

Universidade de São Paulo
Faculdade de Filosofia, Letras e Ciências Humanas
Departamento de Ciência Política
FLS 5028- Métodos Quantitativos e Técnicas de Pesquisa em Ciência Política
FLP 0406 - Métodos e Técnicas de Pesquisa em Ciência Política
1º Semestre/ 2018

Profº Dr. Glauco Peres da Silva

LISTA DE EXERCÍCIO 09

Data de entrega: 11/06/2018 (noturno) e 13/06/2018 (vespertino)

Exercício 01 (2 pontos)

Nas alternativas abaixo, marque “Verdadeiro” (V) ou “Falso” (F) e justifique as suas escolhas em, no máximo, 5 linhas. Quando marcar falso, justifique porque discorda da afirmação, e quando marcar verdadeiro, justifique o raciocínio no qual sua escolha se baseia.

() – Ao elaborar um *survey* a ser aplicado a uma amostra da população da cidade de São Paulo, o único viés com que a pesquisadora deve se preocupar é o viés (tendenciosidade) amostral, que está relacionado com a aleatoriedade e/ou a cobertura desta amostra.

() – Em pesquisas de n-grande nas Ciências Sociais utilizamos dados observacionais ou dados experimentais. Contudo, só podemos apontar relações de causalidade em pesquisas que utilizam dados experimentais.

() – Duas variáveis contínuas (X e Y), independentes entre si, possuem correlação igual a zero. Contudo, diante somente da informação de que a correlação entre duas variáveis (X e Y) é igual a zero, não podemos afirmar a inexistência de causalidade entre elas.

() – King, Keohane e Verba (1994) apontam dois pressupostos para a estimação de efeitos causais. O primeiro destes pressupostos (*unit homogeneity*) afirma que duas unidades são homogêneas se os valores esperados da variável dependente para cada uma delas são iguais quando o valor da variável explicativa é o mesmo. Uma versão menos rigorosa desse pressuposto (o pressuposto de efeitos constantes) assume que oscilações pequenas na variável explicativa não afetam a possibilidade de se obter valores esperados idênticos da variável dependente.

Exercício 02 (3,5 pontos)

Pesquisas quantitativas em Ciências Sociais estão primordialmente preocupadas em avaliar o efeito de uma variável explicativa (X) em outra dependente (Y). Dessa forma, a relação entre elas é comumente ilustrada como $X \rightarrow Y$. Contudo, conforme demonstram os textos obrigatórios e complementares deste tópico, a preocupação central destas pesquisas na atualidade é com o estabelecimento de relações causais. Nesse sentido, na quase totalidade dos casos, a relação ilustrada anteriormente não é suficiente para estabelecer relações causais em nossa “realidade multivariada”.

- a) **(0,5 ponto)** Tendo em vista o texto acima, explique porque a relação $X \rightarrow Y$ apresentada não é suficiente para estabelecer relações causais em pesquisas nas Ciências Sociais (**Tamanho máximo: 5 linhas**).
- b) **(2 pontos)** De que maneiras outras variáveis (Z) podem afetar a relação entre $X \rightarrow Y$? Dê ao menos um exemplo concreto para cada forma elencada na resposta (**Tamanho máximo: 20 linhas**)

[Dica: reformule a representação $X \rightarrow Y$, de forma a incluir as variáveis Z nestas ilustrações, explicando como Z está afetando a relação original]

- c) **(1 ponto)** O seguinte experimento será conduzido para avaliar o efeito da inclusão de uma questão na média dos resultados da prova de alunos de uma turma de Métodos do curso de Ciências Sociais da USP:

Os 300 alunos de uma turma de Métodos do curso de Ciências Sociais da USP realizarão uma prova sobre seus conhecimentos ao final do semestre. Desejamos avaliar o impacto da inclusão de uma nova questão na média padronizada das notas desta avaliação. O tempo de prova oferecido aos dois grupos é o mesmo. Partindo do pressuposto de que é possível classificar a ordem de dificuldade das questões, todas as questões da prova possuem o mesmo grau de dificuldade. Os alunos da turma são divididos em 2 grupos, com 150 indivíduos em cada. O primeiro grupo realizará uma prova de 5 questões e o outro grupo uma avaliação de 6 questões. A nota final da avaliação para os dois grupos será idêntica: 10 pontos. O efeito da inclusão de uma questão será avaliado pela diferença entre a nota média dos dois grupos.

Como podemos garantir, no desenho deste experimento, que outras variáveis Z não interferirão na estimação do efeito causal investigado? (**Tamanho máximo: 10 linhas**)

Exercício 03 (4,5 pontos)

Interpretar os resultados de estudos quantitativos é uma atividade comum para diversos cientistas sociais, seja na academia ou nos setores público, privado e terceiro. Suponha que você é um gestor de políticas públicas de saúde no nível federal e sua secretaria está muito preocupada com os índices de automedicação no Brasil. Você precisa decidir se sua secretaria fará ou não uma campanha em vídeo contra automedicação para os brasileiros adultos. Infelizmente, a secretaria responsável pelo planejamento do projeto, sem informá-lo anteriormente, já preparou o vídeo da campanha e encomendou quatro estudos de efetividade de campanha para quatro institutos de pesquisa diferentes, sob a alegação de que “quantos mais estudos, melhor”. Essas pesquisas (hipotéticas) já estão prontas. Em **todas** as pesquisas, a pergunta feita aos participantes foi “Você usaria o medicamento X sem anteriormente consultar um médico?”, que iremos chamar de **Pergunta I**. Eles podiam responder “Sim” ou “Não” e não houve indivíduos que deixaram de responder alguma das pesquisas. Responda as questões abaixo. Atente-se ao tamanho máximo das respostas, pois as linhas excedentes a ele **não** serão corrigidas.

- a) **(0,5 ponto)** Qual a população de interesse para sua secretaria? (**Tamanho máximo: 1 linha**).

A Tabela 1 mostra os resultados das quatro pesquisas (em um formato razoavelmente similar ao que geralmente usamos em publicações científicas). As duas primeiras linhas são as proporções dos indivíduos que responderam “Sim” à Pergunta I nos grupos de tratamento e controle, respectivamente. A diferença de médias se refere à diferença entre as médias (proporções) dos grupos de tratamento e de controle. O erro-padrão é o erro-padrão da diferença das médias. A estatística t é a divisão da diferença das médias pelo erro-padrão. IC 95% é o intervalo de 95% de confiança para a diferença das médias. O p-valor é referente ao teste da hipótese nula de não diferença de médias entre grupos de controle e tratamento, contra a hipótese alternativa de diferença de médias entre controle e tratamento. Finalmente, a última linha mostra o tamanho da amostra.

[Dica: O aluno pode revisar a seção “Robustez para violações da suposição de normalidade” em Agresti e Finlay, p. 146, para verificar as consequências de amostras não aleatórias e pequenas. Os resultados são apresentados no contexto de um intervalo de confiança de uma média, mas se mantêm, de maneira geral, para os testes de hipóteses de diferenças de média, à exceção do tamanho. Uma amostra de 15 observações pode ser ainda muito pequena para um teste de diferenças de médias. Outra fonte sobre testes de diferenças de médias que o aluno pode usar é a seção 7.4.2 de Kellstedt e Whitten (p. 177-182).]

Tabela 1

	Pesquisa A	Pesquisa B	Pesquisa C	Pesquisa D
Proporção (Tratamento)	0,12	0,32	0,34	0,30
Proporção (Controle)	0,15	0,70	0,41	0,36
Diferença de médias	-0,03	-0,38	-0,07	-0,06
Erro-padrão	0,02	0,10	0,02	0,02
Estatística t	-1,24	-3,68	-3,24	-2,56
P-valor	0,214	0,000	0,001	0,011
IC 95%	[-0,08; 0,02]	[-0,58; -0,18]	[-0,11; -0,03]	[-0,11; -0,01]
Tamanho da amostra	800	80	2.000	1.600

Pesquisa A. A consultoria responsável pela pesquisa A enviou um e-mail de convite para todos os quase 60.000 estudantes da Universidade de São Paulo e 5.000 voluntários se interessaram em participar. Então, a consultoria escolheu aleatoriamente 800 voluntários para comporem a amostra. Destes, 400 foram escolhidos aleatoriamente para o grupo de tratamento e assistiram o vídeo da campanha, ao passo que os outros 400 foram para o grupo de controle e assistiram um vídeo placebo, sem relação qualquer com temas de saúde. Após assistirem o vídeo, todos os indivíduos responderam a Pergunta I e essas respostas foram documentadas.

- b) **(1 ponto)** A diferença de médias indica que a pesquisa A estimou o efeito de redução de 3 pontos percentuais (na Tabela 1, -0,03), associado ao vídeo da campanha, na proporção de indivíduos que responderam que se automedicariam com o medicamento X. Sobre a pesquisa A, responda às seguintes perguntas. A diferença de médias é significativa ao nível de 95%? A amostra é grande? A amostra é representativa da população de interesse? O desenho de pesquisas dá motivos para supor que os grupos de tratamento e controle estão balanceados? Você acredita que o efeito estimado de menos 13 pontos percentuais é causal? Justifique as respostas brevemente. **(Tamanho máximo: 10 linhas)**.

Pesquisa B. A fundação contratada fez um grupo focal com 40 indivíduos, selecionados para a amostra por ordem de inscrição em um formulário postado publicamente na página do Facebook do Ministério da Saúde (MS). Todos os 40 indivíduos responderam a Pergunta I antes de assistirem o vídeo da campanha, respostas estas que a fundação considerou como compo do grupo de controle. Para construir o grupo de tratamento, a fundação registrou a resposta à Pergunta I após os indivíduos assistirem o vídeo. Note que o tamanho da amostra é 80, pois cada indivíduo foi contado duas vezes.

- c) **(1 ponto)** Sobre a pesquisa B, responda às seguintes perguntas. Qual o efeito estimado? O efeito estimado é significativo ao nível de 95%? A amostra é grande? A amostra é representativa da população de interesse? O desenho de pesquisas dá motivos para supor que os grupos de tratamento e controle estão balanceados? Você acredita que o efeito estimado é causal? Justifique as respostas brevemente. **(Tamanho máximo: 15 linhas)**.

Pesquisa C. O instituto de pesquisa responsável selecionou uma amostra aleatória estratificada em multiestágios que contemplava toda a população brasileira, com 2000 indivíduos, entrevistados em suas respectivas casas. O grupo de tratamento foi composto por 1.000 indivíduos que quiseram, voluntariamente, ver o vídeo da campanha. Os outros 1.000 indivíduos viram um vídeo placebo sem relação com saúde e compuseram o grupo de controle. Após assistirem os vídeos, todos os indivíduos responderam a Pergunta I.

- d) **(0,5 ponto)** Sobre a pesquisa C, responda às seguintes perguntas. Qual o efeito estimado? O efeito estimado é significativo ao nível de 95%? A amostra é grande? A amostra é representativa da população de interesse? O desenho de pesquisas dá motivos para supor que os grupos de tratamento e controle estão balanceados? Você acredita que o efeito estimado é causal? Justifique as respostas brevemente. **(Tamanho máximo: 10 linhas)**.

Pesquisa D. A consultoria contratada também utilizou uma amostra aleatória estratificada em multiestágios que contemplava toda a população brasileira. A amostra tinha 1600 indivíduos, divididos aleatoriamente em dois grupos de igual tamanho, tratamento e controle. O grupo de tratamento recebeu um panfleto que afirmava que a automedicação é muito perigosa (e **não** assistiu

o vídeo da campanha). O grupo de controle recebeu um panfleto placebo, sem relação com saúde. Após lerem os panfletos, ambos os grupos responderam a Pergunta I.

- e) **(0,5 ponto)** Sobre a pesquisa D, responda às seguintes perguntas. Qual o efeito estimado? O efeito estimado é significativo ao nível de 95%? A amostra é grande? A amostra é representativa da população de interesse? O desenho de pesquisas dá motivos para supor que os grupos de tratamento e controle estão balanceados? Você acredita que o efeito estimado é causal? Justifique as respostas brevemente. **(Tamanho máximo: 10 linhas).**
- f) **(1 ponto)** Você considera que a reputação das quatro consultorias é a mesma e sabe que as pesquisas foram executadas aproximadamente no mesmo período e usaram exatamente a mesma pergunta de pesquisa (Pergunta I), de maneira que não parece problemático comparar seus resultados, desde que se considere as particularidades de cada desenho de pesquisa. Além disso, você sabe que a campanha de vídeo planejada é bastante cara e, caso seja inócua, terá desperdiçado recursos que poderiam ter sido dirigidos a serviços essenciais aos cidadãos, recursos estes que serão retirados do orçamento da sua secretaria. *Com base nos resultados das quatro pesquisas e tendo em vista a população de interesse, você decide promover a campanha ou não? Justifique.* Obs.: Como o objetivo da questão é estimular a reflexão do aluno sobre interpretação de resultados, argumentos de textos sobre comunicação de políticas públicas e mesmos citações de materiais do curso não serão admitidas como corretas nessa questão. Reflita com base nos conhecimentos do curso e argumente. **(Tamanho máximo: 20 linhas).**

Exercício 04 (5 pontos): Apenas para a Pós-Graduação

Neste exercício, iremos estudar o desenho de pesquisa e alguns resultados do Working Paper “*Avis, E., Ferraz, C., & Finan, F. (2017). Do Government Audits Reduce Corruption?*”. Como o trabalho é extenso (82 páginas), selecionamos alguns excertos, que foram disponibilizados no Moodle. Leia a seleção disponibilizada para responder às questões propostas. Recomendamos que você não leia o trabalho inteiro para responder os exercícios, pois as questões foram pensadas especificamente para a seleção. Além disso, note que o paper fala em estimações estruturais e em formas reduzidas, mas abordaremos apenas os resultados em forma reduzida, que já podem ser lidos com os conhecimentos do curso até agora (ainda que você provavelmente não saiba os significados desses termos econométricos). Esperamos que esse exercício ajude os alunos que ainda não tiveram contato com métodos quantitativos a começar a ler resultados desse tipo de trabalho. Aproveite o exercício para se familiarizar com a análise de desenhos de pesquisa e dos resultados e construa respostas breves. Você só precisa justificar os pontos que explicitamente exigirem isso.

- a) **(1 ponto)** Qual a principal pergunta teórica da pesquisa? Expresse sua resposta como uma única relação monocausal entre conceitos (não entre as variáveis), como no exemplo a seguir: “maiores investimentos em infraestrutura causam crescimento econômico”. Os autores expressam essa pergunta de maneira determinística ou probabilística? Quais as

variáveis usadas para operacionalizar esses conceitos? Há uma relação clara entre medida e conceito? (**Tamanho máximo: 10 linhas**).

b) (2 pontos) Qual a população de interesse do estudo? Qual foi a amostra selecionada? Essa amostra é grande? A amostra é representativa da população de interesse? Por quê? Quais unidades formam o grupo de controle? Quais unidades foram o grupo de tratamento? (**Tamanho máximo: 15 linhas**).

c) (1 ponto) O desenho de pesquisa é observacional, experimental ou quase-experimental? Há aleatorização do tratamento? Se houver, os autores dão argumentos contextuais para que o leitor acredite que o tratamento foi aleatorizado? Se sim, cite dois argumentos. Os autores mostram evidências de estatística inferencial que indiquem que os grupos estão balanceados e que, portanto, a aleatorização parece ter sido bem feita (isto é, que não houve manipulação arbitrária da atribuição do tratamento)?

[Dica: Veja a tabela 1. Note que as notas de rodapé das tabelas trazem informações importantes. Os autores citam que fizeram controles por certos efeitos fixos. Por ora, você pode ler os resultados como testes de diferenças de médias (exatamente como fizemos no exercício 3), tendo em mente que houve o cuidado de controlar por outras variáveis. Aprenderemos mais sobre o que significa “controlar” até o fim do curso]. (**Tamanho máximo: 10 linhas**).

d) (1 ponto) Há evidências de um efeito do tratamento significativo para a pergunta principal de pesquisa? Quanto é esse efeito estimado? Se sim, os autores dão algum indicativo de que o efeito estimado é substantivo para a política pública? (Um efeito de 0,0001% no crescimento econômico, por exemplo, ainda que estatisticamente significativo, não tem muita importância substantiva). Você acredita que o efeito é causal?

[Dica: Veja a tabela 2. Como é normal na apresentação desse tipo de resultados, a primeira linha tem o efeito estimado para o tratamento e as outras têm os efeitos estimados dos controles, com os quais você não precisa se preocupar, por ora. Novamente, você pode ler a primeira linha como um teste de diferença de médias (sabendo que houve controles). Note que a Tabela 2 responde a pelo menos duas perguntas de pesquisa e que o enunciado deste exercício apenas te indaga algo apenas sobre a pergunta principal – supostamente a mesma que você respondeu no item 4-a.] (**Tamanho máximo: 10 linhas**).