



A COMPOSIÇÃO PROBABILÍSTICA DE PREFERÊNCIAS COM MEDIDAS DE DESIGUALDADE: CORRELAÇÕES COM OS PONTOS DE VISTA PROGRESSISTA E CONSERVADOR

Luiz Octávio Gavião
luiz.gaviao67@gmail.com

Annibal Parracho Sant'Anna
annibal.parracho@gmail.com

Gilson Brito Alves Lima
Universidade Federal Fluminense
Rua Passo da Pátria, 156, Bloco D, São Domingos, Niterói-RJ, CEP 24.220-240
glima@id.uff.br

RESUMO

A Composição Probabilística de Preferências (CPP) é um método probabilístico de apoio à decisão multicritério. As avaliações das alternativas segundo cada critério são substituídas por medidas de probabilidades conjuntas de cada alternativa ser superior ou inferior às demais. Essas probabilidades são compostas por diferentes pontos de vista, que refletem o contexto de decisão em relação a extremos progressista e conservador. Existem, entretanto, problemas de decisão não diretamente relacionados a esses extremos. A escolha de atletas multifuncionais na composição de equipes esportivas ilustra esse caso. Essa pesquisa propõe nova composição ao CPP, com base em medidas de desigualdade. A base de dados foi simulada por cem matrizes de decisão, com avaliações geradas aleatoriamente a partir de distribuição uniforme. Os resultados da proposta e os dos extremos progressista e conservador foram correlacionados com os índices de Kendall e Spearman. As correlações fracas e muito fracas confirmaram a relevância conceitual da proposta.

PALAVRAS CHAVE. CPP; Medidas de Desigualdade; Análise de Correlação.
Tópico: OA – Outras Aplicações; MP – Modelos Probabilísticos.

ABSTRACT

The Composition of Probabilistic Preferences (CPP) is a probabilistic multicriteria decision aid (MCDA) method. The alternative evaluations according to each criterion are replaced by probabilities for each alternative being higher or lower than the other. These probabilities are composed by different points of view, associated with the decision-making context. These points of view reflect a progressive extreme, focusing on maximizing gains in decision-making, or a conservative side, on minimizing losses. However, there are decision problems not directly related to the extremes of ratings. The recruitment of multifunctional athletes illustrates this case. A new composition to CPP is based on performance regularity. The database was simulated by a hundred decision matrices, with random ratings generated by uniform distribution. These data were modeled by CPP, considering the regularity composition and the progressive-conservative points of view. The weak correlation by the Kendall and Spearman indices confirmed the relevance of the proposal.

KEYWORDS.

Paper topics – OA; MP



1. Introdução

Um método de apoio à decisão multicritério (MCDA) agrega as avaliações sob os múltiplos critérios por algum algoritmo. No caso da Composição Probabilística de Preferências (CPP), as avaliações das alternativas segundo cada critério são substituídas por medidas de probabilidades conjuntas de cada alternativa ser superior ou inferior às demais. Em seguida, essas probabilidades são compostas por diferentes pontos de vista, associados ao contexto da tomada de decisão. A ordenação dos resultados dessas composições indica as preferências das alternativas para cada ponto de vista. As composições do CPP refletem as preferências em relação a dois extremos: o “progressista”, com foco na maximização de ganhos na tomada de decisão, enquanto o “conservador” se refere à minimização de perdas [Sant’Anna 2015].

Existem, entretanto, problemas de decisão não diretamente relacionados aos extremos das avaliações. A escolha de atletas de alta performance na composição de equipes esportivas, por exemplo, ilustra esse caso. Treinadores bem-sucedidos em diferentes modalidades esportivas revelam a importância de plantéis com jogadores versáteis e multifuncionais [Anderson & Sally 2013; Lewis 2004; Rezende 2006]. Essas características estão associadas ao desempenho regular nos diversos fundamentos da modalidade. Por vezes, jogadores da mesma posição se encontram em patamar técnico equivalente, porém treinadores e gestores precisam escolher um jogador para compor uma seleção nacional, para escalar na partida decisiva ou mesmo para sugerir a contratação ou dispensa do elenco. Nesse contexto, uma composição do CPP com base na regularidade de desempenho pode ser útil à tomada de decisão.

A modelagem do CPP com medidas de desigualdade é original, embora essas medidas tenham sido registradas com outros métodos MCDA [Ogryczak 2009]. De fato, essas medidas são normalmente exploradas em contexto socioeconômico, que se reflete em sua própria definição. Segundo [Cowell 2011], uma medida de desigualdade é uma representação escalar e numérica das diferenças de renda pessoal dos integrantes de uma população. Assim, uma adaptação conceitual é necessária para o uso das medidas de desigualdade com o CPP. Por analogia à definição, os valores das avaliações nos critérios podem substituir a “renda pessoal” e o conjunto de critérios a “população”. Dessa forma, é possível efetuar a composição das avaliações segundo os múltiplos critérios por medidas de desigualdade. Resta, então, a seguinte pergunta de pesquisa: a composição por medidas de desigualdade e as composições por pontos de vista apresentam resultados significativamente diferentes entre si?

Este artigo tem por objetivo discutir a correlação entre os resultados da composição por medidas de desigualdade com os resultados das composições progressista e conservadora, para as mesmas bases de dados. Foram simuladas cem matrizes de decisão, compostas por vinte alternativas e dez critérios. As avaliações no intervalo (0, 10) foram aleatoriamente geradas a partir de distribuição uniforme, com auxílio do software R [R-Core-Team 2016].

Os resultados indicam que as diferentes formas de composição produzem ordenações significativamente diferentes. As correlações, pelos índices ordinais de Kendall e Spearman, foram classificadas como fracas e muito fracas para quatro das cinco medidas de desigualdade utilizadas e moderada para a medida divergente. Esses graus de correlação demonstraram que os resultados da proposta não se refletem nas composições progressistas e conservadoras. Assim, conclui-se que o uso do CPP com medidas de desigualdade pode ser relevante ao apoio à decisão em determinados contextos.

O artigo foi elaborado em cinco Seções. A Seção 2 efetua uma revisão bibliográfica, com ênfase ao CPP, com as aplicações de interesse indicadas. A Seção 3 apresenta a metodologia proposta nessa pesquisa. A Seção 4 apresenta o desenvolvimento da análise. Por fim, a Seção 5 traz as considerações finais.



2. Revisão Bibliográfica

2.1 O método CPP

A própria denominação do CPP indica a abordagem probabilística ao problema MCDA. O CPP assume a ideia-chave da “aleatorização” das alternativas. Esse conceito deriva da imprecisão das medidas, da natureza aleatória do problema ou mesmo da incerteza acerca do contexto que envolve a tomada de decisão. Caso esses elementos não estivessem presentes, outros métodos determinísticos poderiam ser aplicados ao problema, com soluções satisfatórias [Sant’Anna 2015]. Porém, o tratamento probabilístico de um problema sob tais condições confere maior realismo à modelagem.

Nesse primeiro passo do método CPP, as avaliações exatas das alternativas são interpretadas como medidas de posição (como média, moda ou mediana) de um conjunto de dados. Essa transformação permite o tratamento estatístico das avaliações por funções de probabilidade. A escolha da função probabilística que melhor se ajusta aos dados pode advir de informações *a priori*, de aplicações bem-sucedidas a problemas similares, do ajuste dos dados a distribuições conhecidas por técnicas de “goodness-of-fit” ou, na ausência dessas informações, arbitrada ao contexto. De maneira geral, o uso de distribuições triangulares têm suprido as necessidades da modelagem assimétrica dos dados, com resultados satisfatórios, sendo assim utilizada nesta pesquisa [Sant’Anna 2015].

$$M_{ij} = \int_{D_{X_i}} \left[\prod F_{X_{-i}}(x_{-i}) \right] f_{X_i}(x_i) dx_i \quad (1)$$

$$m_{ij} = \int_{D_{X_i}} \left[\prod (1 - F_{X_{-i}}(x_{-i})) \right] f_{X_i}(x_i) dx_i \quad (2)$$

Após o processo de “aleatorização”, as probabilidades da *i*-ésima alternativa ser superior (M_{ij}) e inferior (m_{ij}) às demais, conjuntamente, são calculadas para cada *j*-ésimo critério. Nas equações (1) e (2), um vetor aleatório, “X”, simboliza as avaliações das alternativas do problema. As funções relativas à *i*-ésima alternativa são indexadas por “*i*” e as demais por “*-i*”. As notações F_X , f_X e D_{X_i} representam, respectivamente, a função distribuição cumulativa (cdf), a função densidade de probabilidade (pdf) e o domínio da variável aleatória “X” [Sant’Anna et al. 2012].

Para a ordenação das alternativas, as probabilidades “ M_{ij} ” e “ m_{ij} ” são compostas em escores de preferência global, para diferentes pontos de vista do decisor. As composições com base em “ M_{ij} ” são denominadas progressistas, por associar a decisão à maximização de ganhos. Nesse caso, os ganhos estão associados às probabilidades de maximizar as preferências. As composições com base em “ m_{ij} ” são denominadas conservadoras, por associar a decisão à minimização de perdas, representadas pelas probabilidades de minimizar as preferências. As composições que requerem a maximização “ M_{ij} ” e a minimização “ m_{ij} ” em todos os critérios são denominadas pessimistas, enquanto as otimistas são satisfeitas com a máxima ou mínima preferência em ao menos um critério. Dessa forma, é possível associar a tomada de decisão a quatro pontos de vista, referentes à interseção desses dois eixos progressista-conservador e otimista-pessimista.

$$PP_i = \prod_{j=1, \dots, n} M_{ij} \quad (3)$$

$$PO_i = 1 - \prod_{j=1, \dots, n} (1 - M_{ij}) \quad (4)$$

$$CP_i = \prod_{j=1, \dots, n} (1 - m_{ij}) \quad (5)$$

$$CO_i = 1 - \prod_{j=1, \dots, n} m_{ij} \quad (6)$$



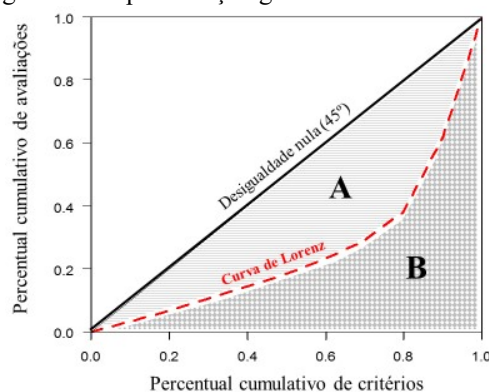
Os cálculos das composições podem ainda considerar a independência ou outras hipóteses de dependência das avaliações. Na hipótese mais simples de independência das avaliações, aplicada nessa pesquisa, o decisor objetiva selecionar a alternativa com a melhor performance segundo o conjunto de critérios. Essa composição, denominada Progressista-Pessimista (PP), é obtida com o produto dos “ M_{ij} ” da i -ésima alternativa na totalidade dos critérios, com a seleção do maior escore global, conforme a Equação (3). A composição Progressista-Otimista (PO), descrita na Equação (4), permite ordenar a alternativa com a melhor performance em ao menos um critério. A composição Conservadora-Pessimista (CP), descrita na Equação (5), permite ordenar a alternativa que não apresente a pior performance segundo o conjunto de critérios. Por fim, a composição Conservadora-Otimista (CO), conforme a Equação (6), permite ordenar a alternativa que não apresente a pior performance em ao menos um critério [Sant’Anna 2015].

2.2 Medidas de Desigualdade

Diversas medidas de desigualdade são aplicadas em pesquisas científicas. Em avaliações socioeconômicas com base na distribuição de renda populacional, por exemplo, o índice de Gini tem sido amplamente explorado [Gini 1921; Giorgi & Gigliarano 2016]. Outras medidas, porém, também podem ser úteis ao apoio à decisão. Nessa pesquisa foram usados os coeficientes de Ricci-Schutz, Atkinson, Kolm e Coeficiente de Variação, além do próprio índice de Gini. Dada a natureza exploratória do tema, a escolha desses índices priorizou a simplicidade de cálculo, proporcionada pelo pacote “ineq” do software R [R-Core-Team 2016].

As medidas de desigualdade requerem a padronização das unidades de avaliação. Da mesma forma que a moeda padroniza as medidas de renda, a normalização das avaliações é essencial para adaptar a medida de desigualdade ao MCDA. No caso específico do CPP, as probabilidades de maximizar e minimizar das avaliações já estão sob escala única, entre 0% e 100%, não sendo necessário procedimento adicional para a padronização dos dados.

Figura 1 – Representação gráfica do índice de Gini



Fonte: Adaptado de [Allison 1978]

A Figura 1 traduz o índice de Gini graficamente. As escalas do gráfico representam os percentuais cumulativos da quantidade de critérios e das avaliações. A diagonal identifica os pontos de desigualdade nula, enquanto a curva de Lorenz reflete a desigualdade da amostra de dados. Quanto maior a “folga” da curva de Lorenz, maior a desigualdade, indicando que uma quantidade considerável de critérios recebeu avaliações bem diferentes entre si. As regiões “A” e “B” são identificadas, respectivamente, pela área entre as curvas e pela área inferior à curva de Lorenz. O índice de Gini é definido pelo quociente entre “A” e “A+B”. Tendo em vista que “A+B” é igual a 0,5, é possível deduzir que o índice também equivale a “2A” ou “1-2B”. Considerando que “B” equivale à integral definida da curva de Lorenz, no intervalo (0, 1), o cálculo do índice de Gini pode ser generalizado à Equação (7) [Allison 1978].



$$G = 1 - 2 \int_0^1 L(X) dX \quad (7)$$

Sem a função que define a curva de Lorenz, o índice de Gini pode ser simplificado a partir da Equação (8), para uma amostra de “n” dados [Dixon et al. 1987]. O índice pode variar no intervalo [0, 1], em que “1” representa a máxima desigualdade entre os valores. Por exemplo, se todo componente de uma população perceber o mesmo salário, o índice de Gini para essa população é nulo.

$$G = \frac{1}{n} (n+1 - 2 \frac{\sum_{i=1}^n (n+1-i)y_i}{\sum_{i=1}^n y_i}) \quad (8)$$

Na Equação (8) e no que segue, a notação “ y_i ” representa os valores da amostra, sendo dispostos de forma crescente, de forma que $y_i \leq y_{i+1}$, para $i \in \{1, \dots, n\}$. Na abordagem probabilística, os valores utilizados no cálculo do índice de Gini e dos demais são as probabilidades “ M_{ij} ” e “ m_{ij} ”. Dessa forma, é possível ordenar as alternativas por desigualdade do desempenho probabilístico nos diferentes critérios, a partir do método CPP. A menor desigualdade aponta a alternativa de melhor desempenho quanto à regularidade.

O coeficiente de Ricci-Schutz, também denominado índice de Pietra ou índice Robin Hood, mede a proporção do rendimento total que teria de ser redistribuído dos mais ricos aos mais pobres, para anular a desigualdade socioeconômica [Su 2014]. O coeficiente de Ricci-Schutz equivale à máxima distância vertical entre a curva de Lorenz, que descreve a desigualdade de renda e a linha diagonal, que indica a distribuição de renda perfeita. A Equação (9), em que X representa as avaliações e μ o valor esperado de X , efetua o cálculo do coeficiente RS para $X \in (0, \infty)$.

$$RS = \frac{1}{2\mu} \int_{-\infty}^{\infty} |X - \mu| dF(X) \quad (9)$$

O coeficiente de Atkinson é uma medida da desigualdade de renda desenvolvido por [Atkinson 1970]. A medida é útil para determinar que extremo da distribuição contribuiu mais para a desigualdade observada. Nesse sentido, é possível alterar a variável ε da Equação (10), para ajustar o nível de "aversão à desigualdade" da população analisada. Por exemplo, para $\varepsilon = 0$, assume-se que a redistribuição de renda não apresenta utilidade social e, dessa forma, o coeficiente de Atkinson é nulo. Por outro lado, para $\varepsilon = \infty$, o coeficiente de Atkinson atinge o valor unitário, indicando que a redistribuição de renda é considerada absolutamente importante para a população. Para $\varepsilon = 1$, a Equação (11) deve ser aplicada ao cálculo do coeficiente de Atkinson, conforme apresentado em [Dahlby 1987].

$$A_\varepsilon = 1 - ((1/n) \sum_{i=1}^n (y_i / \mu)^{1-\varepsilon})^{1/1-\varepsilon} ; \varepsilon > 0, \varepsilon \neq 1 \quad (10)$$

$$A_1 = 1 - \exp((1/n) \sum_{i=1}^n \ln(y_i / \mu)) ; \varepsilon = 1 \quad (11)$$

O coeficiente de Kolm foi elaborado por [Kolm 1976]. O índice buscou corrigir uma distorção circunstancial observada no cálculo com o índice de Atkinson. Em decorrência de uma greve na França, em 1968, os salários foram igualmente reajustados em 13% para diferentes categorias de trabalhadores. Por não haver escalonamento percentual, o aumento nominal aos maiores salários foi novamente questionado, por ampliar a desigualdade social. Esse aspecto, não era suficiente para alterar o índice de desigualdade na equação de Atkinson, insensível aos



diferentes aumentos nominais. A inclusão do fator *alfa*, na Equação (12), permitiu ajustar o índice ao contexto. Se todas as rendas forem multiplicadas por uma constante, o índice de Kolm sofrerá um impacto que depende do valor de *alfa*, enquanto o índice de Atkinson permanece invariável. Por outro lado, se uma constante for adicionada a todas as rendas, é o índice de Kolm que permanece constante, enquanto muda o valor do índice de Atkinson.

$$K_{\alpha} = (1/\alpha) \ln\left(\frac{1}{n} \sum_{i=1}^n e^{\alpha(\mu - y_i)}\right); \alpha > 0 \quad (12)$$

O Coeficiente de Variação equivale ao quociente entre o desvio-padrão e a média da amostra, conforme a Equação (13). A divisão pela média é necessária para tornar o índice invariante à escala. Sua fórmula tem considerável simplicidade de cálculo, envolvendo medidas estatísticas usuais e sem a exigência de parâmetros adicionais. Entretanto, esse coeficiente não se limita no intervalo [0, 1], comum à maioria dos índices de desigualdade [Allison 1978].

$$V = \frac{\sigma}{\mu} \quad (13)$$

O pacote “ineq” do software R ainda oferece os índices de Entropia Generalizada e de Theil. Os testes iniciais revelaram que os parâmetros *default* do pacote geravam *rankings* iguais para os índices de Atkinson e de Entropia Generalizada, quando aplicados às cem matrizes de decisão. Além disso, o uso dos parâmetros “0” e “1” nas equações dos índices de Entropia Generalizada e de Theil, respectivamente, resultavam em medidas de desigualdade iguais para a mesma amostra de dados. Assim, dentre essas três medidas, optou-se por manter o índice de Atkinson nesta pesquisa.

2.3 Correlações ordinais

Os coeficientes de Kendall e Spearman estão entre as técnicas de correlação ordinal mais utilizadas em estatística [Gibbons & Chakraborti 2011; Hauke & Kossowski 2011]. Esses métodos não-paramétricos efetuam os cálculos a partir das posições que os valores das variáveis ocupam quando ordenados. Os resultados das correlações variam no intervalo [-1, 1], em que “-1” caracteriza a máxima correlação negativa e “1” a máxima positiva. O coeficiente nulo indica ausência de correlação entre os postos.

$$\tau = \frac{N_c - N_d}{\frac{n(n-1)}{2}} \quad (14)$$

$$\rho = \frac{\text{cov}(R_X, R_Y)}{\sigma_{R_X} \sigma_{R_Y}} \quad (15)$$

O coeficiente de Kendall foi elaborado por [Kendall 1938]. Por convenção, esse coeficiente é identificado com a letra grega τ (*tau*). Na Equação (14), N_c representa a quantidade de variáveis ordenadas dos dois métodos sob análise, cujas posições ordinais sejam concordantes, enquanto N_d indica a quantidade de pares discordantes. O denominador da Equação (14) indica a quantidade de possíveis combinações de pares ordenados.

O coeficiente de Spearman está definido na Equação (15), sendo representado pela letra grega ρ (*rho*). De fato, esse coeficiente utiliza a mesma fórmula do coeficiente de correlação de Pearson, porém com os valores originais das variáveis X e Y substituídos pelos seus postos, aqui representadas por R_X e R_Y . O coeficiente de Spearman é, assim, obtido a partir da divisão entre a

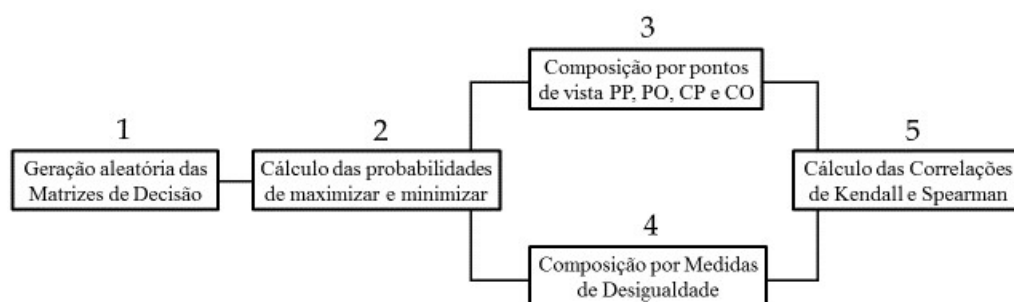


covariância das variáveis ordinais R_X e R_Y e o produto dos seus desvios-padrão [Spearman 1904].

3. Metodologia

A metodologia de análise aqui empregada envolve cinco etapas, conforme apresentado na Figura 2. De maneira geral, o modelo tem por finalidade verificar a correlação entre as ordenações pelos quatro pontos de vista do CPP e pelas medidas de desigualdade.

Figura 2 – Metodologia



Na Etapa 1 foram geradas as matrizes de decisão. A amostra foi gerada de forma aleatória. Foram geradas cem matrizes, compostas por vinte alternativas e dez critérios. Essas quantidades foram arbitradas para atender ao equilíbrio entre a significância estatística e o esforço computacional. Testes com 500 matrizes indicaram um colapso do software R. Por outro lado, uma quantidade reduzida de alternativas e critérios não proporcionaria a desejada variabilidade das ordenações. Isto poderia distorcer as conclusões acerca dos resultados das correlações. Os valores referentes às avaliações foram produzidos a partir da distribuição de probabilidade uniforme, com valores pertencentes ao intervalo [0,10]. O software R foi utilizado nas cinco etapas de cálculo [R-Core-Team 2016].

Na Etapa 2 foram calculadas as probabilidades de cada alternativa maximizar e minimizar sua preferência em relação às demais, conforme as Equações (1) e (2). Os dados foram ajustados a distribuições triangulares. Assim, cada avaliação da matriz de decisão foi transformada no valor modal de uma distribuição triangular, capaz de variar entre os valores mínimo e máximo da escala (i.e. intervalo [0,10]). As distribuições triangulares têm sido amplamente usadas na modelagem do CPP, produzindo resultados satisfatórios em diferentes aplicações [Sant'Anna 2015]. A Etapa 3 completa o CPP, por meio dos cálculos das composições de preferência por pontos de vista PP, PO, CP e CO, conforme as Equações (3) a (6).

Na Etapa 4, foram calculadas as medidas de desigualdade das alternativas, com base nas suas probabilidades de maximizar, obtidas na Etapa 2. Foram efetuados os cálculos para os índices de Gini, Ricci-Schutz (RS), Atkinson (ATK), Kolm (KOL), Variação (VAR). Na Etapa 5, os resultados das ordenações dos pontos de vista e das medidas de desigualdade foram correlacionados. Foram utilizados os índices de correlação de Spearman e Kendall, para comparar os vinte pares ordenados de rankings (i.e. PP-Gini, PP-RS, dentre outros).

4. Análise dos Resultados

Por necessidade de concisão do texto, os resultados das Etapas 1 a 4 se referem à primeira matriz de decisão gerada aleatoriamente. A Etapa 5, entretanto, reúne os resultados da amostra completa. A matriz de decisão #1 encontra-se descrita na Tabela 1, com valores arredondados a duas casas decimais.



Tabela 1 – Matriz de decisão #1

Alt	Crit 1	Crit 2	Crit 3	Crit 4	Crit 5	Crit 6	Crit 7	Crit 8	Crit 9	Crit 10
1	2,93	8,56	2,89	2,51	2,37	0,71	4,59	7,3	1,46	2,73
2	3,82	6,93	1,23	6,32	8,39	7,38	3,64	3,06	0,79	2,2
3	0,64	2,12	9,33	3,81	0,78	0,96	0,99	1,96	1,2	6,26
4	9,28	3,08	7,79	7,31	5,66	0,32	9,17	5,45	9,75	4,67
5	9,99	6,42	8,34	7,41	6,61	5,21	5,94	7,08	8,61	9,35
6	7,67	3,33	4,73	8,14	0,04	6,7	2,43	1,61	5,38	1,63
7	3,38	3,63	5,5	2,64	7,76	0,6	6,93	9,52	5,58	8,62
8	8,21	3,7	4,59	5,03	0,72	7,1	4,93	2,47	2,9	5,74
9	1,84	6,04	9,76	6,2	1,87	2,97	9,65	8,67	0,88	7,09
10	2,23	3,04	9,01	2,14	8,81	7,68	7,93	3,55	4,89	3,57
11	6,11	0,26	3,54	8,79	7,63	0,6	7,78	2,74	9,74	3,09
12	9,81	0,57	0,95	8,63	5,96	8,65	6,53	9,32	0,73	1,47
13	8,25	5,15	3,87	5,94	5,32	2,96	2,44	9,01	8,51	6,84
14	3,36	7,9	0	9,08	9,8	1,89	7,2	6,25	8,59	3,57
15	2,94	5,1	0,46	9,53	1,97	1,09	0,5	3,11	9,13	0,56
16	2,8	5,17	3,55	0,7	0,36	6,23	9,78	7,21	9,63	5,15
17	5,14	1,71	8,01	3,82	3,2	6,74	0,02	0,74	1,6	8,48
18	8,82	8,46	2,6	7,19	7,61	6,2	4,03	1,06	7,89	9,71
19	3,01	8,72	5,43	2,59	4,18	0,28	4,49	6,94	7,85	6,32
20	3,54	2,11	5,83	8,74	9,5	4,88	0,22	4,49	7,01	0,01

A Tabela 2 apresenta as probabilidades de cada alternativa maximizar sua preferência em relação às demais, em cada critério. No Critério 1, por exemplo, é possível verificar que a Alternativa 5 obteve a maior probabilidade de maximizar sua preferência em relação às demais. Esse posto da alternativa 5 corresponde ao valor esperado, tendo em vista que sua avaliação (i.e. 9,9) foi a maior no Critério 1, conforme apresentou a Tabela 1.

Tabela 2 – Probabilidades de maximizar referentes à Matriz de Decisão #1

Alt	Crit 1	Crit 2	Crit 3	Crit 4	Crit 5	Crit 6	Crit 7	Crit 8	Crit 9	Crit 10
1	1,62E-02	1,28E-01	1,94E-02	1,60E-02	1,72E-02	2,29E-02	2,47E-02	5,27E-02	1,10E-02	2,01E-02
2	1,86E-02	6,19E-02	1,56E-02	3,33E-02	8,45E-02	8,59E-02	2,09E-02	1,98E-02	1,02E-02	1,87E-02
3	1,22E-02	2,32E-02	1,51E-01	1,95E-02	1,42E-02	2,35E-02	1,47E-02	1,70E-02	1,07E-02	3,99E-02
4	1,26E-01	2,65E-02	6,52E-02	4,62E-02	3,08E-02	2,19E-02	1,33E-01	3,05E-02	1,50E-01	2,76E-02
5	1,90E-01	5,26E-02	8,57E-02	4,80E-02	3,99E-02	4,53E-02	3,33E-02	4,85E-02	6,94E-02	1,60E-01
6	5,13E-02	2,75E-02	2,63E-02	6,74E-02	1,32E-02	6,73E-02	1,75E-02	1,63E-02	2,06E-02	1,74E-02
7	1,74E-02	2,88E-02	3,10E-02	1,63E-02	6,15E-02	2,26E-02	4,46E-02	1,69E-01	2,16E-02	1,06E-01
8	6,67E-02	2,92E-02	2,56E-02	2,44E-02	1,41E-02	7,72E-02	2,64E-02	1,82E-02	1,33E-02	3,48E-02
9	1,40E-02	4,73E-02	1,91E-01	3,22E-02	1,62E-02	3,04E-02	1,75E-01	1,03E-01	1,03E-02	5,20E-02
10	1,47E-02	2,63E-02	1,25E-01	1,53E-02	1,07E-01	9,73E-02	6,71E-02	2,13E-02	1,86E-02	2,28E-02
11	3,00E-02	1,87E-02	2,14E-02	9,84E-02	5,80E-02	2,26E-02	6,26E-02	1,89E-02	1,49E-01	2,11E-02
12	1,73E-01	1,93E-02	1,51E-02	8,92E-02	3,32E-02	1,54E-01	3,92E-02	1,51E-01	1,02E-02	1,70E-02
13	6,81E-02	3,82E-02	2,25E-02	3,01E-02	2,85E-02	3,04E-02	1,75E-02	1,25E-01	6,52E-02	4,77E-02
14	1,73E-02	9,16E-02	1,37E-02	1,18E-01	1,90E-01	2,63E-02	4,92E-02	3,73E-02	6,85E-02	2,28E-02
15	1,63E-02	3,78E-02	1,43E-02	1,57E-01	1,64E-02	2,39E-02	1,39E-02	1,99E-02	9,80E-02	1,54E-02
16	1,59E-02	3,84E-02	2,14E-02	1,29E-02	1,36E-02	5,84E-02	1,88E-01	5,09E-02	1,38E-01	3,04E-02
17	2,38E-02	2,20E-02	7,23E-02	1,95E-02	1,94E-02	6,82E-02	1,32E-02	1,47E-02	1,12E-02	9,79E-02
18	9,45E-02	1,21E-01	1,86E-02	4,41E-02	5,75E-02	5,79E-02	2,23E-02	1,53E-02	4,65E-02	1,94E-01
19	1,64E-02	1,39E-01	3,05E-02	1,62E-02	2,28E-02	2,18E-02	2,42E-02	4,62E-02	4,56E-02	4,06E-02
20	1,78E-02	2,31E-02	3,36E-02	9,54E-02	1,61E-01	4,23E-02	1,35E-02	2,50E-02	3,24E-02	1,45E-02

A Tabela 3 traz as probabilidades de cada alternativa minimizar sua preferência em relação às demais, em cada critério. Por analogia ao caso da Alternativa 5, na tabela anterior, é esperado que a probabilidade de minimizar a sua preferência no Critério 1 seja a menor em



relação às demais. A probabilidade obtida pela Alternativa 5 (i.e. 1,70E-02) é, de fato, a menor nesse critério.

Tabela 3 – Probabilidades de minimizar referentes à Matriz de Decisão #1

Alt	Crit 1	Crit 2	Crit 3	Crit 4	Crit 5	Crit 6	Crit 7	Crit 8	Crit 9	Crit 10
1	6,05E-02	1,61E-02	4,56E-02	8,34E-02	5,00E-02	1,02E-01	2,62E-02	2,13E-02	9,49E-02	5,10E-02
2	4,57E-02	2,00E-02	1,01E-01	3,16E-02	1,35E-02	1,17E-02	3,33E-02	5,26E-02	1,41E-01	6,37E-02
3	1,83E-01	6,85E-02	1,36E-02	5,36E-02	1,20E-01	8,50E-02	1,09E-01	8,33E-02	1,10E-01	2,15E-02
4	1,83E-02	4,65E-02	1,63E-02	2,72E-02	2,02E-02	1,35E-01	1,29E-02	2,88E-02	1,39E-02	2,90E-02
5	1,70E-02	2,17E-02	1,52E-02	2,68E-02	1,73E-02	1,67E-02	2,01E-02	2,20E-02	1,58E-02	1,43E-02
6	2,22E-02	4,28E-02	2,73E-02	2,43E-02	1,87E-01	1,29E-02	5,09E-02	9,98E-02	2,56E-02	8,49E-02
7	5,20E-02	3,91E-02	2,33E-02	7,91E-02	1,47E-02	1,10E-01	1,71E-02	1,63E-02	2,46E-02	1,55E-02
8	2,07E-02	3,83E-02	2,81E-02	4,00E-02	1,25E-01	1,22E-02	2,43E-02	6,59E-02	4,87E-02	2,35E-02
9	9,70E-02	2,31E-02	1,30E-02	3,22E-02	6,35E-02	2,98E-02	1,22E-02	1,79E-02	1,34E-01	1,89E-02
10	8,04E-02	4,71E-02	1,41E-02	9,79E-02	1,29E-02	1,13E-02	1,49E-02	4,50E-02	2,82E-02	3,84E-02
11	2,80E-02	1,91E-01	3,69E-02	2,25E-02	1,49E-02	1,10E-01	1,52E-02	5,91E-02	1,39E-02	4,47E-02
12	1,73E-02	1,61E-01	1,19E-01	2,29E-02	1,92E-02	9,97E-03	1,82E-02	1,66E-02	1,46E-01	1,43E-02
13	2,06E-02	2,72E-02	3,36E-02	3,37E-02	2,16E-02	2,99E-02	5,07E-02	1,72E-02	1,60E-02	1,96E-02
14	5,23E-02	1,75E-02	2,04E-01	2,17E-02	1,16E-02	4,75E-02	1,65E-02	2,50E-02	1,58E-02	3,84E-02
15	6,03E-02	2,74E-02	1,59E-01	2,07E-02	6,03E-02	7,75E-02	1,46E-01	5,17E-02	1,49E-02	1,57E-01
16	6,35E-02	2,71E-02	3,67E-02	1,98E-01	1,56E-01	1,39E-02	1,21E-02	2,16E-02	1,41E-02	2,62E-02
17	3,35E-02	8,43E-02	1,59E-02	5,35E-02	3,66E-02	1,28E-02	1,92E-01	1,62E-01	8,78E-02	1,57E-02
18	1,92E-02	1,63E-02	5,10E-02	2,76E-02	1,49E-02	1,40E-02	3,00E-02	1,36E-01	1,73E-02	1,37E-02
19	5,88E-02	1,58E-02	2,36E-02	8,07E-02	2,77E-02	1,39E-01	2,68E-02	2,25E-02	1,74E-02	2,13E-02
20	4,95E-02	6,88E-02	2,20E-02	2,26E-02	1,19E-02	1,79E-02	1,72E-01	3,52E-02	1,95E-02	2,11E-01

A Tabela 4 apresenta os resultados das composições dos pontos de vista e dos índices de regularidade. Os resultados foram obtidos a partir das Equações (3) a (6), para o CPP e das Equações (7) a (13), para os índices de regularidade. As probabilidades de maximizar as preferências das alternativas, descritas na Tabela 2, foram utilizadas para os cálculos dos índices de regularidade.

Tabela 4 – Resultados da Matriz de Decisão #1

Alt	CPP (Pontos de Vista)				Medidas de Desigualdade				
	PP	PO	CP	CO	GINI	RS	ATK	KOL	VAR
1	7,3E-17	2,9E-01	5,6E-01	0,9999999999999949	0,4165	0,3499	0,1530	0,0005	1,0189
2	3,4E-16	3,2E-01	5,8E-01	0,9999999999999993	0,3857	0,3289	0,1235	0,0004	0,7505
3	3,0E-17	2,9E-01	4,1E-01	0,9999999999997993	0,4700	0,3860	0,1966	0,0008	1,2359
4	1,1E-13	5,0E-01	7,0E-01	1,0000000000000000	0,3812	0,3214	0,1208	0,0011	0,7284
5	1,3E-12	5,6E-01	8,3E-01	1,0000000000000000	0,3264	0,2639	0,0872	0,0013	0,6645
6	2,3E-16	2,8E-01	5,4E-01	0,999999999999983	0,3235	0,2727	0,0852	0,0002	0,6218
7	6,1E-15	4,2E-01	6,7E-01	0,999999999999999	0,4359	0,3484	0,1510	0,0011	0,9035
8	3,0E-16	2,9E-01	6,4E-01	0,999999999999998	0,3156	0,2417	0,0809	0,0002	0,6242
9	1,9E-14	5,1E-01	6,3E-01	0,999999999999999	0,4868	0,3985	0,1965	0,0020	0,9438
10	4,7E-15	4,2E-01	6,7E-01	0,999999999999999	0,4213	0,3695	0,1511	0,0008	0,7997
11	5,7E-15	4,1E-01	5,7E-01	0,999999999999996	0,4101	0,3350	0,1358	0,0008	0,8256
12	2,3E-14	5,3E-01	5,1E-01	0,999999999999994	0,4736	0,4086	0,1994	0,0019	0,8883
13	1,0E-14	3,9E-01	7,6E-01	1,0000000000000000	0,3239	0,2469	0,0836	0,0005	0,6462
14	3,7E-14	4,9E-01	6,2E-01	0,999999999999999	0,4376	0,3379	0,1520	0,0014	0,8395
15	2,3E-16	3,5E-01	4,4E-01	0,999999999999549	0,4875	0,4177	0,2009	0,0010	1,1044
16	5,4E-15	4,5E-01	5,4E-01	0,999999999999997	0,4850	0,3767	0,1895	0,0015	0,9888
17	2,1E-16	3,1E-01	4,8E-01	0,999999999999951	0,4076	0,3580	0,1384	0,0004	0,8123
18	9,6E-14	5,1E-01	7,0E-01	1,0000000000000000	0,4076	0,3096	0,1344	0,0014	0,7864
19	1,2E-15	3,4E-01	6,4E-01	0,999999999999998	0,3697	0,2727	0,1169	0,0006	0,8565
20	1,4E-15	3,8E-01	5,1E-01	0,999999999999991	0,4484	0,3594	0,1632	0,0010	0,9720



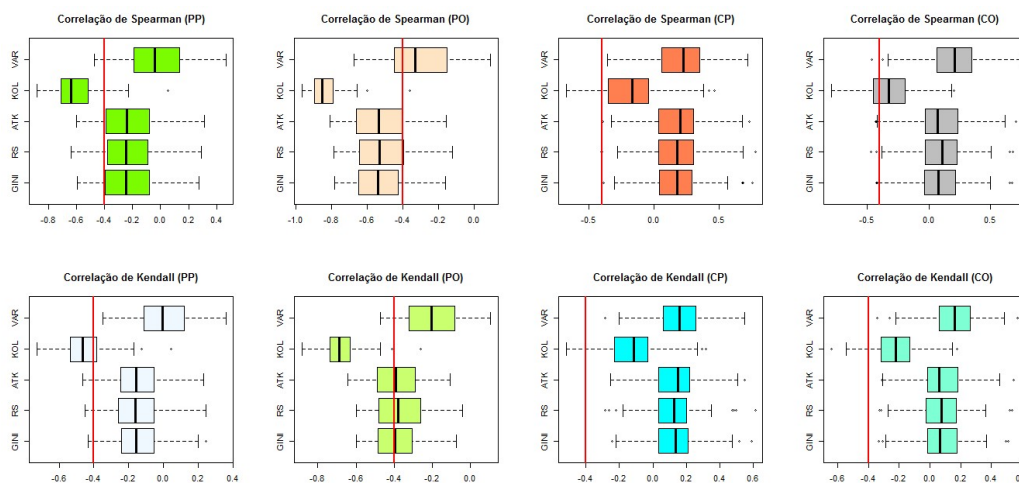
A Tabela 5 ordena os resultados das Etapas 1 a 4, descritos na tabela anterior. Esses rankings são referentes à matriz de decisão #1.

Tabela 5 – Rankings da matriz de decisão #1

Alt	CPP (Pontos de Vista)				Medidas de Desigualdade				
	PP	PO	CP	CO	GINI	RS	ATK	KOL	VAR
1	19	18	13	18	11	12	14	6	18
2	14	15	11	14	7	8	7	3	6
3	20	17	20	20	16	17	18	8	20
4	2	5	4	3	6	7	6	14	5
5	1	1	1	1	4	3	4	15	4
6	16	20	15	16	2	5	3	1	1
7	8	8	6	7	13	11	11	13	14
8	15	19	7	10	1	1	1	2	2
9	6	3	9	8	19	18	17	20	15
10	11	9	5	5,5	12	15	12	10	8
11	9	10	12	12	10	9	9	9	10
12	5	2	16	13	17	19	19	19	13
13	7	11	2	3	3	2	2	5	3
14	4	6	10	5,5	14	10	13	17	11
15	17	13	19	19	20	20	20	12	19
16	10	7	14	11	18	16	16	18	17
17	18	16	18	17	9	13	10	4	9
18	3	4	3	3	8	6	8	16	7
19	13	14	8	9	5	4	5	7	12
20	12	12	17	15	15	14	15	11	16

A Figura 4 traz os resultados da Etapa 5, para a amostra completa. Os resultados estão descritos em oito quadros. A primeira linha de quadros descreve os resultados da correlação de Spearman e a segunda linha os de Kendall. Em cada quadro, os gráficos *boxplot* foram elaborados com base nos dados dos cem coeficientes de correlação. A caixa do *boxplot* é delimitada pelos três quartis da amostra, enquanto as barras verticais foram arbitradas por [Tukey 1970], equivalendo ao intervalo entre o 1º e o 3º quartis, multiplicado por um fator (1,5, seguindo a recomendação de [Tukey 1970]). As correlações foram calculadas entre o ponto de vista indicado no título de cada quadro e as seis medidas de desigualdade. Cada quadro tem um eixo vertical na abscissa “-0,4”, que limita a região entre as correlações fracas (inferiores à referência) e moderadas (superiores à referência), segundo [Evans & Over 2013]. Esse eixo tem a finalidade de facilitar a comparação relativa dos *boxplot*.

Figura 4 – Correlações entre Pontos de Vista e Medidas de Desigualdade da Amostra



Em geral, as correlações entre os resultados das ordenações pelos pontos de vista e pelas medidas de desigualdade se encontram entre os limites das regiões de correlação fraca (i.e. intervalos entre 0,4 e 0,2 ou entre -0,4 e -0,2) e muito fraca (i.e., entre -0,2 e 0,2). A exceção



ocorreu no cálculo do coeficiente de Spearman para o ponto de vista PO (i.e. quadro da primeira linha e segunda coluna), em que a maioria dos gráficos se concentrou na região de correlações moderadas (valores absolutos entre 0,4 e 0,6). Entretanto, os resultados desse ponto de vista, quando aplicado o coeficiente de Kendall, se mantiveram no limite entre as regiões fraca e moderada. Esse aspecto pode ser constatado no quadro da segunda linha e segunda coluna da Figura 2.

Os resultados da aplicação do índice de Kolm também merecem destaque. Por uma questão matemática, seu algoritmo produziu resultados distorcidos em relação aos demais índices. Além disso, as correlações entre os pontos de vista e o índice de Kolm apresentaram variações entre graus “muito fracos” até “fortes”, conforme indicam os quadros da segunda coluna. O contraste das correlações do índice de Kolm, com as das demais medidas de desigualdade, foi evidente. Dessa forma, as aplicações do CPP com o índice de Kolm podem apresentar resultados diferentes das demais medidas, com maior variância e maior correlação com os resultados dos pontos de vista. Assim, as demais medidas de desigualdade devem ser priorizadas ao uso no CPP.

A diferença conceitual entre as composições por pontos de vista e por medidas de desigualdade se confirmou quantitativamente. As composições são calculadas com base em referências extremas, representadas pelas probabilidades de maximizar e minimizar as preferências. Por outro lado, as medidas de desigualdade se fundamentam no equilíbrio entre os dados. As correlações fracas e muito fracas foram significativas para indicar que as medidas de desigualdade não são redundantes com os resultados das composições por pontos de vista. Em outras palavras, os resultados das medidas de desigualdade contêm informação relevante para contextos específicos do apoio à decisão, ainda não considerada nas composições originais do CPP.

5. Considerações Finais

O artigo explorou medidas de desigualdade para adicionar uma nova forma de composição ao método CPP. A relação de preferência das alternativas no método CPP é elaborada por meio dos pontos de vista associados à tomada de decisão. Os pontos de vista estão associados a posições extremas, diferentemente das medidas de desigualdade. Por um lado, o ponto de vista progressista tem por foco a fronteira de máximo desempenho. Por outro lado, o ponto de vista conservador tem por foco o afastamento da fronteira de mínimo desempenho. Entretanto, determinados contextos de problemas MCDA podem indicar a necessidade de avaliar a regularidade das alternativas frente ao conjunto de critérios.

Os resultados das composições com seis medidas de desigualdade foram comparados aos dos pontos de vista PP, PO, CP e CO. As ordenações foram correlacionadas por índices de Kendall e Spearman, indicando uma fraca correlação aos índices de Gini, Ricci-Schutz, Atkinson e Coeficiente de Variação. Apenas o índice de Kolm apresentou correlações moderadas com os pontos de vista. A fraca correlação correspondeu à diferença conceitual das abordagens. Isto indica que, em problemas MCDA que requeiram uma análise de regularidade de desempenho das alternativas, é possível aplicar a composição pela desigualdade com respaldo estatístico.

A pesquisa pode ser estendida em estudos futuros. Existem outras medidas de desigualdade que não foram exploradas no artigo. A amostra de cem matrizes de decisão pode ser novamente gerada ou mesmo ampliada, de forma a comparar os gráficos das correlações. Por fim, estudos em ciência do esporte e administração de recursos humanos, que envolvam a análise de regularidade de atletas ou potenciais contratados, podem ser replicados e comparados à luz da metodologia proposta.

Referências

Afonso, H., LaFleur, M., & Alarcón, D. (2015). *Inequality Measurement: Development Issues*



- No. 2. Disponível em: <http://www.un.org/en/development/desa/policy/wess/wess_dev_issues/dsp_policy_02.pdf>. Acesso em 11 fev 2017.
- Allison, P. D. (1978). Measures of inequality. *American Sociological Review*, 865–880.
- Anderson, C., & Sally, D. (2013). Os números do jogo. Por que tudo o que você sabe sobre futebol está errado. *Editora Paralela. São Paulo*.
- Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2(3), 244–263.
- Cowell, F. (2011). *Measuring inequality*. Oxford University Press.
- Dahlby, B. G. (1987). Interpreting inequality measures in a Harsanyi framework. *Theory and Decision*, 22(3), 187–202.
- Dixon, P. M., Weiner, J., Mitchell-Olds, T., & Woodley, R. (1987). Bootstrapping the Gini coefficient of inequality. *Ecology*, 68(5), 1548–1551.
- Evans, J. S. B. T., & Over, D. E. (2013). *Rationality and reasoning*. Psychology Press.
- Gibbons, J. D., & Chakraborti, S. (2011). *Nonparametric statistical inference*. Springer.
- Gini, C. (1921). Measurement of inequality of incomes. *The Economic Journal*, 31(121), 124–126.
- Giorgi, G. M., & Gigliarano, C. (2016). The Gini Concentration Index: a review of the inference literature. *Journal of Economic Surveys*.
- Hauke, J., & Kossowski, T. (2011). Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae*, 30(2), 87–93.
- Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 30(1/2), 81–93.
- Kolm, S.-C. (1976). Unequal inequalities. I. *Journal of Economic Theory*, 12(3), 416–442.
- Lewis, M. (2004). *Moneyball: The art of winning an unfair game*. WW Norton & Company.
- Ogryczak, W. (2009). Inequality measures and equitable locations. *Annals of Operations Research*, 167(1), 61–86.
- R-Core-Team. (2016). R: A language and environment for statistical computing. [Http://www. R-Project. Org](http://www.R-Project.Org). Vienna, Austria.
- Rezende, B. R. de. (2006). *Transformando suor em ouro*. Rio de Janeiro: Sextante.
- Sant'Anna, A. P. (2015). *Probabilistic Composition of Preferences, Theory and Applications*. NY: Springer.
- Sant'Anna, A. P., Gomes, L. F. A. M., Costa, F. F. da, Rangel, L. A. D., Faria, M. J. da S., Ferreira, R. G., ... Senna, V. de. (2012). Análise multicritério baseada em probabilidades de preferência. In V. F. de Oliveira, V. Cavenaghi, & F. S. Másculo (Eds.), *Tópicos emergentes e desafios metodológicos em Engenharia de Produção: casos, experiências e proposições - Volume V* (p. 258). Rio de Janeiro: ABEPRO.
- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101.
- Su, Y. (2014). *Jackknife Empirical Likelihood Inference For The Pietra Ratio [Master Thesis]*. Georgia State University. Disponível em http://scholarworks.gsu.edu/math_theses/140/. Acesso em 10 fev 2017.
- Tukey, J. W. (1970). *Exploratory Data Analysis*. Addison-Wesley.