

FLS 5028
Métodos Quantitativos e Técnicas de Pesquisa em Ciência Política

FLP0406
Métodos e Técnicas de Pesquisa em Ciência Política
1º semestre / 2016

Glauco Peres da Silva

LISTA DE EXERCÍCIOS 09

Data de entrega: 30/05/2016 (noturno) e 01/06/2016 (vespertino).

Exercício 01 (3 pontos)

Assinale os enunciados abaixo com (V) verdadeiro ou (F) falso. Em caso de falso, indique o erro e o corrija, justificando sua resposta.

- () I. O modelo de regressão linear é determinista: cada valor de x corresponde a um único valor de y .
- () II. A ocorrência de valores *outliers* em uma distribuição afeta a equação da reta de um modelo linear, mas não a correlação entre as variáveis.
- () III. Um modelo de regressão linear pode ser aplicado com sucesso ainda que seja violado o princípio da linearidade, desde que haja aleatorização.
- () IV. O coeficiente r^2 informa sobre os ganhos preditivos adquiridos quando utilizamos a média de y para prever os valores das observações de y . Quanto maior o coeficiente r^2 , mais seguros ficamos de que o conhecimento de x não nos fornece melhores meios para prever y , em comparação à média de y .
- () V. As estimativas dos parâmetros populacionais α e β devem ser estimativas por ponto. Em outras palavras, em regressões lineares, a construção de intervalos de confiança para os parâmetros mencionados não é possível.

Exercício 02 (4 pontos)

Parte da literatura brasileira (AMES, 2003; AMORIM NETO; SANTOS, 2003; PEREIRA & MUELLER, 2002, 2003) argumenta que as emendas individuais orçamentárias – instrumento à disposição dos parlamentares para a alteração do orçamento – são essenciais para a garantia da estabilidade na relação entre o Executivo e o Legislativo. Como o orçamento no Brasil é autorizativo e não impositivo, ou seja, o Executivo é quem decide quais e como os recursos serão gastos, esses autores argumentam que o presidente utilizaria

dessas emendas como uma ferramenta para garantir que sua agenda de políticas seja aprovada. Nesse argumento, os parlamentares somente votariam de acordo com o governo mediante a execução de suas emendas individuais. Figueiredo e Limongi (2008), por outro lado, mostraram empiricamente que a associação entre a liberação das emendas individuais orçamentárias e o apoio ao governo não é tão forte como os autores supracitados argumentam.

Para resolver o exercício, utilize o banco de dados XXX e seu codebook, disponíveis no Moodle.

Mínimo de 3 e máximo de 10 linhas por item (a, b, c, d).

a-) Um primeiro passo para começar a investigar a relação entre duas variáveis é através da representação gráfica. Pensando o problema acima, responda:

I. Quais variáveis você escolheria para testar? Dado o enunciado do exercício, qual é a dependente e qual é a independente? De que tipo elas são?

II. Assumindo a perspectiva de Ames (2003), Santos (2003) e Pereira e Mueller (2002; 2003), o que você esperaria encontrar? Justifique.

III. Qual o melhor gráfico para ver essa relação? Execute-o e interprete-o.

b-) A **covariância** é “um modo estatístico para resumir um padrão de associação geral (ou falta dele) entre duas variáveis contínuas” (Kellstendt&Whitten: 2015, p.183). Seu cálculo se dá pela fórmula:

$$COV_{xy} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{n}$$

Resultados positivos descrevem uma relação positiva e resultados negativos uma relação negativa. O que o resultado nos diz sobre a relação entre as variáveis? Ele é condizente com o esperado a partir do gráfico?

c-) Para saber a força dessa relação precisamos dar mais um passo na nossa análise e olhar o **coeficiente de correlação (r)**. Defina esse coeficiente e sua utilidade na análise, em seguida, calcule-o e interprete-o com base na teoria.

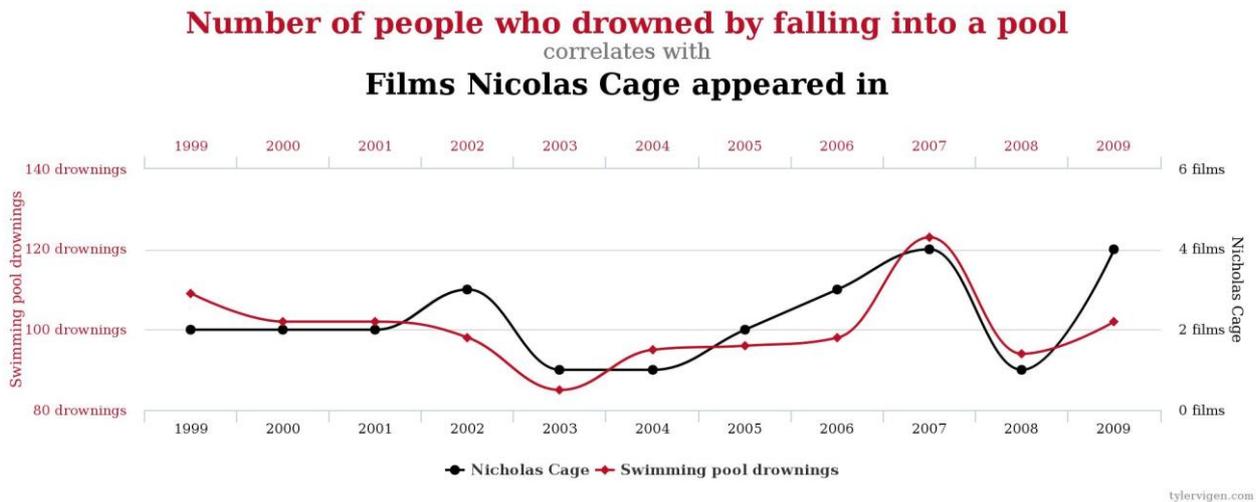
A fórmula para o cálculo do coeficiente de Pearson é:

$$r = \frac{COV_{xy}}{\sqrt{var_x var_y}}$$

Em que var_x e var_y representam a variância de X e a variância de Y, respectivamente.

d-) Por fim, execute o teste de hipóteses para essa correlação. Adote um nível confiança de 95% (significância de 05%). O que podemos concluir?

Exercício 03 (3 pontos)



O Gráfico acima apresenta o relacionamento entre as variáveis “número de pessoas afogadas ao cair em uma piscina” e “filmes em que Nicolas Cage apareceu”. O índice de correlação é de 0.66, indicando um relacionamento positivo moderado entre as variáveis. Um colega, também pesquisador, diz que quer empreender uma pesquisa acadêmica sobre esse relacionamento, uma vez que conseguiu boas evidências estatísticas de causalidade entre as variáveis.

Diante dessa situação, podemos perceber que o pesquisador está duplamente equivocado. Visando o aprimoramento de seu colega como pesquisador, quais críticas você faria? **Mínimo de 10 linhas.**

Exercício 04: Pós-Graduação (5 pontos)

Agresti e Finlay (2012) elaboram alguns passos a serem seguidos quando o objetivo é a construção de um modelo baseado em uma regressão linear. No *Moodle* está disponível o banco de dados que será utilizado neste exercício.

a-) O banco de dados apresenta a taxa de homicídio e o IDH para 58 distritos da cidade de São Paulo, amostrados aleatoriamente. Assuma como variável dependente o número de homicídios e, como variável independente, o IDH (a variável foi convertida de modo a variar entre 1 e 10). Estime a equação linear de previsão e interprete cada um dos coeficientes. Dica: utilize as fórmulas e substitua os valores.

b-) Construa um intervalo de confiança de 99% para o valor estimado de Beta e expresse as hipóteses nula e alternativa. Podemos rejeitar a hipótese nula? Dica: utilize as fórmulas e substitua os valores.

c-) Calcule a soma dos quadrados totais (SQT) e a soma dos quadrados do erro (SQE). Há diferença? O que podemos concluir? Dica: utilize as fórmulas.

d-) Calcule o r^2 . O que o seu valor nos informa sobre o poder preditivo do modelo e sobre a relação entre a taxa de homicídios e o IDH dos distritos da cidade de São Paulo? O valor encontrado é surpreendente? Qual é a relação entre valor obtido e o fato de se tratar de uma análise bivariada?

e-) Explique o que são variabilidade condicional e variabilidade marginal. Como a variabilidade condicional se relaciona ao modelo de regressão linear? Por que a variabilidade condicional tende a ser inferior à variabilidade marginal quando a relação entre as variáveis atende ao princípio da linearidade?

Dica: faça referência ao Teorema do Limite Central.

Boa lista!