A20 - assimétrico "concentrado" n=500 400 comparações, 400 x e 400 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49995	0,47144	0,51015	0,36117
	4	0,49995	0,47523	0,51409	0,37334
	5	0,49995	0,47681	0,51571	0,37634
	6	0,49995	0,47712	0,51603	0,38066
	7	0,49995	0,47663	0,51553	0,37716
	8	0,49995	0,47684	0,51574	0,37865
	9	0,49995	0,46752	0,50618	0,39634
	10	0,49995	0,45968	0,49815	0,41784
	11	0,49995	0,45578	0,49418	0,43762
Desvio Padrão	3	0,03390	0,05132	0,05287	0,04781
	4	0,03390	0,04856	0,04992	0,04724
	5	0,03390	0,04888	0,05023	0,05120
	6	0,03390	0,04855	0,04989	0,05155
	7	0,03390	0,04895	0,05032	0,05222
	8	0,03390	0,04903	0,05040	0,05298
	9	0,03390	0,04730	0,04884	0,04949
	10	0,03390	0,04264	0,04411	0,04532
	11	0,03390	0,03842	0,03977	0,04180
Erro Quadrático	3	0,45851	1,37713	1,15669	8,62136
	4	0,45851	1,18626	1,07371	7,30723
	5	0,45851	1,16830	1,10522	7,16318
	6	0,45851	1,15009	1,09584	6,75669
	7	0,45851	1,17441	1,10687	7,12413
	8	0,45851	1,17362	1,11282	7,01020
	9	0,45851	1,31447	0,96691	5,27587
	10	0,45851	1,37596	0,77767	3,51951
	11	0,45851	1,37105	0,64458	2,25325
Erro Quadrático Médio	3	0,00115	0,00344	0,00289	0,02155
	4	0,00115	0,00297	0,00268	0,01827
	5	0,00115	0,00292	0,00276	0,01791
	6	0,00115	0,00288	0,00274	0,01689
	7	0,00115	0,00294	0,00277	0,01781
	8	0,00115	0,00293	0,00278	0,01753
	9	0,00115	0,00329	0,00242	0,01319
	10	0,00115	0,00344	0,00194	0,00880
	11	0,00115	0,00343	0,00161	0,00563

A21 - Simétrico "suave" n=100 1000 comparações, 1000 x e 1000 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49407	0,46621	0,50372	0,40309
	4	0,49407	0,45849	0,49605	0,43919
	5	0,49407	0,46994	0,50787	0,45163
	6	0,49407	0,46082	0,49859	0,46514
	7	0,49407	0,46440	0,50232	0,47039
	8	0,49407	0,46094	0,49875	0,47701
	9	0,49407	0,45968	0,49746	0,47800
	10	0,49407	0,46544	0,50336	0,47634
	11	0,49407	0,45855	0,49633	0,48276
Desvio Padrão	3	0,07617	0,09527	0,09820	0,08470
	4	0,07617	0,08557	0,08851	0,08208
	5	0,07617	0,08480	0,08747	0,08038
	6	0,07617	0,08054	0,08327	0,07962
	7	0,07617	0,07892	0,08151	0,07830
	8	0,07617	0,07889	0,08162	0,07896
	9	0,07617	0,07836	0,08106	0,07813
	10	0,07617	0,07998	0,08257	0,07903
	11	0,07617	0,07691	0,07957	0,07747
Erro Quadrático	3	5,83110	10,20919	9,64777	16,55871
	4	5,83110	9,03774	7,84134	10,42838
	5	5,83110	8,08740	7,70538	8,79535
	6	5,83110	8,01563	6,92899	7,54821
	7	5,83110	7,48937	6,64334	7,00082
	8	5,83110	7,74278	6,65670	6,75704
	9	5,83110	7,76076	6,56994	6,58245
	10	5,83110	7,58406	6,82289	6,79861
	11	5,83110	7,62839	6,33782	6,29270
Erro Quadrático Médio	3	0,00583	0,01021	0,00965	0,01656
	4	0,00583	0,00904	0,00784	0,01043
	5	0,00583	0,00809	0,00771	0,00880
	6	0,00583	0,00802	0,00693	0,00755
	7	0,00583	0,00749	0,00664	0,00700
	8	0,00583	0,00774	0,00666	0,00676
	9	0,00583	0,00776	0,00657	0,00658
	10	0,00583	0,00758	0,00682	0,00680
	11	0,00583	0,00763	0,00634	0,00629

A22 - Simétrica concentrada n=100 1000 comparações, 1000 x e 1000 y

	Categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49407	0,51151	0,54948	0,36256
	4	0,49407	0,47064	0,50853	0,43926
	5	0,49407	0,52492	0,56304	0,38565
	6	0,49407	0,47910	0,51723	0,45729
	7	0,49407	0,52485	0,56302	0,38535
	8	0,49407	0,47951	0,51768	0,45418
	9	0,49407	0,52608	0,56426	0,38725
	10	0,49407	0,48019	0,51837	0,45476
	11	0,49407	0,52577	0,56395	0,38447
Desvio Padrão	3	0,07617	0,10918	0,11116	0,08917
	4	0,07617	0,08660	0,08924	0,08140
	5	0,07617	0,10886	0,11043	0,09794
	6	0,07617	0,08580	0,08820	0,08200
	7	0,07617	0,10776	0,10941	0,10044
	8	0,07617	0,08505	0,08745	0,08408
	9	0,07617	0,10774	0,10930	0,10366
	10	0,07617	0,08523	0,08761	0,08540
	11	0,07617	0,10750	0,10900	0,10586
Erro Quadrático	3	5,83110	12,04005	14,79221	26,83480
	4	5,83110	8,35452	8,02804	10,30847
	5	5,83110	12,45886	16,15669	22,65938
	6	5,83110	7,79180	8,06797	8,54045
	7	5,83110	12,21747	15,92970	23,22064
	8	5,83110	7,64679	7,95229	9,16108
	9	5,83110	12,27612	16,06497	23,44695
	10	5,83110	7,64936	8,00448	9,33293
	11	5,83110	12,20973	15,95808	24,54249
Erro Quadrático Médio	3	0,00583	0,01204	0,01479	0,02683
	4	0,00583	0,00835	0,00803	0,01031
	5	0,00583	0,01246	0,01616	0,02266
	6	0,00583	0,00779	0,00807	0,00854
	7	0,00583	0,01222	0,01593	0,02322
	8	0,00583	0,00765	0,00795	0,00916
	9	0,00583	0,01228	0,01606	0,02345
	10	0,00583	0,00765	0,00800	0,00933
	11	0,00583	0,01221	0,01596	0,02454

A23 - Uniforme n=100 1000 comparações, 1000 x e 1000 y

	Categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49407	0,45339	0,49046	0,39841
	4	0,49407	0,44810	0,48528	0,43208
	5	0,49407	0,44946	0,48682	0,44654
	6	0,49407	0,44936	0,48679	0,45368
	7	0,49407	0,45125	0,48877	0,46020
	8	0,49407	0,45085	0,48834	0,46305
	9	0,49407	0,45011	0,48764	0,46466
	10	0,49407	0,45136	0,48890	0,46762
	11	0,49407	0,45088	0,48844	0,46869
Desvio Padrão	3	0,07617	0,09693	0,10043	0,08609
	4	0,07617	0,08654	0,08977	0,08374
	5	0,07617	0,08179	0,08487	0,08133
	6	0,07617	0,07884	0,08179	0,07985
	7	0,07617	0,07796	0,08082	0,07929
	8	0,07617	0,07826	0,08114	0,08049
	9	0,07617	0,07604	0,07890	0,07851
	10	0,07617	0,07723	0,08009	0,07991
	11	0,07617	0,07542	0,07821	0,07889
Erro Quadrático	3	5,83110	11,55783	10,16806	17,72389
	4	5,83110	10,17609	8,26755	11,61860
	5	5,83110	9,23701	7,37008	9,46579
	6	5,83110	8,77387	6,85685	8,51578
	7	5,83110	8,44745	6,65232	7,86523
	8	5,83110	8,53421	6,71327	7,83832
	9	5,83110	8,26512	6,37132	7,40695
	10	5,83110	8,32378	6,53122	7,42732
	11	5,83110	8,09538	6,24500	7,19862
Erro Quadrático Médio	3	0,00583	0,01156	0,01017	0,01772
	4	0,00583	0,01018	0,00827	0,01162
	5	0,00583	0,00924	0,00737	0,00947
	6	0,00583	0,00877	0,00686	0,00852
	7	0,00583	0,00845	0,00665	0,00787
	8	0,00583	0,00853	0,00671	0,00784
	9	0,00583	0,00827	0,00637	0,00741
	10	0,00583	0,00832	0,00653	0,00743
	11	0,00583	0,00810	0,00624	0,00720

A24 - Assimétrico "suave" n=100 1000 comparações, 1000 x e 1000 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49407	0,45870	0,49578	0,39637
	4	0,49407	0,45441	0,49154	0,41757
	5	0,49407	0,45980	0,49701	0,41263
	6	0,49407	0,45361	0,49082	0,42718
	7	0,49407	0,45233	0,48949	0,42360
	8	0,49407	0,45177	0,48895	0,42635
	9	0,49407	0,45012	0,48726	0,42644
	10	0,49407	0,44748	0,48463	0,43921
	11	0,49407	0,44693	0,48407	0,43971
Desvio Padrão	3	0,07617	0,10203	0,10559	0,09224
	4	0,07617	0,09631	0,09981	0,09067
	5	0,07617	0,09904	0,10250	0,10257
	6	0,07617	0,09250	0,09594	0,09768
	7	0,07617	0,09263	0,09609	0,09864
	8	0,07617	0,09154	0,09496	0,10021
	9	0,07617	0,09096	0,09439	0,10078
	10	0,07617	0,08682	0,09011	0,09423
	11	0,07617	0,08713	0,09051	0,09438
Erro Quadrático	3	5,83110	12,10527	11,15545	19,23744
	4	5,83110	11,34392	10,02354	15,00865
	5	5,83110	11,41558	10,50472	18,14486
	6	5,83110	10,70051	9,27863	14,83585
	7	5,83110	10,84503	9,33500	15,55679
	8	5,83110	10,69726	9,13050	15,45619
	9	5,83110	10,75438	9,06382	15,55737
	10	5,83110	10,28798	8,34796	12,56560
	11	5,83110	10,39959	8,43707	12,53396
Erro Quadrático Médio	3	0,00583	0,01211	0,01116	0,01924
	4	0,00583	0,01134	0,01002	0,01501
	5	0,00583	0,01142	0,01050	0,01814
	6	0,00583	0,01070	0,00928	0,01484
	7	0,00583	0,01085	0,00933	0,01556
	8	0,00583	0,01070	0,00913	0,01546
	9	0,00583	0,01075	0,00906	0,01556
	10	0,00583	0,01029	0,00835	0,01257
	11	0,00583	0,01040	0,00844	0,01253

A25 - assimétrico “concentrado” n=100 1000 comparações, 1000 x e 1000 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,49407	0,46566	0,50212	0,35650
	4	0,49407	0,46489	0,50138	0,36036
	5	0,49407	0,47051	0,50721	0,37222
	6	0,49407	0,47043	0,50727	0,37393
	7	0,49407	0,46867	0,50540	0,37133
	8	0,49407	0,46929	0,50606	0,37195
	9	0,49407	0,45865	0,49546	0,38929
	10	0,49407	0,45303	0,49004	0,41148
	11	0,49407	0,44759	0,48466	0,43132
Desvio Padrão	3	0,07617	0,12619	0,13063	0,11115
	4	0,07617	0,12488	0,12935	0,11123
	5	0,07617	0,12302	0,12722	0,12121
	6	0,07617	0,11936	0,12336	0,11910
	7	0,07617	0,12086	0,12500	0,12207
	8	0,07617	0,12026	0,12426	0,12241
	9	0,07617	0,11041	0,11440	0,11396
	10	0,07617	0,09869	0,10236	0,10476
	11	0,07617	0,09034	0,09382	0,09657
Erro Quadrático	3	5,83110	17,08604	17,05120	32,93673
	4	5,83110	16,81129	16,71723	31,85858
	5	5,83110	15,98933	16,22115	31,00582
	6	5,83110	15,10662	15,25624	30,06492
	7	5,83110	15,57438	15,63843	31,44077
	8	5,83110	15,39121	15,46151	31,36671
	9	5,83110	13,88849	13,09379	25,22985
	10	5,83110	11,93733	10,56683	18,79978
	11	5,83110	10,89948	9,02944	14,03308
Erro Quadrático Médio	3	0,00583	0,01709	0,01705	0,03294
	4	0,00583	0,01681	0,01672	0,03186
	5	0,00583	0,01599	0,01622	0,03101
	6	0,00583	0,01511	0,01526	0,03006
	7	0,00583	0,01557	0,01564	0,03144
	8	0,00583	0,01539	0,01546	0,03137
	9	0,00583	0,01389	0,01309	0,02523
	10	0,00583	0,01194	0,01057	0,01880
	11	0,00583	0,01090	0,00903	0,01403

A26 - Simétrico "suave" n=500 1000 comparações, 1000 x e 1000 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,50021	0,46975	0,50854	0,40894
	4	0,50021	0,46190	0,50052	0,44503
	5	0,50021	0,47416	0,51316	0,45797
	6	0,50021	0,46468	0,50341	0,47096
	7	0,50021	0,47065	0,50957	0,47715
	8	0,50021	0,46525	0,50401	0,48258
	9	0,50021	0,46673	0,50555	0,48500
	10	0,50021	0,47028	0,50918	0,48068
	11	0,50021	0,46427	0,50301	0,48851
Desvio Padrão	3	0,03392	0,04259	0,04391	0,03898
	4	0,03392	0,03780	0,03902	0,03657
	5	0,03392	0,03654	0,03760	0,03534
	6	0,03392	0,03531	0,03643	0,03550
	7	0,03392	0,03451	0,03556	0,03468
	8	0,03392	0,03393	0,03499	0,03434
	9	0,03392	0,03391	0,03498	0,03441
	10	0,03392	0,03510	0,03616	0,03496
	11	0,03392	0,03369	0,03478	0,03451
Erro Quadrático	3	1,14940	2,72773	1,99902	9,81013
	4	1,14940	2,87911	1,52166	4,35740
	5	1,14940	2,00162	1,58526	3,01377
	6	1,14940	2,49300	1,33767	2,10212
	7	1,14940	2,05161	1,35471	1,72376
	8	1,14940	2,35735	1,23940	1,48145
	9	1,14940	2,25526	1,25319	1,40800
	10	1,14940	2,11451	1,39046	1,59383
	11	1,14940	2,41080	1,21745	1,32200
Erro Quadrático Médio	3	0,00115	0,00273	0,00200	0,00981
	4	0,00115	0,00288	0,00152	0,00436
	5	0,00115	0,00200	0,00159	0,00301
	6	0,00115	0,00249	0,00134	0,00210
	7	0,00115	0,00205	0,00135	0,00172
	8	0,00115	0,00236	0,00124	0,00148
	9	0,00115	0,00226	0,00125	0,00141
	10	0,00115	0,00211	0,00139	0,00159
	11	0,00115	0,00241	0,00122	0,00132

A27 - Simétrica concentrada n=500 1000 comparações, 1000 x e 1000 y

	Categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,50021	0,51247	0,55214	0,36580
	4	0,50021	0,47221	0,51114	0,44444
	5	0,50021	0,52703	0,56689	0,38706
	6	0,50021	0,48115	0,52034	0,46067
	7	0,50021	0,52836	0,56824	0,38752
	8	0,50021	0,48229	0,52150	0,45791
	9	0,50021	0,52979	0,56968	0,38959
	10	0,50021	0,48310	0,52233	0,45833
	11	0,50021	0,52983	0,56973	0,38667
Desvio Padrão	3	0,03392	0,04774	0,04853	0,04024
	4	0,03392	0,03775	0,03885	0,03600
	5	0,03392	0,04669	0,04721	0,04278
	6	0,03392	0,03699	0,03796	0,03538
	7	0,03392	0,04653	0,04701	0,04443
	8	0,03392	0,03703	0,03798	0,03653
	9	0,03392	0,04673	0,04719	0,04626
	10	0,03392	0,03713	0,03807	0,03732
	11	0,03392	0,04651	0,04698	0,04662
Erro Quadrático	3	1,14940	2,43256	5,07086	19,62727
	4	1,14940	2,19591	1,63228	4,38194
	5	1,14940	2,90819	6,70130	14,58406
	6	1,14940	1,72226	1,85283	2,79718
	7	1,14940	2,96694	6,86514	14,62391
	8	1,14940	1,68361	1,90326	3,10461
	9	1,14940	3,06850	7,07938	14,32841
	10	1,14940	1,66308	1,94661	3,12782
	11	1,14940	3,05117	7,06654	15,01580
Erro Quadrático Médio	3	0,00115	0,00243	0,00507	0,01963
	4	0,00115	0,00220	0,00163	0,00438
	5	0,00115	0,00291	0,00670	0,01458
	6	0,00115	0,00172	0,00185	0,00280
	7	0,00115	0,00297	0,00687	0,01462
	8	0,00115	0,00168	0,00190	0,00310
	9	0,00115	0,00307	0,00708	0,01433
	10	0,00115	0,00166	0,00195	0,00313
	11	0,00115	0,00305	0,00707	0,01502

A28 - Uniforme n=500 1000 comparações, 1000 x e 1000 y

	Categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,50021	0,46147	0,50000	0,40757
	4	0,50021	0,45230	0,49058	0,43802
	5	0,50021	0,45415	0,49253	0,45275
	6	0,50021	0,45541	0,49384	0,46143
	7	0,50021	0,45560	0,49404	0,46641
	8	0,50021	0,45658	0,49506	0,47005
	9	0,50021	0,45703	0,49553	0,47252
	10	0,50021	0,45713	0,49564	0,47423
	11	0,50021	0,45737	0,49588	0,47554
Desvio Padrão	3	0,03392	0,04307	0,04447	0,03903
	4	0,03392	0,03807	0,03944	0,03792
	5	0,03392	0,03581	0,03709	0,03622
	6	0,03392	0,03513	0,03636	0,03626
	7	0,03392	0,03438	0,03558	0,03571
	8	0,03392	0,03399	0,03517	0,03589
	9	0,03392	0,03404	0,03522	0,03563
	10	0,03392	0,03364	0,03480	0,03534
	11	0,03392	0,03324	0,03438	0,03535
Erro Quadrático	3	1,14940	3,33786	1,97563	10,06585
	4	1,14940	3,72304	1,64239	5,27849
	5	1,14940	3,38323	1,42982	3,54258
	6	1,14940	3,22134	1,35858	2,80133
	7	1,14940	3,15233	1,30027	2,40206
	8	1,14940	3,03911	1,25988	2,18355
	9	1,14940	3,00363	1,25888	2,02362
	10	1,14940	2,96795	1,22884	1,91134
	11	1,14940	2,92127	1,19788	1,84635
Erro Quadrático Médio	3	0,00115	0,00334	0,00198	0,01007
	4	0,00115	0,00372	0,00164	0,00528
	5	0,00115	0,00338	0,00143	0,00354
	6	0,00115	0,00322	0,00136	0,00280
	7	0,00115	0,00315	0,00130	0,00240
	8	0,00115	0,00304	0,00126	0,00218
	9	0,00115	0,00300	0,00126	0,00202
	10	0,00115	0,00297	0,00123	0,00191
	11	0,00115	0,00292	0,00120	0,00185

A29 - Assimétrico "suave" n=500 1000 comparações, 1000 x e 1000 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,50021	0,46110	0,49961	0,40058
	4	0,50021	0,46099	0,49953	0,42466
	5	0,50021	0,46485	0,50350	0,41932
	6	0,50021	0,45964	0,49816	0,43357
	7	0,50021	0,45912	0,49763	0,43036
	8	0,50021	0,45909	0,49759	0,43427
	9	0,50021	0,45778	0,49626	0,43455
	10	0,50021	0,45512	0,49351	0,44665
	11	0,50021	0,45469	0,49307	0,44693
Desvio Padrão	3	0,03392	0,04362	0,04506	0,03954
	4	0,03392	0,04082	0,04217	0,03861
	5	0,03392	0,04221	0,04353	0,04330
	6	0,03392	0,03890	0,04020	0,04087
	7	0,03392	0,03874	0,04004	0,04089
	8	0,03392	0,03897	0,04028	0,04251
	9	0,03392	0,03824	0,03955	0,04260
	10	0,03392	0,03718	0,03849	0,03998
	11	0,03392	0,03712	0,03843	0,03974
Erro Quadrático	3	1,14940	3,41377	2,02809	11,44523
	4	1,14940	3,18693	1,77710	7,16626
	5	1,14940	3,01537	1,90495	8,38148
	6	1,14940	3,14092	1,61740	6,08089
	7	1,14940	3,17075	1,60723	6,52030
	8	1,14940	3,19123	1,62638	6,12531
	9	1,14940	3,24299	1,57643	6,09647
	10	1,14940	3,39510	1,52186	4,44347
	11	1,14940	3,42981	1,52351	4,39354
Erro Quadrático Médio	3	0,00115	0,00341	0,00203	0,01145
	4	0,00115	0,00319	0,00178	0,00717
	5	0,00115	0,00302	0,00190	0,00838
	6	0,00115	0,00314	0,00162	0,00608
	7	0,00115	0,00317	0,00161	0,00652
	8	0,00115	0,00319	0,00163	0,00613
	9	0,00115	0,00324	0,00158	0,00610
	10	0,00115	0,00340	0,00152	0,00444
	11	0,00115	0,00343	0,00152	0,00439

A30 - assimétrico "concentrado" n=500 1000 comparações, 1000 x e 1000 y

	categorias	Pearson	Mardia	Chambers	Pear.cat.
Média	3	0,50021	0,47277	0,51152	0,36273
	4	0,50021	0,47144	0,51018	0,36457
	5	0,50021	0,47780	0,51672	0,37808
	6	0,50021	0,47752	0,51644	0,38156
	7	0,50021	0,47741	0,51632	0,37888
	8	0,50021	0,47747	0,51638	0,38004
	9	0,50021	0,46831	0,50701	0,39762
	10	0,50021	0,45999	0,49849	0,41915
	11	0,50021	0,45532	0,49370	0,43872
Desvio Padrão	3	0,03392	0,05142	0,05292	0,04706
	4	0,03392	0,04973	0,05120	0,04599
	5	0,03392	0,04943	0,05078	0,05122
	6	0,03392	0,04890	0,05024	0,05055
	7	0,03392	0,04915	0,05050	0,05153
	8	0,03392	0,04918	0,05054	0,05210
	9	0,03392	0,04604	0,04747	0,04821
	10	0,03392	0,04163	0,04301	0,04420
	11	0,03392	0,03835	0,03969	0,04076
Erro Quadrático	3	1,14940	3,38270	2,92998	21,05512
	4	1,14940	3,28573	2,72194	20,45461
	5	1,14940	2,93390	2,85579	17,48422
	6	1,14940	2,89416	2,79207	16,58079
	7	1,14940	2,92373	2,81403	17,32211
	8	1,14940	2,92418	2,82009	17,10135
	9	1,14940	3,12221	2,30062	12,80390
	10	1,14940	3,33264	1,85065	8,48771
	11	1,14940	3,46617	1,61334	5,41504
Erro Quadrático Médio	3	0,00115	0,00338	0,00293	0,02106
	4	0,00115	0,00329	0,00272	0,02045
	5	0,00115	0,00293	0,00286	0,01748
	6	0,00115	0,00289	0,00279	0,01658
	7	0,00115	0,00292	0,00281	0,01732
	8	0,00115	0,00292	0,00282	0,01710
	9	0,00115	0,00312	0,00230	0,01280
	10	0,00115	0,00333	0,00185	0,00849
	11	0,00115	0,00347	0,00161	0,00542

A31- Distribuição de X, Simétrica suave, versus diferentes formas distribucionais de Y para 5 categorias, n=500 e 200 repetições

	Forma marginal de Y	Pearson	Mardia	Chambers	Pear.cat.
Média	simétrico concentrado	0,50010	0,50818	0,54784	0,41457
	simétrico suave	0,50010	0,46839	0,50724	0,46145
	uniforme	0,50010	0,46124	0,49986	0,45743
	assimétrico concentrado	0,50010	0,47865	0,51765	0,37264
	assimétrico suave	0,50010	0,46738	0,50617	0,41786
Desvio Padrão	simétrico concentrado	0,03261	0,04473	0,04554	0,03735
	simétrico suave	0,03261	0,03530	0,03643	0,03434
	uniforme	0,03261	0,03608	0,03731	0,03584
	assimétrico concentrado	0,03261	0,04515	0,04630	0,03610
	assimétrico suave	0,03261	0,03762	0,03881	0,03301
Erro Quadrático	simétrico concentrado	0,21162	0,41163	0,87033	1,73711
	simétrico suave	0,21162	0,44781	0,27464	0,53190
	uniforme	0,21162	0,55954	0,27704	0,61791
	assimétrico concentrado	0,21162	0,49679	0,48887	3,50355
	assimétrico suave	0,21162	0,49447	0,30731	1,56608

A32- Distribuição de X, Simétrica concentrado, versus diferentes formas distribucionais de Y para 5 categorias, n=500 e 200 repetições

	Vs (y)	Pearson	Mardia	Chambers	Pear.cat.
Média	simétrico concentrado	0,50010	0,52738	0,56725	0,38836
	simétrico suave	0,50010	0,50542	0,54510	0,41119
	uniforme	0,50010	0,49871	0,53824	0,40311
	assimétrico concentrado	0,50010	0,51445	0,55413	0,33967
	assimétrico suave	0,50010	0,50919	0,54888	0,37530
Desvio Padrão	simétrico concentrado	0,03261	0,04750	0,04814	0,04286
	simétrico suave	0,03261	0,04015	0,04091	0,03405
	uniforme	0,03261	0,04204	0,04292	0,03457
	assimétrico concentrado	0,03261	0,04864	0,04933	0,03884
	assimétrico suave	0,03261	0,04343	0,04413	0,03619
Erro Quadrático	simétrico concentrado	0,21162	0,59887	1,36551	2,85809
	simétrico suave	0,21162	0,32672	0,73984	1,80810
	uniforme	0,21162	0,35198	0,65894	2,11525
	assimétrico concentrado	0,21162	0,51245	1,07024	5,44128
	assimétrico suave	0,21162	0,39218	0,86535	3,37087