

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 de 2019  
Profº. Dr. Glauco Peres da Silva

**LISTA DE EXERCÍCIOS 08**

**Aula 09 – “Inferência Estatística”**

**Data de entrega: 20/05/2019 (noturno) e 22/05/2019 (vespertino)**

**Nome:**

**Período: ( ) Vespertino, ( ) Noturno**

**Exercício 01**

Com base na bibliografia obrigatória indicada, marque “Verdadeiro” (V) ou “Falso” (F) para as afirmações a seguir e, quando falsas, justifique sua escolha em 5 linhas, no máximo. Cada questão tem o peso de 0,75 pontos.

A. (F) A interpretação do intervalo de confiança de 95% da média de uma variável qualquer é que 95% dos valores observados para aquela variável na população pertencerão ao intervalo de confiança.

**A interpretação do intervalo de confiança de 95% é que se fosse possível calcular a média amostral de cada uma de infinitas amostras aleatórias do mesmo tamanho, retiradas da mesma população, 95% dessas médias amostrais pertenceriam ao intervalo de confiança (AGRESTI & FINLAY, 2012, p. 139).**

B. (V) A amplitude dos intervalos de confiança aumenta quando se adotam níveis de confiança maiores e diminui com o aumento do tamanho das amostras.

C. (V) Amostras aleatórias grandes permitem a utilização da distribuição normal padrão em substituição à distribuição t para intervalos de confiança de médias.

D. (F) A violação do pressuposto da distribuição da população ser normal afeta a robustez da construção dos intervalos de confiança para uma média.

**A construção de intervalos de confiança para médias utilizando a distribuição t é robusta contra violações de normalidade da população, uma vez que, mesmo em casos em que a distribuição da população não seja normal, os intervalos de confiança baseados na distribuição t funcionam bem. Quanto maior a amostra, menor a importância da suposição de normalidade, seguindo o Teorema Central do Limite (AGRESTI & FINLAY, 2012, p. 146).**

E. (V) Um estimador é usado na inferência estatística para estimar um parâmetro populacional. A distribuição amostral de um bom estimador possui duas características: não tendenciosidade e eficiência.

F. (V) Em intervalos de confiança para proporções o tamanho da amostra é um fator importante a ser considerado. Para que o método de inferência seja válido é necessário que o n da amostra seja grande: 15 casos na categoria de interesse e também nas demais mobilizadas.

G. (F) Um exemplo de estimativa pontual é o intervalo de confiança – ao nível de significância de 95% – de uma proporção estimada.

**Uma estimativa pontual é uma avaliação de um parâmetro como um número único. O intervalo de confiança é um conjunto de números e, portanto, é uma estimativa intervalar (AGRESTI & FINLAY, 2012, p. 131).**

H. (V) Apesar de os intervalos de confiança utilizando a distribuição t serem robustos à violação de normalidade, eles não o são para a violação do pressuposto de aleatorização na produção dos dados.

## Exercício 02

*Obs.: descreva cada etapa das suas respostas, inclusive se utilizar o Excel ou algum programa estatístico. Utilize apenas o teste-Z.*

A tabela abaixo exhibe os resultados de uma pesquisa do Instituto de Pesquisas Datafolha referente ao nível de aprovação popular de diferentes instituições brasileiras em abril de 2019.

Avaliação / Instituições	Presidência da República		Congresso Nacional (Câmara dos Deputados e Senado Federal)		Supremo Tribunal Federal (STF)		Forças Armadas		Partidos políticos	
	%	n	%	n	%	n	%	n	%	n

<b>Confia muito</b>	29%	605	8%	167	18%	375	46%	960	5%	104
<b>Confia um pouco</b>	41%	855	49%	1022	48%	1001	36%	751	40%	834
<b>Não confia</b>	30%	626	43%	897	34%	709	18%	375	55%	1147
<b>Total</b>	100%	2086	100%	2086	100%	2086	100%	2086	100%	2086

Obs.: A tabela foi adaptada para o exercício. Retiramos “não sabe” e arredondamos alguns valores.

Disponível em:

<http://media.folha.uol.com.br/datafolha/2019/04/15/e4dfasfas453434vfa423vavsxfd429b35922gci.pdf>

Com base nos dados apresentados, faça as questões A, B e C.

A. Admitindo-se que as amostras sejam aleatórias, calcule a média de aprovação de cada instituição. Em seguida, monte uma tabela com as médias em ordem decrescente. (1,0 ponto)

*Orientação: ao transformar as variáveis categóricas em numéricas, considere “confia muito” = 10, “confia um pouco” = 5, “não confia” = 0.*

**RESPOSTA:**

**Primeiramente calculamos a média ponderada:**

Avaliação / Instituições	Transformação em variáveis numéricas (notas)	Presidência da República			Congresso Nacional (Câmara dos Deputados e Senado Federal)			Supremo Tribunal Federal (STF)			Forças Armadas			Partidos políticos		
		%	n	Ponderação (% * pesos)	%	n	Ponderação (% * peso)	%	n	Ponderação (% * peso)	%	n	Ponderação (% * peso)	%	n	Ponderação (% * peso)
<b>Confia muito</b>	10	29%	605	2,90	8%	167	0,80	18%	376	1,80	46%	960	4,60	5%	104	0,50
<b>Confia um pouco</b>	5	41%	855	2,05	49%	1022	2,45	48%	1001	2,40	36%	751	1,80	40%	835	2,00
<b>Não confia</b>	0	30%	626	0,00	43%	897	0,00	34%	709	0,00	18%	375	0,00	55%	1147	0,00
<b>Total</b>	-	100%	2086	<b>4,95</b>	100%	2086	<b>3,25</b>	100%	2086	<b>4,20</b>	100%	2086	<b>6,40</b>	100%	2086	<b>2,50</b>

Em seguida, filtramos apenas pela média das amostras e ordenamos os valores em outra tabela.

Instituição	Média
Forças Armadas	6,40
Presidência da República	4,95
Supremo Tribunal Federal (STF)	4,20
Congresso Nacional	3,25
Partidos políticos	2,50

B. Calcule o erro padrão e o intervalo de confiança das médias da Presidência da República e do STF aos níveis de 95% e 99%. Em algum dos níveis os intervalos das médias amostrais da Presidência e do STF se interpolam? (1,0 ponto)

*Dica: ao efetuar os cálculos, não esqueça de multiplicar o desvio<sup>2</sup> pelo número de casos.*

**RESPOSTA:**

**Inicialmente calculamos os valores da Presidência.**

Presidência da República						
Nota	%	n	Ponderação	Desvio (nota - 4,95)	Desvio <sup>2</sup>	Desvio <sup>2</sup> * número de casos
10	0,29	605	2,9	5,05	25,5025	15427,48235
5	0,41	855	2,05	0,05	0,0025	2,13815
0	0,3	626	0	-4,95	24,5025	15333,6645
Soma dos quadrados dos desvios (SQD)						30763,285
$s^2 = \text{Variância amostral (SQD)/(n-1)}$						14,75457314
$s = \text{Desvio-padrão amostral (Raiz de } s^2)$						3,8411682
Erro-padrão da média ( $s/\sqrt{n}$ )						0,08410197077

**INTERVALOS DA MÉDIA AMOSTRAL DA PRESIDÊNCIA**

$IC\alpha = \bar{y} \pm z\alpha s \sqrt{n}$ , em que  $\alpha$  é o nível de confiança

$IC95\% = 4,95 \pm 1,96 * 0,084 = [4,79 ; 5,11]$

$IC99\% = 4,95 \pm 2,576 * 0,084 = [4,73 ; 5,17]$

Em seguida, procedemos aos valores do STF.

Supremo Tribunal Federal (STF)						
Nota	%	n	Ponderação (% * peso)	Desvio (nota - 4,20)	Desvio <sup>2</sup>	Desvio <sup>2</sup> * número de casos
10	0,18	376	1,8	5,8	33,64	12648,64
5	0,48	1001	2,4	0,8	0,64	640,64
0	0,34	709	0	-4,2	17,64	12506,76
Soma dos quadrados dos desvios (SQD)						25796,04
$s^2 = \text{Variância amostral (SQD)/(n-1)}$						12,37220144
$s = \text{Desvio-padrão amostral (Raiz de } s^2)$						3,517414027
Erro-padrão da média ( $s/\sqrt{n}$ )						0,0770134074 5

### INTERVALOS DA MÉDIA AMOSTRAL DO STF

$IC\alpha = \bar{y} \pm z\alpha s \sqrt{n}$ , em que  $\alpha$  é o nível de confiança

$IC95\% = 4,20 \pm 1,96 * 0,077 = [4,05; 4,35]$

$IC99\% = 4,20 \pm 2,576 * 0,077 = [4,00; 4,40]$

Mesmo nos níveis de confiança mais altos, não há interpolação entre as médias da Presidência e do STF.

C. Quais características populacionais o Instituto Datafolha incorporou à amostragem por estratificação? (1,0 ponto)

### RESPOSTA:

O Instituto Datafolha incorporou as seguintes características populacionais à amostragem: sexo, idade, média de escolaridade, partido, região, natureza do município, porte do município, religião, ocupação principal, cor e renda familiar mensal.

D. Uma pesquisa divulgada pelo instituto Levada-Center em abril de 2019 evidenciou a seguinte percepção da população russa em relação à figura de Joseph Stalin:

Faixa etária	18-24	25-39	40-54	55-	Total
<i>n</i>	147	498	412	582	1.638
<b>Pergunta: no geral, qual é a sua atitude em relação a Stalin? (percentuais da população)</b>					
Admiração	5	6	5	3	4
Respeito	27	39	43	47	41
Simpatia	6	6	7	6	6
Indiferença	39	29	25	22	26
Antipatia	6	6	6	7	6
Medo	4	3	5	6	5
Ódio	3	2	2	3	3
Não sei quem é Stalin	2	2	<1	<1	1
Não respondeu	8	7	8	7	7

Disponível em: <https://www.levada.ru/en/2019/04/19/dynamic-of-stalin-s-perception/>

Ignorando os que responderam “indiferença”, “não sei quem é Stalin” ou não responderam e admitindo-se que a amostragem tenha sido aleatória, crie uma tabela dicotômica que aponte apenas atitude “positiva” e “negativa” por faixa etária. Em seguida, calcule os intervalos de confiança ao nível de 95% apenas para atitude “positiva”. Levando em consideração esses intervalos, qual faixa da amostra tem a atitude mais positiva? Os dados revelados no Brasil e na Rússia apresentam alguma semelhança no que concerne aos valores e instituições democráticas? Explique. (1,0 ponto)

**RESPOSTA:**

<b>Faixa etária</b>	<b>18-24</b>	<b>25-39</b>	<b>40-54</b>	<b>55-</b>	<b>Total</b>
<b><i>n</i></b>	<b>147</b>	<b>498</b>	<b>412</b>	<b>582</b>	<b>1.638</b>
<b>Pergunta: no geral, qual é a sua atitude em relação à Stalin? (percentuais da população)</b>					
<b>Positiva</b>	<b>38</b>	<b>51</b>	<b>55</b>	<b>56</b>	<b>51</b>
<b>Negativa</b>	<b>13</b>	<b>11</b>	<b>13</b>	<b>16</b>	<b>14</b>

Seguimos na resolução a notação de Agresti e Finlay.

**FAIXA DE 18 A 24 ANOS**

**$n = 147$**

**Positiva = 0,38**

**$ep = \sqrt{\pi^{\wedge}(1 - \pi^{\wedge})n}$**

**$ep = \sqrt{0,38(1 - 0,38)147}$**

**$ep = 0,040$**

$IC\alpha = \bar{y} \pm z\alpha s \sqrt{n}$ , em que  $\alpha$  é o nível de confiança

$IC95\% = 0,38 \pm (1,96*0,040) = [0,3015 ; 0,4585]$  ou  $[30,15\% ; 45,85\%]$ .

#### FAIXA DE 25 A 39 ANOS

$n = 498$

Positiva = 0,51

$$ep = \sqrt{\pi^{\wedge}(1 - \pi^{\wedge})n}$$

$$ep = \sqrt{0,51(1 - 0,51)498}$$

$ep = 0,022$

$IC\alpha = \bar{y} \pm z\alpha s \sqrt{n}$ , em que  $\alpha$  é o nível de confiança

$IC95\% = 0,51 \pm (1,96*0,022) = [0,4661 ; 0,5539]$  ou  $[46,61\% ; 55,39\%]$

#### FAIXA DE 40 A 54 ANOS

$n = 412$

Positiva = 0,55

$$ep = \sqrt{\pi^{\wedge}(1 - \pi^{\wedge})n}$$

$$ep = \sqrt{0,55(1 - 0,55)412}$$

$ep = 0,025$

$IC\alpha = \bar{y} \pm z\alpha s \sqrt{n}$ , em que  $\alpha$  é o nível de confiança

$IC95\% = 0,55 \pm (1,96*0,025) = [0,5019 ; 0,5980]$  ou  $[50,19\% ; 59,80\%]$ .

#### FAIXA DE 55 OU MAIS ANOS

$n = 582$

Positiva = 0,56

$$ep = \sqrt{\pi^{\wedge}(1 - \pi^{\wedge})n}$$

$$ep = \sqrt{0,56(1 - 0,56)582}$$

$ep = 0,020$

$IC95\% = 0,56 \pm (1,96*0,025) = [0,5197 ; 0,6003]$  ou  $[51,97\% ; 60,03\%]$ .

FAIXAS ETÁRIAS	INTERVALO DE 95%
18 A 24 ANOS	30,15% ; 45,85%
25 A 39 ANOS	46,61% ; 55,39%
40 A 54 ANOS	50,19% ; 59,80%
55 OU MAIS ANOS	51,97% ; 60,03%

Uma vez que há interpolação entre os intervalos das três últimas faixas, não podemos afirmar com confiança que uma delas apresenta atitude mais positiva que as demais, especialmente entre as duas últimas. A resposta correta, portanto, é “25 A 39 ANOS”, “40 A 54 ANOS” e “55 OU MAIS ANOS”; ou ainda “40 A 54 ANOS” e “55 OU MAIS ANOS”.

Tanto os baixos índices de aprovação de instituições representativas, como o Congresso e os partidos políticos, no Brasil, quanto a atitude positiva na Rússia em relação a um dos maiores

**líderes totalitários do século XX, dão indícios de uma crise na percepção da população quanto aos valores e instituições democráticas.**

**[Outras interpretações são possíveis.]**