

Gabarito - Lista 7

Inferência: Comparação de Grupos

Gabarito

Exercício 1 [2,5 pontos]. De acordo com Agresti e Finlay, podemos reduzir o teste de duas amostras pareadas a um teste de uma amostra devido ao seguinte resultado: $\mu_d = \mu_2 - \mu_1$

- (a) Explique o que são μ_d , μ_2 e μ_1
- (b) Dadas as hipóteses $H_0 : \mu_2 - \mu_1 = 0$ e $H_0 : \mu_1 = \mu_2$, identifique a que trata de uma amostra e a que trata de duas amostras. Justifique sua resposta.
- (c) Ao reduzirmos o teste de duas amostras pareadas a uma amostra, por que é importante saber que “a diferença entre a média de dois grupos é igual média da diferença dos valores”?

(a) μ_1 e μ_2 são as médias das duas amostras que estão sendo comparadas. μ_d é a média da diferença entre essas variáveis.

(b) $H_0 : \mu_1 = \mu_2$ é a hipótese que compara as médias de duas variáveis, relativamente às amostras 1 e 2. A distinção importante aqui é que, para um teste de duas variáveis, o erro-padrão é uma composição dos erros das duas variáveis originais. Por isso, testes de amostras independentes geram erros-padrões maiores do que testes de amostras pareadas.

$H_0 : \mu_2 - \mu_1 = 0$ é um teste para apenas uma amostra: aquela que diz respeito à diferença entre as duas variáveis. Como temos apenas uma amostra, o cálculo do erro-padrão é o mesmo que vimos para uma variável qualquer. Assim, testes para amostras pareadas geram erros-padrões menores do que aqueles para amostras independentes.

(c) De acordo com a exposição de Agresti na seção 7.4, esse é o resultado que nos permite fazer o cálculo do erro-padrão para amostras pareadas como se elas fossem apenas uma variável.

Exercício 2 [2,5 pontos]. Uma das consultas feitas no questionário da pesquisa “Brasil, as Américas e o Mundo” em 2011 foi a avaliação de diversos líderes latino-americanos. Os respondentes deveriam responder à seguinte pergunta:

(37) Agora vou lhe pedir sua opinião sobre alguns líderes políticos. Você pode dar uma nota de 0 a 100, sendo que 100 é a nota mais favorável sobre esse líder na sua opinião. Se você não tem opinião sobre esse líder ou nunca ouviu falar sobre ele, por favor, me diga.

	Nota	NS	NR	Nunca ouviu falar
Cristina Kirchner	_____	998	999	997
Raúl Castro	_____	998	999	997
Luís Inácio "Lula" da Silva	_____	998	999	997
Michelle Bachelet	_____	998	999	997
Hugo Chávez	_____	998	999	997
George Bush	_____	998	999	997
Alvaro Colon	_____	998	999	997
Alan García	_____	998	999	997
Evo Morales	_____	998	999	997
Miguel Ángel López	_____	998	999	997
Alvaro Uribe	_____	998	999	997
Felipe Calderón	_____	998	999	997
José Luis Rodríguez Zapatero	_____	998	999	997

Um dos testes possíveis de serem feitos com essa pergunta apresenta os seguintes resultados:

. ttest Kirchner == Chavez, unpaired						
Two-sample t test with equal variances						
Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Kirchner	647	41.91963	1.121729	28.53253	39.71695	44.1223
Chavez	1188	25.99832	.7820407	26.95489	24.46398	27.53265
combined	1835	31.61199	.6663985	28.54645	30.30501	32.91897
diff		15.92131	1.3447		13.28401	18.55862
diff = mean(Kirchner) - mean(Chavez)				t =	11.8401	
Ho: diff = 0				degrees of freedom =	1833	
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0		
Pr(T < t) = 1.0000		Pr(T > t) = 0.0000		Pr(T > t) = 0.0000		

- Indique que hipótese está sendo testada acima
- Identifique, para cada hipótese alternativa, se o teste é bicaudal ou monocaudal
- Interprete os resultados desse teste, articulando a estatística-teste, os p-valores e a decisão sobre a hipótese nula

- $H_0: \mu_{Kirchner} = \mu_{Chavez}$
 $H_a: \mu_{Kirchner} < \mu_{Chavez}$
 $H_a: \mu_{Kirchner} \neq \mu_{Chavez}$
 $H_a: \mu_{Kirchner} > \mu_{Chavez}$

- Exercício 3 [2,5 pontos].** Os testes abaixo indicam a avaliação de Cristina Kirchner e Hugo Chávez agregadas por gênero. Comente os resultados (não se esqueça de indicar as hipóteses nula e alternativa, o nível de significância, o p-valor correspondente ao teste, a decisão sobre rejeição ou não, e uma conclusão) e indique se eles correspondem à sua expectativa

```
. ttest kirchner, by(sexo) unequal
```

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Masculin	364	42.82143	1.505526	28.7236	39.86078	45.78208
Feminino	283	40.75972	1.681861	28.29328	37.44912	44.07031
combined	647	41.91963	1.121729	28.53253	39.71695	44.1223
diff		2.061711	2.257269		-2.371243	6.494666

diff = mean(Masculin) - mean(Feminino) t = 0.9134
Ho: diff = 0 Satterthwaite's degrees of freedom = 610.486

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 0.8193 Pr(|T| > |t|) = 0.3614 Pr(T > t) = 0.1807

. ttest Chavez, by(sexo) unequal						
Two-sample t test with unequal variances						
Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Masculin	655	26.49008	1.078783	27.60927	24.37178	28.60837
Feminino	533	25.394	1.132311	26.14143	23.16965	27.61835
combined	1188	25.99832	.7820407	26.95489	24.46398	27.53265
diff		1.09608	1.563938		-1.972386	4.164546
diff = mean(Masculin) - mean(Feminino) t = 0.7008						
Ho: diff = 0 Satterthwaite's degrees of freedom = 1159.2						
Ha: diff < 0 Ha: diff != 0 Ha: diff > 0						
Pr(T < t) = 0.7582 Pr(T > t) = 0.4835 Pr(T > t) = 0.2418						

As hipóteses sendo testadas são:

- $H_0: \mu_{Homens} = \mu_{Mulheres}$
- $H_a: \mu_{Homens} < \mu_{Mulheres}$
- $H_a: \mu_{Homens} \neq \mu_{Mulheres}$
- $H_a: \mu_{Homens} > \mu_{Mulheres}$

Os resultados demonstram que a avaliação da opinião pública brasileira sobre a Cristina Kirschner é estatisticamente indistinguível para homens e mulheres. Os resultados são confirmados pela análise da estatística-teste (que é menor do que o valor crítico a 90% de significância), do intervalo de confiança (cujos valores se cruzam) e dos p-valores (que são menores do que α).

Exercício 4 [2,5 pontos]. A pergunta abaixo diz respeito à avaliação dos entrevistados sobre diversos atores internacionais. Comente os resultados abaixo, tanto do ponto de vista estatístico quanto do ponto de vista substantivo.

(44) E na mesma escala de 0 a 100, sendo 0 uma opinião muito desfavorável, 100 uma opinião muito favorável e 50 uma opinião nem favorável nem desfavorável, qual seria sua opinião sobre as seguintes organizações internacionais? Se você não tem opinião a respeito ou nunca ouviu falar dessa organização, por favor, me diga. (MOSTRAR CARTÃO) (44_7) 61_7				
	Nota	NS	NR	Nunca ouviu falar
A Organização das Nações Unidas (ONU)		998	999	9999
As empresas multinacionais		998	999	9999
O Fundo Monetário Internacional		998	999	9999
As organizações não governamentais internacionais		998	999	9999
A Organização dos Estados Americanos (OEA)		998	999	9999
O Tratado de Livre Comércio da América do Norte (TLCAN)		998	999	9999
O MERCOSUL (Mercado Comum do Sul)		998	999	9999

. ttest Multis == ONGs , unpaired						
Two-sample t test with equal variances						
Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Multis	1618	51.99876	.7339051	29.52087	50.55926	53.43827
ONGs	1369	47.1008	.7993387	29.57553	45.53274	48.66887
combined	2987	49.75393	.5423565	29.64165	48.6905	50.81736
diff		4.89796	1.084987		2.770563	7.025358
diff = mean(Multis) - mean(ONGs)				t =	4.5143	
Ho: diff = 0				degrees of freedom =	2985	
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0		
Pr(T < t) = 1.0000		Pr(T > t) = 0.0000		Pr(T > t) = 0.0000		

As hipóteses sendo testadas são:

- $H_0 : \mu_{Multinacionais} = \mu_{ONGs}$
- $H_a : \mu_{Multinacionais} < \mu_{ONGs}$
- $H_a : \mu_{Multinacionais} \neq \mu_{ONGs}$
- $H_a : \mu_{Multinacionais} > \mu_{ONGs}$

Os resultados demonstram que a avaliação dos brasileiros sobre as empresas multinacionais é mais positiva do que a avaliação sobre ONGs internacionais. O teste de $H_a : \mu_{Multinacionais} > \mu_{ONGs}$ demonstra que esse resultado é estatisticamente significativo com um α de 0,01, mas a avaliação substantiva dos resultados indica que, na prática, essa diferença é muito baixa. É importante não confundir significância estatística com significância substantiva: resultados podem ser estatisticamente significantes puramente por conta de erros-padrões baixos.