

Cálculo do Índice de Similaridade de Krugman para as regiões intermediárias do país, com o emprego dividido em setores segundo a classificação da CNAE 95

Luís Zucchi de Souza
Victor Ramos de Oliveira

NEREUS - USP
2023

1. Introdução

No presente relatório, iremos explicar a metodologia empregada para o cálculo do Índice de Similaridade de Krugman para o emprego classificado por setores, segundo a categorização da CNAE 95. Este índice é uma medida que busca quantificar a semelhança entre duas regiões de uma economia com base em suas estruturas setoriais. Isso é feito através da comparação das participações relativas de cada setor econômico nas duas regiões. Além disso, o índice varia negativamente com o grau de similaridade. Dessa forma, quanto mais similares setorialmente forem duas regiões, menor será o índice.

2. Metodologia

A partir da base de dados da RAIS (Relação Anual de Informações Sociais), fornecida pelo Ministério do Trabalho e Emprego, obtivemos os valores anuais do nível de emprego para cada município de 2002 a 2021, segmentado por setor da economia segundo classificação por Seção da CNAE 95, sendo o nível de emprego calculado como a soma de trabalhadores com vínculo ativo em 31 de dezembro de cada ano, para cada setor.

Os 17 setores da classificação da CNAE 95 por Seção são os seguintes:

- 1 - Agricultura, pecuária, silvicultura e exploração florestal;
- 2 - Pesca;
- 3 - Indústrias extrativas;
- 4 - Indústrias de transformação;
- 5 - Produção e distribuição de eletricidade, gás e água;
- 6 - Construção;
- 7 - Comércio, reparação de veículos automotores, objetos pessoais e domésticos;
- 8 - Alojamento e alimentação;
- 9 - Transporte, armazenagem e comunicações;
- 10 - Intermediação financeira, seguros, previdência complementar e serviços relacionados;
- 11 - Atividades imobiliárias, aluguéis e serviços prestados às empresas;
- 12 - Administração pública, defesa e seguridade social;
- 13 - Educação;
- 14 - Saúde e serviços sociais;
- 15 - Outros serviços coletivos, sociais e pessoais;
- 16 - Serviços domésticos;
- 17 - Organismos internacionais e outras instituições extraterritoriais.

Após isso, agregamos os dados obtidos por Região Geográfica Intermediária (RGInt), uma classificação do Instituto Brasileiro de Geografia e Estatística (IBGE) que divide o país em 133 regiões. Em seguida, calculamos os percentuais anuais do emprego para em cada setor para cada RGInt.

Finalmente, calculamos o Índice de Similaridade de Krugman (KSI) anualmente para cada região intermediária da seguinte forma:

$$KSI_{A,B,t} = \sum_{i=1}^S \left| \frac{E_{a\ i,t}}{E_{a\ t}} - \frac{E_{b\ i,t}}{E_{b\ t}} \right|$$

Isto é, o Índice de Similaridade de Krugman entre as regiões A e B no ano t ($KSI_{A,B,t}$) é dado pela soma em todos os setores do módulo da diferença entre a participação relativa do emprego do setor i na região A ($E_{a\ i,t} / E_{a\ t}$) e da participação relativa do emprego do setor i na região B ($E_{b\ i,t} / E_{b\ t}$).

Pode-se perceber facilmente que o valor do índice varia entre 0 e 2. Num caso extremo, em que as duas regiões possuem a mesma estrutura setorial com bases nas participações relativas o índice será 0. Com efeito, o índice de uma região com ela mesma será zero. Num outro caso extremo, se duas regiões são completamente diferentes setorialmente (uma delas tem 100% do emprego no setor i e a outra tem 100% do emprego no setor j) o módulo da diferença será 1 para os setores i e j e será 0 caso contrário, resultando em um índice com valor 2. Os resultados estão apresentados em formato de matriz para cada ano entre 2002 e 2021.